Graduate Texts in Mathematics

John M. Lee

Introduction to Smooth Manifolds

Springer

To my students

Preface

Manifolds are everywhere. These generalizations of curves and surfaces to arbitrarily many dimensions provide the mathematical context for understanding "space" in all of its manifestations. Today, the tools of manifold theory are indispensable in most major subfields of pure mathematics, and outside of pure mathematics they are becoming increasingly important to scientists in such diverse fields as genetics, robotics, econometrics, computer graphics, biomedical imaging, and, of course, the undisputed leader among consumers (and inspirers) of mathematics—theoretical physics. No longer a specialized subject that is studied only by differential geometers, manifold theory is now one of the basic skills that all mathematics students should acquire as early as possible.

Over the past few centuries, mathematicians have developed a wondrous collection of conceptual machines designed to enable us to peer ever more deeply into the invisible world of geometry in higher dimensions. Once their operation is mastered, these powerful machines enable us to think geometrically about the 6-dimensional zero set of a polynomial in four complex variables, or the 10-dimensional manifold of 5×5 orthogonal matrices, as easily as we think about the familiar 2-dimensional sphere in $\mathbb{R}^3$. The price we pay for this power, however, is that the machines are built out of layer upon layer of abstract structure. Starting with the familiar raw materials of Euclidean spaces, linear algebra, and multivariable calculus, one must progress through topological spaces, smooth atlases, tangent bundles, cotangent bundles, immersed and embedded submanifolds, tensors, Riemannian metrics, differential forms, vector fields, flows, foliations, Lie derivatives, Lie groups, Lie algebras, and more—just to get to the point where one can even think about studying specialized applications of manifold theory such as gauge theory or symplectic topology.

This book is designed as a first-year graduate text on manifold theory, for students who already have a solid acquaintance with general topology, the fundamental group, and covering spaces, as well as basic undergraduate linear algebra and real analysis. The book is similar in philosophy and scope to the first volume of Spivak's classic text [Spi79], though perhaps a bit more dense. I have tried neither to write an encyclopedic introduction to manifold theory in its utmost generality, nor to write a simplified introduction that gives students a "feel" for the subject without the struggle that is required to master the tools. Instead, I have tried to find a middle path by introducing and using all of the standard tools of manifold theory, and proving all of its fundamental theorems, while avoiding unnecessary generalization or specialization. I try to keep the approach as concrete as possible, with pictures and intuitive discussions of how one should think geometrically about the abstract concepts, but without shying away from the powerful tools that modern mathematics has to offer. To fit in all of the basics and still maintain a reasonably sane pace, I have had to omit a number of important topics entirely, such as complex manifolds, infinite-dimensional manifolds, connections, geodesics, curvature, fiber bundles, sheaves, characteristic classes, and Hodge theory. Think of them as dessert, to be savored after completing this book as the main course.

The goal of my choice of topics is to cover those portions of smooth manifold theory that most people who will go on to use manifolds in mathematical or scientific research will need. To convey the book's compass, it is easiest to describe where it starts and where it ends.

The starting line is drawn just after topology: I assume that the reader has had a rigorous course in topology at the beginning graduate or advanced undergraduate level, including a treatment of the fundamental group and covering spaces. One convenient source for this material is my *Introduction to Topological Manifolds* [Lee00], which I wrote two years ago precisely with the intention of providing the necessary foundation for this book. There are other books that cover similar material well; I am especially fond of Sieradski's *An Introduction to Topology and Homotopy* [Sie92] and the new edition of Munkres's *Topology* [Mun00].

The finish line is drawn just after a broad and solid background has been established, but before getting into the more specialized aspects of any particular subject. For example, I introduce Riemannian metrics, but I do not go into connections or curvature. There are many Riemannian geometry books for the interested student to take up next, including one that I wrote five years ago [Lee97] with the goal of moving expediently in a one-quarter course from basic smooth manifold theory to some nontrivial geometric theorems about curvature and topology. For more ambitious readers, I recommend the beautiful recent books by Petersen [Pet98], Sharpe [Sha97], and Chavel [Cha93].

This subject is often called “differential geometry.” I have deliberately avoided using that term to describe what this book is about, however, because the term applies more properly to the study of smooth manifolds endowed with some extra structure—such as Lie groups, Riemannian manifolds, symplectic manifolds, vector bundles, foliations—and of their properties that are invariant under structure-preserving maps. Although I do give all of these geometric structures their due (after all, smooth manifold theory is pretty sterile without some geometric applications), I felt that it was more honest not to suggest that the book is primarily about one or all of these geometries. Instead, it is about developing the general tools for working with smooth manifolds, so that the reader can go on to work in whatever field of differential geometry or its cousins he or she feels drawn to.

One way in which this emphasis makes itself felt is in the organization of the book. Instead of gathering the material about a geometric structure together in one place, I visit each structure repeatedly, each time delving as deeply as is practical with the tools that have been developed so far. Thus, for example, there are no chapters whose main subjects are Riemannian manifolds or symplectic manifolds. Instead, Riemannian metrics are introduced in Chapter 11 right after tensors; they then return to play major supporting roles in the chapters on orientations and integration, followed by cameo appearances in the chapters on de Rham cohomology and Lie derivatives. Similarly, symplectic structures make their first appearance at the end of the chapter on differential forms, and can be seen lurking in an occasional problem or two for a while, until they come into prominence at the end of the chapter on Lie derivatives. To be sure, there are two chapters (9 and 20) whose sole subject matter is Lie groups and/or Lie algebras, but my goals in these chapters are less to give a comprehensive introduction to Lie theory than to develop some of the more general tools that everyone who studies manifolds needs to use, and to demonstrate some of the amazing things one can do with those tools.

The book is organized roughly as follows. The twenty chapters fall into four major sections, characterized by the kinds of tools that are used.

The first major section comprises Chapters 1 through 6. In these chapters I develop as much of the theory of smooth manifolds as one can do using, essentially, only the tools of topology, linear algebra, and advanced calculus. I say “essentially” because, as the reader will soon find out, there are a great many definitions here that will be unfamiliar to most readers and will make the material seem very new. The reader’s main job in these first six chapters is to absorb all the definitions and learn to think about familiar objects in new ways. It is the bane of this subject that there are so many definitions that must be piled on top of one another before anything interesting can be said, much less proved. I have tried, nonetheless, to bring in significant applications as early and as often as possible. By the end of these six chapters, the reader will have been introduced to topological manifolds, smooth manifolds, the tangent and cotangent bundles, and abstract vector bundles.

The next major section comprises Chapters 7 through 10. Here the main tools are the inverse function theorem and its corollaries. This is the first of four foundational theorems on which all of smooth manifold theory rests. It is applied primarily to the study of submanifolds (including Lie subgroups and vector subbundles), quotients of manifolds by group actions, embeddings of smooth manifolds into Euclidean spaces, and approximation of continuous maps by smooth ones.

The third major section, consisting of Chapters 11 through 16, uses tensors and tensor fields as its primary tools. Beginning with the definition (or, rather, two different definitions) of tensors, I introduce Riemannian metrics, differential forms, integration, Stokes's theorem (the second of the four foundational theorems), and de Rham cohomology. The section culminates in the de Rham theorem, which relates differential forms on a smooth manifold to its topology via its singular cohomology groups.

The last major section, Chapters 17 through 20, explores the circle of ideas surrounding integral curves and flows of vector fields, which are the smooth-manifold version of systems of ordinary differential equations. The main tool here is the fundamental theorem on flows, the third foundational theorem. It is a consequence of the basic existence, uniqueness, and smoothness theorem for ordinary differential equations. Both of these theorems are proved in Chapter 17. Flows are used to define Lie derivatives and describe some of their applications (most notably to symplectic geometry), to study tangent distributions and foliations, and to explore in some detail the relationship between Lie groups and their Lie algebras. Along the way, we meet the fourth foundational theorem, the Frobenius theorem, which is essentially a corollary of the inverse function theorem and the fundamental theorem on flows.

The Appendix (which most readers should read, or at least skim, first) contains a cursory summary of the prerequisite material on topology, linear algebra, and calculus that is used throughout the book. Although no student who has not seen this material before is going to learn it from reading the Appendix, I like having all of the background material collected in one place. Besides giving me a convenient way to refer to results that I want to assume as known, it also gives the reader a splendid opportunity to brush up on topics that were once (hopefully) well understood but may have faded a bit.

I should say something about my choices of conventions and notations. The old joke that "differential geometry is the study of properties that are invariant under change of notation" is funny primarily because it is alarmingly close to the truth. Every geometer has his or her favorite system of notation, and while the systems are all in some sense formally isomorphic, the transformations required to get from one to another are often not at all obvious to the student. Because one of my central goals is to prepare students to read advanced texts and research articles in differential geometry, I have tried to choose notation and conventions that are as close to the mainstream as I can make them without sacrificing too much internal consistency. When there are multiple conventions or notations in common use (such as the two common conventions for the wedge product or the Laplace operator), I explain what the alternatives are and alert the student to be aware of which convention is in use by any given writer. Striving for too much consistency in this subject can be a mistake, however, and I have eschewed absolute consistency whenever I felt it would get in the way of ease of understanding. I have also introduced some common shortcuts at an early stage, such as the Einstein summation convention and the systematic confounding of maps with their coordinate representations, both of which tend to drive students crazy at first, but pay off enormously in efficiency later.

This book has a rather large number of exercises and problems for the student to work out. Embedded in the text of each chapter are questions labeled as “exercises.” These are (mostly) short opportunities to fill in the gaps in the text. Many of them are routine verifications that would be tedious to write out in full, but are not quite trivial enough to warrant tossing off as obvious. I hope that conscientious readers will take the time at least to stop and convince themselves that they fully understand what is involved in doing each exercise, if not to write out a complete solution, because it will make their reading of the text far more fruitful. At the end of each chapter is a collection of (mostly) longer and harder questions labeled as “problems.” These are the ones from which I select written homework assignments when I teach this material, and many of them will take hours for students to work through. It is really only in doing these problems that one can hope to absorb this material deeply. I have tried insofar as possible to choose problems that are enlightening in some way and have interesting consequences in their own right. The results of many of them are used in the text.

I welcome corrections or suggestions from readers. I plan to keep an up-to-date list of corrections on my Web site, *www.math.washington.edu/~lee*. If that site becomes unavailable for any reason, the publisher will know where to find me.

Happy reading!

Acknowledgments. There are many people who have contributed to the development of this book in indispensable ways. I would like to mention especially Judith Arms and Tom Duchamp, both of whom generously shared their own notes and ideas about teaching this subject; Jim Isenberg and Steve Mitchell, who had the courage to teach from early drafts of this book, and who have provided spectacularly helpful suggestions for improvement; and Gary Sandine, who found a draft on the Web, and not only read it with incredible thoroughness and made more helpful suggestions than any-

one else, but also created more than a third of the illustrations in the book, with no compensation other than the satisfaction of contributing to our communal quest for knowledge while gaining a deeper understanding for himself. In addition, I would like to thank the many other people who read the draft and sent their corrections and suggestions to me, especially Jaejeong Lee. (In the Internet age, textbook writing becomes ever a more collaborative venture.) Most of all, I would like to thank all of my students past, present, and future, to whom this book is dedicated. It is a cliché in the mathematical community that the only way to really learn a subject is to teach it; but I have come to appreciate much more deeply over the years how much feedback from students shapes and hones not only my teaching and my writing, but also my very understanding of what mathematics is all about. This book could not have come into being without them.

Finally, I am deeply indebted to my beloved family—Pm, Nathan, and Jeremy—who once again have endured my preoccupation and extended absences with generosity and grace. This time I plan to thank them by not writing a book for a while.

John M. Lee
Seattle, Washington
July, 2002

Contents

1
Smooth Manifolds

This book is about smooth manifolds. In the simplest terms, these are spaces that locally look like some Euclidean space $\mathbb{R}^n$, and on which one can do calculus. The most familiar examples, aside from Euclidean spaces themselves, are smooth plane curves such as circles and parabolas, and smooth surfaces such as spheres, tori, paraboloids, ellipsoids, and hyperboloids. Higher-dimensional examples include the set of unit vectors in $\mathbb{R}^{n+1}$ (the n-sphere) and graphs of smooth maps between Euclidean spaces.

The simplest examples of manifolds are the topological manifolds, which are topological spaces with certain properties that encode what we mean when we say that they "locally look like" $\mathbb{R}^n$. Such spaces are studied intensively by topologists.

However, many (perhaps most) important applications of manifolds involve calculus. For example, most applications of manifold theory to geometry involve the study of such properties as volume and curvature. Typically, volumes are computed by integration, and curvatures are computed by formulas involving second derivatives, so to extend these ideas to manifolds would require some means of making sense of differentiation and integration on a manifold. The applications of manifold theory to classical mechanics involve solving systems of ordinary differential equations on manifolds, and the applications to general relativity (the theory of gravitation) involve solving a system of partial differential equations.

The first requirement for transferring the ideas of calculus to manifolds is some notion of "smoothness." For the simple examples of manifolds we described above, all of which are subsets of Euclidean spaces, it is fairly easy to describe the meaning of smoothness on an intuitive level. For ex-

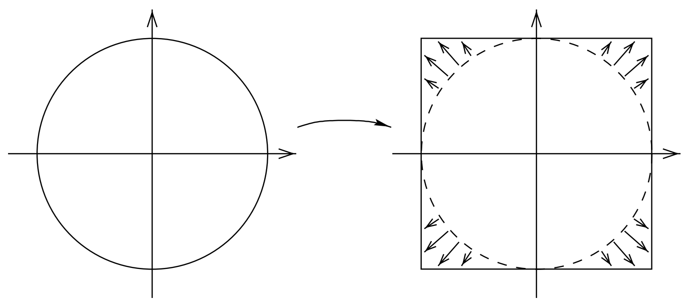


Figure 1.1. A homeomorphism from a circle to a square.

ample, we might want to call a curve “smooth” if it has a tangent line that varies continuously from point to point, and similarly a “smooth surface” should be one that has a tangent plane that varies continuously from point to point. But for more sophisticated applications it is an undue restriction to require smooth manifolds to be subsets of some ambient Euclidean space. The ambient coordinates and the vector space structure of $\mathbb{R}^n$ are superfluous data that often have nothing to do with the problem at hand. It is a tremendous advantage to be able to work with manifolds as abstract topological spaces, without the excess baggage of such an ambient space. For example, in general relativity, spacetime is thought of as a 4-dimensional smooth manifold that carries a certain geometric structure, called a *Lorentz metric*, whose curvature results in gravitational phenomena. In such a model there is no physical meaning that can be assigned to any higher-dimensional ambient space in which the manifold lives, and including such a space in the model would complicate it needlessly. For such reasons, we need to think of smooth manifolds as abstract topological spaces, not necessarily as subsets of larger spaces.

It is not hard to see that there is no way to define a purely topological property that would serve as a criterion for “smoothness,” because it cannot be invariant under homeomorphisms. For example, a circle and a square in the plane are homeomorphic topological spaces (Figure 1.1), but we would probably all agree that the circle is “smooth,” while the square is not. Thus topological manifolds will not suffice for our purposes. As a consequence, we will think of a smooth manifold as a set with two layers of structure: first a topology, then a smooth structure.

In the first section of this chapter we describe the first of these structures. A topological manifold is a topological space with three special properties that express the notion of being locally like Euclidean space. These properties are shared by Euclidean spaces and by all of the familiar geometric objects that look locally like Euclidean spaces, such as curves and surfaces.

We then prove some important topological properties of manifolds that we will use throughout the book.

In the next section we introduce an additional structure, called a smooth structure, that can be added to a topological manifold to enable us to make sense of derivatives.

Following the basic definitions, we introduce a number of examples of manifolds, so you can have something concrete in mind as you read the general theory. At the end of the chapter we introduce the concept of a smooth manifold with boundary, an important generalization of smooth manifolds that will be important in our study of integration in Chapters 14–16.

Topological Manifolds

In this section we introduce topological manifolds, the most basic type of manifolds. We assume that the reader is familiar with the basic properties of topological spaces, as summarized in the Appendix.

Suppose M is a topological space. We say that M is a *topological manifold of dimension* n or a *topological n-manifold* if it has the following properties:

- M is a *Hausdorff space*: For every pair of points $p, q \in M$, there are disjoint open subsets $U, V \subset M$ such that $p \in U$ and $q \in V$.
- M is *second countable*: There exists a countable basis for the topology of M.
- M is *locally Euclidean of dimension* n: Every point of M has a neighborhood that is homeomorphic to an open subset of $\mathbb{R}^n$.

The locally Euclidean property means, more specifically, that for each $p \in M$, we can find the following:

- an open set $U \subset M$ containing p;
- an open set $\widetilde{U} \subset \mathbb{R}^n$; and
- a homeomorphism $\varphi\colon U \to \widetilde{U}$.

◇ **Exercise 1.1.** Show that equivalent definitions of locally Euclidean spaces are obtained if instead of requiring U to be homeomorphic to an open subset of $\mathbb{R}^n$, we require it to be homeomorphic to an open ball in $\mathbb{R}^n$, or to $\mathbb{R}^n$ itself.

If M is a topological manifold, we often abbreviate the dimension of M as $\dim M$. In informal writing, one sometimes writes "Let M^n be a manifold" as shorthand for "Let M be a manifold of dimension n." The superscript n is not part of the name of the manifold, and is usually not included in the notation after the first occurrence.

The basic example of a topological n-manifold is, of course, $\mathbb{R}^n$. It is Hausdorff because it is a metric space, and it is second countable because the set of all open balls with rational centers and rational radii is a countable basis.

Requiring that manifolds share these properties helps to ensure that manifolds behave in the ways we expect from our experience with Euclidean spaces. For example, it is easy to verify that in a Hausdorff space, one-point sets are closed and limits of convergent sequences are unique (see Exercise A.5 in the Appendix). The motivation for second countability is a bit less evident, but it will have important consequences throughout the book, mostly based on the existence of partitions of unity (see Chapter 2).

In practice, both the Hausdorff and second countability properties are usually easy to check, especially for spaces that are built out of other manifolds, because both properties are inherited by subspaces and products (Lemmas A.5 and A.8). In particular, it follows easily that any open subset of a topological n-manifold is itself a topological n-manifold (with the subspace topology, of course).

The way we have defined topological manifolds, the empty set is a topological n-manifold for every n. For the most part, we will ignore this special case (sometimes without remembering to say so). But because it is useful in certain contexts to allow the empty manifold, we have chosen not to exclude it from the definition.

We should note that some authors choose to omit the Hausdorff property or second countability or both from the definition of manifolds. However, most of the interesting results about manifolds do in fact require these properties, and it is exceedingly rare to encounter a space "in nature" that would be a manifold except for the failure of one or the other of these hypotheses. For a couple of simple examples, see Problems 1-1 and 1-2; for a more involved example (a connected, locally Euclidean, Hausdorff space that is not second countable), see [Lee00, Problem 4-6].

Coordinate Charts

Let M be a topological n-manifold. A *coordinate chart* (or just a *chart*) on M is a pair (U, φ), where U is an open subset of M and $\varphi\colon U \to \widetilde{U}$ is a homeomorphism from U to an open subset $\widetilde{U} = \varphi(U) \subset \mathbb{R}^n$ (Figure 1.2). By definition of a topological manifold, each point $p \in M$ is contained in the domain of some chart (U, φ). If $\varphi(p) = 0$, we say that the chart is *centered at* p. If (U, φ) is any chart whose domain contains p, it is easy to obtain a new chart centered at p by subtracting the constant vector $\varphi(p)$.

Given a chart (U, φ), we call the set U a *coordinate domain*, or a *coordinate neighborhood* of each of its points. If in addition $\varphi(U)$ is an open ball in $\mathbb{R}^n$, then U is called a *coordinate ball*. The map φ is called a (*local*) *coordinate map*, and the component functions $(x^1, \ldots, x^n)$ of φ, defined by

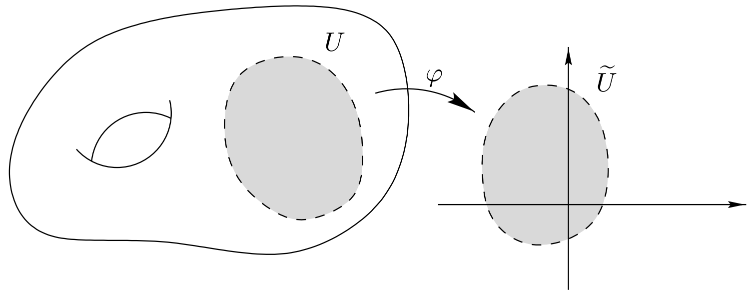


Figure 1.2. A coordinate chart.

$\varphi(p) = (x^1(p), \ldots, x^n(p))$, are called *local coordinates* on U. We will sometimes write things like "(U, φ) is a chart containing p" as shorthand for "(U, φ) is a chart whose domain U contains p." If we wish to emphasize the coordinate functions $(x^1, \ldots, x^n)$ instead of the coordinate map φ, we will sometimes denote the chart by $(U, (x^1, \ldots, x^n))$ or $(U, (x^i))$.

Examples of Topological Manifolds

Here are some simple examples of topological manifolds.

Example 1.1 (Graphs of Continuous Functions). Let $U \subset \mathbb{R}^n$ be an open set, and let $F\colon U \to \mathbb{R}^k$ be a continuous function. The *graph* of F is the subset of $\mathbb{R}^n \times \mathbb{R}^k$ defined by

$$\Gamma(F) = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^k : x \in U \text{ and } y = F(x)\},$$

with the subspace topology. Let $\pi_1\colon \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^n$ denote the projection onto the first factor, and let $\varphi_F\colon \Gamma(F) \to U$ be the restriction of π_1 to $\Gamma(F)$:

$$\varphi_F(x, y) = x, \quad (x, y) \in \Gamma(F).$$

Because φ_F is the restriction of a continuous map, it is continuous; and it is a homeomorphism because it has a continuous inverse given by

$$(\varphi_F)^{-1}(x) = (x, F(x)).$$

Thus $\Gamma(F)$ is a topological manifold of dimension n. In fact, $\Gamma(F)$ is homeomorphic to U itself, and $(\Gamma(F), \varphi_F)$ is a global coordinate chart, called *graph coordinates*. The same observation applies to any subset of $\mathbb{R}^{n+k}$ defined by setting any k of the coordinates (not necessarily the last k) equal to some continuous function of the other n, which are restricted to lie in an open subset of $\mathbb{R}^n$.

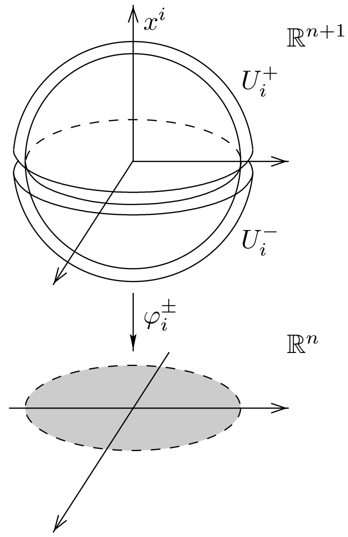


Figure 1.3. Charts for $\mathbb{S}^n$.

Example 1.2 (Spheres). Let $\mathbb{S}^n$ denote the (*unit*) *n-sphere*, which is the set of unit vectors in $\mathbb{R}^{n+1}$:

$$\mathbb{S}^n = \left\{x \in \mathbb{R}^{n+1} : |x| = 1\right\},$$

with the subspace topology. It is Hausdorff and second countable because it is a topological subspace of $\mathbb{R}^n$. To show that it is locally Euclidean, for each index $i = 1, \dots, n+1$ let U_i^+ denote the subset of $\mathbb{S}^n$ where the ith coordinate is positive:

$$U_i^+ = \left\{\left(x^1, \dots, x^{n+1}\right) \in \mathbb{S}^n : x^i > 0\right\}.$$

(See Figure 1.3.) Similarly, U_i^- is the set where $x^i < 0$.

Let $\mathbb{B}^n = \{x \in \mathbb{R}^n : |x| < 1\}$ denote the open unit ball in $\mathbb{R}^n$, and let $f\colon \mathbb{B}^n \to \mathbb{R}$ be the continuous function

$$f(u) = \sqrt{1 - |u|^2}.$$

Then for each $i = 1, \dots, n+1$, it is easy to check that $U_i^+ \cap \mathbb{S}^n$ is the graph of the function

$$x^i = f\left(x^1, \dots, \widehat{x^i}, \dots, x^{n+1}\right),$$

where the hat over x^i indicates that x^i is omitted. Similarly, $U_i^- \cap \mathbb{S}^n$ is the graph of

$$x^i = -f\left(x^1, \dots, \widehat{x^i}, \dots, x^{n+1}\right).$$

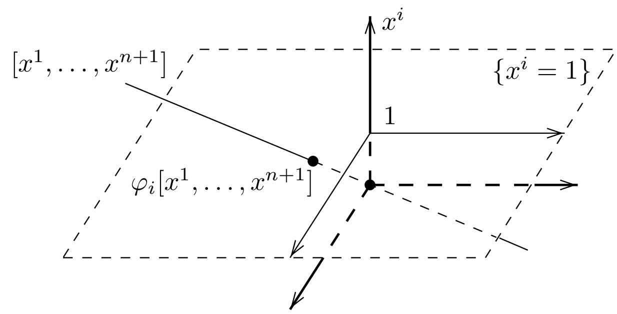


Figure 1.4. A chart for $\mathbb{RP}^n$.

Thus each set $U_i^{\pm} \cap \mathbb{S}^n$ is locally Euclidean of dimension n, and the maps $\varphi_i^{\pm} \colon U_i^{\pm} \cap \mathbb{S}^n \to \mathbb{B}^n$ given by

$$\varphi_i^{\pm}\left(x^1, \ldots, x^{n+1}\right) = \left(x^1, \ldots, \widehat{x^i}, \ldots, x^{n+1}\right)$$

are graph coordinates for $\mathbb{S}^n$. Since every point in $\mathbb{S}^n$ is in the domain of at least one of these $2n+2$ charts, $\mathbb{S}^n$ is a topological n-manifold.

Example 1.3 (Projective Spaces). The n-dimensional *real projective space*, denoted by $\mathbb{RP}^n$ (or sometimes just $\mathbb{P}^n$), is defined as the set of 1-dimensional linear subspaces of $\mathbb{R}^{n+1}$. We give it the quotient topology determined by the natural map $\pi \colon \mathbb{R}^{n+1} \smallsetminus \{0\} \to \mathbb{RP}^n$ sending each point $x \in \mathbb{R}^{n+1} \smallsetminus \{0\}$ to the subspace spanned by x. For any point $x \in \mathbb{R}^{n+1} \smallsetminus \{0\}$, let $[x] = \pi(x)$ denote the equivalence class of x in $\mathbb{RP}^n$.

For each $i = 1, \ldots, n+1$, let $\widetilde{U}_i \subset \mathbb{R}^{n+1} \smallsetminus \{0\}$ be the set where $x^i \neq 0$, and let $U_i = \pi(\widetilde{U}_i) \subset \mathbb{RP}^n$. Since $\widetilde{U}_i$ is a saturated open set, U_i is open and $\pi|_{\widetilde{U}_i} \colon \widetilde{U}_i \to U_i$ is a quotient map (see Lemma A.10). Define a map $\varphi_i \colon U_i \to \mathbb{R}^n$ by

$$\varphi_i\left[x^1, \ldots, x^{n+1}\right] = \left(\frac{x^1}{x^i}, \ldots, \frac{x^{i-1}}{x^i}, \frac{x^{i+1}}{x^i}, \ldots, \frac{x^{n+1}}{x^i}\right).$$

This map is well-defined because its value is unchanged by multiplying x by a nonzero constant. Because $\varphi_i \circ \pi$ is continuous, φ_i is continuous by the characteristic property of quotient maps (Lemma A.10). In fact, φ_i is a homeomorphism, because its inverse is given by

$$\varphi_i^{-1}\left(u^1, \ldots, u^n\right) = \left[u^1, \ldots, u^{i-1}, 1, u^i, \ldots, u^n\right],$$

as you can easily check. Geometrically, if we identify $\mathbb{R}^n$ in the obvious way with the affine subspace where $x^i = 1$, then $\varphi_i[x]$ can be interpreted as the point where the line $[x]$ intersects this subspace (Figure 1.4). Because the sets U_i cover $\mathbb{RP}^n$, this shows that $\mathbb{RP}^n$ is locally Euclidean of dimension n. The Hausdorff and second countability properties are left as exercises.

◇ **Exercise 1.2.** Show that $\mathbb{RP}^n$ is Hausdorff and second countable, and is therefore a topological n-manifold.

◇ **Exercise 1.3.** Show that $\mathbb{RP}^n$ is compact. [Hint: Show that the restriction of π to $\mathbb{S}^n$ is surjective.]

Example 1.4 (Product Manifolds). Suppose $M_1, \ldots, M_k$ are topological manifolds of dimensions $n_1, \ldots, n_k$, respectively. We will show that the product space $M_1 \times \cdots \times M_k$ is a topological manifold of dimension $n_1 + \cdots + n_k$. It is Hausdorff and second countable by Lemmas A.5 and A.8, so only the locally Euclidean property needs to be checked. Given any point $(p_1, \ldots, p_k) \in M_1 \times \cdots \times M_k$, we can choose a coordinate chart (U_i, φ_i) for each M_i with $p_i \in U_i$. The product map

$$\varphi_1 \times \cdots \times \varphi_k \colon U_1 \times \cdots \times U_k \to \mathbb{R}^{n_1 + \cdots + n_k}$$

is a homeomorphism onto its image, which is an open subset of $\mathbb{R}^{n_1 + \cdots + n_k}$. Thus $M_1 \times \cdots \times M_k$ is a topological manifold of dimension $n_1 + \cdots + n_k$, with charts of the form $(U_1 \times \cdots \times U_k, \varphi_1 \times \cdots \times \varphi_k)$.

Example 1.5 (Tori). For any positive integer n, the *n-torus* is the product space $\mathbb{T}^n = \mathbb{S}^1 \times \cdots \times \mathbb{S}^1$. By the discussion above, it is an n-dimensional topological manifold. (The 2-torus is usually called simply "the torus.")

Topological Properties of Manifolds

As topological spaces go, manifolds are quite special, because they share so many important properties with Euclidean spaces. In this section we discuss a few such properties that will be of use to us throughout the book.

The first property we need is that every manifold has a particularly well behaved basis for its topology. If X is a topological space, a subset $K \subset X$ is said to be *precompact* (or *relatively compact*) in X if its closure in X is compact.

Lemma 1.6. *Every topological manifold has a countable basis of precompact coordinate balls.*

Proof. Let M be a topological n-manifold. First we will prove the lemma in the special case in which M can be covered by a single chart. Suppose $\varphi \colon M \to \widetilde{U} \subset \mathbb{R}^n$ is a global coordinate map, and let $\mathcal{B}$ be the collection of all open balls $B_r(x) \subset \mathbb{R}^n$ such that r is rational, x has rational coordinates, and $\overline{B}_r(x) \subset \widetilde{U}$. Each such ball is precompact in $\widetilde{U}$, and it is easy to check that $\mathcal{B}$ is a countable basis for the topology of $\widetilde{U}$. Because φ is a homeomorphism, it follows that the collection of sets of the form $\varphi^{-1}(B)$ for $B \in \mathcal{B}$ is a countable basis for the topology of M, consisting of precompact coordinate balls, with the restrictions of φ as coordinate maps.

Now let M be an arbitrary n-manifold. By definition, every point of M is in the domain of a chart. Because every open cover of a second countable space has a countable subcover (Lemma A.4), M is covered by countably many charts $\{(U_i, \varphi_i)\}$. By the argument in the preceding paragraph, each coordinate domain U_i has a countable basis of precompact coordinate balls, and the union of all these countable bases is a countable basis for the topology of M. If $V \subset U_i$ is one of these precompact balls, then the closure of V in U_i is compact, hence closed in M. It follows that the closure of V in M the same as its closure in U_i, so V is precompact in M as well. □

A topological space M is said to be *locally compact* if every point has a neighborhood contained in a compact subset of M. If M is Hausdorff, this is equivalent to the requirement that M have a basis of precompact open sets (see [Lee00, Proposition 4.27]). The following corollary is immediate.

Corollary 1.7. *Every topological manifold is locally compact.*

Connectivity

The existence of a basis of coordinate balls has important consequences for the connectivity properties of manifolds. Recall that a topological space X is said to be

- *connected* if there do not exist two disjoint, nonempty, open subsets of X whose union is X;
- *path connected* if every pair of points in X can be joined by a path in X; and
- *locally path connected* if X has a basis of path connected open sets.

(See the Appendix, pages 550–552, for a review of these concepts.) The following proposition shows that connectivity and path connectivity coincide for manifolds.

Proposition 1.8. *Let M be a topological manifold.*

(*a*) *M is locally path connected.*

(*b*) *M is connected if and only if it is path connected.*

(*c*) *The components of M are the same as its path components.*

(*d*) *M has at most countably many components, each of which is an open subset of M and a connected topological manifold.*

Proof. Since every coordinate ball is path connected, part (a) follows from the fact that M has a basis of coordinate balls (Lemma 1.6). Parts (b) and (c) are immediate consequences of (a) (see Lemma A.16). To prove (d), note that each component is open in M by Lemma A.16, so the collection of components is an open cover of M. Because M is second countable, this cover must have a countable subcover. But since the components are all disjoint, the cover must have been countable to begin with, which is to say that M has only countably many components. □

Fundamental Groups of Manifolds

The following result about fundamental groups of manifolds will be important in our study of covering manifolds in Chapters 2 and 9. For a brief review of the fundamental group, see the Appendix, pages 553–555.

Proposition 1.9. *The fundamental group of any topological manifold is countable.*

Proof. Let M be a topological manifold. By Lemma 1.6, there is a countable collection $\mathcal{B}$ of coordinate balls covering M. For any pair of coordinate balls $B, B' \in \mathcal{B}$, the intersection $B \cap B'$ has at most countably many components, each of which is path connected. Let $\mathcal{X}$ be a countable set containing one point from each component of $B \cap B'$ for each $B, B' \in \mathcal{B}$ (including $B = B'$). For each $B \in \mathcal{B}$ and each $x, x' \in \mathcal{X}$ such that $x, x' \in B$, let $p^B_{x,x'}$ be some path from x to x' in B.

Since the fundamental groups based at any two points in the same component of M are isomorphic, and $\mathcal{X}$ contains at least one point in each component of M, we may as well choose a point $q \in \mathcal{X}$ as base point. Define a *special loop* to be a loop based at q that is equal to a finite product of paths of the form $p^B_{x,x'}$. Clearly, the set of special loops is countable, and each special loop determines an element of $\pi_1(M, q)$. To show that $\pi_1(M, q)$ is countable, therefore, it suffices to show that every element of $\pi_1(M, q)$ is represented by a special loop.

Suppose $f\colon [0, 1] \to M$ is any loop based at q. The collection of components of sets of the form $f^{-1}(B)$ as B ranges over $\mathcal{B}$ is an open cover of $[0, 1]$, so by compactness it has a finite subcover. Thus there are finitely many numbers $0 = a_0 < a_1 < \cdots < a_k = 1$ such that $[a_{i-1}, a_i] \subset f^{-1}(B)$ for some $B \subset \mathcal{B}$. For each i, let f_i be the restriction of f to the interval $[a_{i-1}, a_i]$, reparametrized so that its domain is $[0, 1]$, and let $B_i \in \mathcal{B}$ be a coordinate ball containing the image of f_i. For each i, we have $f(a_i) \in B_i \cap B_{i+1}$, and there is some $x_i \in \mathcal{X}$ that lies in the same component of $B_i \cap B_{i+1}$ as $f(a_i)$. Let g_i be a path in $B_i \cap B_{i+1}$ from x_i to $f(a_i)$ (Figure 1.5), with the understanding that $x_0 = x_k = q$, and g_0 and g_k are both equal to the constant path c_q based at q. Then, because $g_i^{-1} \cdot g_i$ is path homotopic to a constant path,

$$\begin{aligned} f &\sim f_1 \cdot \cdots \cdot f_k \\ &\sim g_0 \cdot f_1 \cdot g_1^{-1} \cdot g_1 \cdot f_2 \cdot g_2^{-1} \cdot \cdots \cdot g_{k-1}^{-1} \cdot g_{k-1} \cdot f_k \cdot g_k^{-1} \\ &\sim \widetilde{f}_1 \cdot \widetilde{f}_2 \cdot \cdots \cdot \widetilde{f}_n, \end{aligned}$$

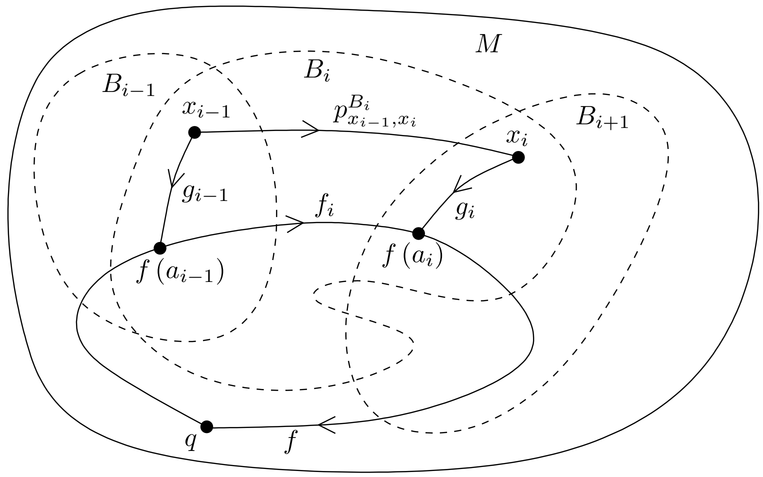


Figure 1.5. The fundamental group of a manifold is countable.

where $\widetilde{f}_i = g_{i-1} \cdot f_i \cdot g_i^{-1}$. For each i, $\widetilde{f}_i$ is a path in B_i from x_{i-1} to x_i. Since B_i is simply connected, $\widetilde{f}_i$ is path homotopic to $p^{B_i}_{x_{i-1},x_i}$. It follows that f is path homotopic to a special loop, as claimed. □

Smooth Structures

The definition of manifolds that we gave in the preceding section is sufficient for studying topological properties of manifolds, such as compactness, connectedness, simple connectedness, and the problem of classifying manifolds up to homeomorphism. However, in the entire theory of topological manifolds there is no mention of calculus. There is a good reason for this: However we might try to make sense of derivatives of functions on a manifold, such derivatives cannot be invariant under homeomorphisms. For example, the map $\varphi\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by $\varphi(u,v) = \left(u^{1/3}, v^{1/3}\right)$ is a homeomorphism, and it is easy to construct differentiable functions $f\colon \mathbb{R}^2 \to \mathbb{R}$ such that $f \circ \varphi$ is not differentiable at the origin. (The function $f(x,y) = x$ is one such.)

To make sense of derivatives of real-valued functions, curves, or maps between manifolds, we will need to introduce a new kind of manifold called a "smooth manifold." It will be a topological manifold with some extra structure in addition to its topology, which will allow us to decide which functions on the manifold are smooth.

The definition will be based on the calculus of maps between Euclidean spaces, so let us begin by reviewing some basic terminology about such maps. If U and V are open subsets of Euclidean spaces $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, a function $F\colon U \to V$ is said to be *smooth* (or C^∞, or *infinitely differentiable*) if each of its component functions has continuous partial derivatives of all orders. If in addition F is bijective and has a smooth inverse map, it is called a *diffeomorphism*. A diffeomorphism is, in particular, a homeomorphism. A review of some of the most important properties of smooth maps is given in the Appendix. (You should be aware that some authors use the word "smooth" in somewhat different senses, for example to mean continuously differentiable or merely differentiable. On the other hand, some use the word "differentiable" to mean what we call "smooth." Throughout this book, "smooth" will for us be synonymous with C^∞.)

To see what additional structure on a topological manifold might be appropriate for discerning which maps are smooth, consider an arbitrary topological n-manifold M. Each point in M is in the domain of a coordinate map $\varphi\colon U \to \widetilde{U} \subset \mathbb{R}^n$. A plausible definition of a smooth function on M would be to say that $f\colon M \to \mathbb{R}$ is smooth if and only if the composite function $f \circ \varphi^{-1}\colon \widetilde{U} \to \mathbb{R}$ is smooth in the sense of ordinary calculus. But this will make sense only if this property is independent of the choice of coordinate chart. To guarantee this independence, we will restrict our attention to "smooth charts." Since smoothness is not a homeomorphism-invariant property, the way to do this is to consider the collection of all smooth charts as a new kind of structure on M.

With this motivation in mind, we now describe the details of the construction.

Let M be a topological n-manifold. If (U,φ), (V,ψ) are two charts such that $U \cap V \neq \varnothing$, the composite map $\psi \circ \varphi^{-1}\colon \varphi(U \cap V) \to \psi(U \cap V)$ is called the *transition map* from φ to ψ (Figure 1.6). It is a composition of homeomorphisms, and is therefore itself a homeomorphism. Two charts (U,φ) and (V,ψ) are said to be *smoothly compatible* if either $U \cap V = \varnothing$ or the transition map $\psi \circ \varphi^{-1}$ is a diffeomorphism. (Since $\varphi(U \cap V)$ and $\psi(U \cap V)$ are open subsets of $\mathbb{R}^n$, smoothness of this map is to be interpreted in the ordinary sense of having continuous partial derivatives of all orders.)

We define an *atlas* for M to be a collection of charts whose domains cover M. An atlas $\mathcal{A}$ is called a *smooth atlas* if any two charts in $\mathcal{A}$ are smoothly compatible with each other.

It often happens in practice that we can prove for *every pair* of coordinate maps φ and ψ in a given atlas that the transition map $\psi \circ \varphi^{-1}$ is smooth. Once we have done this, it is unnecessary to verify directly that $\psi \circ \varphi^{-1}$ is a diffeomorphism, because its inverse $(\psi \circ \varphi^{-1})^{-1} = \varphi \circ \psi^{-1}$ is one of the transition maps we have already shown to be smooth. We will use this observation without further comment when appropriate.

Our plan is to define a "smooth structure" on M by giving a smooth atlas, and to define a function $f\colon M \to \mathbb{R}$ to be smooth if and only if $f \circ \varphi^{-1}$ is smooth in the sense of ordinary calculus for each coordinate chart (U,φ) in the atlas. There is one minor technical problem with this approach: In

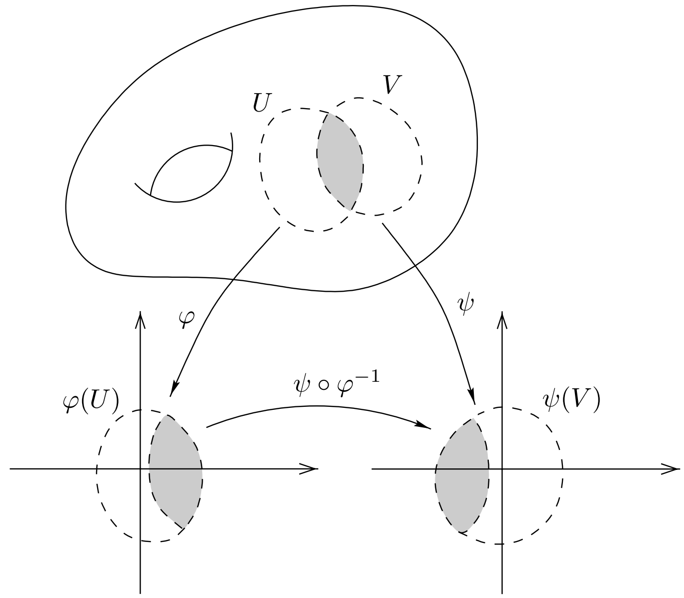


Figure 1.6. A transition map.

general, there will be many possible choices of atlas that give the "same" smooth structure, in that they all determine the same collection of smooth functions on M. For example, consider the following pair of atlases on $\mathbb{R}^n$:

$$\mathcal{A}_1 = \{(\mathbb{R}^n, \mathrm{Id}_{\mathbb{R}^n})\}$$
$$\mathcal{A}_2 = \left\{\left(B_1(x), \mathrm{Id}_{B_1(x)}\right) : x \in \mathbb{R}^n\right\}.$$

Although these are different smooth atlases, clearly a function $f\colon \mathbb{R}^n \to \mathbb{R}$ is smooth with respect to either atlas if and only if it is smooth in the sense of ordinary calculus.

We could choose to define a smooth structure as an equivalence class of smooth atlases under an appropriate equivalence relation. However, it is more straightforward to make the following definition: A smooth atlas $\mathcal{A}$ on M is *maximal* if it is not contained in any strictly larger smooth atlas. This just means that any chart that is smoothly compatible with every chart in $\mathcal{A}$ is already in $\mathcal{A}$. (Such a smooth atlas is also said to be *complete*.)

Now we can define the main concept of this chapter. A *smooth structure* on a topological n-manifold M is a maximal smooth atlas. A *smooth manifold* is a pair $(M, \mathcal{A})$, where M is a topological manifold and $\mathcal{A}$ is a smooth structure on M. When the smooth structure is understood, we usually omit mention of it and just say "M is a smooth manifold." Smooth structures are also called *differentiable structures* or C^∞ *structures* by some authors. We will use the term *smooth manifold structure* to mean a manifold topology together with a smooth structure.

We emphasize that a smooth structure is an additional piece of data that must be added to a topological manifold before we are entitled to talk about a "smooth manifold." In fact, a given topological manifold may have many different smooth structures (see Example 1.14 and Problem 1-3). And it should be noted that it is not always possible to find a smooth structure on a given topological manifold: There exist topological manifolds that admit no smooth structures at all. (The first example was a compact 10-dimensional manifold found in 1960 by Michel Kervaire [Ker60].)

It is generally not very convenient to define a smooth structure by explicitly describing a maximal smooth atlas, because such an atlas contains very many charts. Fortunately, we need only specify *some* smooth atlas, as the next lemma shows.

Lemma 1.10. *Let M be a topological manifold.*

(*a*) *Every smooth atlas for M is contained in a unique maximal smooth atlas.*

(*b*) *Two smooth atlases for M determine the same maximal smooth atlas if and only if their union is a smooth atlas.*

Proof. Let $\mathcal{A}$ be a smooth atlas for M, and let $\overline{\mathcal{A}}$ denote the set of all charts that are smoothly compatible with every chart in $\mathcal{A}$. To show that $\overline{\mathcal{A}}$ is a smooth atlas, we need to show that any two charts of $\overline{\mathcal{A}}$ are smoothly compatible with each other, which is to say that for any (U, φ), $(V, \psi) \in \overline{\mathcal{A}}$, $\psi \circ \varphi^{-1} \colon \varphi(U \cap V) \to \psi(U \cap V)$ is smooth.

Let $x = \varphi(p) \in \varphi(U \cap V)$ be arbitrary. Because the domains of the charts in $\mathcal{A}$ cover M, there is some chart $(W, \theta) \in \mathcal{A}$ such that $p \in W$ (Figure 1.7). Since every chart in $\overline{\mathcal{A}}$ is smoothly compatible with (W, θ), both of the maps $\theta \circ \varphi^{-1}$ and $\psi \circ \theta^{-1}$ are smooth where they are defined. Since $p \in U \cap V \cap W$, it follows that $\psi \circ \varphi^{-1} = (\psi \circ \theta^{-1}) \circ (\theta \circ \varphi^{-1})$ is smooth on a neighborhood of x. Thus $\psi \circ \varphi^{-1}$ is smooth in a neighborhood of each point in $\varphi(U \cap V)$. Therefore, $\overline{\mathcal{A}}$ is a smooth atlas. To check that it is maximal, just note that any chart that is smoothly compatible with every chart in $\overline{\mathcal{A}}$ must in particular be smoothly compatible with every chart in $\mathcal{A}$, so it is already in $\overline{\mathcal{A}}$. This proves the existence of a maximal smooth atlas containing $\mathcal{A}$. If $\mathcal{B}$ is any other maximal smooth atlas containing $\mathcal{A}$, each of its charts is smoothly compatible with each chart in $\mathcal{A}$, so $\mathcal{B} \subset \overline{\mathcal{A}}$. By maximality of $\mathcal{B}$, $\mathcal{B} = \overline{\mathcal{A}}$.

The proof of (b) is left as an exercise. □

◇ **Exercise 1.4.** Prove Lemma 1.10(b).

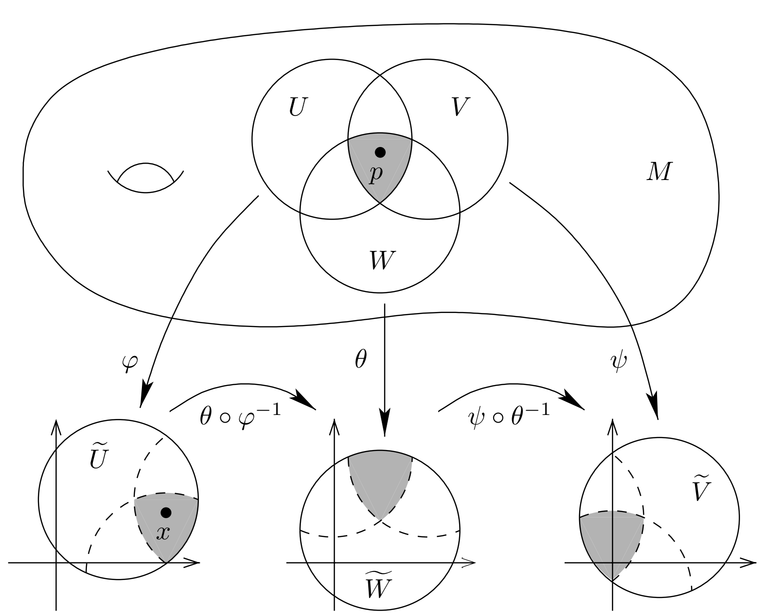


Figure 1.7. Proof of Lemma 1.10(a).

For example, if a topological manifold M can be covered by a single chart, the smooth compatibility condition is trivially satisfied, so any such chart automatically determines a smooth structure on M.

It is worth mentioning that the notion of smooth structure can be generalized in several different ways by changing the compatibility requirement for charts. For example, if we replace the requirement that charts be smoothly compatible by the weaker requirement that each transition map $\psi \circ \varphi^{-1}$ (and its inverse) be of class C^k, we obtain the definition of a *C^k structure*. Similarly, if we require that each transition map be real-analytic (i.e., expressible as a convergent power series in a neighborhood of each point), we obtain the definition of a *real-analytic structure*, also called a *C^ω structure*. If M has even dimension $n = 2m$, we can identify $\mathbb{R}^{2m}$ with $\mathbb{C}^m$ and require that the transition maps be complex-analytic; this determines a *complex-analytic structure*. A manifold endowed with one of these structures is called a *C^k manifold*, *real-analytic manifold*, or *complex manifold*, respectively. (Note that a C^0 manifold is just a topological manifold.) We will not treat any of these other kinds of manifolds in this book, but they play important roles in analysis, so it is useful to know the definitions.

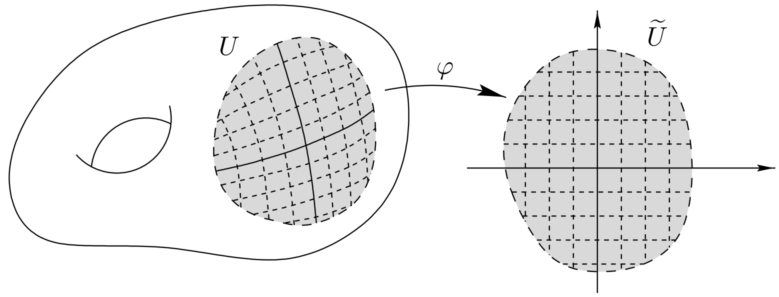


Figure 1.8. A coordinate grid.

Local Coordinate Representations

If M is a smooth manifold, any chart (U, φ) contained in the given maximal smooth atlas will be called a *smooth chart*, and the corresponding coordinate map φ will be called a *smooth coordinate map*. It is useful also to introduce the terms *smooth coordinate domain* or *smooth coordinate neighborhood* for the domain of a smooth coordinate chart. A *smooth coordinate ball* will mean a smooth coordinate domain whose image under a smooth coordinate map is a ball in Euclidean space.

The next lemma gives a slight improvement on Lemma 1.6 for smooth manifolds. Its proof is a straightforward adaptation of the proof of that lemma.

Lemma 1.11. *Every smooth manifold has a countable basis of precompact smooth coordinate balls.*

◇ **Exercise 1.5.** Prove Lemma 1.11.

Here is how one usually thinks about coordinate charts on a smooth manifold. Once we choose a smooth chart (U, φ) on M, the coordinate map $\varphi\colon U \to \widetilde{U} \subset \mathbb{R}^n$ can be thought of as giving an *identification* between U and $\widetilde{U}$. Using this identification, we can think of U simultaneously as an open subset of M and (at least temporarily while we work with this chart) as an open subset of $\mathbb{R}^n$. You can visualize this identification by thinking of a "grid" drawn on U representing the inverse images of the coordinate lines under φ (Figure 1.8). Under this identification, we can represent a point $p \in U$ by its coordinates $(x^1, \dots, x^n) = \varphi(p)$, and think of this n-tuple as *being* the point p. We will typically express this by saying "$(x^1, \dots, x^n)$ is the (local) coordinate representation for p" or "$p = (x^1, \dots, x^n)$ in local coordinates."

Another way to look at it is that by means of our identification $U \leftrightarrow \widetilde{U}$, we can think of φ as the identity map and suppress it from the notation. This takes a bit of getting used to, but the payoff is a huge simplification of the notation in many situations. You just need to remember that the identification is in general only local, and depends heavily on the choice of coordinate chart.

For example, if $M = \mathbb{R}^2$, let $U = \{(x, y) : x > 0\} \subset M$ be the open right half-plane, and let $\varphi\colon U \to \mathbb{R}^2$ be the *polar coordinate map* $\varphi(x, y) = (r, \theta) = \left(\sqrt{x^2 + y^2}, \tan^{-1} y/x\right)$. We can write a given point $p \in U$ either as $p = (x, y)$ in standard coordinates or as $p = (r, \theta)$ in polar coordinates, where the two coordinate representations are related by $(r, \theta) = \left(\sqrt{x^2 + y^2}, \tan^{-1} y/x\right)$ and $(x, y) = (r \cos\theta, r \sin\theta)$.

Examples of Smooth Manifolds

Before proceeding further with the general theory, let us survey some examples of smooth manifolds.

Example 1.12 (Zero-Dimensional Manifolds). A zero-dimensional topological manifold M is just a countable discrete space. For each point $p \in M$, the only neighborhood of p that is homeomorphic to an open subset of $\mathbb{R}^0$ is $\{p\}$ itself, and there is exactly one coordinate map $\varphi\colon \{p\} \to \mathbb{R}^0$. Thus the set of all charts on M trivially satisfies the smooth compatibility condition, and every zero-dimensional manifold has a unique smooth structure.

Example 1.13 (Euclidean Spaces). $\mathbb{R}^n$ is a smooth n-manifold with the smooth structure determined by the atlas consisting of the single chart $(\mathbb{R}^n, \mathrm{Id}_{\mathbb{R}^n})$. We call this the *standard smooth structure*, and the resulting coordinate map *standard coordinates*. Unless we explicitly specify otherwise, we will always use this smooth structure on $\mathbb{R}^n$.

Example 1.14 (Another Smooth Structure on the Real Line). Consider the homeomorphism $\psi\colon \mathbb{R} \to \mathbb{R}$ given by

$$\psi(x) = x^3. \tag{1.1}$$

The atlas consisting of the single chart $(\mathbb{R}, \psi)$ defines a smooth structure on $\mathbb{R}$. This chart is not smoothly compatible with the standard smooth structure, because the transition map $\mathrm{Id}_{\mathbb{R}^n} \circ \psi^{-1}(y) = y^{1/3}$ is not smooth at the origin. Therefore, the smooth structure defined on $\mathbb{R}$ by ψ is not the same as the standard one. Using similar ideas, it is not hard to construct many distinct smooth structures on any given positive-dimensional topological manifold, as long as it has one smooth structure to begin with (see Problem 1-3).

Example 1.15 (Finite-Dimensional Vector Spaces). Let V be a finite-dimensional vector space. Any norm on V determines a topology, which is independent of the choice of norm (Exercise A.53). With this topology, V has a natural smooth manifold structure defined as follows. Any (ordered) basis $(E_1, \dots, E_n)$ for V defines a basis isomorphism $E\colon \mathbb{R}^n \to V$ by

$$E(x) = \sum_{i=1}^{n} x^i E_i.$$

This map is a homeomorphism, so the atlas consisting of the single chart (V, E^{-1}) defines a smooth structure. To see that this smooth structure is independent of the choice of basis, let $\big(\widetilde{E}_1, \dots, \widetilde{E}_n\big)$ be any other basis and let $\widetilde{E}(x) = \sum_j x^j \widetilde{E}_j$ be the corresponding isomorphism. There is some invertible matrix $\big(A_i^j\big)$ such that $E_i = \sum_j A_i^j \widetilde{E}_j$ for each i. The transition map between the two charts is then given by $\widetilde{E}^{-1} \circ E(x) = \widetilde{x}$, where $\widetilde{x} = \big(\widetilde{x}^1, \dots, \widetilde{x}^n\big)$ is determined by

$$\sum_{j=1}^{n} \widetilde{x}^j \widetilde{E}_j = \sum_{i=1}^{n} x^i E_i = \sum_{i,j=1}^{n} x^i A_i^j \widetilde{E}_j.$$

It follows that $\widetilde{x}^j = \sum_i A_i^j x^i$. Thus the map from x to $\widetilde{x}$ is an invertible linear map and hence a diffeomorphism, so the two charts are smoothly compatible. This shows that the union of the two charts determined by any two bases is still a smooth atlas, and thus all bases determine the same smooth structure. We will call this the *standard smooth structure* on V.

The Einstein Summation Convention

This is a good place to pause and introduce an important notational convention that we will use throughout the book. Because of the proliferation of summations such as $\sum_i x^i E_i$ in this subject, we will often abbreviate such a sum by omitting the summation sign, as in

$$E(x) = x^i E_i.$$

We interpret any such expression according to the following rule, called the *Einstein summation convention*: If the same index name (such as i in the expression above) appears exactly twice in any monomial term, once as an upper index and once as a lower index, that term is understood to be summed over all possible values of that index, generally from 1 to the dimension of the space in question. This simple idea was introduced by Einstein to reduce the complexity of the expressions arising in the study of smooth manifolds by eliminating the necessity of explicitly writing summation signs.

Another important aspect of the summation convention is the positions of the indices. We will always write basis vectors (such as E_i) with lower indices, and components of a vector with respect to a basis (such as x^i) with upper indices. These index conventions help to ensure that, in summations that make mathematical sense, any index to be summed over will typically appear twice in any given term, once as a lower index and once as an upper index. Any index that is implicitly summed over is a "dummy index," meaning that the value of such an expression is unchanged if a different name is substituted for each dummy index. For example, $x^i E_i$ and $x^j E_j$ mean exactly the same thing.

Since the coordinates of a point $\left(x^1, \dots, x^n\right) \in \mathbb{R}^n$ are also its components with respect to the standard basis, in order to be consistent with our convention of writing components of vectors with upper indices, we need to use upper indices for these coordinates, and we will do so throughout this book. Although this may seem awkward at first, in combination with the summation convention it offers enormous advantages when we work with complicated indexed sums, not the least of which is that expressions that are not mathematically meaningful often betray themselves quickly by violating the index convention. (The main exceptions are expressions involving the Euclidean dot product $x \cdot y = \sum_i x^i y^i$, in which the same index appears twice in the upper position, and the standard symplectic form on $\mathbb{R}^{2n}$, which we will define in Chapter 12. We will always explicitly write summation signs in such expressions.)

More Examples

Now we continue with our examples of smooth manifolds.

Example 1.16 (Matrices). Let $\mathrm{M}(m \times n, \mathbb{R})$ denote the space of $m \times n$ matrices with real entries. It is a vector space of dimension mn under matrix addition and scalar multiplication. Thus $\mathrm{M}(m \times n, \mathbb{R})$ is a smooth mn-dimensional manifold. Similarly, the space $\mathrm{M}(m \times n, \mathbb{C})$ of $m \times n$ complex matrices is a vector space of dimension $2mn$ over $\mathbb{R}$, and thus a smooth manifold of dimension $2mn$. In the special case $m = n$ (square matrices), we will abbreviate $\mathrm{M}(n \times n, \mathbb{R})$ and $\mathrm{M}(n \times n, \mathbb{C})$ by $\mathrm{M}(n, \mathbb{R})$ and $\mathrm{M}(n, \mathbb{C})$, respectively.

Example 1.17 (Open Submanifolds). Let U be any open subset of $\mathbb{R}^n$. Then U is a topological n-manifold, and the single chart (U, Id_U) defines a smooth structure on U.

More generally, let M be a smooth n-manifold and let $U \subset M$ be any open subset. Define an atlas on U by

$$\mathcal{A}_U = \{\text{smooth charts } (V, \varphi) \text{ for } M \text{ such that } V \subset U\}.$$

Any point $p \in U$ is contained in the domain of some chart (W, φ) for M; if we set $V = W \cap U$, then $(V, \varphi|_V)$ is a chart in A_U whose domain contains p. Therefore, U is covered by the domains of charts in $\mathcal{A}_U$, and it is easy to verify that this is a smooth atlas for U. Thus any open subset of M is itself a smooth n-manifold in a natural way. Endowed with this smooth structure, we call any open subset an *open submanifold* of M. (We will define a more general class of submanifolds in Chapter 8.)

Example 1.18 (The General Linear Group). The *general linear group* $\mathrm{GL}(n,\mathbb{R})$ is the set of invertible $n \times n$ matrices with real entries. It is a smooth n^2-dimensional manifold because it is an open subset of the n^2-dimensional vector space $\mathrm{M}(n,\mathbb{R})$, namely the set where the (continuous) determinant function is nonzero.

Example 1.19 (Matrices of Maximal Rank). The previous example has a natural generalization to rectangular matrices of maximal rank. Suppose $m < n$, and let $\mathrm{M}_m(m \times n,\mathbb{R})$ denote the subset of $\mathrm{M}(m \times n,\mathbb{R})$ consisting of matrices of rank m. If A is an arbitrary such matrix, the fact that $\operatorname{rank} A = m$ means that A has some nonsingular $m \times m$ minor. By continuity of the determinant function, this same minor has nonzero determinant on some neighborhood of A in $\mathrm{M}(m \times n,\mathbb{R})$, which implies that A has a neighborhood contained in $\mathrm{M}_m(m \times n,\mathbb{R})$. Thus $\mathrm{M}_m(m \times n,\mathbb{R})$ is an open subset of $\mathrm{M}(m \times n,\mathbb{R})$, and therefore is itself a smooth mn-dimensional manifold. A similar argument shows that $\mathrm{M}_n(m \times n,\mathbb{R})$ is a smooth mn-manifold when $n < m$.

Example 1.20 (Spheres). We showed in Example 1.2 that the n-sphere $\mathbb{S}^n \subset \mathbb{R}^{n+1}$ is a topological n-manifold. Now we put a smooth structure on $\mathbb{S}^n$ as follows. For each $i = 1,\dots,n+1$, let $(U_i^\pm,\varphi_i^\pm)$ denote the graph coordinate charts we constructed in Example 1.2. For any distinct indices i and j, the transition map $\varphi_i^\pm \circ (\varphi_j^\pm)^{-1}$ is easily computed. In the case $i < j$, we get

$$\varphi_i^\pm \circ (\varphi_j^\pm)^{-1}\left(u^1,\dots,u^n\right) = \left(u^1,\dots,\widehat{u^i},\dots,\pm\sqrt{1-|u|^2},\dots,u^n\right),$$

and a similar formula holds when $i > j$. When $i = j$, an even simpler computation gives $\varphi_i^\pm \circ (\varphi_i^\pm)^{-1} = \mathrm{Id}_{\mathbb{B}^n}$. Thus the collection of charts $\left\{\left(U_i^\pm,\varphi_i^\pm\right)\right\}$ is a smooth atlas, and so defines a smooth structure on $\mathbb{S}^n$. We call this its *standard smooth structure*.

Example 1.21 (Projective Spaces). The n-dimensional real projective space $\mathbb{RP}^n$ is a topological n-manifold by Example 1.3. We will show that the coordinate charts (U_i,φ_i) constructed in that example are all smoothly compatible. Assuming for convenience that $i > j$, it is straightforward to compute that

$$\varphi_j \circ \varphi_i^{-1}(u^1,\dots,u^n) = \left(\frac{u^1}{u^j},\dots,\frac{u^{j-1}}{u^j},\frac{u^{j+1}}{u^j},\dots,\frac{u^{i-1}}{u^j},\frac{1}{u^j},\frac{u^i}{u^j},\dots,\frac{u^n}{u^j}\right),$$

which is a diffeomorphism from $\varphi_i(U_i \cap U_j)$ to $\varphi_j(U_i \cap U_j)$.

Example 1.22 (Smooth Product Manifolds). If $M_1,\dots,M_k$ are smooth manifolds of dimensions $n_1,\dots,n_k$, respectively, we showed in Example 1.4 that the product space $M_1 \times \cdots \times M_k$ is a topological manifold of dimension $n_1 + \cdots + n_k$, with charts of the form $(U_1 \times \cdots \times U_k, \varphi_1 \times \cdots \times \varphi_k)$.

Any two such charts are smoothly compatible because, as is easily verified,

$$(\psi_1 \times \cdots \times \psi_k) \circ (\varphi_1 \times \cdots \times \varphi_k)^{-1} = (\psi_1 \circ \varphi_1^{-1}) \times \cdots \times (\psi_k \circ \varphi_k^{-1}),$$

which is a smooth map. This defines a natural smooth manifold structure on the product, called the *product smooth manifold structure*. For example, this yields a smooth manifold structure on the n-torus $\mathbb{T}^n = \mathbb{S}^1 \times \cdots \times \mathbb{S}^1$.

In each of the examples we have seen so far, we have constructed a smooth manifold structure in two stages: We started with a topological space and checked that it was a topological manifold, and then we specified a smooth structure. It is often more convenient to combine these two steps into a single construction, especially if we start with a set that is not already equipped with a topology. The following lemma provides a shortcut.

Lemma 1.23 (Smooth Manifold Construction Lemma). *Let M be a set, and suppose we are given a collection $\{U_\alpha\}$ of subsets of M, together with an injective map $\varphi_\alpha\colon U_\alpha \to \mathbb{R}^n$ for each α, such that the following properties are satisfied:*

(i) *For each α, $\varphi_\alpha(U_\alpha)$ is an open subset of $\mathbb{R}^n$.*

(ii) *For each α and β, $\varphi_\alpha(U_\alpha \cap U_\beta)$ and $\varphi_\beta(U_\alpha \cap U_\beta)$ are open in $\mathbb{R}^n$.*

(iii) *Whenever $U_\alpha \cap U_\beta \neq \varnothing$, $\varphi_\alpha \circ \varphi_\beta^{-1}\colon \varphi_\beta(U_\alpha \cap U_\beta) \to \varphi_\alpha(U_\alpha \cap U_\beta)$ is a diffeomorphism.*

(iv) *Countably many of the sets U_α cover M.*

(v) *Whenever p, q are distinct points in M, either there exists some U_α containing both p and q or there exist disjoint sets U_α, U_β with $p \in U_\alpha$ and $q \in U_\beta$.*

Then M has a unique smooth manifold structure such that each $(U_\alpha, \varphi_\alpha)$ is a smooth chart.

Proof. We define the topology by taking all sets of the form $\varphi_\alpha^{-1}(V)$, with V an open subset of $\mathbb{R}^n$, as a basis. To prove that this is a basis for a topology, we need to show that for any point p in the intersection of two basis sets $\varphi_\alpha^{-1}(V)$ and $\varphi_\beta^{-1}(W)$, there is a third basis set containing p and contained in the intersection. It suffices to show that $\varphi_\alpha^{-1}(V) \cap \varphi_\beta^{-1}(W)$ is itself a basis set (Figure 1.9). To see this, observe that (iii) implies that $\varphi_\alpha \circ \varphi_\beta^{-1}(W)$ is an open subset of $\varphi_\alpha(U_\alpha \cap U_\beta)$, and (ii) implies that this set is also open in $\mathbb{R}^n$. It follows that

$$\varphi_\alpha^{-1}(V) \cap \varphi_\beta^{-1}(W) = \varphi_\alpha^{-1}\left(V \cap \varphi_\alpha \circ \varphi_\beta^{-1}(W)\right)$$

is also a basis set, as claimed.

Each of the maps φ_α is then a homeomorphism (essentially by definition), so M is locally Euclidean of dimension n. If $\{U_{\alpha_i}\}$ is a countable collection of the sets U_α covering M, each of the sets U_{α_i} has a countable basis, and

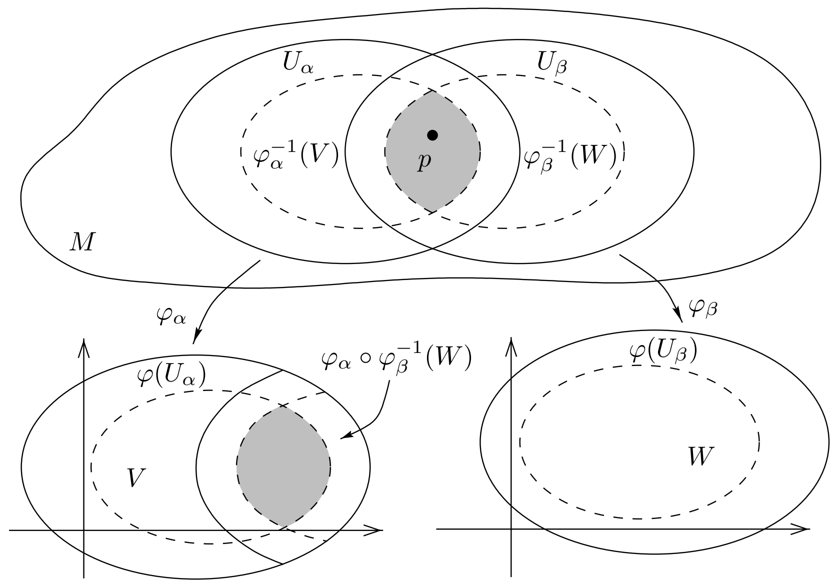


Figure 1.9. The smooth manifold construction lemma.

the union of all these is a countable basis for M, so M is second countable, and the Hausdorff property follows easily from (v). Finally, (iii) guarantees that the collection $\{(U_\alpha, \varphi_\alpha)\}$ is a smooth atlas. It is clear that this topology and smooth structure are the unique ones satisfying the conclusions of the lemma. □

Example 1.24 (Grassmann Manifolds). Let V be an n-dimensional real vector space. For any integer $0 \leq k \leq n$, we let $\mathrm{G}_k(V)$ denote the set of all k-dimensional linear subspaces of V. We will show that $\mathrm{G}_k(V)$ can be naturally given the structure of a smooth manifold of dimension $k(n-k)$. The construction is somewhat more involved than the ones we have done so far, but the basic idea is just to use linear algebra to construct charts for $\mathrm{G}_k(V)$, and then apply the smooth manifold construction lemma (Lemma 1.23). Since we will give a more straightforward proof that $\mathrm{G}_k(V)$ is a smooth manifold in Chapter 9 (Example 9.32), you may wish to skip the hard part of this construction (the verification that the charts are smoothly compatible) on first reading.

Let P and Q be any complementary subspaces of V of dimensions k and $(n-k)$, respectively, so that V decomposes as a direct sum: $V = P \oplus Q$. The graph of any linear map $A\colon P \to Q$ is a k-dimensional subspace $\Gamma(A) \subset V$, defined by

$$\Gamma(A) = \{x + Ax : x \in P\}.$$

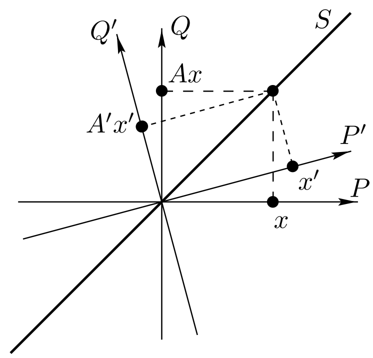


Figure 1.10. Smooth compatibility of coordinates on $\mathrm{G}_k(V)$.

Any such subspace has the property that its intersection with Q is the zero subspace. Conversely, any subspace with this property is easily seen to be the graph of a unique linear map $A\colon P \to Q$.

Let $L(P,Q)$ denote the vector space of linear maps from P to Q, and let U_Q denote the subset of $\mathrm{G}_k(V)$ consisting of k-dimensional subspaces whose intersection with Q is trivial. Define a map $\psi\colon L(P,Q) \to U_Q$ by

$$\psi(A) = \Gamma(A).$$

The discussion above shows that ψ is a bijection. Let $\varphi = \psi^{-1}\colon U_Q \to L(P,Q)$. By choosing bases for P and Q, we can identify $L(P,Q)$ with $\mathrm{M}((n-k)\times k, \mathbb{R})$ and hence with $\mathbb{R}^{k(n-k)}$, and thus we can think of (U_Q, φ) as a coordinate chart. Since the image of each chart is all of $L(P,Q)$, condition (i) of Lemma 1.23 is clearly satisfied.

Now let (P', Q') be any other such pair of subspaces, and let ψ', φ' be the corresponding maps. The set $\varphi(U_Q \cap U_{Q'}) \subset L(P,Q)$ consists of all $A \in L(P,Q)$ whose graphs intersect Q' trivially, which is easily seen to be an open set, so (ii) holds. We need to show that the transition map $\varphi' \circ \varphi^{-1} = \varphi' \circ \psi$ is smooth on this set. This is the trickiest part of the argument.

Suppose $A \in \varphi(U_Q \cap U_{Q'}) \subset L(P,Q)$ is arbitrary, and let S denote the subspace $\psi(A) = \Gamma(A) \subset V$. If we put $A' = \varphi' \circ \psi(A)$, then by definition A' is the unique linear map from P' to Q' whose graph is equal to S. To identify this map, let $x' \in P'$ be arbitrary, and note that $A'x'$ is the unique element of Q' such that $x' + A'x' \in S$, which is to say that

$$x' + A'x' = x + Ax \quad \text{for some } x \in P. \tag{1.2}$$

(See Figure 1.10.) There is in fact a unique $x \in P$ for which this holds, characterized by the property that

$$x + Ax - x' \in Q'.$$

If we let $I_A\colon P \to V$ denote the map $I_A(x) = x + Ax$ and let $\pi_{P'}\colon V \to P'$ be the projection onto P' with kernel Q', then x satisfies

$$0 = \pi_{P'}(x + Ax - x') = \pi_{P'} \circ I_A(x) - x'.$$

As long as A stays in the open subset of linear maps whose graphs intersect Q' trivially, $\pi_{P'} \circ I_A\colon P \to P'$ is invertible, and thus we can solve this last equation for x to obtain $x = (\pi_{P'} \circ I_A)^{-1}(x')$. Therefore, A' is given in terms of A by

$$A'x' = I_A x - x' = I_A \circ (\pi_{P'} \circ I_A)^{-1}(x') - x'. \tag{1.3}$$

If we choose bases (E'_i) for P' and (F'_j) for Q', the columns of the matrix representation of A' are the components of $A'E'_i$. By (1.3), this can be written

$$A'E'_i = I_A \circ (\pi_{P'} \circ I_A)^{-1}(E'_i) - E'_i.$$

The matrix entries of I_A clearly depend smoothly on those of A, and thus so also do those of $\pi_{P'} \circ I_A$. By Cramer's rule, the components of the inverse of a matrix are rational functions of the matrix entries, so the expression above shows that the components of $A'E'_i$ depend smoothly on the components of A. This proves that $\varphi' \circ \varphi^{-1}$ is a smooth map, so the charts we have constructed satisfy condition (iii) of Lemma 1.23.

To check the countability condition (iv), we just note that $\mathrm{G}_k(V)$ can in fact be covered by *finitely* many of the sets U_Q: For example, if $(E_1, \dots, E_n)$ is any fixed basis for V, any partition of the basis elements into two subsets containing k and $n-k$ elements determines appropriate subspaces P and Q, and any subspace S must have trivial intersection with Q for at least one of these partitions (see Exercise A.34). Thus $\mathrm{G}_k(V)$ is covered by the finitely many charts determined by all possible partitions of a fixed basis. Finally, the Hausdorff condition (v) is easily verified by noting that for any two k-dimensional subspaces $P, P' \subset V$, it is possible to find a subspace Q of dimension $n-k$ whose intersections with both P and P' are trivial, and then P and P' are both contained in the domain of the chart determined by, say, (P, Q).

The smooth manifold $\mathrm{G}_k(V)$ is called the *Grassmann manifold* of k-planes in V, or simply a *Grassmannian.* In the special case $V = \mathbb{R}^n$, the Grassmannian $\mathrm{G}_k(\mathbb{R}^n)$ is often denoted by some simpler notation such as $\mathrm{G}_{k,n}$ or $\mathrm{G}(k,n)$. Note that $\mathrm{G}_1(\mathbb{R}^{n+1})$ is exactly the n-dimensional projective space $\mathbb{RP}^n$.

Manifolds with Boundary

In many important applications of manifolds, most notably those involving integration, we will encounter spaces that would be smooth manifolds except that they have a "boundary" of some sort. Simple examples of such

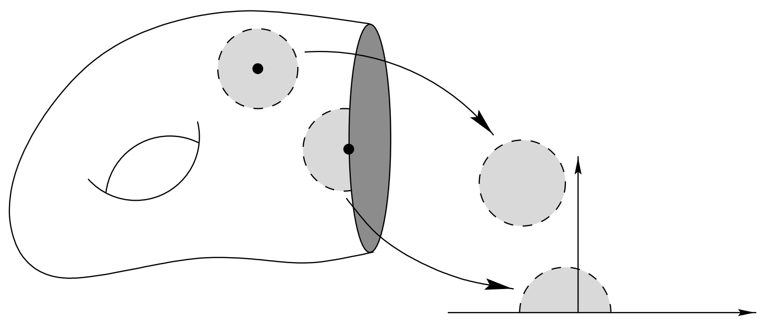


Figure 1.11. A manifold with boundary.

spaces include the closed unit ball in $\mathbb{R}^n$ and the closed upper hemisphere in $\mathbb{S}^n$. To accommodate such spaces, we need to generalize our definition of manifolds.

The model for these spaces will be the closed n-dimensional *upper half-space* $\mathbb{H}^n \subset \mathbb{R}^n$, defined as

$$\mathbb{H}^n = \left\{ \left(x^1, \dots, x^n\right) \in \mathbb{R}^n : x^n \geq 0 \right\}.$$

We will use $\operatorname{Int} \mathbb{H}^n$ and $\partial \mathbb{H}^n$ to denote the interior and boundary of $\mathbb{H}^n$, respectively, as a subset of $\mathbb{R}^n$:

$$\begin{aligned} \operatorname{Int} \mathbb{H}^n &= \left\{ \left(x^1, \dots, x^n\right) \in \mathbb{R}^n : x^n > 0 \right\}, \\ \partial \mathbb{H}^n &= \left\{ \left(x^1, \dots, x^n\right) \in \mathbb{R}^n : x^n = 0 \right\}. \end{aligned}$$

An n-dimensional *topological manifold with boundary* is a second-countable Hausdorff space M in which every point has a neighborhood homeomorphic to a (relatively) open subset of $\mathbb{H}^n$ (Figure 1.11). An open subset $U \subset M$ together with a homeomorphism φ from U to an open subset of $\mathbb{H}^n$ will be called a chart, just as in the case of manifolds. When it is necessary to make the distinction, we will call (U, φ) an *interior chart* if $\varphi(U) \subset \operatorname{Int} \mathbb{H}^n$, and a *boundary chart* if $\varphi(U) \cap \partial \mathbb{H}^n \neq \varnothing$.

To see how to define a smooth structure on a manifold with boundary, recall that a smooth map from an arbitrary subset $A \subset \mathbb{R}^n$ to $\mathbb{R}^k$ is defined to be a map that admits a smooth extension to an open neighborhood of each point (see the Appendix, page 587). Thus if U is an open subset of $\mathbb{H}^n$, a map $F \colon U \to \mathbb{R}^k$ is smooth if for each $x \in U$, there exists an open set $V \subset \mathbb{R}^n$ and a smooth map $\widetilde{F} \colon V \to \mathbb{R}^k$ that agrees with F on $V \cap \mathbb{H}^n$ (Figure 1.12). If F is such a map, the restriction of F to $U \cap \operatorname{Int} \mathbb{H}^n$ is smooth in the usual sense. By continuity, all the partial derivatives of F at points of $U \cap \partial \mathbb{H}^n$ are determined by their values in $\operatorname{Int} \mathbb{H}^n$, and therefore in particular are independent of the choice of extension. It is a fact (which we will neither prove nor use) that $F \colon U \to \mathbb{R}^k$ is smooth in this sense if and

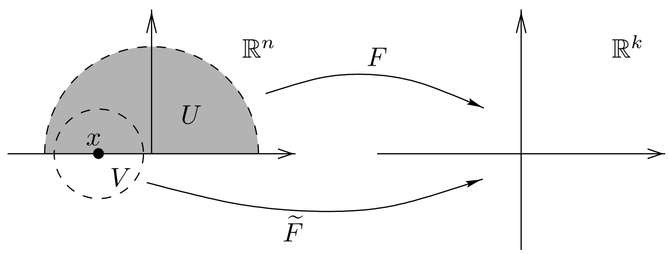


Figure 1.12. Smoothness of maps on open subsets of $\mathbb{H}^n$.

only if F is continuous, $F|_{U\cap \operatorname{Int}\mathbb{H}^n}$ is smooth, and the partial derivatives of $F|_{U\cap \operatorname{Int}\mathbb{H}^n}$ of all orders have continuous extensions to all of U.

For example, let $\mathbb{B}^2 \subset \mathbb{R}^2$ be the open unit disk, let $U = \mathbb{B}^2 \cap \mathbb{H}^2$, and define $f\colon U \to \mathbb{R}$ by $f(x,y) = \sqrt{1-x^2-y^2}$. Because f extends smoothly to all of $\mathbb{B}^2$ (by the same formula), f is a smooth function on U. On the other hand, although $g(x,y) = \sqrt{y}$ is continuous on U and smooth in $U \cap \operatorname{Int}\mathbb{H}^2$, it has no smooth extension to any neighborhood of the origin in $\mathbb{R}^2$ because $\partial g/\partial y \to \infty$ as $y \to 0$. Thus g is not smooth on U.

Now let M be a topological manifold with boundary. Just as in the manifold case, a smooth structure for M is defined to be a maximal smooth atlas—a collection of charts whose domains cover M and whose transition maps (and their inverses) are smooth in the sense just described. With such a structure, M is called a *smooth manifold with boundary.* A point $p \in M$ is called a *boundary point* if its image under some smooth chart is in $\partial\mathbb{H}^n$, and an *interior point* if its image under some smooth chart is in $\operatorname{Int}\mathbb{H}^n$. The *boundary* of M (the set of all its boundary points) is denoted by ∂M; similarly, its *interior*, the set of all its interior points, is denoted by $\operatorname{Int} M$. Once we have developed a bit more machinery, you will be able to show that M is the disjoint union of ∂M and $\operatorname{Int} M$ (see Problem 7-7).

Be careful to observe the distinction between these new definitions of the terms "boundary" and "interior" and their usage to refer to the boundary and interior of a subset of a topological space. A manifold M with boundary may have nonempty boundary in this new sense, irrespective of whether it has a boundary as a subset of some other topological space. If we need to emphasize the difference between the two notions of boundary, we will use the terms *topological boundary* and *manifold boundary* as appropriate. For example, the closed unit disk $\overline{\mathbb{B}^2}$ is a smooth manifold with boundary (as you will be asked to show in Problem 1-9), whose manifold boundary is the circle. Its topological boundary as a subspace of $\mathbb{R}^2$ happens to be the circle as well. However, if we think of $\overline{\mathbb{B}^2}$ as a topological space in its own right, then as a subset of itself, it has empty topological boundary. And if we think of it as a subset of $\mathbb{R}^3$ (considering $\mathbb{R}^2$ as a subset of $\mathbb{R}^3$ in the obvious way), its topological boundary is all of $\overline{\mathbb{B}^2}$. Note that $\mathbb{H}^n$ is itself a smooth manifold with boundary, and its manifold boundary is the same as its topological boundary as a subset of $\mathbb{R}^n$.

Every smooth n-manifold can be considered as a smooth n-manifold with boundary in a natural way: By composing with a diffeomorphism from $\mathbb{R}^n$ to $\mathbb{H}^n$ such as $\left(x^1, \dots, x^{n-1}, x^n\right) \mapsto \left(x^1, \dots, x^{n-1}, e^{x^n}\right)$, we can modify any manifold chart to take its values in $\operatorname{Int} \mathbb{H}^n$ without affecting the smooth compatibility condition. On the other hand, if M is a smooth n-manifold with boundary, any interior point $p \in \operatorname{Int} M$ is by definition in the domain of a smooth chart (U, φ) such that $\varphi(p) \in \operatorname{Int} \mathbb{H}^n$. Replacing U by the (possibly smaller) open set $\varphi^{-1}(\operatorname{Int} \mathbb{H}^n) \subset U$, we may assume that (U, φ) is an interior chart. Because open sets in $\operatorname{Int} \mathbb{H}^n$ are also open in $\mathbb{R}^n$, each interior chart is a chart in the ordinary manifold sense. Thus $\operatorname{Int} M$ is a topological n-manifold, and the set of all smooth interior charts is easily seen to be a smooth atlas, turning it into a smooth n-manifold. In particular, a smooth manifold with boundary whose boundary happens to be empty is a smooth manifold. However, manifolds with boundary are not manifolds in general.

Even though the term "manifold with boundary" encompasses manifolds as well, for emphasis we will sometimes use the phrase "manifold without boundary" when we are talking about manifolds in the original sense, and "manifold with or without boundary" when we are working in the broader class that includes both cases. In the literature, you will also encounter the terms *closed manifold* to mean a compact manifold without boundary, and *open manifold* to mean a noncompact manifold without boundary.

The topological properties of manifolds that we proved earlier in the chapter have natural extensions to manifolds with boundary. For the record, we state them here.

Proposition 1.25. *Let M be a topological manifold with boundary.*

(*a*) *M is locally path connected.*

(*b*) *M has at most countably many components, each of which is a connected topological manifold with boundary.*

(*c*) *The fundamental group of M is countable.*

◇ **Exercise 1.6.** Prove Proposition 1.25.

Many of the results that we will prove about smooth manifolds throughout the book have natural analogues for manifolds with boundary. We will mention the most important of these as we go along.

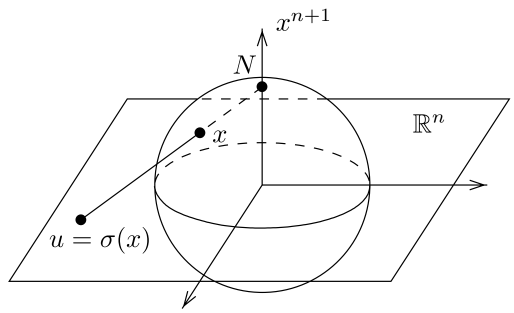


Figure 1.13. Stereographic projection.

Problems

1-1. Let X be the set of all points $(x, y) \in \mathbb{R}^2$ such that $y = \pm 1$, and let M be the quotient of X by the equivalence relation generated by $(x, -1) \sim (x, 1)$ for all $x \neq 0$. Show that M is locally Euclidean and second countable, but not Hausdorff. (This space is called the *line with two origins*.)

1-2. Show that the disjoint union of uncountably many copies of $\mathbb{R}$ is locally Euclidean and Hausdorff, but not second countable.

1-3. Let M be a nonempty topological manifold of dimension $n \geq 1$. If M has a smooth structure, show that it has uncountably many distinct ones. [Hint: Begin by constructing homeomorphisms from $\mathbb{B}^n$ to itself that are smooth on $\mathbb{B}^n \smallsetminus \{0\}$.]

1-4. If k is an integer between 0 and $\min(m, n)$, show that the set of $m \times n$ matrices whose rank is at least k is an open submanifold of $\mathrm{M}(m \times n, \mathbb{R})$. Show that this is *not* true if "at least k" is replaced by "equal to k."

1-5. Let $N = (0, \dots, 0, 1)$ be the "north pole" in $\mathbb{S}^n \subset \mathbb{R}^{n+1}$, and let $S = -N$ be the "south pole." Define *stereographic projection* $\sigma\colon \mathbb{S}^n \smallsetminus \{N\} \to \mathbb{R}^n$ by

$$\sigma\left(x^1, \dots, x^{n+1}\right) = \frac{\left(x^1, \dots, x^n\right)}{1 - x^{n+1}}.$$

Let $\widetilde{\sigma}(x) = -\sigma(-x)$ for $x \in \mathbb{S}^n \smallsetminus \{S\}$.

(a) For any $x \in \mathbb{S}^n \smallsetminus \{N\}$, show that $\sigma(x)$ is the point where the line through N and x intersects the linear subspace where $x^{n+1} = 0$, identified with $\mathbb{R}^n$ in the obvious way (Figure 1.13). Similarly, show that $\widetilde{\sigma}(x)$ is the point where the line through S

and x intersects the same subspace. (For this reason, $\widetilde{\sigma}$ is called *stereographic projection from the south pole.*)

(b) Show that σ is bijective, and

$$\sigma^{-1}\left(u^1,\dots,u^n\right)=\frac{\left(2u^1,\dots,2u^n,|u|^2-1\right)}{|u|^2+1}.$$

(c) Compute the transition map $\widetilde{\sigma}\circ\sigma^{-1}$ and verify that the atlas consisting of the two charts $(\mathbb{S}^n\smallsetminus\{N\},\sigma)$ and $(\mathbb{S}^n\smallsetminus\{S\},\widetilde{\sigma})$ defines a smooth structure on $\mathbb{S}^n$. (The coordinates defined by σ or $\widetilde{\sigma}$ are called *stereographic coordinates.*)

(d) Show that this smooth structure is the same as the one defined in Example 1.20.

1-6. By identifying $\mathbb{R}^2$ with $\mathbb{C}$ in the usual way, we can think of the unit circle $\mathbb{S}^1$ as a subset of the complex plane. An *angle function* on a subset $U\subset\mathbb{S}^1$ is a continuous function $\theta\colon U\to\mathbb{R}$ such that $e^{i\theta(p)}=p$ for all $p\in U$. Show that there exists an angle function θ on an open subset $U\subset\mathbb{S}^1$ if and only if $U\neq\mathbb{S}^1$. For any such angle function, show that (U,θ) is a smooth coordinate chart for $\mathbb{S}^1$ with its standard smooth structure.

1-7. *Complex projective n-space*, denoted by $\mathbb{CP}^n$, is the set of 1-dimensional complex-linear subspaces of $\mathbb{C}^{n+1}$, with the quotient topology inherited from the natural projection $\pi\colon\mathbb{C}^{n+1}\smallsetminus\{0\}\to\mathbb{CP}^n$. Show that $\mathbb{CP}^n$ is a compact $2n$-dimensional topological manifold, and show how to give it a smooth structure analogous to the one we constructed for $\mathbb{RP}^n$. (We identify $\mathbb{C}^{n+1}$ with $\mathbb{R}^{2n+2}$ via $\left(x^1+iy^1,\dots,x^{n+1}+iy^{n+1}\right)\leftrightarrow\left(x^1,y^1,\dots,x^{n+1},y^{n+1}\right)$.)

1-8. Let k and n be integers such that $0<k<n$, and let $P,Q\subset\mathbb{R}^n$ be the subspaces spanned by $(e_1,\dots,e_k)$ and $(e_{k+1},\dots,e_n)$, respectively, where e_i is the ith standard basis vector. For any k-dimensional subspace $S\subset\mathbb{R}^n$ that has trivial intersection with Q, show that the coordinate representation $\varphi(S)$ constructed in Example 1.24 is the unique $(n-k)\times k$ matrix B such that S is spanned by the columns of the matrix $\binom{I_k}{B}$, where I_k denotes the $k\times k$ identity matrix.

1-9. Let $M=\overline{\mathbb{B}^n}$, the closed unit ball in $\mathbb{R}^n$. Show that M is a topological manifold with boundary, and that it can be given a natural smooth structure in which each point in $\mathbb{S}^{n-1}$ is a boundary point and each point in $\mathbb{B}^n$ is an interior point.

2
Smooth Maps

The main reason for introducing smooth structures was to enable us to define smooth functions on manifolds and smooth maps between manifolds. In this chapter, we carry out that project.

Although the terms "function" and "map" are technically synonymous, when studying smooth manifolds it is often convenient to make a slight distinction between them. Throughout this book, we will generally reserve the term "function" for a map whose range is $\mathbb{R}$ (a *real-valued function*) or $\mathbb{R}^k$ for some $k > 1$ (a *vector-valued function*). The word "map" or "mapping" can mean any type of map, such as a map between arbitrary manifolds.

We begin by defining smooth real-valued and vector-valued functions, and then generalize this to smooth maps between manifolds. We then study diffeomorphisms, which are bijective smooth maps with smooth inverses. If there is a diffeomorphism between two smooth manifolds, we say they are diffeomorphic. The main objects of study in smooth manifold theory are properties that are invariant under diffeomorphisms.

Later in the chapter, we introduce Lie groups, which are smooth manifold that are also groups in which multiplication and inversion are smooth maps, and we study smooth covering maps and their relationship to the continuous covering maps studied in topology.

At the end of the chapter, we introduce a powerful tool for smoothly piecing together local smooth objects, called partitions of unity. They will be used throughout the book for building global smooth objects out of ones that are initially defined only locally.

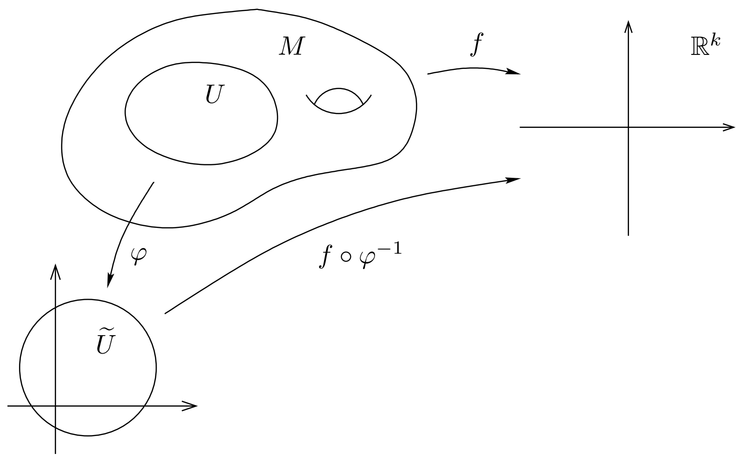


Figure 2.1. Definition of smooth functions.

Smooth Functions and Smooth Maps

If M is a smooth n-manifold, a function $f\colon M\to\mathbb{R}^k$ is said to be *smooth* if for every $p\in M$, there exists a smooth chart (U,φ) for M whose domain contains p and such that the composite function $f\circ\varphi^{-1}$ is smooth on the open subset $\widetilde{U}=\varphi(U)\subset\mathbb{R}^n$ (Figure 2.1). The most important special case is that of smooth real-valued functions $f\colon M\to\mathbb{R}$; the set of all such functions is denoted by $C^\infty(M)$. Because sums and constant multiples of smooth functions are smooth, $C^\infty(M)$ is a vector space.

◇ **Exercise 2.1.** Show that pointwise multiplication turns $C^\infty(M)$ into a commutative ring and a commutative and associative algebra over $\mathbb{R}$ (see the Appendix, page 564, for the definition of an algebra).

◇ **Exercise 2.2.** Let $U\subset\mathbb{R}^n$ be an open set with its usual smooth manifold structure. Show that a map $f\colon U\to\mathbb{R}^k$ is smooth in the sense just defined if and only if it is smooth in the sense of ordinary calculus.

◇ **Exercise 2.3.** Suppose M is a smooth manifold and $f\colon M\to\mathbb{R}^k$ is a smooth function. Show that $f\circ\varphi^{-1}\colon\varphi(U)\to\mathbb{R}^k$ is smooth for *every* smooth chart (U,φ) for M.

Given a function $f\colon M\to\mathbb{R}^k$ and a chart (U,φ) for M, the function $\widehat{f}\colon\varphi(U)\to\mathbb{R}^k$ defined by $\widehat{f}(x)=f\circ\varphi^{-1}(x)$ is called the *coordinate representation* of f. By definition, f is smooth if and only if its coordinate representation is smooth in some smooth chart around each point. By the preceding exercise, smooth maps have smooth coordinate representations in *every* smooth chart.

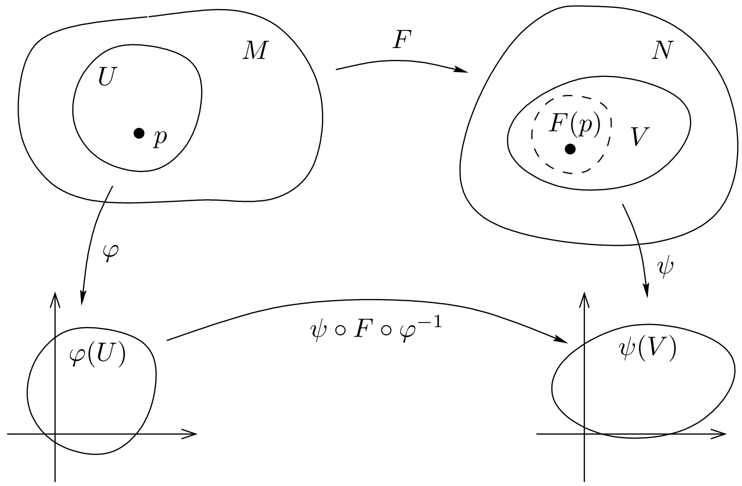


Figure 2.2. Definition of smooth maps.

For example, consider the real-valued function $f(x,y) = x^2 + y^2$ on the plane. In polar coordinates on the set $U = \{(x,y) : x > 0\}$, it has the coordinate representation $\widehat{f}(r,\theta) = r^2$. In keeping with our practice of using local coordinates to identify an open subset of a manifold with an open subset of Euclidean space, in cases where it will cause no confusion we will often not even observe the distinction between $\widehat{f}$ and f itself, and write $f(r,\theta) = r^2$ in polar coordinates. Thus we might say, "f is smooth on U because its coordinate representation $f(r,\theta) = r^2$ is smooth."

The definition of smooth functions generalizes easily to maps between manifolds. Let M, N be smooth manifolds, and let $F\colon M \to N$ be any map. We say F is a *smooth map* if for any $p \in M$, there exist smooth charts (U,φ) containing p and (V,ψ) containing $F(p)$ such that $F(U) \subset V$ and the composite map $\psi \circ F \circ \varphi^{-1}$ is smooth from $\varphi(U)$ to $\psi(V)$ (Figure 2.2). Note that our previous definition of smoothness of real-valued functions can be viewed as a special case of this one, by taking $N = V = \mathbb{R}^k$ and $\psi = \mathrm{Id}\colon \mathbb{R}^k \to \mathbb{R}^k$.

◇ **Exercise 2.4 (Smoothness Is Local).** Let M and N be smooth manifolds, and let $F\colon M \to N$ be a map. If every point $p \in M$ has a neighborhood U such that the restriction $F|_U$ is smooth, show that F is smooth. Conversely, if F is smooth, show that its restriction to any open subset is smooth.

The next lemma is really just a restatement of the previous exercise, but it gives a highly useful way of constructing smooth maps.

Lemma 2.1. *Let M and N be smooth manifolds, and let $\{U_\alpha\}_{\alpha\in A}$ be an open cover of M. Suppose that for each $\alpha \in A$, we are given a smooth map $F_\alpha\colon U_\alpha \to N$ such that the maps agree on overlaps: $F_\alpha|_{U_\alpha\cap U_\beta} = F_\beta|_{U_\alpha\cap U_\beta}$ for all α and β. Then there exists a unique smooth map $F\colon M \to N$ such that $F|_{U_\alpha} = F_\alpha$ for each $\alpha \in A$.*

◇ **Exercise 2.5.** Prove the preceding lemma.

One important observation about our definition of smooth maps is that, as one might expect, smoothness implies continuity.

Lemma 2.2. *Every smooth map between smooth manifolds is continuous.*

Proof. Suppose $F\colon M \to N$ is smooth. For each $p \in M$, the definition of smoothness guarantees that we can choose smooth charts (U, φ) containing p and (V, ψ) containing $F(p)$ such that $F(U) \subset V$ and $\psi \circ F \circ \varphi^{-1}\colon \varphi(U) \to \varphi(V)$ is a smooth map, hence continuous. Since $\varphi\colon U \to \varphi(U)$ and $\psi\colon V \to \psi(V)$ are homeomorphisms, this implies in turn that

$$F|_U = \psi^{-1} \circ (\psi \circ F \circ \varphi^{-1}) \circ \varphi\colon U \to V,$$

which is a composition of continuous maps. Since F is continuous in a neighborhood of each point, it is continuous on M. □

If $F\colon M \to N$ is a smooth map, and (U, φ) and (V, ψ) are any smooth charts for M and N, respectively, we call $\widehat{F} = \psi \circ F \circ \varphi^{-1}$ the coordinate representation of F with respect to the given coordinates.

◇ **Exercise 2.6.** Suppose $F\colon M \to N$ is a smooth map between smooth manifolds. Show that the coordinate representation of F with respect to *any* pair of smooth charts for M and N is smooth.

As with real-valued or vector-valued functions, once we have chosen specific local coordinates in both the domain and range, we can often ignore the distinction between F and $\widehat{F}$.

To prove that a map $F\colon M \to N$ is smooth directly from the definition requires, in part, that for each $p \in M$ we prove the existence of coordinate domains U containing p and V containing $F(p)$ such that $F(U) \subset V$. This requirement is included in the definition precisely so that smoothness will automatically imply continuity as in Lemma 2.2. However, if F is known a priori to be continuous, then smoothness can be checked somewhat more easily by examining its coordinate representations in the charts of particular smooth atlases for M and N, as the next lemma shows.

Lemma 2.3. *Let M and N be smooth manifolds, and let $F\colon M \to N$ be a continuous map. If $\{(U_\alpha, \varphi_\alpha)\}$ and $\{(V_\beta, \psi_\beta)\}$ are smooth atlases for M and N, respectively, and if for each α and β, $\psi_\beta \circ F \circ \varphi_\alpha^{-1}$ is smooth on its domain of definition, then F is smooth.*

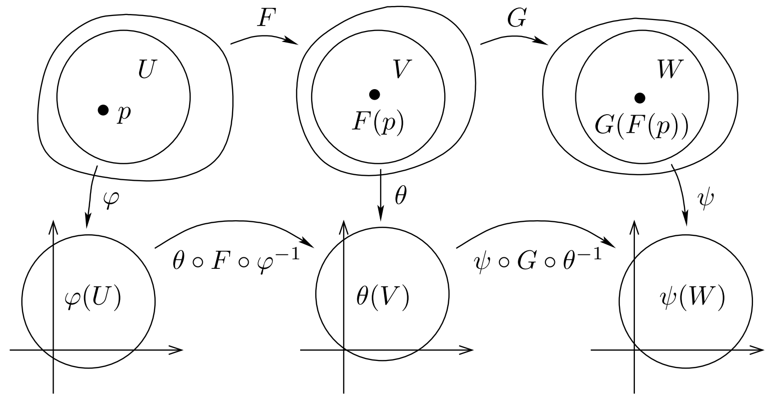


Figure 2.3. A composition of smooth maps is smooth.

Proof. Given $p \in M$, choose a pair of charts $(U_\alpha, \varphi_\alpha)$ and (V_β, ψ_β) from the given atlases such that $p \in U_\alpha$ and $F(p) \in V_\beta$. By continuity of F, the set $U = F^{-1}(V_\beta) \cap U_\alpha$ is open in M, and $F(U) \subset V_\beta$. Therefore the charts $(U, \varphi_\alpha|_U)$ and (V_β, ψ_β) satisfy the conditions required in the definition of smoothness. □

Lemma 2.4. *Any composition of smooth maps between smooth manifolds is smooth.*

Proof. Let $F\colon M \to N$ and $G\colon N \to P$ be smooth maps, and let $p \in M$ be arbitrary. By definition of smoothness of G, there exist smooth charts (V, θ) containing $F(p)$ and (W, ψ) containing $G(F(p))$ such that $G(V) \subset W$ and $\psi \circ G \circ \theta^{-1}\colon \theta(V) \to \psi(W)$ is smooth. Since F is continuous, $F^{-1}(V)$ is an open neighborhood of p in M, so there is a smooth chart (U, φ) for M such that $p \in U \subset F^{-1}(V)$ (Figure 2.3). By Exercise 2.6, $\theta \circ F \circ \varphi^{-1}$ is smooth from $\varphi(U)$ to $\theta(V)$. Then we have $G \circ F(U) \subset G(V) \subset W$, and $\psi \circ (G \circ F) \circ \varphi^{-1} = (\psi \circ G \circ \theta^{-1}) \circ (\theta \circ F \circ \varphi^{-1})\colon \varphi(U) \to \psi(W)$ is smooth because it is a composition of smooth maps between open subsets of Euclidean spaces. □

Although most of our efforts in this book will be devoted to the study of smooth manifolds and smooth maps, we will also need to work with topological manifolds and continuous maps on occasion. For the sake of consistency, we adopt the following convention: Without further qualification, a "manifold" will always be understood to be a topological manifold, and a "coordinate chart" will be understood in the topological sense, as a homeomorphism from an open set in the manifold to an open set in Euclidean space. Similarly, our default assumption for most functions, maps, and other geometric objects will be merely continuity; smoothness will not be assumed unless explicitly specified. This convention requires a certain discipline, in that we have to remember to state the smoothness hypothesis whenever it is needed; but its advantage is that it frees us from having to remember which types of maps are assumed to be smooth and which are not. The main exceptions will be a few concepts that require smoothness for their very definitions.

We now have enough information to produce a number of interesting examples of smooth maps. In spite of the apparent complexity of the definition, it is usually not hard to prove that a particular map is smooth. There are basically only three common ways to do so:

- Write the map in smooth local coordinates and recognize its component functions as compositions of smooth elementary functions;
- Exhibit the map as a composition of known smooth maps; or
- Apply some special-purpose theorem that applies to the particular case under consideration.

Example 2.5 (Smooth maps).

(a) Any map from a zero-dimensional manifold into a smooth manifold is automatically smooth.

(b) Consider the n-sphere $\mathbb{S}^n$ with its standard smooth structure. The inclusion map $\iota\colon \mathbb{S}^n \hookrightarrow \mathbb{R}^{n+1}$ is certainly continuous, because it is the inclusion map of a topological subspace. It is a smooth map because its coordinate representation with respect to any of the graph coordinates of Example 1.20 is

$$\begin{aligned}\widehat{\iota}(u^1,\dots,u^n) &= \iota\circ(\varphi_i^{\pm})^{-1}(u^1,\dots,u^n)\\ &= \left(u^1,\dots,u^{i-1},\pm\sqrt{1-|u|^2},u^i,\dots,u^n\right),\end{aligned}$$

which is smooth on its domain (the set where $|u|^2<1$).

(c) The quotient map $\pi\colon \mathbb{R}^{n+1}\smallsetminus\{0\}\to\mathbb{RP}^n$ is smooth, because its coordinate representation in terms of any of the coordinates for $\mathbb{RP}^n$ constructed in Example 1.21 and standard coordinates on $\mathbb{R}^{n+1}\smallsetminus\{0\}$ is

$$\begin{aligned}\widehat{\pi}(x^1,\dots,x^{n+1}) &= \varphi_i\circ\pi(x^1,\dots,x^{n+1}) = \varphi_i[x^1,\dots,x^{n+1}]\\ &= \left(\frac{x^1}{x^i},\dots,\frac{x^{i-1}}{x^i},\frac{x^{i+1}}{x^i},\dots,\frac{x^{n+1}}{x^i}\right).\end{aligned}$$

(d) Define $p\colon\mathbb{S}^n\to\mathbb{RP}^n$ as the restriction of $\pi\colon\mathbb{R}^{n+1}\smallsetminus\{0\}\to\mathbb{RP}^n$ to $\mathbb{S}^n\subset\mathbb{R}^{n+1}\smallsetminus\{0\}$. It is a smooth map, because it is the composition $p=\pi\circ\iota$ of the maps in the preceding two examples. Because p is surjective, $\mathbb{RP}^n$ is the continuous image of a compact set and therefore compact.

◇ **Exercise 2.7.** Let $M_1, \dots, M_k$ and N be smooth manifolds. Show that a map $F\colon N \to M_1 \times \cdots \times M_k$ is smooth if and only if each of the component maps $F_i = \pi_i \circ F\colon N \to M_i$ is smooth. (Here $\pi_i\colon M_1 \times \cdots \times M_k \to M_i$ is the projection onto the ith factor.)

We can also define smooth maps to and from smooth manifolds with boundary, with the understanding that a map whose domain is a subset of the half space $\mathbb{H}^n$ is smooth if it admits an extension to a smooth map in an open neighborhood of each point, and a map whose range is $\mathbb{H}^n$ is smooth if it is smooth as a map into $\mathbb{R}^n$. You can work out the details for yourself.

Diffeomorphisms

A *diffeomorphism* between smooth manifolds M and N is a smooth bijective map $F\colon M \to N$ that has a smooth inverse. We say M and N are *diffeomorphic* if there exists a diffeomorphism between them. Sometimes this is symbolized by $M \approx N$. For example, if $\mathbb{B}^n$ denotes the open unit ball in $\mathbb{R}^n$, the map $F\colon \mathbb{B}^n \to \mathbb{R}^n$ given by $F(x) = x/(1 - |x|^2)$ is easily seen to be a diffeomorphism, so $\mathbb{B}^n \approx \mathbb{R}^n$. If M is any smooth manifold and (U, φ) is a smooth coordinate chart on M, then $\varphi\colon U \to \varphi(U) \subset \mathbb{R}^n$ is a diffeomorphism. (In fact, its coordinate representation is the identity.)

◇ **Exercise 2.8.** Show that "diffeomorphic" is an equivalence relation.

More generally, $F\colon M \to N$ is called a *local diffeomorphism* if every point $p \in M$ has a neighborhood U such that $F(U)$ is open in N and $F\colon U \to F(U)$ is a diffeomorphism. It is clear from the definition that a local diffeomorphism is, in particular, a local homeomorphism and therefore an open map.

◇ **Exercise 2.9.** Show that a map between smooth manifolds is a diffeomorphism if and only if it is a bijective local diffeomorphism.

Just as two topological spaces are considered to be "the same" if they are homeomorphic, two smooth manifolds are essentially indistinguishable if they are diffeomorphic. The central concern of smooth manifold theory is the study of properties of smooth manifolds that are preserved by diffeomorphisms.

It is very natural to wonder whether the smooth structure on a given topological manifold is unique in some sense. We observed in Example 1.12 that every zero-dimensional manifold has a unique smooth structure. However, as Problem 1-3 showed, any positive-dimensional topological manifold M admits many distinct smooth structures.

A more subtle and interesting question is whether a given topological manifold admits smooth structures that are not diffeomorphic to each other. For example, let $\widetilde{\mathbb{R}}$ denote the topological manifold $\mathbb{R}$, but endowed with the smooth structure described in Example 1.14 (defined by the global chart $\psi(x) = x^3$). It turns out that $\widetilde{\mathbb{R}}$ is diffeomorphic to $\mathbb{R}$ with its standard smooth structure. Define a map $F\colon \mathbb{R} \to \widetilde{\mathbb{R}}$ by $F(x) = x^{1/3}$. The coordinate representation of this map is $\widehat{F}(t) = \psi \circ F \circ \mathrm{Id}_{\mathbb{R}}^{-1}(t) = t$, which is clearly smooth. Moreover, the coordinate representation of its inverse is $\widehat{F^{-1}}(y) = \mathrm{Id}_{\mathbb{R}} \circ F^{-1} \circ \psi^{-1}(y) = y$, which is also smooth, so F is a diffeomorphism. (This is one case in which it *is* important to maintain the distinction between a map and its coordinate representation!)

It turns out, as you will see later, that there is only one smooth structure on $\mathbb{R}$ up to diffeomorphism (see Problem 17-8). More precisely, if $\mathcal{A}_1$ and $\mathcal{A}_2$ are any two smooth structures on $\mathbb{R}$, there exists a diffeomorphism $F\colon (\mathbb{R}, \mathcal{A}_1) \to (\mathbb{R}, \mathcal{A}_2)$. In fact, it follows from work of James Munkres [Mun60] and Edwin Moise [Moi77] that every topological manifold of dimension less than or equal to 3 has a smooth structure that is unique up to diffeomorphism. The analogous question in higher dimensions turns out to be quite deep, and is still largely unanswered. Even for Euclidean spaces, the problem was not completely solved until late in the twentieth century. The answer is somewhat surprising: As long as $n \neq 4$, $\mathbb{R}^n$ has a unique smooth structure (up to diffeomorphism); but $\mathbb{R}^4$ has uncountably many distinct smooth structures, no two of which are diffeomorphic to each other! The existence of nonstandard smooth structures on $\mathbb{R}^4$ (called *fake* $\mathbb{R}^4$*s*) was first proved by Simon Donaldson and Michael Freedman in 1984 as a consequence of their work on the geometry and topology of compact 4-manifolds; the results are described in [DK90] and [FQ90].

For compact manifolds, the situation is even more fascinating. For example, in 1963, Michel Kervaire and John Milnor [KM63] showed that, up to diffeomorphism, $\mathbb{S}^7$ has exactly 28 non-diffeomorphic smooth structures. On the other hand, in all dimensions greater than 3 there are compact topological manifolds that have no smooth structures at all. The problem of identifying the number of smooth structures (if any) on topological 4-manifolds is an active subject of current research.

Lie Groups

A *Lie group* is a smooth manifold G that is also a group in the algebraic sense, with the property that the multiplication map $m\colon G \times G \to G$ and inversion map $i\colon G \to G$, given by

$$m(g,h) = gh, \qquad i(g) = g^{-1},$$

are both smooth. Because smooth maps are continuous, a Lie group is, in particular, a *topological group* (a topological space with a group structure such that the multiplication and inversion maps are continuous).

The group operation in an arbitrary Lie group will be denoted by juxtaposition, except in certain abelian groups such as $\mathbb{R}^n$ in which the operation is usually written additively. It is traditional to denote the identity element of an arbitrary Lie group by the symbol e (for German *Einselement*, "unit element"), and we will follow this convention, except in specific examples in which there are more common notations (such as I_n for the identity matrix in a matrix group, or 0 for the identity element in $\mathbb{R}^n$).

The following alternative characterization of the smoothness condition is sometimes useful.

Lemma 2.6. *If G is a smooth manifold with a group structure such that the map $G \times G \to G$ given by $(g, h) \mapsto gh^{-1}$ is smooth, then G is a Lie group.*

◇ **Exercise 2.10.** Prove Lemma 2.6.

Example 2.7 (Lie Groups). Each of the following manifolds is a Lie group with the indicated group operation.

(a) The *general linear group* $\mathrm{GL}(n, \mathbb{R})$ is the set of invertible $n \times n$ matrices with real entries. It is a group under matrix multiplication, and it is an open submanifold of the vector space $\mathrm{M}(n, \mathbb{R})$, as we observed in Chapter 1. Multiplication is smooth because the matrix entries of a product matrix AB are polynomials in the entries of A and B. Inversion is smooth because Cramer's rule expresses the entries of A^{-1} as rational functions of the entries of A.

(b) The *complex general linear group* $\mathrm{GL}(n, \mathbb{C})$ is the group of complex $n \times n$ matrices under matrix multiplication. It is an open submanifold of $\mathrm{M}(n, \mathbb{C})$ and thus a $2n^2$-dimensional smooth manifold, and it is a Lie group because matrix products and inverses are smooth functions of the real and imaginary parts of the matrix entries.

(c) If V is any real or complex vector space, we let $\mathrm{GL}(V)$ denote the set of invertible linear transformations from V to itself. It is a group under composition. If V is finite-dimensional, any basis for V determines an isomorphism of $\mathrm{GL}(V)$ with $\mathrm{GL}(n, \mathbb{R})$ or $\mathrm{GL}(n, \mathbb{C})$, with $n = \dim V$, so $\mathrm{GL}(V)$ is a Lie group. The transition map between any two such isomorphisms is given by a map of the form $A \mapsto BAB^{-1}$ (where B is the transition matrix between the two bases), which is smooth. Thus the smooth manifold structure on $\mathrm{GL}(V)$ is independent of the choice of basis.

(d) The real number field $\mathbb{R}$ and Euclidean space $\mathbb{R}^n$ are Lie groups under addition, because the coordinates of $x - y$ are smooth (linear!) functions of (x, y).

(e) The set $\mathbb{R}^*$ of nonzero real numbers is a 1-dimensional Lie group under multiplication. (In fact, it is exactly $\mathrm{GL}(1,\mathbb{R})$, if we identify a 1×1 matrix with the corresponding real number.) The subset $\mathbb{R}^+$ of positive real numbers is an open subgroup, and is thus itself a 1-dimensional Lie group.

(f) The set $\mathbb{C}^*$ of nonzero complex numbers is a 2-dimensional Lie group under complex multiplication, which can be identified with $\mathrm{GL}(1,\mathbb{C})$.

(g) The circle $\mathbb{S}^1 \subset \mathbb{C}^*$ is a smooth manifold and a group under complex multiplication. Using appropriate angle functions as local coordinates on open subsets of $\mathbb{S}^1$ (see Problem 1-6), multiplication and inversion have the smooth coordinate expressions $(\theta_1, \theta_2) \mapsto \theta_1 + \theta_2$ and $\theta \mapsto -\theta$, and therefore $\mathbb{S}^1$ is a Lie group, called the *circle group*.

(h) Any product of Lie groups is a Lie group with the product smooth manifold structure and the direct product group structure, as you can easily check.

(i) The n-torus $\mathbb{T}^n = \mathbb{S}^1 \times \cdots \times \mathbb{S}^1$ is an n-dimensional abelian Lie group.

(j) Any finite or countably infinite group with the discrete topology is a zero-dimensional Lie group. We will call any such group a *discrete group*.

If G and H are Lie groups, a *Lie group homomorphism* from G to H is a smooth map $F\colon G \to H$ that is also a group homomorphism. It is called a *Lie group isomorphism* if it is also a diffeomorphism, which implies that it has an inverse that is also a Lie group homomorphism. In this case, we say that G and H are *isomorphic Lie groups*.

Example 2.8 (Lie Group Homomorphisms).

(a) The inclusion map $\mathbb{S}^1 \hookrightarrow \mathbb{C}^*$ is a Lie group homomorphism.

(b) The map $\exp\colon \mathbb{R} \to \mathbb{R}^*$ given by $\exp(t) = e^t$ is smooth, and is a Lie group homomorphism because $e^{(s+t)} = e^s e^t$. (Note that $\mathbb{R}$ is considered as a Lie group under addition, while $\mathbb{R}^*$ is a Lie group under multiplication.) The image of exp is the open subgroup $\mathbb{R}^+$ consisting of positive real numbers, and $\exp\colon \mathbb{R} \to \mathbb{R}^+$ is a Lie group isomorphism with inverse $\log\colon \mathbb{R}^+ \to \mathbb{R}$.

(c) Similarly, $\exp\colon \mathbb{C} \to \mathbb{C}^*$ given by $\exp(z) = e^z$ is a Lie group homomorphism. It is surjective but not injective, because its kernel consists of the complex numbers of the form $2\pi i k$ where k is an integer.

(d) The map $\varepsilon\colon \mathbb{R} \to \mathbb{S}^1$ defined by $\varepsilon(t) = e^{2\pi i t}$ is a Lie group homomorphism whose kernel is the set $\mathbb{Z}$ of integers. Similarly the map $\varepsilon^n\colon \mathbb{R}^n \to \mathbb{T}^n$ defined by $\varepsilon^n(t_1, \ldots, t_n) = (e^{2\pi i t_1}, \ldots, e^{2\pi i t_n})$ is a Lie group homomorphism whose kernel is $\mathbb{Z}^n$.

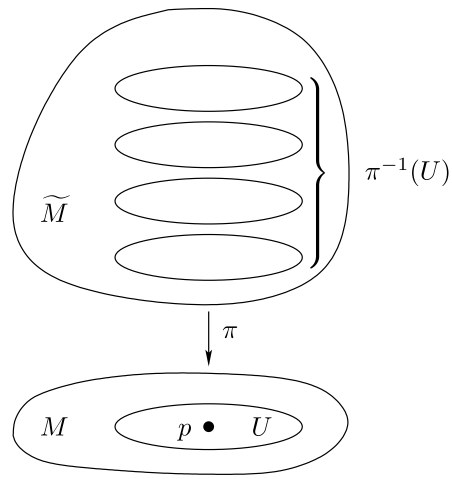


Figure 2.4. A covering map.

(e) The determinant function $\det\colon \mathrm{GL}(n,\mathbb{R}) \to \mathbb{R}^*$ is smooth because $\det A$ is a polynomial in the matrix entries of A. It is a Lie group homomorphism because $\det(AB) = (\det A)(\det B)$. Similarly, $\det\colon \mathrm{GL}(n,\mathbb{C}) \to \mathbb{C}^*$ is a Lie group homomorphism.

(f) If G is any Lie group and $g \in G$, define $C_g\colon G \to G$ to be conjugation by g: $C_g(h) = ghg^{-1}$. Then C_g is smooth because group multiplication is smooth, and a simple computation shows that it is a group homomorphism.

Smooth Covering Maps

You are probably already familiar with the notion of a *covering map* between topological spaces: This is a surjective continuous map $\pi\colon \widetilde{M} \to M$ between connected, locally path connected spaces, with the property that every point $p \in M$ has a neighborhood U that is *evenly covered*, meaning that U is connected and each component of $\pi^{-1}(U)$ is mapped homeomorphically onto U by π (Figure 2.4). The basic properties of covering maps are summarized in the Appendix (pages 556–557).

In the context of smooth manifolds, it is useful to introduce a slightly more restrictive type of covering map. If $\widetilde{M}$ and M are connected smooth manifolds, a *smooth covering map* $\pi\colon \widetilde{M} \to M$ is a smooth surjective map with the property that every $p \in M$ has a connected neighborhood U such that each component of $\pi^{-1}(U)$ is mapped *diffeomorphically* onto U by π. In this context, we will also say that U is evenly covered. The manifold M

is called the *base* of the covering, and $\widetilde{M}$ is called a *covering manifold* of M.

To distinguish this new definition from the previous one, we will often call an ordinary (not necessarily smooth) covering map a *topological covering map*. A smooth covering map is, in particular, a topological covering map. However, it is important to bear in mind that a smooth covering map is more than just a topological covering map that happens to be smooth—the definition of smooth covering map requires in addition that the restriction of π to each component of the inverse image of an evenly covered set be a diffeomorphism, not just a smooth homeomorphism.

Proposition 2.9 (Properties of Smooth Coverings).

(*a*) *Any smooth covering map is a local diffeomorphism and an open map.*

(*b*) *An injective smooth covering map is a diffeomorphism.*

(*c*) *A topological covering map is a smooth covering map if and only if it is a local diffeomorphism.*

◇ **Exercise 2.11.** Prove Proposition 2.9.

◇ **Exercise 2.12.** If $\pi_1 \colon \widetilde{M_1} \to M_1$ and $\pi_2 \colon \widetilde{M_2} \to M_2$ are smooth covering maps, show that $\pi_1 \times \pi_2 \colon \widetilde{M_1} \times \widetilde{M_2} \to M_1 \times M_2$ is a smooth covering map.

◇ **Exercise 2.13.** Suppose $\pi \colon \widetilde{M} \to M$ is a smooth covering map. Since π is also a topological covering map, there is a potential ambiguity about what it means for a subset $U \subset M$ to be evenly covered: Does π map the components of $\pi^{-1}(U)$ diffeomorphically onto U, or merely homeomorphically? Show that the two concepts are in fact equivalent: If $U \subset M$ is evenly covered in the topological sense, then π maps each component of $\pi^{-1}(U)$ diffeomorphically onto U.

If $\pi \colon \widetilde{M} \to M$ is any continuous map, a *section* of π is a continuous map $\sigma \colon M \to \widetilde{M}$ such that $\pi \circ \sigma = \mathrm{Id}_M$:

$$
\begin{tikzcd}
\widetilde{M} \arrow[d, "\pi"'] \\
M \arrow[u, bend right, "\sigma"']
\end{tikzcd}
$$

A *local section* is a continuous map $\sigma \colon U \to \widetilde{M}$ defined on some open set $U \subset M$ and satisfying the analogous relation $\pi \circ \sigma = \mathrm{Id}_U$. Many of the important properties of smooth covering maps follow from the existence of smooth local sections.

Lemma 2.10 (Local Sections of Smooth Coverings). *Suppose $\pi \colon \widetilde{M} \to M$ is a smooth covering map. Every point of $\widetilde{M}$ is in the image of a smooth local section of π. More precisely, for any $q \in \widetilde{M}$, there is* *a neighborhood U of $p = \pi(q)$ and a smooth local section $\sigma\colon U \to \widetilde{M}$ such that $\sigma(p) = q$.*

Proof. Let $U \subset M$ be an evenly covered neighborhood of p. If $\widetilde{U}$ is the component of $\pi^{-1}(U)$ containing q, then $\pi|_{\widetilde{U}}\colon \widetilde{U} \to U$ is by hypothesis a diffeomorphism. It follows that $\sigma = (\pi|_{\widetilde{U}})^{-1}\colon U \to \widetilde{U}$ is a smooth local section of π such that $\sigma(p) = q$. □

One important application of local sections is the following proposition, which gives a very simple criterion for deciding which maps out of the base of a covering are smooth.

Proposition 2.11. *Suppose $\pi\colon \widetilde{M} \to M$ is a smooth covering map and N is any smooth manifold. A map $F\colon M \to N$ is smooth if and only if $F \circ \pi\colon \widetilde{M} \to N$ is smooth:*

$$\begin{tikzcd} \widetilde{M} \arrow[d, "\pi"'] \arrow[dr, "F \circ \pi"] & \\ M \arrow[r, "F"'] & N. \end{tikzcd}$$

Proof. One direction is obvious by composition. Suppose conversely that $F \circ \pi$ is smooth, and let $p \in M$ be arbitrary. By the preceding lemma, there is a neighborhood U of p and a smooth local section $\sigma\colon U \to \widetilde{M}$, so that $\pi \circ \sigma = \mathrm{Id}_U$. Then the restriction of F to U satisfies

$$F|_U = F \circ \mathrm{Id}_U = F \circ (\pi \circ \sigma) = (F \circ \pi) \circ \sigma,$$

which is a composition of smooth maps. Thus F is smooth on U. Since F is smooth in a neighborhood of each point, it is smooth. □

The next proposition shows that every covering space of a connected smooth manifold is itself a smooth manifold.

Proposition 2.12. *If M is a connected smooth n-manifold and $\pi\colon \widetilde{M} \to M$ is a topological covering map, then $\widetilde{M}$ is a topological n-manifold and has a unique smooth structure such that π is a smooth covering map.*

Proof. Because π is, in particular, a local homeomorphism, it is clear that $\widetilde{M}$ is locally Euclidean.

Let p and q be distinct points in $\widetilde{M}$. If $\pi(p) = \pi(q)$ and $U \subset M$ is an evenly covered open set containing $\pi(p)$, then the components of $\pi^{-1}(U)$ containing p and q are disjoint open subsets of $\widetilde{M}$ separating p and q. On the other hand, if $\pi(p) \neq \pi(q)$, there are disjoint open sets $U, V \subset M$ containing $\pi(p)$ and $\pi(q)$, respectively, and then $\pi^{-1}(U)$ and $\pi^{-1}(V)$ are open subsets of $\widetilde{M}$ separating p and q. Thus $\widetilde{M}$ is Hausdorff.

To show that $\widetilde{M}$ is second countable, we will show first that each fiber of π is countable. Given $q \in M$ and an arbitrary point $\widetilde{q}_0 \in \pi^{-1}(q)$, we will construct a surjective map $\beta\colon \pi_1(M,q) \to \pi^{-1}(q)$; since $\pi_1(M,q)$ is countable, this suffices. Let $[f] \in \pi_1(M,q)$ be the path class of an arbitrary loop $f\colon [0,1] \to M$ based at q. The path lifting property of covering maps (Proposition A.26(b) in the Appendix) guarantees that there is a lift $\widetilde{f}\colon [0,1] \to \widetilde{M}$ of f starting at $\widetilde{q}_0$, and the homotopy lifting property (Proposition A.26(c)) shows that the lifts of path-homotopic loops are themselves path-homotopic. Thus the endpoint $\widetilde{f}(1) \in \pi^{-1}(q)$ depends only on the path class of f, so it makes sense to define $\beta[f] = \widetilde{f}(1)$. To see that β is surjective, just note that any point $\widetilde{q} \in \pi^{-1}(q)$ can be joined to $\widetilde{q}_0$ by some path $\widetilde{f}\colon [0,1] \to \widetilde{M}$, and then $\widetilde{q} = \beta[\pi \circ \widetilde{f}]$.

Now, the collection of all evenly covered open sets is an open cover of M, and therefore has a countable subcover $\{U_i\}$. For any given i, each component of $\pi^{-1}(U_i)$ contains exactly one point in each fiber over U_i, so $\pi^{-1}(U_i)$ has countably many components. Since each component is homeomorphic to U_i, it has a countable basis. The union of all of these countable bases forms a countable basis for the topology of $\widetilde{M}$, so $\widetilde{M}$ is second countable.

Given any point $\widetilde{q} \in \widetilde{M}$, let U be an evenly covered neighborhood of $\pi(\widetilde{q})$. Shrinking U if necessary, we may assume also that it is the domain of a smooth coordinate map $\varphi\colon U \to \mathbb{R}^n$. Letting $\widetilde{U}$ be the component of $\pi^{-1}(U)$ containing $\widetilde{q}$, and $\widetilde{\varphi} = \varphi \circ \pi\colon \widetilde{U} \to \mathbb{R}^n$, it is clear that $(\widetilde{U}, \widetilde{\varphi})$ is a chart on $\widetilde{M}$ (see Figure 2.5). If two such charts $(\widetilde{U}, \widetilde{\varphi})$ and $(\widetilde{V}, \widetilde{\psi})$ overlap, the transition map can be written

$$\begin{aligned}
\widetilde{\psi} \circ \widetilde{\varphi}^{-1} &= (\psi \circ \pi|_{\widetilde{U}\cap\widetilde{V}}) \circ (\varphi \circ \pi|_{\widetilde{U}\cap\widetilde{V}})^{-1} \\
&= \psi \circ \pi|_{\widetilde{U}\cap\widetilde{V}} \circ (\pi|_{\widetilde{U}\cap\widetilde{V}})^{-1} \circ \varphi^{-1} \\
&= \psi \circ \varphi^{-1},
\end{aligned}$$

which is smooth. Thus the collection of all such charts defines a smooth structure on $\widetilde{M}$. The uniqueness of this smooth structure is left to the reader (Problem 2-7). □

The next result is an important application of the preceding proposition.

Theorem 2.13 (Existence of a Universal Covering Group). *Let G be a connected Lie group. There exist a simply connected Lie group $\widetilde{G}$ (called the universal covering group of G) and a smooth covering map $\pi\colon \widetilde{G} \to G$ that is also a Lie group homomorphism.*

Proof. By Proposition A.29 in the Appendix, there exist a simply connected topological space $\widetilde{G}$ and a (topological) covering map $\pi\colon \widetilde{G} \to G$. Proposition 2.12 shows that $\widetilde{G}$ has a unique smooth manifold structure such that π is a smooth covering map. By Exercise 2.12, $\pi \times \pi\colon \widetilde{G} \times \widetilde{G} \to G \times G$ is also a smooth covering map.

Let $m\colon G\times G \to G$ and $i\colon G \to G$ denote the multiplication and inversion maps of G, respectively, and let $\widetilde{e}$ be an arbitrary element of the fiber

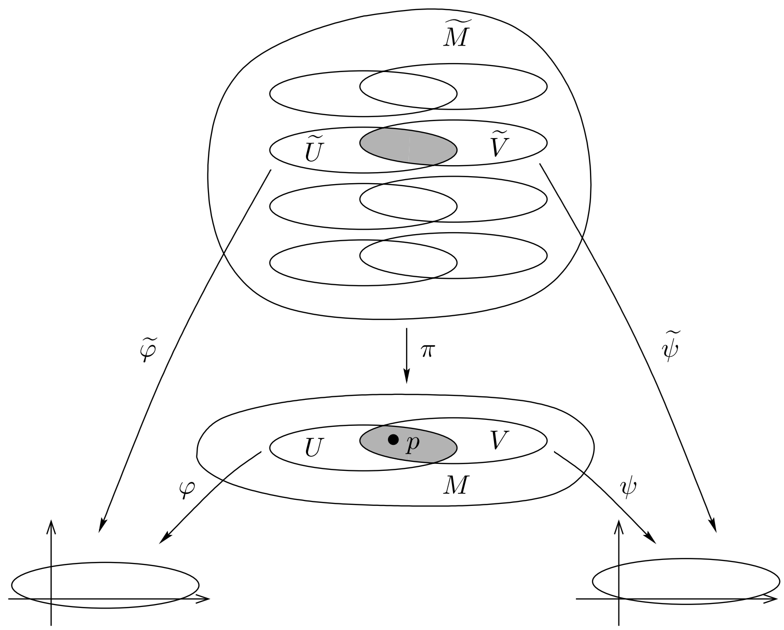


Figure 2.5. Smooth compatibility of charts on a covering manifold.

$\pi^{-1}(e) \subset \widetilde{G}$. Since $\widetilde{G}$ is simply connected, the lifting criterion for covering maps (Proposition A.27 in the Appendix) guarantees that the map $m \circ (\pi \times \pi)\colon \widetilde{G} \times \widetilde{G} \to G$ has a unique continuous lift $\widetilde{m}\colon \widetilde{G} \times \widetilde{G} \to \widetilde{G}$ satisfying $\widetilde{m}(\widetilde{e}, \widetilde{e}) = \widetilde{e}$ and $\pi \circ \widetilde{m} = m \circ (\pi \times \pi)$:

$$\begin{tikzcd} \widetilde{G} \times \widetilde{G} \arrow[r, "\widetilde{m}"] \arrow[d, "\pi \times \pi"'] & \widetilde{G} \arrow[d, "\pi"] \\ G \times G \arrow[r, "m"'] & G. \end{tikzcd} \tag{2.1}$$

Because $\widetilde{m}$ can be expressed locally as $\sigma \circ m \circ (\pi \times \pi)$ for a smooth local section σ of π, it follows that $\widetilde{m}$ is smooth. By the same reasoning, $i \circ \pi\colon \widetilde{G} \to G$ has a smooth lift $\widetilde{i}\colon \widetilde{G} \to \widetilde{G}$ satisfying $\widetilde{i}(\widetilde{e}) = \widetilde{e}$ and $\pi \circ \widetilde{i} = i \circ \pi$:

$$\begin{tikzcd} \widetilde{G} \arrow[r, "\widetilde{i}"] \arrow[d, "\pi"'] & \widetilde{G} \arrow[d, "\pi"] \\ G \arrow[r, "i"'] & G. \end{tikzcd} \tag{2.2}$$

We define multiplication and inversion in $\widetilde{G}$ by $xy = \widetilde{m}(x, y)$ and $x^{-1} = \widetilde{i}(x)$ for all $x, y \in \widetilde{G}$. Then (2.1) and (2.2) can be rewritten as

$$\pi(xy) = \pi(x)\pi(y), \tag{2.3}$$

$$\pi(x^{-1}) = \pi(x)^{-1}. \tag{2.4}$$

It remains only to show that $\widetilde{G}$ is a group with these operations, for then it is a Lie group because $\widetilde{m}$ and $\widetilde{i}$ are smooth, and (2.3) shows that π is a Lie group homomorphism.

First we show that $\widetilde{e}$ is an identity for multiplication in $\widetilde{G}$. Consider the map $f\colon \widetilde{G} \to \widetilde{G}$ defined by $f(x) = \widetilde{e}x$. Then (2.3) implies that $\pi \circ f(x) = \pi(\widetilde{e})\pi(x) = e\pi(x) = \pi(x)$, so f is a lift of $\pi\colon \widetilde{G} \to G$. The identity map $\mathrm{Id}_{\widetilde{G}}$ is another lift of π, and agrees with f at a point because $f(\widetilde{e}) = m(\widetilde{e}, \widetilde{e}) = \widetilde{e}$, so the unique lifting property of covering maps (Proposition A.26(a)) implies that $f = \mathrm{Id}_{\widetilde{G}}$, or equivalently $\widetilde{e}x = x$ for all $x \in \widetilde{G}$.

Next, to show that multiplication in $\widetilde{G}$ is associative, consider the two maps $\alpha_L, \alpha_R\colon \widetilde{G} \times \widetilde{G} \times \widetilde{G} \to \widetilde{G}$ defined by

$$\begin{aligned} \alpha_L(x, y, z) &= (xy)z, \\ \alpha_R(x, y, z) &= x(yz). \end{aligned}$$

Then (2.3) applied repeatedly implies that

$$\begin{aligned} \pi \circ \alpha_L(x, y, z) &= (\pi(x)\pi(y))\pi(z) \\ &= \pi(x)(\pi(y)\pi(z)) \\ &= \pi \circ \alpha_R(x, y, z), \end{aligned}$$

so α_L and α_R are both lifts of the same map $\alpha(x, y, z) = \pi(x)\pi(y)\pi(z)$. Because α_L and α_R agree at $(\widetilde{e}, \widetilde{e}, \widetilde{e})$, they are equal. A similar argument shows that $x^{-1}x = xx^{-1} = \widetilde{e}$, so $\widetilde{G}$ is a group. □

◇ **Exercise 2.14.** Complete the proof of the preceding theorem by showing that $x^{-1}x = xx^{-1} = \widetilde{e}$.

Proper Maps

There are not many simple criteria for determining whether a given surjective map is a covering map, even if it is known to be a local diffeomorphism. In this section, we describe one such criterion, called properness. It will have many other applications throughout the book.

If M and N are topological spaces, a map $F\colon M \to N$ (continuous or not) is said to be *proper* if for any compact set $K \subset N$, the inverse image $F^{-1}(K)$ is compact. The next three lemmas give useful sufficient conditions for a map to be proper.

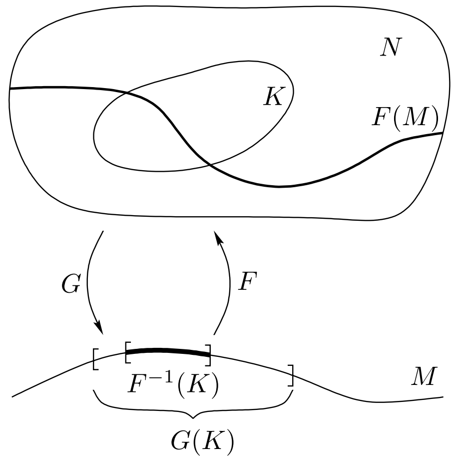


Figure 2.6. A map with a left inverse is proper.

Lemma 2.14. *Suppose M is a compact space and N is a Hausdorff space. Then any continuous map $F: M \to N$ is proper.*

Proof. If $K \subset N$ is compact, then it is closed in N because N is Hausdorff. By continuity, $F^{-1}(K)$ is closed in M and therefore compact. $\square$

Recall that a subset $A \subset M$ is said to be saturated with respect to a map $F: M \to N$ if $A = F^{-1}(F(A))$ (see page 548).

Lemma 2.15. *Suppose $F: M \to N$ is a proper map between topological spaces, and $A \subset M$ is any subset that is saturated with respect to F. Then $F|_A: A \to F(A)$ is proper.*

Proof. Let $K \subset F(A)$ be compact. The fact that A is saturated means that $(F|_A)^{-1}(K) = F^{-1}(K)$, which is compact because F is proper. $\square$

Lemma 2.16. *Let $F: M \to N$ be a continuous map between Hausdorff spaces. If there exists a left inverse for F (i.e., a continuous map $G: N \to M$ such that $G \circ F = \mathrm{Id}_M$), then F is proper.*

Proof. If $K \subset N$ is any compact set, then any point $x \in F^{-1}(K)$ satisfies $x = G(F(x)) \in G(K)$. Since K is closed in N, it follows that $F^{-1}(K)$ is a closed subset of the compact set $G(K)$ (Figure 2.6), so it is compact. $\square$

For continuous maps between topological manifolds, there is an alternative characterization of properness in terms of divergent sequences, which is somewhat easier to visualize. If X is a topological space, a sequence $\{p_i\}$ in X is said to *escape to infinity* if for any compact set $K \subset X$, there are at most finitely many values of i for which $p_i \in K$.

◇ **Exercise 2.15.** Show that a sequence of points in a topological manifold escapes to infinity if and only if it has no convergent subsequence.

Proposition 2.17 (Sequential Characterization of Proper Maps). *Suppose M and N are topological manifolds. A continuous map $F\colon M \to N$ is proper if and only if for every sequence $\{p_i\}$ in M that escapes to infinity, $\{F(p_i)\}$ escapes to infinity in N.*

Proof. First suppose that F is proper, and let $\{p_i\}$ be a sequence in M that escapes to infinity. If $\{F(p_i)\}$ does not escape to infinity, then there is a compact subset $K \subset N$ that contains $F(p_i)$ for infinitely many values of i. It follows that p_i lies in the compact set $F^{-1}(K)$ for these values of i, which contradicts the assumption that $\{p_i\}$ escapes to infinity.

Conversely, suppose every sequence escaping to infinity in M is taken by F to a sequence escaping to infinity in N. Let $K \subset N$ be a compact set, and let $L = F^{-1}(K) \subset M$. To show that L is compact, we will show that every sequence in L has a convergent subsequence. Suppose on the contrary that $\{p_i\}$ is a sequence in L with no convergent subsequence. Then by the hypothesis and Exercise 2.15, $\{F(p_i)\}$ has no convergent subsequence; but this is impossible because $F(p_i)$ lies in the compact set K for all i. □

It is often extremely useful to be able to show that a given continuous map is a closed map. For example, Lemma A.13 in the Appendix shows that a closed continuous map that is also surjective, injective, or bijective is automatically a quotient map, a topological embedding, or a homeomorphism, respectively. One situation in which this condition is automatically fulfilled is when the domain is compact: The closed map lemma (Lemma A.19) asserts that any continuous map from a compact topological space to a Hausdorff space is closed. But there are also plenty of interesting manifolds that are not compact and therefore are not covered by this result. The following is a powerful generalization of the closed map lemma.

Proposition 2.18 (Proper Continuous Maps Are Closed). *Suppose $F\colon M \to N$ is a proper continuous map between topological manifolds. Then F is closed.*

Proof. If K is a closed subset of M, we need to prove that $F(K)$ is closed in N. Let y be an arbitrary point of $\overline{F(K)}$. By Exercise A.8 in the Appendix, there exists a sequence $\{y_i\}$ of points in $F(K)$ such that $y_i \to y$. The fact that $y_i \in F(K)$ means that there exists $x_i \in K$ such that $F(x_i) = y_i$.

Let U be a precompact neighborhood of y in N. For all large enough i, $y_i \in U \subset \overline{U}$, and therefore $x_i \in F^{-1}(\overline{U})$. The hypothesis that F is proper implies that $F^{-1}(\overline{U})$ is compact, and thus $\{x_i\}$ has a subsequence $\{x_{i_k}\}$ converging to a point $x \in M$. Because K is closed, $x \in K$. By continuity,

$$F(x) = \lim_{k\to\infty} F(x_{i_k}) = \lim_{k\to\infty} y_{i_k} = y,$$

which implies $y \in F(K)$ as desired. □

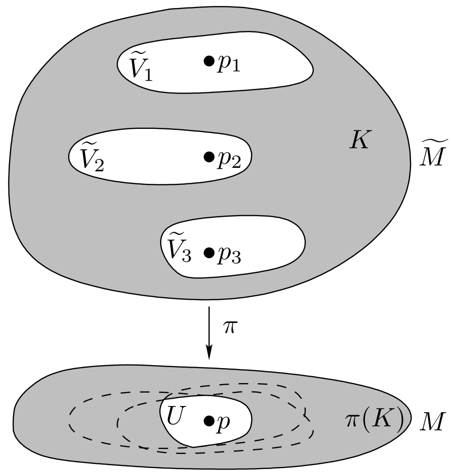


Figure 2.7. A proper local diffeomorphism is a covering map.

The next proposition is the main result of this section.

Proposition 2.19. *Suppose $\widetilde{M}$ and M are connected smooth manifolds and $\pi\colon \widetilde{M} \to M$ is a proper local diffeomorphism. Then π is a smooth covering map.*

Proof. Because π is a local diffeomorphism, it is an open map, and because it is proper, it is a closed map. Thus $\pi(\widetilde{M})$ is both open and closed in M. Since it is obviously nonempty, it is all of M, so π is surjective.

Let $p \in M$ be arbitrary. Since π is a local diffeomorphism, each point of $\pi^{-1}(p)$ has a neighborhood on which π is injective, so $\pi^{-1}(p)$ is a discrete set. Since π is proper, $\pi^{-1}(p)$ is also compact, so it is finite. Write $\pi^{-1}(p) = \{\widetilde{p}_1, \dots, \widetilde{p}_k\}$. For each i, there exists a neighborhood $\widetilde{V}_i$ of $\widetilde{p}_i$ on which π is a diffeomorphism onto an open set $V_i \subset M$. Shrinking each $\widetilde{V}_i$ if necessary, we may assume also that $\widetilde{V}_i \cap \widetilde{V}_j = \varnothing$ for $i \neq j$.

Set $U = V_1 \cap \dots \cap V_k$ (Figure 2.7), which is a neighborhood of p. Then U obviously satisfies

$$U \subset V_i \text{ for each } i. \tag{2.5}$$

Because $K = \widetilde{M} \smallsetminus (\widetilde{V}_1 \cup \dots \cup \widetilde{V}_k)$ is closed in $\widetilde{M}$ and π is a closed map, $\pi(K)$ is closed in M. Replacing U by $U \smallsetminus \pi(K)$, we can assume that U also satisfies

$$\pi^{-1}(U) \subset \widetilde{V}_1 \cup \dots \cup \widetilde{V}_k. \tag{2.6}$$

Finally, after replacing U by the connected component of U containing p, we can assume U is connected and still satisfies (2.5) and (2.6). We will show that U is evenly covered.

Let $\widetilde{U}_i = \pi^{-1}(U) \cap \widetilde{V}_i$. By virtue of (2.6), $\pi^{-1}(U) = \widetilde{U}_1 \cup \cdots \cup \widetilde{U}_k$. Because $\pi\colon \widetilde{V}_i \to V_i$ is a diffeomorphism, (2.5) implies that $\pi\colon \widetilde{U}_i \to U$ is still a diffeomorphism, and in particular $\widetilde{U}_i$ is connected. Because $\widetilde{U}_1, \dots, \widetilde{U}_k$ are disjoint connected open subsets of $\pi^{-1}(U)$, they are exactly the components of $\pi^{-1}(U)$. □

Partitions of Unity

One of the more useful tools in topology is the gluing lemma (Lemma A.7 in the Appendix), which shows how to construct continuous maps by "gluing together" maps defined on subspaces. For smooth manifolds, however, the gluing lemma is of limited usefulness, because the map it produces is rarely smooth, even when it is built out of smooth maps on subspaces. For example, the two functions $f_+ \colon [0,\infty) \to \mathbb{R}$ and $f_- \colon (-\infty, 0] \to \mathbb{R}$ defined by

$$\begin{aligned} f_+(x) &= +x, \qquad x \in [0,\infty) \\ f_-(x) &= -x, \qquad x \in (-\infty, 0] \end{aligned}$$

are both smooth and agree at the point 0 where they overlap, but the continuous map $f\colon \mathbb{R} \to \mathbb{R}$ that they define, namely $f(x) = |x|$, is not smooth at the origin.

In this section, we introduce partitions of unity, which are tools for patching together local smooth objects into global ones. They are indispensable in smooth manifold theory and will reappear throughout the book.

All of our constructions in this section are based on the existence of smooth functions that are positive in a specified part of a manifold and identically zero in some other part. We begin by defining a smooth function on the real line that is zero for $t \le 0$ and positive for $t > 0$.

Lemma 2.20. *The function $f\colon \mathbb{R} \to \mathbb{R}$ defined by*

$$f(t) = \begin{cases} e^{-1/t} & t > 0, \\ 0 & t \le 0 \end{cases}$$

is smooth.

Proof. The function in question is pictured in Figure 2.8. It is clearly smooth on $\mathbb{R} \smallsetminus \{0\}$, so we need only show that all derivatives of f exist and are continuous at the origin. We begin by noting that f is continuous because $\lim_{t \searrow 0} e^{-1/t} = 0$. In fact, a standard application of l'Hôpital's rule and induction shows that for any integer $k \ge 0$,

$$\lim_{t \searrow 0} \frac{e^{-1/t}}{t^k} = \lim_{t \searrow 0} \frac{t^{-k}}{e^{1/t}} = 0. \tag{2.7}$$

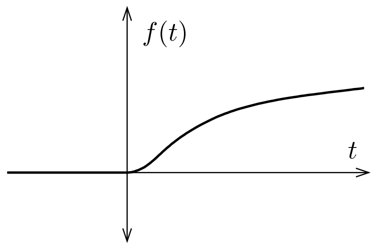


Figure 2.8. $f(t) = e^{-1/t}$.

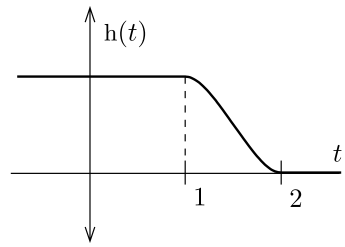


Figure 2.9. A cutoff function.

We will show by induction that for $t > 0$, the kth derivative of f is of the form

$$f^{(k)}(t) = \frac{p_k(t)}{t^{2k}} e^{-1/t} \tag{2.8}$$

for some polynomial $p_k(t)$. It is clearly true (with $p_0(t) = 1$) for $k = 0$, so suppose it is true for some $k \geq 0$. By the product rule,

$$\begin{aligned} f^{(k+1)}(t) &= \frac{p_k'(t)}{t^{2k}} e^{-1/t} - \frac{2kp_k(t)}{t^{2k+1}} e^{-1/t} + \frac{p_k(t)}{t^{2k}} \frac{1}{t^2} e^{-1/t} \\ &= \frac{t^2 p_k'(t) - 2ktp_k(t) + p_k(t)}{t^{2k+2}} e^{-1/t}, \end{aligned}$$

which is of the required form. Note that (2.8) and (2.7) imply

$$\lim_{t \searrow 0} f^{(k)}(t) = 0, \tag{2.9}$$

since a polynomial is continuous.

Finally, we prove that for each $k \geq 0$,

$$f^{(k)}(0) = 0.$$

For $k = 0$, this is true by definition, so assume that it is true for some $k \geq 0$. It suffices to show that f has one-sided derivatives from both sides and that they are equal. Clearly the derivative from the left is zero. Using (2.7) again, we compute

$$f^{(k+1)}(0) = \lim_{t \searrow 0} \frac{\dfrac{p_k(t)}{t^{2k}} e^{-1/t} - 0}{t} = \lim_{t \searrow 0} \frac{p_k(t)}{t^{2k+1}} e^{-1/t} = 0.$$

By (2.9), this implies each $f^{(k)}$ is continuous, so f is smooth. $\square$

Lemma 2.21. *There exists a smooth function* $h\colon \mathbb{R} \to [0,1]$ *such that* $h(t) \equiv 1$ *for* $t \leq 1$, $0 < h(t) < 1$ *for* $1 < t < 2$, *and* $h(t) \equiv 0$ *for* $t \geq 2$.

Proof. Let f be the function of the previous lemma, and set

$$h(t) = \frac{f(2-t)}{f(2-t)+f(t-1)}.$$

(See Figure 2.9.) Note that the denominator is positive for all t, because at least one of the expressions $2-t$ or $t-1$ is always positive. Since $f \geq 0$ always, it is easy to check that $h(t)$ is always between 0 and 1, and is zero when $t \geq 2$. When $t \leq 1$, $f(t-1) = 0$, so $h(t) \equiv 1$ there. □

A function with the properties of h in this lemma is usually called a *cutoff function.*

If f is any real-valued or vector-valued function on a topological space M, the *support* of f, denoted by $\operatorname{supp} f$, is the closure of the set of points where f is nonzero:

$$\operatorname{supp} f = \overline{\{p \in M : f(p) \neq 0\}}.$$

If $\operatorname{supp} f$ is contained in some set U, we say f is *supported in* U. A function f is said to be *compactly supported* if $\operatorname{supp} f$ is a compact set. Clearly every function on a compact space is compactly supported.

Lemma 2.22. *There is a smooth function $H\colon \mathbb{R}^n \to [0,1]$ such that $H \equiv 1$ on $\overline{B}_1(0)$ and $\operatorname{supp} H = \overline{B}_2(0)$.*

Proof. Just set $H(x) = h(|x|)$, where h is the function of the preceding lemma. Clearly H is smooth on $\mathbb{R}^n \smallsetminus \{0\}$, because it is a composition of smooth functions there. Since it is identically equal to 1 on $B_1(0)$, it is smooth there too. □

The function H constructed in this lemma is an example of a *smooth bump function*—a smooth real-valued function that is equal to 1 on a specified closed set (in this case $\overline{B}_1(0)$) and is supported in a specified open set (in this case any open set containing $\overline{B}_2(0)$). Later, we will generalize this notion to manifolds.

Paracompactness

To use bump functions effectively on a manifold, we will need to construct some open covers with special properties. Let X be a topological space. A collection $\mathcal{U}$ of subsets of X is said to be *locally finite* if each point of X has a neighborhood that intersects at most finitely many of the sets in $\mathcal{U}$.

◇ **Exercise 2.16.** If $\mathcal{U}$ is an *open* cover of X such that each set in $\mathcal{U}$ intersects only finitely many others, show that $\mathcal{U}$ is locally finite.

Given an open cover $\mathcal{U}$ of X, another open cover $\mathcal{V}$ is called a *refinement* of $\mathcal{U}$ if for each $V \in \mathcal{V}$ there exists some $U \in \mathcal{U}$ such that $V \subset U$. We say X is *paracompact* if every open cover of X admits a locally finite refinement.

A key topological result that we will need is the fact that every manifold is paracompact. This is a consequence of second countability, and in fact is one of the most important reasons why second countability is included in the definition of manifolds.

The following lemma will be a first step in the proof of paracompactness.

Lemma 2.23. *Every topological manifold admits a countable, locally finite cover by precompact open sets.*

Proof. Let M be a topological manifold. We will construct the desired cover in three steps.

Start with a countable cover $\{B_j\}_{j=1}^\infty$ by precompact open sets (such as the countable basis whose existence was proved in Lemma 1.6). Next, we will show that M admits a countable cover $\{U_j\}_{j=1}^\infty$ satisfying

(i) U_j is a precompact open subset of M.

(ii) $\overline{U}_{j-1} \subset U_j$ for $j \geq 2$.

(iii) $B_j \subset U_j$.

Begin with $U_1 = B_1$. Assume by induction that open sets U_j have been defined for $j = 1, \ldots, k$ satisfying (i)–(iii). Because $\overline{U}_k$ is compact and covered by $\{B_j\}$, there is some m_k such that

$$\overline{U}_k \subset B_1 \cup \cdots \cup B_{m_k}.$$

If we let $U_{k+1} = B_1 \cup \cdots \cup B_{m_k}$, then clearly (i) and (ii) are satisfied with $j = k+1$. Moreover, by increasing m_k if necessary, we may assume that $m_k \geq k+1$, so that $B_{k+1} \subset U_{k+1}$. Thus by induction we obtain a countable sequence of open sets $\{U_j\}$ satisfying (i)–(iii). Since $\{B_j\}$ is a cover of M, (iii) guarantees that $\{U_j\}$ is a cover as well.

Finally, to obtain a locally finite cover, we just set $V_j = U_j \smallsetminus \overline{U}_{j-2}$ (Figure 2.10). Since $\overline{V}_j$ is a closed subset of the compact set $\overline{U}_j$, it is compact. If $p \in M$ is arbitrary, then $p \in V_k$, where k is the smallest integer such that $p \in U_k$. Clearly V_k has nonempty intersection only with V_{k-1} and V_{k+1}, so the cover $\{V_j\}$ is locally finite. □

Now we are ready to show that every smooth manifold is paracompact. In fact, for future use, we will show something stronger—that every open cover admits a locally finite refinement of a particularly nice type. If M is a smooth manifold, we say an open cover $\{W_i\}$ of M is *regular* if it satisfies the following properties:

(i) The cover $\{W_i\}$ is countable and locally finite.

(ii) Each W_i is the domain of a smooth coordinate map $\varphi_i\colon W_i \to \mathbb{R}^n$ whose image is $B_3(0) \subset \mathbb{R}^n$.

(iii) The collection $\{U_i\}$ still covers M, where $U_i = \varphi_i^{-1}(B_1(0))$.

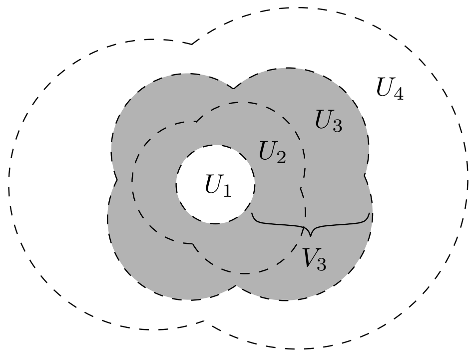


Figure 2.10. Constructing a locally finite cover.

Proposition 2.24. *Let M be a smooth manifold. Every open cover of M has a regular refinement. In particular, M is paracompact.*

Proof. Let $\mathcal{X}$ be any open cover of M, and let $\{V_j\}$ be a countable, locally finite cover of M by precompact open sets. For each $p \in M$, let W_p be a neighborhood of p that intersects only finitely many of the sets V_j. Replacing W_p by its intersection with those V_js that contain p, we may assume that

- If $p \in V_j$, then $W_p \subset V_j$ as well.

Since $\mathcal{X}$ is an open cover of M, $p \in X$ for some set $X \in \mathcal{X}$. Shrinking W_p further if necessary, we may also assume that

- W_p is contained in one of the open sets of $\mathcal{X}$.

Shrinking once more, we can assume W_p is a smooth coordinate ball, and by choosing the coordinate map judiciously, we can arrange that

- W_p is the domain of a smooth coordinate map $\varphi_p \colon W_p \to B_3(0)$ centered at p.

Let $U_p = \varphi_p^{-1}(B_1(0))$.

For each k, the collection $\{U_p : p \in \overline{V}_k\}$ is an open cover of $\overline{V}_k$. By compactness, $\overline{V}_k$ is covered by finitely many of these sets. Call the sets $U_k^1, \dots, U_k^{m_k}$, and let $(W_k^1, \varphi_k^1), \dots, (W_k^{m_k}, \varphi_k^{m_k})$ denote the corresponding coordinate charts. The collection of all the sets $\{W_k^i\}$ as k and i vary is clearly a countable open cover of M that refines $\mathcal{X}$ and satisfies (ii) and (iii) in the definition of a regular cover. To show it is a regular cover, we need only show it is locally finite.

For any given k, each set W_k^i is by construction contained in some V_j, which obviously satisfies $\overline{V}_k \cap V_j \neq 0$. The compact set $\overline{V}_k$ is covered by finitely many V_js, and each such V_j intersects at most finitely many others;

it follows that there are only finitely many values of k' for which W^i_k and $W^{i'}_{k'}$ can have nonempty intersection. Since there are only finitely many sets $W^{i'}_{k'}$ for each k', the cover we have constructed is locally finite. □

Now let M be a topological space, and let $\mathfrak{X} = \{X_\alpha\}_{\alpha\in A}$ be an arbitrary open cover of M. A *partition of unity subordinate to* $\mathfrak{X}$ is a collection of continuous functions $\{\psi_\alpha \colon M \to \mathbb{R}\}_{\alpha\in A}$, with the following properties:

(i) $0 \le \psi_\alpha(x) \le 1$ for all $\alpha \in A$ and all $x \in M$.

(ii) $\operatorname{supp}\psi_\alpha \subset X_\alpha$.

(iii) The set of supports $\{\operatorname{supp}\psi_\alpha\}_{\alpha\in A}$ is locally finite.

(iv) $\sum_{\alpha\in A}\psi_\alpha(x) = 1$ for all $x \in M$.

Because of the local finiteness condition (iii), the sum in (iv) actually has only finitely many nonzero terms in a neighborhood of each point, so there is no issue of convergence. If M is a smooth manifold, a *smooth partition of unity* is one for which each of the functions ψ_α is smooth.

Theorem 2.25 (Existence of Partitions of Unity). *If M is a smooth manifold and $\mathfrak{X} = \{X_\alpha\}_{\alpha\in A}$ is any open cover of M, there exists a smooth partition of unity subordinate to $\mathfrak{X}$.*

Proof. Let $\{W_i\}$ be a regular refinement of $\mathfrak{X}$ (see Proposition 2.24). For each i, let $\varphi_i \colon W_i \to B_3(0)$ be the smooth coordinate map whose existence is guaranteed by the definition of a regular cover, and let

$$\begin{aligned} U_i &= \varphi_i^{-1}(B_1(0)), \\ V_i &= \varphi_i^{-1}(B_2(0)). \end{aligned}$$

For each i, define a function $f_i \colon M \to \mathbb{R}$ by

$$f_i = \begin{cases} H \circ \varphi_i & \text{on } W_i, \\ 0 & \text{on } M \smallsetminus \overline{V}_i, \end{cases}$$

where $H \colon \mathbb{R}^n \to \mathbb{R}$ is the smooth bump function of Lemma 2.22. On the set $W_i \smallsetminus \overline{V}_i$ where the two definitions overlap, both definitions yield the zero function, so f_i is well-defined and smooth, and $\operatorname{supp} f_i \subset W_i$.

Define new functions $g_i \colon M \to \mathbb{R}$ by

$$g_i(x) = \frac{f_i(x)}{\sum_j f_j(x)}.$$

Because of the local finiteness of the cover $\{W_i\}$, the sum in the denominator has only finitely many nonzero terms in a neighborhood of each point and thus defines a smooth function. Because $f_i \equiv 1$ on U_i and every point of M is in some U_i, the denominator is always positive, so g_i is a smooth function on M. It is immediate from the definition that $0 \le g_i \le 1$ and $\sum_i g_i \equiv 1$.

Finally, we need to re-index our functions so that they are indexed by the same set A as our open cover. Because the cover $\{W_i\}$ is a refinement of $\mathcal{X}$, for each i we can choose some index $a(i) \in A$ such that $W_i \subset X_{a(i)}$. For each $\alpha \in A$, define $\psi_\alpha \colon M \to \mathbb{R}$ by

$$\psi_\alpha = \sum_{i:a(i)=\alpha} g_i.$$

If there are no indices i for which $a(i) = \alpha$, then this sum should be interpreted as the zero function. Each ψ_α is smooth and satisfies $0 \le \psi_\alpha \le 1$ and $\operatorname{supp}\psi_\alpha \subset X_\alpha$. Moreover, the set of supports $\{\operatorname{supp}\psi_\alpha\}_{\alpha\in A}$ is still locally finite, and $\sum_\alpha \psi_\alpha \equiv \sum_i g_i \equiv 1$, so this is the desired partition of unity. □

◇ **Exercise 2.17.** Let M be a smooth manifold with boundary. Show that M is paracompact, and that every open cover of M admits a subordinate smooth partition of unity.

◇ **Exercise 2.18.** Let M be a topological manifold with or without boundary. Show that M is paracompact, and that every open cover of M admits a subordinate partition of unity.

As our first application of partitions of unity, we will extend the notion of bump functions to arbitrary closed sets in manifolds. If M is a smooth manifold, $A \subset M$ is a closed subset, and $U \subset M$ is an open set containing A, a continuous function $\psi \colon M \to \mathbb{R}$ is called a *bump function for A supported in U* if $0 \le \psi \le 1$ on M, $\psi \equiv 1$ on A, and $\operatorname{supp}\psi \subset U$.

Proposition 2.26 (Existence of Bump Functions). *Let M be a smooth manifold. For any closed set $A \subset M$ and any open set U containing A, there exists a smooth bump function for A supported in U.*

Proof. Let $U_0 = U$ and $U_1 = M \smallsetminus A$, and let $\{\psi_0, \psi_1\}$ be a smooth partition of unity subordinate to the open cover $\{U_0, U_1\}$. Because $\psi_1 \equiv 0$ on A and therefore $\psi_0 = \sum_i \psi_i = 1$ there, the function ψ_0 has the required properties. □

Our second application is an important result concerning the possibility of extending smooth functions from closed subsets. Suppose M and N are smooth manifolds, and $A \subset M$ is an arbitrary subset. We say that a map $F \colon A \to N$ is smooth if it has a smooth extension in a neighborhood of every point; or more precisely if for every $p \in A$, there is an open set $W \subset M$ containing p and a smooth map $\widetilde{F} \colon W \to N$ whose restriction to $W \cap A$ agrees with F.

Lemma 2.27 (Extension Lemma). *Let M be a smooth manifold, let $A \subset M$ be a closed subset, and let $f \colon A \to \mathbb{R}^k$ be a smooth function. For any open set U containing A, there exists a smooth function $\widetilde{f} \colon M \to \mathbb{R}^k$ such that $\widetilde{f}|_A = f$ and $\operatorname{supp}\widetilde{f} \subset U$.*

Proof. For each $p \in A$, choose a neighborhood W_p of p and a smooth function $\widetilde{f}_p \colon W_p \to \mathbb{R}^k$ that agrees with f on $W_p \cap A$. Replacing W_p by $W_p \cap U$, we may assume that $W_p \subset U$. The collection of sets $\{W_p : p \in A\} \cup \{M \setminus A\}$ is an open cover of M. Let $\{\psi_p : p \in A\} \cup \{\psi_0\}$ be a smooth partition of unity subordinate to this cover, with $\operatorname{supp} \psi_p \subset W_p$ and $\operatorname{supp} \psi_0 \subset M \setminus A$.

For each $p \in A$, the product $\psi_p \widetilde{f}_p$ is smooth on W_p, and has a smooth extension to all of M if we interpret it to be zero on $M \setminus \operatorname{supp} \psi_p$. (The extended function is smooth because the two definitions agree on the open set $W_p \setminus \operatorname{supp} \psi_p$ where they overlap.) Thus we can define $\widetilde{f} \colon M \to \mathbb{R}^k$ by

$$\widetilde{f}(x) = \sum_{p \in A} \psi_p(x) \widetilde{f}_p(x).$$

Because the collection of supports $\{\operatorname{supp} \psi_p\}$ is locally finite, this sum actually has only a finite number of nonzero terms in a neighborhood of any point of M, and therefore defines a smooth function. If $x \in A$, then $\widetilde{f}_p(x) = f(x)$ for each p and $\psi_0(x) = 0$, and thus

$$\widetilde{f}(x) = \sum_{p \in A} \psi_p(x) f(x) = \left(\psi_0(x) + \sum_{p \in A} \psi_p(x) \right) f(x) = f(x),$$

so $\widetilde{f}$ is indeed an extension of f. Finally, suppose $x \in \operatorname{supp} \widetilde{f}$. Then x has a neighborhood on which at most finitely many of the functions ψ_p are nonzero, and x must be in $\operatorname{supp} \psi_p$ for at least one $p \in A$, which implies that $x \in W_p \subset U$. □

The extension lemma, by the way, illustrates an essential difference between smooth manifolds and real-analytic manifolds. The analogue of the extension lemma for real-analytic functions on real-analytic manifolds is decidedly false, because a real-analytic function that is defined on a connected domain and vanishes on an open set must be identically zero.

As our final application of partitions of unity, we will construct a special kind of smooth function. If M is a topological space, an *exhaustion function* for M is a continuous function $f \colon M \to \mathbb{R}$ with the property that the set $M_c = \{x \in M : f(x) \leq c\}$ is compact for every $c \in \mathbb{R}$. The name comes from the fact that the compact sets M_c exhaust M as c increases to positive infinity. For example, the functions $f \colon \mathbb{R}^n \to \mathbb{R}$ and $g \colon \mathbb{B}^n \to \mathbb{R}$ given by

$$\begin{aligned} f(x) &= |x|, \\ g(x) &= \frac{1}{1 - |x|^2} \end{aligned}$$

are exhaustion functions. Of course, if M is compact, any continuous real-valued function is an exhaustion function, so such functions are interesting only for noncompact manifolds.

Proposition 2.28 (Existence of Exhaustion Functions). *Every smooth manifold admits a smooth positive exhaustion function.*

Proof. Let M be a smooth manifold, let $\{V_j\}_{j=1}^{\infty}$ be any countable open cover of M by precompact open sets, and let $\{\psi_j\}$ be a smooth partition of unity subordinate to this cover. Define $f \in C^\infty(M)$ by

$$f(p) = \sum_{j=1}^{\infty} j\psi_j(p).$$

Then f is smooth because only finitely many terms are nonzero in a neighborhood of any point, and positive because $f(p) \geq \sum_j \psi_j(p) = 1$. For any positive integer N, if $p \notin \bigcup_{j=1}^{N} \overline{V}_j$, then $\psi_j(p) = 0$ for $1 \leq j \leq N$, so

$$f(p) = \sum_{j=N+1}^{\infty} j\psi_j(p) > \sum_{j=N+1}^{\infty} N\psi_j(p) = N\sum_{j=1}^{\infty} \psi_j(p) = N.$$

Equivalently, if $f(p) \leq N$, then $p \in \bigcup_{j=1}^{N} \overline{V}_j$. Thus for any $c \leq N$, M_c is a closed subset of the compact set $\bigcup_{j=1}^{N} \overline{V}_j$ and is therefore compact. □

Problems

2-1. Compute the coordinate representation for each of the following maps in stereographic coordinates (see Problem 1-5), and use this to prove that each map is smooth.

(a) For each $n \in \mathbb{Z}$, the *nth power map* $p_n \colon \mathbb{S}^1 \to \mathbb{S}^1$ is given in complex notation by $p_n(z) = z^n$.

(b) $\alpha \colon \mathbb{S}^n \to \mathbb{S}^n$ is the *antipodal map* $\alpha(x) = -x$.

(c) $F \colon \mathbb{S}^3 \to \mathbb{S}^2$ is given by $F(z,w) = (z\overline{w}+w\overline{z}, iw\overline{z}-iz\overline{w}, z\overline{z}-w\overline{w})$, where we think of $\mathbb{S}^3$ as the subset $\{(w,z) : |w|^2 + |z|^2 = 1\}$ of $\mathbb{C}^2$.

2-2. Show that the inclusion map $\overline{\mathbb{B}^n} \hookrightarrow \mathbb{R}^n$ is smooth when $\overline{\mathbb{B}^n}$ is regarded as a smooth manifold with boundary.

2-3. Let $\mathbb{R}$ denote the real line with its standard smooth structure, and let $\widetilde{\mathbb{R}}$ denote the same topological manifold with the smooth structure defined in Example 1.14. Let $f \colon \mathbb{R} \to \mathbb{R}$ be any function. Determine necessary and sufficient conditions on f so that it will be:

(a) a smooth map from $\mathbb{R}$ to $\widetilde{\mathbb{R}}$;

(b) a smooth map from $\widetilde{\mathbb{R}}$ to $\mathbb{R}$.

2-4. Let $P \colon \mathbb{R}^{n+1} \smallsetminus \{0\} \to \mathbb{R}^{k+1} \smallsetminus \{0\}$ be a smooth map, and suppose that for some $d \in \mathbb{Z}$, $P(\lambda x) = \lambda^d P(x)$ for all $\lambda \in \mathbb{R} \smallsetminus \{0\}$ and $x \in \mathbb{R}^{n+1} \smallsetminus \{0\}$. (Such a map is said to be *homogeneous of degree* *d.*) Show that the map $\widetilde{P}\colon \mathbb{RP}^n \to \mathbb{RP}^k$ defined by $\widetilde{P}[x] = [P(x)]$ is well-defined and smooth.

2-5. Let M be a nonempty smooth manifold of dimension $n \geq 1$. Show that $C^\infty(M)$ is infinite-dimensional.

2-6. For any topological space M, let $C(M)$ denote the algebra of continuous functions $f\colon M \to \mathbb{R}$. If $F\colon M \to N$ is a continuous map, define $F^*\colon C(N) \to C(M)$ by $F^*(f) = f \circ F$.

(a) Show that F^* is a linear map.
(b) If M and N are smooth manifolds, show that F is smooth if and only if $F^*(C^\infty(N)) \subset C^\infty(M)$.
(c) If $F\colon M \to N$ is a homeomorphism between smooth manifolds, show that it is a diffeomorphism if and only if F^* restricts to an isomorphism from $C^\infty(N)$ to $C^\infty(M)$.

[Remark: This result shows that in a certain sense, the entire smooth structure of M is encoded in the space $C^\infty(M)$. In fact, some authors *define* a smooth structure on a topological manifold M to be a subalgebra of $C(M)$ with certain properties.]

2-7. Let M be a connected smooth manifold, and let $\pi\colon \widetilde{M} \to M$ be a topological covering map. Show that there is only one smooth structure on $\widetilde{M}$ such that π is a smooth covering map (see Proposition 2.12). [Hint: Use the existence of smooth local sections.]

2-8. Show that the map $\varepsilon^n\colon \mathbb{R}^n \to \mathbb{T}^n$ defined in Example 2.8(d) is a smooth covering map.

2-9. Show that the map $p\colon \mathbb{S}^n \to \mathbb{RP}^n$ defined in Example 2.5(d) is a smooth covering map.

2-10. Let $\mathbb{CP}^n$ denote n-dimensional complex projective space, as defined in Problem 1-7.

(a) Show that the quotient map $\pi\colon \mathbb{C}^{n+1} \smallsetminus \{0\} \to \mathbb{CP}^n$ is smooth.
(b) Show that $\mathbb{CP}^1$ is diffeomorphic to $\mathbb{S}^2$.

2-11. Let G be a connected Lie group, and let $U \subset G$ be any neighborhood of the identity. Show that U generates G, i.e., that every element of G can be written as a finite product of elements of U.

2-12. Let G be a Lie group, and let G_0 denote the connected component of the identity (called the *identity component* of G).

(a) Show that G_0 is the only connected open subgroup of G.
(b) Show that each connected component of G is diffeomorphic to G_0.

2-13. Let G be a connected Lie group. Show that the universal covering group $\widetilde{G}$ constructed in Theorem 2.13 is unique in the following sense:

If G' is any other simply connected Lie group that admits a smooth covering map $\pi'\colon G' \to G$ that is also a Lie group homomorphism, then there exists a Lie group isomorphism $\Phi\colon \widetilde{G} \to G'$ such that $\pi' \circ \Phi = \pi$.

2-14. Let M be a topological manifold, and let $\mathcal{U}$ be a cover of M by precompact open sets. Show that $\mathcal{U}$ is locally finite if and only if each set in $\mathcal{U}$ intersects only finitely many other sets in $\mathcal{U}$. Give a counterexample to show that the conclusion is false if either precompactness or openness is omitted from the hypotheses.

2-15. Suppose M is a locally Euclidean Hausdorff space. Show that M is second countable if and only if it is paracompact and has countably many connected components. [Hint: If M is paracompact, show that each component of M has a locally finite cover by precompact coordinate balls, and extract from this a countable subcover.]

2-16. Suppose M is a topological space with the property that for every open cover $\mathcal{X}$ of M, there exists a partition of unity subordinate to $\mathcal{X}$. Show that M is paracompact.

2-17. Show that the assumption that A is closed is necessary in the extension lemma (Lemma 2.27), by giving an example of a smooth real-valued function on a nonclosed subset of a smooth manifold that admits no smooth extension to the whole manifold.

2-18. Let M be a smooth manifold, let $B \subset M$ be a closed subset, and let $\delta\colon M \to \mathbb{R}$ be a positive continuous function. Show that there is a continuous function $\psi\colon M \to \mathbb{R}$ that is smooth and positive on $M \smallsetminus B$, zero on B, and satisfies $\psi(x) < \delta(x)$ everywhere. [Hint: Consider $1/(1 + f)$, where $f\colon M \smallsetminus B \to \mathbb{R}$ is a positive exhaustion function.]

3
Tangent Vectors

One of the key tools in our study of smooth manifolds will be the idea of *linear approximation*. This is a familiar notion from calculus in Euclidean spaces, where for example a function of one variable can be approximated by its tangent line, a parametrized curve in $\mathbb{R}^n$ by its tangent vector, a surface in $\mathbb{R}^3$ by its tangent plane, or a map from $\mathbb{R}^n$ to $\mathbb{R}^m$ by its total derivative (see the Appendix).

In order to make sense of linear approximations on manifolds, we need to introduce the notion of the tangent space to a manifold at a point, which we can think of as a sort of "linear model" for the manifold near the point. Because of the abstractness of the definition of a smooth manifold, this takes some work, which we carry out in this chapter.

We begin by studying a much more concrete object: geometric tangent vectors in $\mathbb{R}^n$, which can be thought of as "arrows" attached to a particular point in $\mathbb{R}^n$. Because the definition of smooth manifolds is built around the idea of identifying which functions are smooth, the property of a geometric tangent vector that is amenable to generalization is its action on smooth functions as a "directional derivative." The key observation about geometric tangent vectors, which we prove in the first section of this chapter, is that the process of taking directional derivatives gives a natural one-to-one correspondence between geometric tangent vectors and linear maps from $C^\infty(\mathbb{R}^n)$ to $\mathbb{R}$ satisfying the product rule. (Such maps are called "derivations.") With this as motivation, we then *define* a tangent vector on a smooth manifold as a derivation of $C^\infty(M)$ at a point.

In the second section of the chapter, we show how tangent vectors can be "pushed forward" by smooth maps. Using this, we connect the abstract definition of tangent vectors to our concrete geometric picture by showing that any smooth coordinate chart (U, φ) gives a natural isomorphism from the space of tangent vectors to M at p to the space of tangent vectors to $\mathbb{R}^n$ at $\varphi(p)$, which in turn is isomorphic to the geometric tangent vectors at $\varphi(p)$. Thus any smooth coordinate chart yields a basis for each tangent space. Using this isomorphism, we describe how to do concrete computations in such a basis.

The last part of the chapter is devoted to showing how a smooth curve in a smooth manifold has a tangent vector at each point, which can be regarded as the derivation of $C^\infty(M)$ that takes the derivative of each function along the curve. In the final section, we discuss and compare several alternative approaches to defining tangent spaces.

Tangent Vectors

Imagine a manifold in Euclidean space. For concreteness, let us take it to be the unit sphere $\mathbb{S}^{n-1} \subset \mathbb{R}^n$. What do we mean by a "tangent vector" at a point of $\mathbb{S}^{n-1}$? Before we can answer this question, we have to come to terms with a dichotomy in the way we think about an element of $\mathbb{R}^n$. On one hand, we usually think of it as a *point* in space, whose only property is its location, expressed by the coordinates $(x^1, \ldots, x^n)$. On the other hand, when doing calculus we sometimes think of it instead as a *vector*, which is an object that has magnitude and direction, but whose location is irrelevant. A vector $v = v^i e_i$ (where e_i denotes the ith standard basis vector) can be visualized as an arrow with its initial point anywhere in $\mathbb{R}^n$; what is relevant from the vector point of view is only which direction it points and how long it is.

What we really have in mind when we work with tangent vectors is a separate copy of $\mathbb{R}^n$ at each point. When we talk about the set of vectors tangent to the sphere at a point a, for example, we are imagining them as living in a copy of $\mathbb{R}^n$ with its origin translated to a.

Geometric Tangent Vectors

Here is a preliminary definition of tangent vectors in Euclidean space. Let us define the *geometric tangent space* to $\mathbb{R}^n$ at the point $a \in \mathbb{R}^n$, denoted by $\mathbb{R}^n_a$, to be the set $\{a\} \times \mathbb{R}^n$. More explicitly,

$$\mathbb{R}^n_a = \{(a, v) : v \in \mathbb{R}^n\}.$$

A *geometric tangent vector* in $\mathbb{R}^n$ is an element of this space. As a matter of notation, we will abbreviate (a, v) as v_a (or sometimes $v|_a$ if it is clearer, for example if v itself has a subscript). We think of v_a as the vector v with its initial point at a (Figure 3.1). This set $\mathbb{R}^n_a$ is a real vector space (obviously

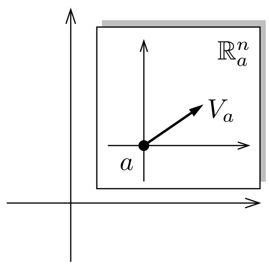


Figure 3.1. Geometric tangent space.

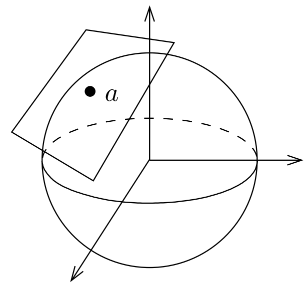


Figure 3.2. Tangent space to $\mathbb{S}^n$.

isomorphic to $\mathbb{R}^n$ itself) under the natural operations

$$v_a + w_a = (v + w)_a,$$
$$c(v_a) = (cv)_a.$$

The vectors $e_i|_a$, $i = 1, \dots, n$, are a basis for $\mathbb{R}^n_a$. In fact, as a vector space, $\mathbb{R}^n_a$ is essentially the same as $\mathbb{R}^n$ itself; the only reason we add the index a is so that the geometric tangent spaces $\mathbb{R}^n_a$ and $\mathbb{R}^n_b$ at distinct points a and b will be disjoint sets.

With this definition we could, for example, think of the tangent space to $\mathbb{S}^{n-1}$ at a point $a \in \mathbb{S}^{n-1}$ as a certain subspace of $\mathbb{R}^n_a$ (Figure 3.2), namely the space of vectors that are orthogonal to the radial unit vector through a, noting that the geometric tangent space $\mathbb{R}^n_a$ inherits an inner product from $\mathbb{R}^n$ via the natural isomorphism $\mathbb{R}^n \cong \mathbb{R}^n_a$.

The problem with this definition, however, is that it gives us no clue as to how we might set about defining tangent vectors on an arbitrary smooth manifold, where there is no ambient Euclidean space. So we need to look for another characterization of tangent vectors that might make sense on a manifold.

The only things we have to work with on smooth manifolds so far are smooth functions, smooth maps, and smooth coordinate charts. Now one thing that a Euclidean tangent vector provides is a means of taking "directional derivatives" of functions. For example, any geometric tangent vector $v_a \in \mathbb{R}^n_a$ yields a map $D_v|_a : C^\infty(\mathbb{R}^n) \to \mathbb{R}$, which takes the directional derivative in the direction v at a:

$$D_v|_a f = D_v f(a) = \left.\frac{d}{dt}\right|_{t=0} f(a + tv). \tag{3.1}$$

This operation is linear and satisfies the product rule:

$$D_v|_a (fg) = f(a)\, D_v|_a g + g(a)\, D_v|_a f.$$

If $v_a = v^i e_i\big|_a$ in terms of the standard basis, then by the chain rule $D_v|_a f$ can be written more concretely as

$$D_v|_a f = v^i \frac{\partial f}{\partial x^i}(a).$$

(Here we are using the summation convention as usual, so the expression on the right-hand side is understood to be summed over $i = 1$ to n. This sum is consistent with our index convention if we stipulate that an upper index "in the denominator" is to be regarded as a lower index.) For example, if $v_a = e_j|_a$, then

$$D_v|_a f = \frac{\partial f}{\partial x^j}(a).$$

With this construction in mind, we make the following definition. If a is a point of $\mathbb{R}^n$, a linear map $X\colon C^\infty(\mathbb{R}^n) \to \mathbb{R}$ is called a *derivation at a* if it satisfies the following product rule:

$$X(fg) = f(a)Xg + g(a)Xf. \tag{3.2}$$

Let $T_a(\mathbb{R}^n)$ denote the set of all derivations of $C^\infty(\mathbb{R}^n)$ at a. Clearly $T_a(\mathbb{R}^n)$ is a vector space under the operations

$$(X+Y)f = Xf + Yf$$
$$(cX)f = c(Xf).$$

The most important (and perhaps somewhat surprising) fact about this space is that it is finite dimensional, and in fact is naturally isomorphic to the geometric tangent space $\mathbb{R}^n_a$ that we defined above. The proof will be based on the following lemma.

Lemma 3.1 (Properties of Derivations). *Suppose $a \in \mathbb{R}^n$ and $X \in T_a(\mathbb{R}^n)$.*

(*a*) *If f is a constant function, then $Xf = 0$.*

(*b*) *If $f(a) = g(a) = 0$, then $X(fg) = 0$.*

Proof. It suffices to prove (a) for the constant function $f_1(x) \equiv 1$, for then $f(x) \equiv c$ implies $Xf = X(cf_1) = cXf_1 = 0$ by linearity. For f_1, it follows from the product rule:

$$Xf_1 = X(f_1 f_1) = f_1(a)Xf_1 + f_1(a)Xf_1 = 2Xf_1,$$

which implies that $Xf_1 = 0$. Similarly, (b) also follows from the product rule:

$$X(fg) = f(a)Xg + g(a)Xf = 0 + 0 = 0. \qquad \square$$

Now let $v_a \in \mathbb{R}^n_a$ be a geometric tangent vector at a. By the product rule, the map $D_v|_a : C^\infty(\mathbb{R}^n) \to \mathbb{R}$ defined by (3.1) is a derivation at a. As the following proposition shows, every derivation at a is of this form.

Proposition 3.2. *For any $a \in \mathbb{R}^n$, the map $v_a \mapsto D_v|_a$ is an isomorphism from $\mathbb{R}^n_a$ onto $T_a(\mathbb{R}^n)$.*

Proof. The map $v_a \mapsto D_v|_a$ is linear, as is easily checked. To see that it is injective, suppose $v_a \in \mathbb{R}^n_a$ has the property that $D_v|_a$ is the zero derivation. Writing $v_a = v^i e_i|_a$ in terms of the standard basis, and taking f to be the jth coordinate function $x^j : \mathbb{R}^n \to \mathbb{R}$, thought of as a smooth function on $\mathbb{R}^n$, we find

$$0 = D_v|_a (x^j) = v^i \frac{\partial}{\partial x^i}(x^j)\bigg|_{x=a} = v^j.$$

Since this is true for each j, it follows that v_a is the zero vector.

To prove surjectivity, let $X \in T_a(\mathbb{R}^n)$ be arbitrary. Motivated by the computation in the preceding paragraph, we define real numbers $v^1, \dots, v^n$ by

$$v^i = X(x^i).$$

We will show that $X = D_v|_a$, where $v = v^i e_i$.

To see this, let f be any smooth real-valued function on $\mathbb{R}^n$. By Taylor's formula with remainder (Theorem A.58 in the Appendix), there are smooth functions $g_1, \dots, g_n$ defined on $\mathbb{R}^n$ such that $g_i(a) = 0$ and

$$f(x) = f(a) + \sum_{i=1}^n \frac{\partial f}{\partial x^i}(a)(x^i - a^i) + \sum_{i=1}^n g_i(x)(x^i - a^i). \tag{3.3}$$

Note that the last term in (3.3) is a sum of functions, each of which is a product of two functions $g_i(x)$ and $(x^i - a^i)$ that vanish when $x = a$. Applying X to this formula and using Lemma 3.1, we obtain

$$\begin{aligned} Xf &= X(f(a)) + \sum_{i=1}^n X\left(\frac{\partial f}{\partial x^i}(a)(x^i - a^i)\right) + X\big(g_i(x)(x^i - a^i)\big) \\ &= 0 + \sum_{i=1}^n \frac{\partial f}{\partial x^i}(a)(X(x^i) - X(a^i)) + 0 \\ &= \sum_{i=1}^n \frac{\partial f}{\partial x^i}(a) v^i \\ &= D_v|_a f. \end{aligned}$$

This shows that $X = D_v|_a$. □

Corollary 3.3. *For any $a \in \mathbb{R}^n$, the n derivations*

$$\frac{\partial}{\partial x^1}\bigg|_a, \dots, \frac{\partial}{\partial x^n}\bigg|_a,$$

defined by

$$\left.\frac{\partial}{\partial x^i}\right|_a f = \frac{\partial f}{\partial x^i}(a),$$

form a basis for $T_a(\mathbb{R}^n)$*, which therefore has dimension* n*.*

Proof. This follows immediately from the preceding proposition, once we note that $\partial/\partial x^i|_a = D_{e_i}|_a$. □

Tangent Vectors on a Manifold

Now we are in a position to define tangent vectors on a manifold. Let M be a smooth manifold and let p be a point of M. A linear map $X : C^\infty(M) \to \mathbb{R}$ is called a *derivation at p* if it satisfies

$$X(fg) = f(p)Xg + g(p)Xf \tag{3.4}$$

for all $f, g \in C^\infty(M)$. The set of all derivations of $C^\infty(M)$ at p is a vector space called the *tangent space* to M at p, and is denoted by T_pM. An element of T_pM is called a *tangent vector* at p.

The following lemma is the analogue of Lemma 3.1 for manifolds.

Lemma 3.4 (Properties of Tangent Vectors on Manifolds). *Suppose* M *is a smooth manifold,* $p \in M$*, and* $X \in T_pM$*.*

(*a*) *If* f *is a constant function, then* $Xf = 0$*.*

(*b*) *If* $f(p) = g(p) = 0$*, then* $X(fg) = 0$*.*

◇ **Exercise 3.1.** Prove Lemma 3.4.

In the special case $M = \mathbb{R}^n$, Proposition 3.2 shows that $T_a\mathbb{R}^n$ is naturally isomorphic to the geometric tangent space $\mathbb{R}^n_a$, and thus also to $\mathbb{R}^n$ itself. For this reason, you should visualize tangent vectors to an abstract smooth manifold M as "arrows" that are tangent to M and whose base points are attached to M at the given point. Theorems about tangent vectors must always be proved using the abstract definition in terms of derivations, but your intuition should be guided as much as possible by the geometric picture.

Push-Forwards

To relate the abstract tangent spaces we have defined on a manifold to geometric tangent spaces in $\mathbb{R}^n$, we have to explore the way tangent vectors behave under smooth maps. In the case of a smooth map between Euclidean spaces, the total derivative of the map at a point (represented

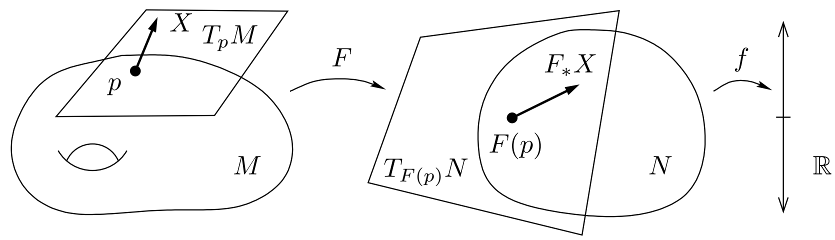


Figure 3.3. The push-forward.

by its Jacobian matrix) is a linear map that represents the "best linear approximation" to the map near the given point. In the manifold case, there is a similar linear map, but now it acts between tangent spaces.

If M and N are smooth manifolds and $F\colon M \to N$ is a smooth map, for each $p \in M$ we define a map $F_*\colon T_pM \to T_{F(p)}N$, called the *push-forward* associated with F (Figure 3.3), by

$$(F_*X)(f) = X(f \circ F).$$

Note that if $f \in C^\infty(N)$, then $f \circ F \in C^\infty(M)$, so $X(f \circ F)$ makes sense. The operator F_*X is clearly linear, and is a derivation at $F(p)$ because

$$\begin{aligned}(F_*X)(fg) &= X((fg)\circ F)\\ &= X((f\circ F)(g\circ F))\\ &= f\circ F(p)X(g\circ F) + g\circ F(p)X(f\circ F)\\ &= f(F(p))(F_*X)(g) + g(F(p))(F_*X)(f).\end{aligned}$$

Because the notation F_* does not explicitly mention the point p, we will have to be careful to specify it when necessary to avoid confusion.

Lemma 3.5 (Properties of Push-Forwards). *Let $F\colon M \to N$ and $G\colon N \to P$ be smooth maps, and let $p \in M$.*

(*a*) $F_*\colon T_pM \to T_{F(p)}N$ *is linear.*

(*b*) $(G\circ F)_* = G_* \circ F_*\colon T_pM \to T_{G\circ F(p)}P$.

(*c*) $(\mathrm{Id}_M)_* = \mathrm{Id}_{T_pM}\colon T_pM \to T_pM$.

(*d*) *If F is a diffeomorphism, then $F_*\colon T_pM \to T_{F(p)}N$ is an isomorphism.*

◇ **Exercise 3.2.** Prove Lemma 3.5.

Our first important application of the push-forward will be to use coordinate charts to relate the tangent space to a point on a manifold with the Euclidean tangent space. But there is an important technical issue that we must address first: While the tangent space is defined in terms of smooth functions on the whole manifold, coordinate charts are in general defined only on open subsets. The key point, expressed in the next lemma, is that the tangent space is really a purely local construction.

Proposition 3.6. *Suppose M is a smooth manifold, $p \in M$, and $X \in T_pM$. If f and g are smooth functions on M that agree on some neighborhood of p, then $Xf = Xg$.*

Proof. Setting $h = f - g$, by linearity it suffices to show that $Xh = 0$ whenever h vanishes in a neighborhood W of p. Let $\psi \in C^\infty(M)$ be a smooth bump function that is identically equal to 1 on the closed subset $M \smallsetminus W$ and supported in $M \smallsetminus \{p\}$. Because $\psi \equiv 1$ where h is nonzero, the product ψh is identically equal to h. Since $h(p) = \psi(p) = 0$, Lemma 3.4 implies that $Xh = X(\psi h) = 0$. □

Using this proposition, the tangent space to an open submanifold can be naturally identified with the tangent space to the whole manifold.

Proposition 3.7. *Let M be a smooth manifold, let $U \subset M$ be an open submanifold, and let $\iota\colon U \hookrightarrow M$ be the inclusion map. For any $p \in U$, $\iota_*\colon T_pU \to T_pM$ is an isomorphism.*

Proof. Let B be a small neighborhood of p such that $\overline{B} \subset U$. First suppose $X \in T_pU$ and $\iota_* X = 0 \in T_pM$. If $f \in C^\infty(U)$ is arbitrary, the extension lemma guarantees that there is a smooth function $\widetilde{f} \in C^\infty(M)$ such that $\widetilde{f} \equiv f$ on $\overline{B}$. Then by Proposition 3.6,

$$Xf = X(\widetilde{f}|_U) = X(\widetilde{f} \circ \iota) = (\iota_* X)\widetilde{f} = 0.$$

Since this holds for every $f \in C^\infty(U)$, it follows that $X = 0$, so ι_* is injective.

On the other hand, suppose $Y \in T_pM$ is arbitrary. Define an operator $X\colon C^\infty(U) \to \mathbb{R}$ by setting $Xf = Y\widetilde{f}$, where $\widetilde{f}$ is any function on all of M that agrees with f on $\overline{B}$. By Proposition 3.6, Xf is independent of the choice of $\widetilde{f}$, so X is well-defined, and it is easy to check that it is a derivation of $C^\infty(U)$ at p. For any $g \in C^\infty(M)$,

$$(\iota_* X)g = X(g \circ \iota) = Y(\widetilde{g \circ \iota}) = Yg,$$

where the last two equalities follow from the facts that $g \circ \iota$, $\widetilde{g \circ \iota}$, and g all agree on B. Therefore, ι_* is also surjective. □

If U is an open set in a smooth manifold M, the isomorphism ι_* between T_pU and T_pM is canonically defined, independently of any choices. From now on we will *identify* T_pU with T_pM for any point $p \in U$. This identification just amounts to the observation that $\iota_* X$ is the the *same derivation* as X, thought of as acting on functions on the bigger manifold M instead of functions on U. Since the action of a derivation on a function depends only on the values of the function in an arbitrarily small neighborhood, this is a harmless identification. In particular, this means that any tangent vector $X \in T_pM$ can be unambiguously applied to functions defined only in a neighborhood of p, not necessarily on all of M.

◇ **Exercise 3.3.** If $F\colon M \to N$ is a local diffeomorphism, show that $F_*\colon T_pM \to T_{F(p)}N$ is an isomorphism for every $p \in M$.

Recall from Chapter 1 that every finite-dimensional vector space has a natural smooth manifold structure that is independent of any choice of basis or norm. The following proposition shows that the tangent space to a vector space can be naturally identified with the vector space itself. Compare this with the isomorphism between $T_a\mathbb{R}^n$ and $\mathbb{R}^n_a$ that we proved in the preceding section.

Proposition 3.8 (The Tangent Space to a Vector Space). *For each finite-dimensional vector space V and each point $a \in V$, there is a natural (basis-independent) isomorphism $V \to T_aV$ such that for any linear map $L\colon V \to W$ the following diagram commutes:*

$$\begin{tikzcd} V \arrow[r, "\cong"] \arrow[d, "L"'] & T_aV \arrow[d, "L_*"] \\ W \arrow[r, "\cong"'] & T_{La}W \end{tikzcd} \tag{3.5}$$

Proof. As we did in the case of $\mathbb{R}^n$, for any vector $v \in V$, we define a derivation $D_v|_a$ of $C^\infty(V)$ at a by

$$D_v|_a f = D_v f(a) = \left.\frac{d}{dt}\right|_{t=0} f(a+tv).$$

Clearly this is independent of any choice of basis. Once we choose a basis for V, we can use the same arguments we used in the case of $\mathbb{R}^n$ to show that $D_v|_a$ is indeed a derivation at a, and that the map $v \mapsto D_v|_a$ is an isomorphism.

Now suppose $L\colon V \to W$ is a linear map. Because its components with respect to any choices of bases for V and W are linear functions, L is smooth. Unwinding the definitions and using the linearity of L, we compute

$$\begin{aligned} (L_*\, D_v|_a)f &= D_v|_a\,(f \circ L) \\ &= \left.\frac{d}{dt}\right|_{t=0} f(L(a+tv)) \\ &= \left.\frac{d}{dt}\right|_{t=0} f(La + tLv) \\ &= D_{Lv}|_{La}\, f, \end{aligned}$$

which is 3.5. □

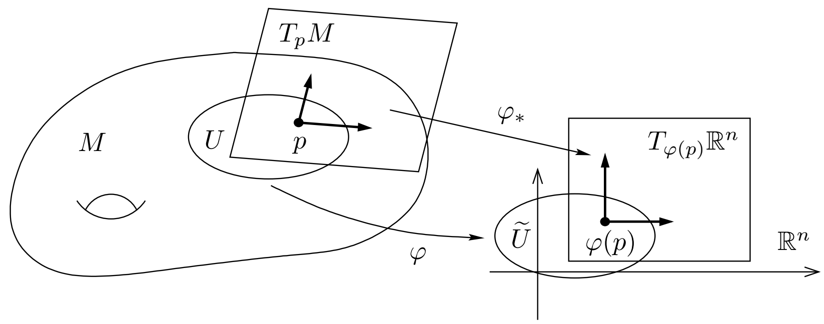


Figure 3.4. Tangent vectors in coordinates.

Using this proposition, we will routinely identify tangent vectors to a finite-dimensional vector space with elements of the space itself.

Computations in Coordinates

Our treatment of the tangent space to a manifold so far might seem hopelessly abstract. To bring it down to earth, we will show how to do computations with tangent vectors and push-forwards in local coordinates.

Let (U, φ) be a smooth coordinate chart on M. Note that φ is, in particular, a diffeomorphism from U to an open subset $\widetilde{U} \subset \mathbb{R}^n$. Thus, combining the results of Proposition 3.7 and Lemma 3.5(d) above, $\varphi_* \colon T_pM \to T_{\varphi(p)}\mathbb{R}^n$ is an isomorphism.

By Corollary 3.3, $T_{\varphi(p)}\mathbb{R}^n$ has a basis consisting of the derivations $\partial/\partial x^i|_{\varphi(p)}$, $i = 1, \ldots, n$. Therefore, the push-forwards of these vectors under $(\varphi^{-1})_*$ form a basis for T_pM (Figure 3.4). In keeping with our standard practice of treating coordinate maps as identifications, we will use the following notation for these push-forwards:

$$\left.\frac{\partial}{\partial x^i}\right|_p = (\varphi^{-1})_* \left.\frac{\partial}{\partial x^i}\right|_{\varphi(p)}.$$

Unwinding the definitions, we see that $\partial/\partial x^i|_p$ acts on a smooth function $f \colon U \to \mathbb{R}$ by

$$\begin{aligned}\left.\frac{\partial}{\partial x^i}\right|_p f &= \left.\frac{\partial}{\partial x^i}\right|_{\varphi(p)} (f \circ \varphi^{-1}) \\ &= \frac{\partial \widehat{f}}{\partial x^i}(\widehat{p}),\end{aligned}$$

where $\widehat{f} = f \circ \varphi^{-1}$ is the coordinate representation of f and $\widehat{p} = (p^1, \ldots, p^n) = \varphi(p)$ is the coordinate representation of p. In other words,

$\partial/\partial x^i|_p$ is just the derivation that takes the ith partial derivative of (the coordinate representation of) f at (the coordinate representation of) p. The vectors $\partial/\partial x^i|_p$ are called the *coordinate vectors* at p associated with the given coordinate system. In the special case of standard coordinates on $\mathbb{R}^n$, the coordinate basis vectors $\partial/\partial x^i|_a$ are literally the partial derivative operators, which correspond to the standard basis vectors $e_i|_a$ under the isomorphism $T_a\mathbb{R}^n \leftrightarrow \mathbb{R}^n_a$.

The following lemma summarizes the discussion so far.

Lemma 3.9. *Let M be a smooth n-manifold. For any $p \in M$, T_pM is an n-dimensional vector space. If $(U,(x^i))$ is any smooth chart containing p, the coordinate vectors $(\partial/\partial x^1|_p, \ldots, \partial/\partial x^n|_p)$ form a basis for T_pM.*

Thus any tangent vector $X \in T_pM$ can be written uniquely as a linear combination

$$X = X^i \left.\frac{\partial}{\partial x^i}\right|_p$$

where we are using the summation convention as usual. The numbers $(X^1, \ldots, X^n)$ are called the *components* of X with respect to the given coordinate system. If X is known, its components can be computed easily from its action on the coordinate functions. For each j, thinking of x^j as a smooth real-valued function on U, we have

$$\begin{aligned} X(x^j) &= \left(X^i \left.\frac{\partial}{\partial x^i}\right|_p \right)(x^j) \\ &= X^i \frac{\partial x^j}{\partial x^i}(p) \\ &= X^j, \end{aligned}$$

where the last line follows because $\partial x^j/\partial x^i = 0$ except when $i = j$, in which case it is equal to 1. Thus the components of X are given by $X^j = X(x^j)$.

Next we explore how push-forwards look in coordinates. We begin by considering the special case of a smooth map $F\colon U \to V$, where $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ are open subsets of Euclidean spaces. For any $p \in \mathbb{R}^n$, we will determine the matrix of $F_*\colon T_p\mathbb{R}^n \to T_{F(p)}\mathbb{R}^m$ in terms of the standard coordinate bases. Using $(x^1, \ldots, x^n)$ to denote the coordinates in the domain and $(y^1, \ldots, y^m)$ to denote those in the range, we use the chain rule to compute the action of F_* on a typical basis vector as follows:

$$\begin{aligned} \left(F_* \left.\frac{\partial}{\partial x^i}\right|_p \right) f &= \left.\frac{\partial}{\partial x^i}\right|_p (f \circ F) \\ &= \frac{\partial f}{\partial y^j}(F(p)) \frac{\partial F^j}{\partial x^i}(p) \\ &= \left(\frac{\partial F^j}{\partial x^i}(p) \left.\frac{\partial}{\partial y^j}\right|_{F(p)} \right) f. \end{aligned}$$

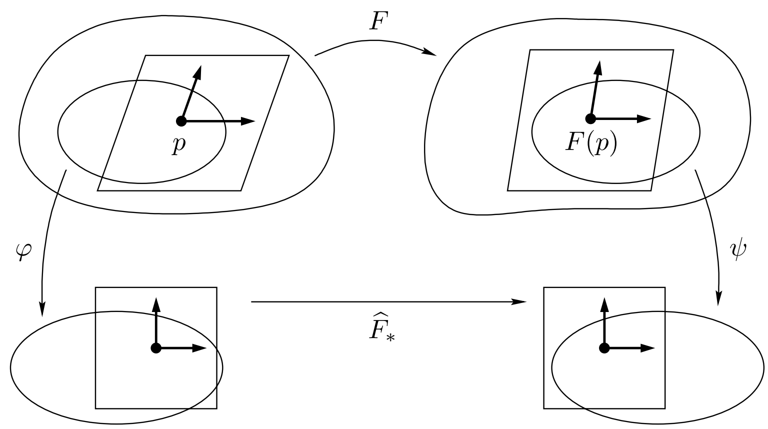


Figure 3.5. The push-forward in coordinates.

Thus

$$F_* \left.\frac{\partial}{\partial x^i}\right|_p = \frac{\partial F^j}{\partial x^i}(p) \left.\frac{\partial}{\partial y^j}\right|_{F(p)}. \tag{3.6}$$

In other words, the matrix of F_* in terms of the standard coordinate bases is

$$\begin{pmatrix} \dfrac{\partial F^1}{\partial x^1}(p) & \cdots & \dfrac{\partial F^1}{\partial x^n}(p) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial F^m}{\partial x^1}(p) & \cdots & \dfrac{\partial F^m}{\partial x^n}(p) \end{pmatrix}.$$

(Recall that the columns of the matrix of a linear map are the components of the images of the basis vectors.) This matrix is none other than the *Jacobian matrix* of F, which is the matrix representation of the total derivative $DF(p)\colon \mathbb{R}^n \to \mathbb{R}^m$. Therefore, in this special case, $F_*\colon T_p\mathbb{R}^n \to T_{F(p)}\mathbb{R}^m$ corresponds to the total derivative $DF(p)\colon \mathbb{R}^n \to \mathbb{R}^m$, under our usual identification of Euclidean space with its tangent space.

Now consider the more general case of a smooth map $F\colon M \to N$ between smooth manifolds. Choosing smooth coordinate charts (U, φ) for M near p and (V, ψ) for N near $F(p)$, we obtain the coordinate representation $\widehat{F} = \psi \circ F \circ \varphi^{-1}\colon \varphi(U \cap F^{-1}(V)) \to \psi(V)$ (Figure 3.5). By the computation above, $\widehat{F}_*$ is represented with respect to the standard coordinate bases by the Jacobian matrix of $\widehat{F}$. Using the fact that $F \circ \varphi^{-1} = \psi^{-1} \circ \widehat{F}$, we compute

$$\begin{aligned} F_* \left.\frac{\partial}{\partial x^i}\right|_p &= F_*\left((\varphi^{-1})_* \left.\frac{\partial}{\partial x^i}\right|_{\varphi(p)} \right) \\ &= (\psi^{-1})_*\left(\widehat{F}_* \left.\frac{\partial}{\partial x^i}\right|_{\varphi(p)} \right) \\ &= (\psi^{-1})_*\left(\frac{\partial \widehat{F}^j}{\partial x^i}(\widehat{p}) \left.\frac{\partial}{\partial y^j}\right|_{\widehat{F}(\varphi(p))} \right) \\ &= \frac{\partial \widehat{F}^j}{\partial x^i}(\widehat{p}) \left.\frac{\partial}{\partial y^j}\right|_{F(p)}. \end{aligned} \tag{3.7}$$

Thus F_* is represented in terms of these coordinate bases by the Jacobian matrix of (the coordinate representative of) F. In fact, the definition of the push-forward was cooked up precisely to give a coordinate-independent meaning to the Jacobian matrix of a smooth map.

Because of this, in the differential geometry literature the push-forward of a smooth map $F\colon M \to N$ is sometimes called its differential, its total derivative, or just its derivative, and can also be denoted by such symbols as

$$F'(p), \quad dF, \quad DF, \quad dF|_p, \quad DF(p), \quad \text{etc.}$$

We will stick with the notation F_* for the push-forward of a map between manifolds, and reserve $DF(p)$ for the total derivative of a map between finite-dimensional vector spaces, which in the case of Euclidean spaces we identify with the Jacobian matrix of F.

Change of Coordinates

Suppose (U, φ) and (V, ψ) are two smooth charts on M, and $p \in U \cap V$. Let us denote the coordinate functions of φ by (x^i) and those of ψ by $(\widetilde{x}^i)$. Any tangent vector at p can be represented with respect to either basis $(\partial/\partial x^i|_p)$ or $(\partial/\partial \widetilde{x}^i|_p)$. How are the two representations related?

Writing the transition map $\psi \circ \varphi^{-1}\colon \varphi(U \cap V) \to \psi(U \cap V)$ in the shorthand notation

$$\psi \circ \varphi^{-1}(x) = (\widetilde{x}^1(x), \ldots, \widetilde{x}^n(x)),$$

by (3.6) the push-forward by $\psi \circ \varphi^{-1}$ can be written

$$(\psi \circ \varphi^{-1})_* \left.\frac{\partial}{\partial x^i}\right|_{\varphi(p)} = \frac{\partial \widetilde{x}^j}{\partial x^i}(\varphi(p)) \left.\frac{\partial}{\partial \widetilde{x}^j}\right|_{\psi(p)}.$$

(See Figure 3.6.) Using the definition of coordinate vectors, we find

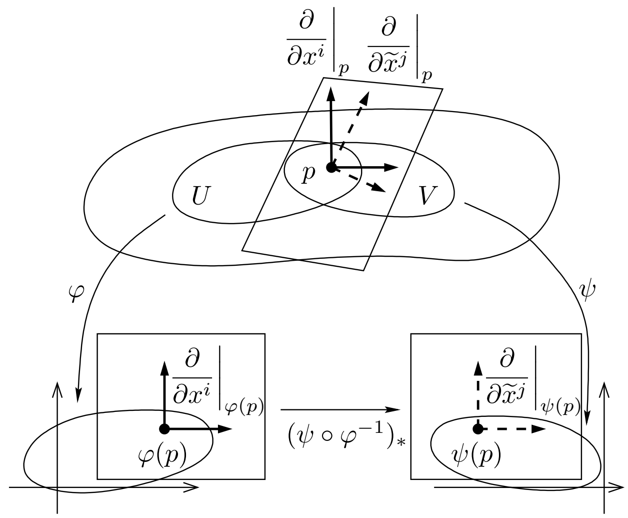


Figure 3.6. Change of coordinates.

$$\begin{aligned}
\left.\frac{\partial}{\partial x^i}\right|_p &= (\varphi^{-1})_* \left.\frac{\partial}{\partial x^i}\right|_{\varphi(p)} \\
&= (\psi^{-1})_* (\psi \circ \varphi^{-1})_* \left.\frac{\partial}{\partial x^i}\right|_{\varphi(p)} \\
&= (\psi^{-1})_* \frac{\partial \widetilde{x}^j}{\partial x^i}(\varphi(p)) \left.\frac{\partial}{\partial \widetilde{x}^j}\right|_{\psi(p)} \\
&= \frac{\partial \widetilde{x}^j}{\partial x^i}(\varphi(p)) (\psi^{-1})_* \left.\frac{\partial}{\partial \widetilde{x}^j}\right|_{\psi(p)} \\
&= \frac{\partial \widetilde{x}^j}{\partial x^i}(\widehat{p}) \left.\frac{\partial}{\partial \widetilde{x}^j}\right|_p ,
\end{aligned} \tag{3.8}$$

where $\widehat{p} = \varphi(p)$ is the representation of p in x^i-coordinates. (This formula is easy to remember, because it looks exactly the same as the chain rule for partial derivatives in $\mathbb{R}^n$.) Applying this to the components of a vector $X = X^i \partial/\partial x^i|_p = \widetilde{X}^j \partial/\partial \widetilde{x}^j|_p$, we find that the components of X transform by the rule

$$\widetilde{X}^j = \frac{\partial \widetilde{x}^j}{\partial x^i}(\widehat{p}) X^i. \tag{3.9}$$

The Tangent Space to a Manifold With Boundary

Suppose M is an n-dimensional manifold with boundary, and p is a boundary point of M. There are a number of ways one might choose to define the

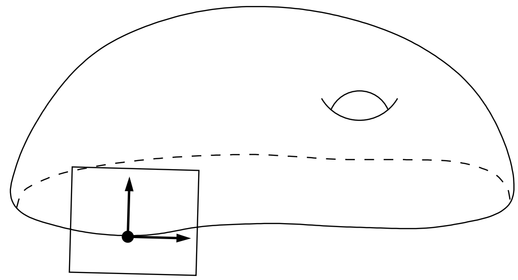


Figure 3.7. The tangent space to a manifold with boundary.

tangent space to M at p. Should it be an n-dimensional vector space, like the tangent space at an interior point? Or should it be $(n-1)$-dimensional, like the boundary? Or should it be an n-dimensional half-space, like the space $\mathbb{H}^n$ on which M is modeled locally? The standard choice is to define T_pM to be an n-dimensional vector space (Figure 3.7). This may or may not seem like the most geometrically intuitive choice, but it has the advantage of making most of the definitions of geometric objects on a manifold with boundary look exactly the same as those on a manifold.

Thus if M is a manifold with boundary and $p \in M$ is arbitrary, we define the tangent space to M at p in the same way as we defined it for a manifold: T_pM is the space of derivations of $C^\infty(M)$ at p. Similarly, if $F\colon M \to N$ is a smooth map between manifolds with boundary, we define the *push-forward* by F at $p \in M$ to be the linear map $F_*\colon T_pM \to T_{F(p)}N$ defined by the same formula as in the manifold case:

$$(F_*X)f = X(f \circ F).$$

The most important fact about these definitions is expressed in the following lemma.

Lemma 3.10. *If M is an n-dimensional manifold with boundary and p is a boundary point of M, then T_pM is an n-dimensional vector space, with basis given by the coordinate vectors $(\partial/\partial x^1|_p, \dots, \partial/\partial x^n|_p)$ in any smooth chart.*

Proof. It is obvious from the definition that T_pM is a vector space. For any smooth coordinate map φ, the push-forward $\varphi_*\colon T_pM \to T_{\varphi(p)}\mathbb{H}^n$ is an isomorphism by the same argument as in the manifold case; thus it suffices to show that for any $a \in \partial\mathbb{H}^n$, $T_a\mathbb{H}^n$ is n-dimensional and spanned by the standard coordinate vectors.

Consider the inclusion map $\iota\colon \mathbb{H}^n \hookrightarrow \mathbb{R}^n$. We will show that $\iota_*\colon T_a\mathbb{H}^n \to T_a\mathbb{R}^n$ is an isomorphism. Suppose $\iota_*X = 0$. Let f be any smooth real-valued function defined on a neighborhood of a in $\mathbb{H}^n$, and let $\widetilde{f}$ be any extension of f to a smooth function on an open subset of $\mathbb{R}^n$. (Such an extension exists by the extension lemma.) Then $\widetilde{f} \circ \iota = f$, so

$$Xf = X(\widetilde{f} \circ \iota) = (\iota_* X)\widetilde{f} = 0,$$

which implies that ι_* is injective. On the other hand, if $Y \in \mathbb{T}_a\mathbb{R}^n$ is arbitrary, define $X \in T_a\mathbb{H}^n$ by

$$Xf = Y\widetilde{f},$$

where $\widetilde{f}$ is any extension of f. Writing $Y = Y^i \partial/\partial x^i|_a$ in terms of the standard basis, this means

$$Xf = Y^i \frac{\partial \widetilde{f}}{\partial x^i}(a).$$

This is well-defined because by continuity the derivatives of $\widetilde{f}$ at a are determined by those of f in $\mathbb{H}^n$. It is easy to check that X is a derivation at a and that $Y = \iota_* X$, so ι_* is surjective. □

Tangent Vectors to Curves

The notion of the tangent vector to a smooth curve in $\mathbb{R}^n$ is familiar from elementary calculus—it is just the vector whose components are the derivatives of the component functions of the curve. In this section, we extend this notion to curves in manifolds.

If M is a manifold, we define a *curve* in M to be a continuous map $\gamma\colon J \to M$, where $J \subset \mathbb{R}$ is an interval. (In this section, we will be primarily interested in curves whose domains are open intervals, but for some purposes it is useful to allow J to have one or two endpoints; the definitions all make sense in that case if we consider J as a manifold with boundary.) Note that in this book the term "curve" will always refer to a map from an interval into M (sometimes called a *parametrized curve*), not just a set of points in M.

Our definition of the tangent space leads to a very natural interpretation of tangent vectors to smooth curves in manifolds. If γ is a smooth curve in a smooth manifold M, we define the *tangent vector to* γ at $t_0 \in J$ (Figure 3.8) to be the vector

$$\gamma'(t_0) = \gamma_* \left(\left. \frac{d}{dt} \right|_{t_0} \right) \in T_{\gamma(t_0)}M,$$

where $d/dt|_{t_0}$ is the standard coordinate basis for $T_{t_0}\mathbb{R}$. (As in ordinary calculus, it is customary to use d/dt instead of $\partial/\partial t$ when the domain is 1-dimensional.) Other common notations for the tangent vector to γ are

$$\dot{\gamma}(t_0), \quad \frac{d\gamma}{dt}(t_0), \quad \text{and} \quad \left. \frac{d\gamma}{dt} \right|_{t=t_0}.$$

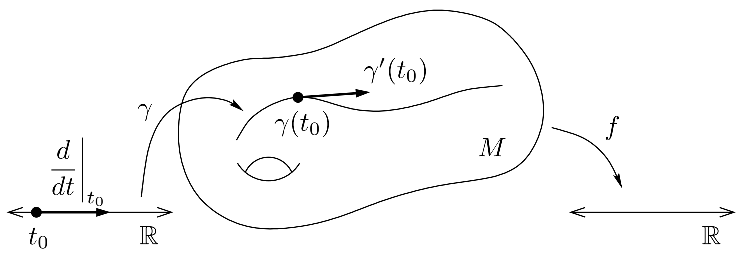


Figure 3.8. The tangent vector to a curve.

This tangent vector acts on functions by

$$\begin{aligned}\gamma'(t_0)f &= \left(\gamma_* \left.\frac{d}{dt}\right|_{t_0}\right) f \\ &= \left.\frac{d}{dt}\right|_{t_0} (f\circ\gamma) \\ &= \frac{d(f\circ\gamma)}{dt}(t_0).\end{aligned}$$

In other words, $\gamma'(t_0)$ is the derivation at $\gamma(t_0)$ obtained by taking the derivative of a function along γ. (If t_0 is an endpoint of J, this still holds provided we interpret the derivative with respect to t as a one-sided derivative.)

Now let (U,φ) be a smooth chart with coordinate functions (x^i). If $\gamma(t_0)\in U$, we can write the coordinate representation of γ as $\gamma(t) = (\gamma^1(t),\dots,\gamma^n(t))$, at least for t sufficiently near t_0, and then the formula for the push-forward in coordinates tells us that

$$\gamma'(t_0) = (\gamma^i)'(t_0) \left.\frac{\partial}{\partial x^i}\right|_{\gamma(t_0)}.$$

This means that $\gamma'(t_0)$ is given by essentially the same formula as it would be in Euclidean space: It is the tangent vector whose components in a coordinate basis are the derivatives of the component functions of γ.

The next lemma shows that every tangent vector on a manifold is the tangent vector to some curve. This gives an alternative and somewhat more geometric way to think about the tangent space: It is just the set of tangent vectors to smooth curves in M.

Lemma 3.11. *Let M be a smooth manifold and $p\in M$. Every $X\in T_pM$ is the tangent vector to some smooth curve in M.*

Proof. Let (U,φ) be a smooth coordinate chart centered at p, and write $X = X^i\partial/\partial x^i|_p$ in terms of the coordinate basis. Define a curve $\gamma\colon(-\varepsilon,\varepsilon)\to U$ by setting $\gamma(t) = (tX^1,\dots,tX^n)$ in these coordinates. (Remember, this really means $\gamma(t) = \varphi^{-1}(tX^1,\dots,tX^n)$.) Clearly this is a smooth curve with $\gamma(0) = p$, and by the computation above $\gamma'(0) = X^i \partial/\partial x^i|_{\gamma(0)} = X$. □

The next proposition shows that tangent vectors to curves behave well under composition with smooth maps.

Proposition 3.12 (The Tangent Vector to a Composite Curve). *Let $F\colon M \to N$ be a smooth map, and let $\gamma\colon J \to M$ be a smooth curve. For any $t_0 \in J$, the tangent vector at $t = t_0$ to the composite curve $F \circ \gamma\colon J \to N$ is given by*

$$(F \circ \gamma)'(t_0) = F_*(\gamma'(t_0)).$$

Proof. Just go back to the definition of the tangent vector to a curve:

$$\begin{aligned}(F \circ \gamma)'(t_0) &= (F \circ \gamma)_* \left.\frac{d}{dt}\right|_{t_0} \\ &= F_* \gamma_* \left.\frac{d}{dt}\right|_{t_0} \\ &= F_*(\gamma'(t_0)).\end{aligned}$$ □

On the face of it, the preceding proposition tells us how to compute the tangent vector to a composite curve in terms of the push-forward map. However, it is often much more useful to turn it around the other way, and use it as a streamlined way to compute push-forwards. Suppose $F\colon M \to N$ is a smooth map, and we need to compute the push-forward map F_* at some point $p \in M$. We can compute $F_* X$ for any $X \in T_p M$ by choosing a smooth curve γ whose tangent vector at $t = 0$ is X, and then

$$F_* X = (F \circ \gamma)'(0). \tag{3.10}$$

This frequently yields a much more succinct computation of F_*, especially if F is presented in some form other than by giving its coordinate functions. We will see many examples of this technique in later chapters.

Alternative Definitions of the Tangent Space

In the literature, you will find tangent vectors to a smooth manifold defined in several different ways. The most common alternative definition is based on the notion of "germs" of smooth functions, which we now define.

A *smooth function element* on a smooth manifold M is an ordered pair (f, U), where U is an open subset of M and $f\colon U \to \mathbb{R}$ is a smooth function. Given a point $p \in M$, let us define an equivalence relation on the set of all smooth function elements whose domains contain p by setting $(f, U) \sim (g, V)$ if $f \equiv g$ on some neighborhood of p. The equivalence class of a function element (f, U) is called the *germ of f at p*. The set of all germs of smooth functions at p is denoted by C_p^∞. It is a real vector space and an associative algebra under the operations

$$\begin{aligned}[(f,U)] + [(g,V)] &= [(f+g, U\cap V)],\\ c[(f,U)] &= [(cf,U)],\\ [(f,U)][(g,V)] &= [(fg, U\cap V)].\end{aligned}$$

(The zero element of this algebra is the equivalence class of the zero function on M.) Let us denote the germ at p of the function element (f,U) simply by $[f]_p$; there is no need to include the domain U in the notation, because the same germ is represented by the restriction of f to any neighborhood of p. To say that two germs $[f]_p$ and $[g]_p$ are equal is simply to say that $f \equiv g$ on some neighborhood of p, however small.

It is common to define T_pM as the vector space of derivations of C_p^∞ at p, that is, the space of all linear maps $X\colon C_p^\infty \to \mathbb{R}$ satisfying the following product rule analogous to (3.4):

$$X[fg]_p = f(p)X[g]_p + g(p)X[f]_p.$$

Thanks to Proposition 3.6, it is a simple matter to prove that this space is naturally isomorphic to the tangent space as we have defined it (see Problem 3-7). The germ definition has a number of advantages. One of the most significant is that it makes the local nature of the tangent space clearer, without requiring the use of bump functions. Because there do not exist analytic bump functions, the germ definition of tangent vectors is the only one available on real-analytic or complex-analytic manifolds. The chief disadvantage of the germ approach is simply that it adds an additional level of complication to an already highly abstract definition.

Another common approach to defining T_pM is to define an intrinsic equivalence relation on the set of smooth curves in M starting at p, which amounts to "having the same tangent vector," and to define a tangent vector as an equivalence class of curves. For example, one such equivalence relation is the following: If $\gamma_1\colon J_1 \to M$ and $\gamma_2\colon J_2 \to M$ are two smooth curves such that $\gamma(0) = \sigma(0) = p$, then we say $\gamma_1 \sim \gamma_2$ if $(f\circ\gamma_1)'(0) = (f\circ\gamma_2)'(0)$ for every smooth real-valued function f defined in a neighborhood of p. Problem 3-8 shows that the set of equivalence classes is in one-to-one correspondence with T_pM. This definition has the advantage of being geometrically more intuitive, but it has the serious drawback that the existence of a vector space structure on T_pM is not at all obvious.

Yet another approach to defining the tangent space is based on the transformation rule (3.9) for the components of tangent vectors in coordinates. One defines a tangent vector at a point $p \in M$ to be a rule that assigns a vector $(X^1,\dots,X^n) \in \mathbb{R}^n$ to each smooth coordinate chart containing p, with the property that the vectors assigned to overlapping charts transform according to (3.9). (This is, in fact, the oldest definition of all, and many physicists are still apt to define tangent vectors this way.)

It is a matter of individual taste which of the various characterizations of T_pM one chooses to take as the definition. The modern definition we have chosen, however abstract it may seem at first, has several advantages: It is relatively concrete (tangent vectors are actual derivations of $C^\infty(M)$, with no equivalence classes involved); it makes the vector space structure on T_pM obvious; and it leads to straightforward coordinate-independent definitions of many of the other geometric objects we will be studying.

Problems

3-1. Suppose M and N are smooth manifolds with M connected, and $F\colon M \to N$ is a smooth map such that $F_*\colon T_pM \to T_{F(p)}N$ is the zero map for each $p \in M$. Show that F is a constant map.

3-2. Let $M_1, \dots, M_k$ be smooth manifolds, and let $\pi_j\colon M_1 \times \cdots \times M_k \to M_j$ be the projection onto the jth factor. For any choices of points $p_i \in M_i$, $i = 1, \dots, k$, show that the map

$$\alpha\colon T_{(p_1,\dots,p_k)}(M_1 \times \cdots \times M_k) \to T_{p_1}M_1 \oplus \cdots \oplus T_{p_k}M_k$$

defined by

$$\alpha(X) = (\pi_{1*}X, \dots, \pi_{k*}X)$$

is an isomorphism. [Remark: Using this isomorphism, we can routinely identify T_pM and T_qN, for example, as subspaces of $T_{(p,q)}(M \times N)$.]

3-3. If a nonempty smooth n-manifold is diffeomorphic to an m-manifold, prove that $n = m$.

3-4. Let $C \subset \mathbb{R}^2$ be the unit circle, and let $S \subset \mathbb{R}^2$ be the boundary of the square of side 2 centered at the origin:

$$S = \{(x, y) : \max(|x|, |y|) = 1\}.$$

Show that there is a homeomorphism $F\colon \mathbb{R}^2 \to \mathbb{R}^2$ such that $F(C) = S$, but there is no *diffeomorphism* with the same property. [Hint: Consider what F does to the tangent vector to a suitable curve in C.]

3-5. Consider $\mathbb{S}^3$ as a subset of $\mathbb{C}^2$ under the usual identification of $\mathbb{C}^2$ with $\mathbb{R}^4$. For each $z = (z^1, z^2) \in \mathbb{S}^3$, define a curve $\gamma_z\colon \mathbb{R} \to \mathbb{S}^3$ by

$$\gamma_z(t) = (e^{it}z^1, e^{it}z^2).$$

(a) Compute the coordinate representation of $\gamma_z(t)$ in stereographic coordinates, and use this to show that γ_z is a smooth curve.

(b) Compute $\gamma_z'(t)$ in stereographic coordinates, and show that it is never zero.

3-6. Let G be a Lie group.

(a) Let $m\colon G \times G \to G$ denote the multiplication map. Identifying $T_{(e,e)}G \times G$ with $T_eG \oplus T_eG$ as in Problem 3-2, show that $m_*\colon T_eG \oplus T_eG \to T_eG$ is given by $m_*(X,Y) = X + Y$ [Hint: Compute $m_*(X,0)$ and $m_*(0,Y)$ separately using (3.10).]

(b) Let $i\colon G \to G$ denote the inversion map. Show that $i_*\colon T_eG \to T_eG$ is given by $i_*X = -X$.

3-7. Let M be a smooth manifold. For any point $p \in M$, let C_p^∞ denote the algebra of germs of smooth real-valued functions at p, and let $\mathcal{D}_p$ denote the vector space of derivations of C_p^∞ at p. Show that T_pM is naturally isomorphic to $\mathcal{D}_p$.

3-8. Let M be a smooth manifold and $p \in M$. A curve $\gamma\colon J \to M$ is said to *start at* p if $0 \in J$ and $\gamma(0) = p$. Define an equivalence relation on the set of smooth curves starting at p by saying $\gamma_1 \sim \gamma_2$ if $(f \circ \gamma_1)'(0) = (f \circ \gamma_2)'(0)$ for every smooth real-valued function f defined in a neighborhood of p. Let $\mathcal{C}_p$ denote the set of equivalence classes. Show that the map $\Phi\colon \mathcal{C}_p \to T_pM$ defined by $\Phi[\gamma] = \gamma'(0)$ is well-defined and yields a one-to-one correspondence between $\mathcal{C}_p$ and T_pM.

4
Vector Fields

Vector fields are familiar objects of study in multivariable calculus. In that setting, a vector field on an open subset $U \subset \mathbb{R}^n$ is simply a continuous map from U to $\mathbb{R}^n$, which can be visualized as attaching an "arrow" to each point of U. In this chapter, we show how to extend this idea to smooth manifolds.

We wish to think of a vector field on an abstract smooth manifold M as a map X that assigns to each point $p \in M$ a tangent vector $X_p \in T_pM$, together with some assumption of continuity or smoothness. But before we can think of such an object as a map, we need to define the set that will be its range. This leads to the definition of the "tangent bundle," which is the disjoint union of all tangent spaces at all points of the manifold. In the first section of the chapter, we show how the tangent bundle can be regarded in a natural way as a smooth manifold in its own right. Then we define vector fields as continuous maps from the manifold to its tangent bundle, and show how vector fields behave under the push-forward by a smooth map.

In the next section, we define the Lie bracket operation, which is a way of combining two smooth vector fields to obtain another. Then we describe the most important application of Lie brackets: The set of all smooth vector fields on a Lie group that are invariant under left multiplication is closed under Lie brackets, and thus forms an algebraic object naturally associated with the group, called the Lie algebra of the Lie group. We describe a few basic properties of Lie algebras, and compute the Lie algebras of a few familiar groups. At the end of the chapter, we show how Lie group homomorphisms induce homomorphisms of their Lie algebras, from which it follows that isomorphic Lie groups have isomorphic Lie algebras.

The Tangent Bundle

For any smooth manifold M, we define the *tangent bundle* of M, denoted by TM, to be the disjoint union of the tangent spaces at all points of M:

$$TM = \coprod_{p \in M} T_pM.$$

We will write an element of this disjoint union as an ordered pair (p, X), with $p \in M$ and $X \in T_pM$ (instead of putting the point p in the second position, as elements of a disjoint union are more commonly written). The tangent bundle comes equipped with a natural *projection map* $\pi\colon TM \to M$, which sends each vector in T_pM to the point p at which it is tangent: $\pi(p, X) = p$. We will often commit the usual mild sin of identifying T_pM with its image under the canonical injection $X \mapsto (p, X)$, and will use any of the notations (p, X), X_p, or X for a tangent vector in T_pM, depending on how much emphasis we wish to give to the point p.

The tangent bundle can be thought of simply as a collection of vector spaces; but it is much more than that. The next lemma shows that TM can be thought of as a smooth manifold in its own right.

Lemma 4.1. *For any smooth n-manifold M, the tangent bundle TM has a natural topology and smooth structure that make it into a $2n$-dimensional smooth manifold. With this structure, $\pi\colon TM \to M$ is a smooth map.*

Proof. We begin by defining the maps that will become our smooth charts. Given any smooth chart (U, φ) for M, let $(x^1, \dots, x^n)$ denote the coordinate functions of φ, and define a map $\widetilde{\varphi}\colon \pi^{-1}(U) \to \mathbb{R}^{2n}$ by

$$\widetilde{\varphi}\left(v^i \frac{\partial}{\partial x^i}\bigg|_p \right) = (x^1(p), \dots, x^n(p), v^1, \dots, v^n).$$

(See Figure 4.1.) Its image set is $\varphi(U) \times \mathbb{R}^n$, which is an open subset of $\mathbb{R}^{2n}$. It is a bijection onto its image, because its inverse can be written explicitly as

$$\widetilde{\varphi}^{-1}(x^1, \dots, x^n, v^1, \dots, v^n) = v^i \frac{\partial}{\partial x^i}\bigg|_{\varphi^{-1}(x)}.$$

Now suppose we are given two smooth charts (U, φ) and (V, ψ) for M, and let $(\pi^{-1}(U), \widetilde{\varphi})$, $(\pi^{-1}(V), \widetilde{\psi})$ be the corresponding charts on TM. The sets $\widetilde{\varphi}(\pi^{-1}(U) \cap \pi^{-1}(V)) = \varphi(U \cap V) \times \mathbb{R}^n$ and $\widetilde{\psi}(\pi^{-1}(U) \cap \pi^{-1}(V)) = \psi(U \cap V) \times \mathbb{R}^n$ are both open in $\mathbb{R}^{2n}$, and the transition map $\widetilde{\psi} \circ \widetilde{\varphi}^{-1}\colon \varphi(U \cap V) \times \mathbb{R}^n \to \psi(U \cap V) \times \mathbb{R}^n$ can be written explicitly using (3.9) as

$$\begin{aligned}
&\widetilde{\psi} \circ \widetilde{\varphi}^{-1}(x^1, \dots, x^n, v^1, \dots, v^n) \\
&\qquad\qquad = \left(\widetilde{x}^1(x), \dots, \widetilde{x}^n(x), \frac{\partial \widetilde{x}^1}{\partial x^j}(x) v^j, \dots, \frac{\partial \widetilde{x}^n}{\partial x^j}(x) v^j \right).
\end{aligned}$$

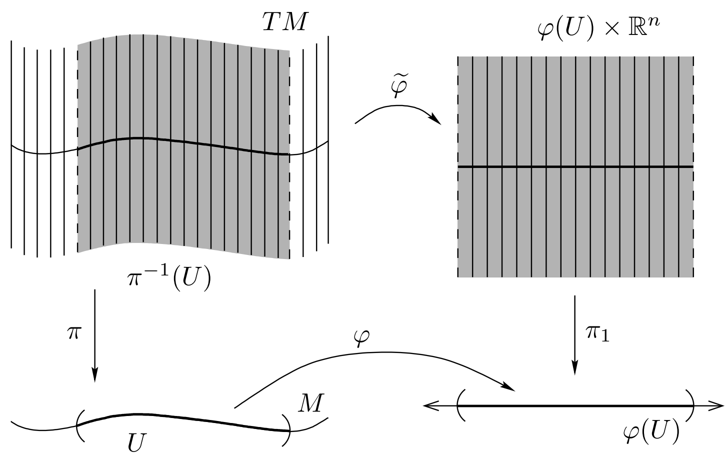


Figure 4.1. Coordinates for the tangent bundle.

This is clearly smooth.

Choosing a countable cover $\{U_i\}$ of M by smooth coordinate domains, we obtain a countable cover of TM by coordinate domains $\{\pi^{-1}(U_i)\}$ satisfying conditions (i)–(iv) of the smooth manifold construction lemma (page 21). To check the Hausdorff condition (v), just note that any two points in the same fiber of π lie in one chart; while if (p, X) and (q, Y) lie in different fibers there exist disjoint smooth coordinate domains U, V for M such that $p \in U$ and $q \in V$, and then the sets $\pi^{-1}(U)$ and $\pi^{-1}(V)$ are disjoint smooth coordinate neighborhoods containing (p, X) and (q, Y), respectively.

To check that π is smooth, we just note that its coordinate representation with respect to charts (U, φ) for M and $(\pi^{-1}(U), \widetilde{\varphi})$ for TM is $\pi(x, v) = x$. □

The coordinates (x^i, v^i) defined in this lemma will be called *standard coordinates* for TM.

◇ **Exercise 4.1.** Show that $T\mathbb{R}^n$ is diffeomorphic to $\mathbb{R}^{2n}$, and $T\mathbb{S}^1$ is diffeomorphic to $\mathbb{S}^1 \times \mathbb{R}$.

◇ **Exercise 4.2.** Suppose $F\colon M \to N$ is a smooth map. By examining the local expression (3.6) for F_* in coordinates, show that $F_*\colon TM \to TN$ is a smooth map.

Vector Fields on Manifolds

Now we can define the main concept of this chapter. If M is a smooth manifold, a *vector field* on M is a section of the map $\pi\colon TM \to M$. More

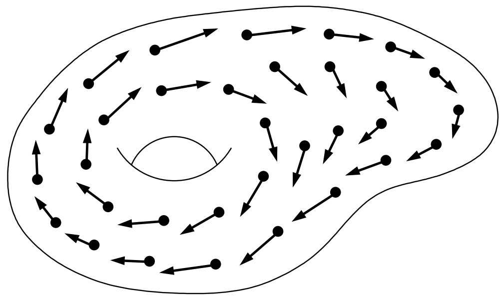


Figure 4.2. A vector field.

concretely, a vector field is a continuous map $Y\colon M \to TM$, usually written $p \mapsto Y_p$, with the property that

$$\pi \circ Y = \mathrm{Id}_M, \tag{4.1}$$

or equivalently $Y_p \in T_pM$ for each $p \in M$. (We write the value of Y at p as Y_p instead of $Y(p)$ to be consistent with our notation for elements of the tangent bundle, as well as to avoid conflict with the notation $Y(f)$ for the action of a vector on a function.) You should think of a vector field on M in the same way as you think of vector fields in Euclidean space: as an arrow attached to each point of M, chosen to be tangent to M and to vary continuously from point to point (Figure 4.2).

We will be primarily interested in *smooth vector fields*, the ones that are smooth as maps from M to TM. In addition, for some purposes it is useful to consider maps from M to TM that would be vector fields except that they might not be continuous. A *rough vector field* on M is a (not necessarily continuous) map $Y\colon M \to TM$ satisfying (4.1).

If $Y\colon M \to TM$ is a rough vector field and $(U,(x^i))$ is any smooth coordinate chart for M, we can write the value of Y at any point $p \in U$ in terms of the coordinate basis vectors:

$$Y_p = Y^i(p) \left.\frac{\partial}{\partial x^i}\right|_p. \tag{4.2}$$

This defines n functions $Y^i\colon U \to \mathbb{R}$, called the *component functions* of Y in the given chart.

Lemma 4.2 (Smoothness Criterion for Vector Fields). *Let M be a smooth manifold, and let $Y\colon M \to TM$ be a rough vector field. If $(U,(x^i))$ is any smooth coordinate chart on M, then Y is smooth on U if and only if its component functions with respect to this chart are smooth.*

Proof. Let (x^i, v^i) be the standard coordinates on $\pi^{-1}(U) \subset TM$ associated with the chart $(U,(x^i))$. By definition of standard coordinates, the coordinate representation of $Y\colon M \to TM$ on U is

$$\widehat{Y}(x) = (x^1, \dots, x^n, Y^1(x), \dots, Y^n(x)),$$

where Y^i is the ith component function of Y in x^i-coordinates. It follows immediately that smoothness of Y in U is equivalent to smoothness of its component functions. $\square$

Example 4.3. If $(U, (x^i))$ is any smooth chart on M, the assignment

$$p \mapsto \left.\frac{\partial}{\partial x^i}\right|_p$$

determines a smooth vector field on U, called the ith *coordinate vector field* and denoted by $\partial/\partial x^i$. (It is smooth because its component functions are constants.)

Example 4.4. Let θ be any angle coordinate on a proper open subset $U \subset \mathbb{S}^1$ (see Problem 1-6), and let $d/d\theta$ denote the corresponding coordinate vector field. Because any other angle coordinate $\widetilde{\theta}$ differs from θ by a constant in a neighborhood of each point, the transformation law for coordinate vector fields (3.8) shows that $d/d\theta = d/d\widetilde{\theta}$ on their common domain. For this reason, there is a globally defined vector field on $\mathbb{S}^1$ whose coordinate representation is $d/d\theta$ with respect to any angle coordinate. It is a smooth vector field because its component function is constant in any such chart. We will denote this global vector field by $d/d\theta$, even though, properly speaking, it cannot be considered as a coordinate vector field on the entire circle at once.

The next lemma shows that every tangent vector at a point can be extended to a smooth global vector field.

Lemma 4.5. *Let M be a smooth manifold. If $p \in M$ and $X \in T_pM$, there is a smooth vector field $\widetilde{X}$ on M such that $\widetilde{X}_p = X$.*

Proof. Let (x^i) be smooth coordinates on a neighborhood U of p, and let $X^i \partial/\partial x^i|_p$ be the coordinate expression for X. If ψ is a smooth bump function supported in U and with $\psi(p) = 1$, the vector field $\widetilde{X}$ defined by

$$\widetilde{X}_q = \begin{cases} \psi(q) X^i \left.\dfrac{\partial}{\partial x^i}\right|_q, & q \in U, \\ 0, & q \notin \operatorname{supp}\psi, \end{cases}$$

is easily seen to be a smooth vector field whose value at p is equal to X. $\square$

Just as for functions, the *support* of a vector field Y is defined to be the closure of the set $\{p \in M : Y_p \neq 0\}$. A vector field is said to be *compactly supported* if its support is a compact set.

If U is any open subset of M, the fact that T_pU is naturally identified with T_pM for each $p \in U$ (Proposition 3.7) allows us to identify TU with the subset $\pi^{-1}(U) \subset TM$. Therefore, a vector field on U can be thought of either as a map from U to TU or as a map from U to TM, whichever is more convenient. If Y is a vector field on M, its restriction $Y|_U$ is a vector field on U, which is smooth if Y is.

We will use the notation $\mathcal{T}(M)$ to denote the set of all smooth vector fields on M. (Some authors use $\mathcal{X}(M)$ instead of $\mathcal{T}(M)$.) It is a vector space under pointwise addition and scalar multiplication:

$$(aY + bZ)_p = aY_p + bZ_p.$$

The zero element of this vector space is the zero vector field, whose value at each $p \in M$ is $0 \in T_pM$. In addition, smooth vector fields can be multiplied by smooth real-valued functions: If $f \in C^\infty(M)$ and $Y \in \mathcal{T}(M)$, we define $fY\colon M \to TM$ by

$$(fY)_p = f(p)Y_p.$$

The next exercise shows that these operations yield smooth vector fields.

◇ **Exercise 4.3.** If Y and Z are smooth vector fields on M and $f, g \in C^\infty(M)$, show that $fY + gZ$ is a smooth vector field.

◇ **Exercise 4.4.** Show that $\mathcal{T}(M)$ is a module over the ring $C^\infty(M)$.

For example, the basis expression (4.2) for a vector field Y can also be written as an equation between vector *fields* instead of an equation between vectors at a point:

$$Y = Y^i \frac{\partial}{\partial x^i},$$

where Y^i is the ith component function of Y in the given coordinates.

An essential property of vector fields is that they define operators on the space of smooth real-valued functions. If $Y \in \mathcal{T}(M)$ and f is a smooth real-valued function defined on an open set $U \subset M$, we obtain a new function $Yf\colon U \to \mathbb{R}$, defined by

$$Yf(p) = Y_p f.$$

(Be careful not to confuse the notations fY and Yf: The former is the smooth *vector field* obtained by multiplying Y by f, while the latter is the real-valued *function* on M obtained by applying the vector field Y to the smooth function f.) Because the action of a tangent vector on a function is determined by the values of the function in an arbitrarily small neighborhood, it follows that Yf is locally determined. In particular, for any open set $V \subset U$, $(Yf)|_V = Y(f|_V)$.

This way of viewing vector fields yields another useful criterion for a vector field to be smooth.

Lemma 4.6. *Let M be a smooth manifold, and let $Y: M \to TM$ be a rough vector field. Then Y is smooth if and only if for every open set $U \subset M$ and every $f \in C^\infty(U)$, the function $Yf: U \to \mathbb{R}$ is smooth.*

Proof. Suppose Y is a rough vector field for which Yf is smooth whenever f is smooth. If (x^i) are any smooth local coordinates on $U \subset M$, we can think of each coordinate x^i as a smooth function on U. Applying Y to one of these functions, we obtain

$$Yx^i = Y^j \frac{\partial}{\partial x^j}(x^i) = Y^i.$$

Because Yx^i is smooth by assumption, it follows that the component functions of Y are smooth, so Y is smooth. Conversely, suppose Y is smooth, and let f be a smooth real-valued function defined in an open set $U \subset M$. For any $p \in U$, we can choose smooth coordinates (x^i) on a neighborhood $W \subset U$ of p. Then for $x \in W$, we can write

$$\begin{aligned} Yf(x) &= \left(Y^i(x) \left. \frac{\partial}{\partial x^i} \right|_x \right) f \\ &= Y^i(x) \frac{\partial f}{\partial x^i}(x). \end{aligned}$$

Since the component functions Y^i are smooth on W by Lemma 4.2, it follows that Yf is smooth on W. Since the same is true in a neighborhood of each point of U, Yf is smooth on U. □

One consequence of the preceding lemma is that a smooth vector field $Y \in \mathcal{T}(M)$ defines a map from $C^\infty(M)$ to itself by $f \mapsto Yf$. This map is clearly linear over $\mathbb{R}$. Moreover, the product rule (3.4) for tangent vectors translates into the following product rule for vector fields:

$$Y(fg) = f\,Yg + g\,Yf, \tag{4.3}$$

as you can easily check by evaluating both sides at an arbitrary point $p \in M$. In general, a map $Y: C^\infty(M) \to C^\infty(M)$ is called a *derivation* (as distinct from a derivation at p, defined in Chapter 3) if it is linear over $\mathbb{R}$ and satisfies (4.3) for all $f, g \in C^\infty(M)$.

The next proposition shows that derivations of $C^\infty(M)$ can be identified with smooth vector fields.

Proposition 4.7. *Let M be a smooth manifold. A map $\mathcal{Y}: C^\infty(M) \to C^\infty(M)$ is a derivation if and only if it is of the form $\mathcal{Y}f = Yf$ for some smooth vector field $Y \in \mathcal{T}(M)$.*

Proof. We just showed that every smooth vector field induces a derivation. Conversely, suppose $\mathcal{Y}: C^\infty(M) \to C^\infty(M)$ is a derivation. We need to concoct a vector field Y so that $\mathcal{Y}f = Yf$ for all f. From the discussion above, it is clear that if there is such a vector field, its value at $p \in M$ must

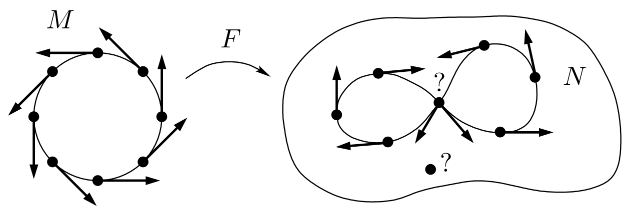


Figure 4.3. Vector fields do not always push forward.

be the derivation at p whose action on any smooth real-valued function f is given by

$$Y_p f = (\mathcal{Y} f)(p).$$

The linearity of $\mathcal{Y}$ guarantees that this expression depends linearly on f, and evaluating (4.3) at p yields the product rule (3.4) for tangent vectors. Thus the map $Y_p \colon C^\infty(M) \to \mathbb{R}$ so defined is indeed a tangent vector, i.e., a derivation of $C^\infty(M)$ at p.

To show that the assignment $p \mapsto Y_p$ is a smooth vector field, we will use Lemma 4.6. If $f \in C^\infty(M)$ is a *globally-defined* smooth function, then $Yf = \mathcal{Y} f$ is certainly smooth; we need to show the same thing holds for a smooth function defined only on an open subset of M. Suppose therefore that $U \subset M$ is open and $f \in C^\infty(U)$. For any $p \in U$, let ψ be a smooth bump function that is equal to 1 in a neighborhood of p and supported in U, and define $\widetilde{f} = \psi f$, extended to be zero on $M \smallsetminus \operatorname{supp} \psi$. Then $Y\widetilde{f} = \mathcal{Y}\widetilde{f}$ is smooth, and is equal to Yf in a neighborhood of p by Proposition 3.6. This shows that Yf is smooth in a neighborhood of each point of U. □

Because of this result, we will sometimes *identify* smooth vector fields on M with derivations of $C^\infty(M)$, using the same letter for both the vector field (thought of as a smooth map from M to TM) and the derivation (thought of as a linear map from $C^\infty(M)$ to itself).

Push-Forwards of Vector Fields

If $F \colon M \to N$ is a smooth map and Y is a vector field on M, then for each point $p \in M$, we obtain a vector $F_* Y_p \in T_{F(p)} N$ by pushing forward Y_p. However, this does not in general define a vector field on N. For example, if F is not surjective, there is no way to decide what vector to assign to a point $q \in N \smallsetminus F(M)$ (Figure 4.3). If F is not injective, then for some points of N there may be several different vectors obtained as push-forwards of Y from different points of M.

If $F \colon M \to N$ is smooth and Y is a vector field on M, suppose there happens to be a vector field Z on N with the property that for each $p \in M$,

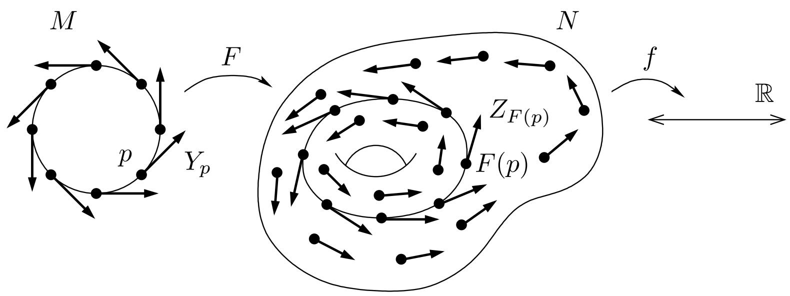


Figure 4.4. F-related vector fields.

$F_*Y_p = Z_{F(p)}$. In this case, we say the vector fields Y and Z are *F-related*. (See Figure 4.4.)

Here is a useful criterion for checking that two vector fields are F-related.

Lemma 4.8. *Suppose $F\colon M \to N$ is a smooth map, $Y \in \mathfrak{T}(M)$, and $Z \in \mathfrak{T}(N)$. Then Y and Z are F-related if and only if for every smooth real-valued function f defined on an open subset of N,*

$$Y(f \circ F) = (Zf) \circ F. \tag{4.4}$$

Proof. For any $p \in M$,

$$\begin{aligned} Y(f \circ F)(p) &= Y_p(f \circ F) \\ &= (F_*Y_p)f, \end{aligned}$$

while

$$\begin{aligned} (Zf) \circ F(p) &= (Zf)(F(p)) \\ &= Z_{F(p)}f. \end{aligned}$$

Thus (4.4) is true for all f if and only if $F_*Y_p = Z_{F(p)}$ for all p, i.e., if and only if Y and Z are F-related. □

Example 4.9. Let $F\colon \mathbb{R} \to \mathbb{R}^2$ be the smooth map $F(t) = (\cos t, \sin t)$. Then $d/dt \in \mathfrak{T}(\mathbb{R})$ is F-related to the vector field $Z \in \mathfrak{T}(\mathbb{R}^2)$ defined by

$$Z = x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}.$$

◇ **Exercise 4.5.** Prove the claim in the preceding example in two ways: directly from the definition, and by using Lemma 4.8.

It is important to remember that for a given smooth map $F\colon M \to N$ and vector field $Y \in \mathfrak{T}(M)$, there may not be *any* vector field on N that is F-related to Y. There is one special case, however, in which there is always such a vector field, as the next proposition shows.

Proposition 4.10. *Suppose $F\colon M \to N$ is a diffeomorphism. For every $Y \in \mathcal{T}(M)$, there is a unique smooth vector field on N that is F-related to Y.*

Proof. For $Z \in \mathcal{T}(N)$ to be F-related to Y means that $F_* Y_p = Z_{F(p)}$ for every $p \in M$. If F is a diffeomorphism, therefore, we *define* Z by

$$Z_q = F_*(Y_{F^{-1}(q)}).$$

It is clear that Z, so defined, is the unique (rough) vector field that is F-related to Y. To see that it is smooth, we just expand the definition in smooth local coordinates, using formula (3.6) for the push-forward:

$$Z_q = \frac{\partial F^j}{\partial x^i}(F^{-1}(q))Y^i(F^{-1}(q)) \left.\frac{\partial}{\partial y^j}\right|_q .$$

The component functions of Z are smooth by composition. □

In the situation of the preceding lemma, we will denote the unique vector field that is F-related to Y by F_*Y, and call it the *push-forward* of Y by F. Remember, it is only when F is a diffeomorphism that F_*Y is defined.

Vector Fields on a Manifold with Boundary

If M is a smooth manifold with boundary, the tangent bundle TM is defined in exactly the same way as on a manifold, as the disjoint union of the tangent spaces at all points of M. An argument entirely analogous to that of Lemma 4.1 shows that TM has a natural topology and smooth structure making it into a smooth manifold with boundary; if $(U, (x^i))$ is any smooth boundary chart for M, it is easy to verify that the standard coordinate chart $(\pi^{-1}(U), (x^i, v^i))$ is a boundary chart for TM. Just as in the manifold case, a vector field on M is a smooth section of $\pi\colon TM \to M$. All of the results of this section hold equally well in that case, although for simplicity we have stated them only for manifolds.

Lie Brackets

In this section, we introduce an important way of combining two smooth vector fields to obtain another vector field.

Let V and W be smooth vector fields on a smooth manifold M. Given a smooth function $f\colon M \to \mathbb{R}$, we can apply V to f and obtain another smooth function Vf (cf. Lemma 4.6). In turn, we can apply W to this function, and obtain yet another smooth function $WVf = W(Vf)$. The operation $f \mapsto WVf$, however, does not in general satisfy the product rule and thus cannot be a vector field, as the following example shows.

Example 4.11. Let $V = \partial/\partial x$ and $W = \partial/\partial y$ on $\mathbb{R}^2$, and let $f(x,y) = x$, $g(x,y) = y$. Then direct computation shows that $VW(fg) = 1$, while $f\,VWg + g\,VWf = 0$, so VW is not a derivation of $C^\infty(\mathbb{R}^2)$.

Now, we can also apply the same two vector fields in the opposite order, obtaining a (usually different) function WVf. Applying both of these operators to f and subtracting, we obtain an operator $[V,W]\colon C^\infty(M) \to C^\infty(M)$, called the *Lie bracket* of V and W, defined by

$$[V,W]f = VWf - WVf.$$

The key fact is that this operation *is* a vector field.

Lemma 4.12. *The Lie bracket of any pair of smooth vector fields is a smooth vector field.*

Proof. By Proposition 4.7, it suffices to show that $[V,W]$ is a derivation of $C^\infty(M)$. For arbitrary $f, g \in C^\infty(M)$, we compute

$$\begin{aligned}
[V,W](fg) &= V(W(fg)) - W(V(fg)) \\
&= V(f\,Wg + g\,Wf) - W(f\,Vg + g\,Vf) \\
&= Vf\,Wg + f\,VWg + Vg\,Wf + g\,VWf \\
&\qquad - Wf\,Vg - f\,WVg - Wg\,Vf - g\,WVf \\
&= f\,VWg + g\,VWf - f\,WVg - g\,WVf \\
&= f[V,W]g + g[V,W]f.
\end{aligned}$$

□

We will describe one significant application of Lie brackets later in this chapter, and we will see others in Chapters 12, 18, 19, and 20. Unfortunately, we are not yet in a position to give Lie brackets a geometric interpretation; that will have to wait until Chapter 18. For now, we develop some of their basic properties.

The value of the vector field $[V,W]$ at a point $p \in M$ is the derivation at p given by the formula

$$[V,W]_p f = V_p(Wf) - W_p(Vf).$$

However, this formula is of limited usefulness for practical computations, because it requires one to compute terms involving second derivatives of f that will always cancel each other out. The next lemma gives an extremely useful coordinate formula for the Lie bracket, in which the cancellations have already been accounted for.

Lemma 4.13. *Let V, W be smooth vector fields on a smooth manifold M, and let $V = V^i\partial/\partial x^i$ and $W = W^j\partial/\partial x^j$ be the coordinate expressions for V and W in terms of some smooth local coordinates (x^i) for M. Then $[V,W]$ has the following coordinate expression:*

$$[V,W] = \left(V^i \frac{\partial W^j}{\partial x^i} - W^i \frac{\partial V^j}{\partial x^i} \right) \frac{\partial}{\partial x^j}, \tag{4.5}$$

or more concisely,

$$[V, W] = (VW^j - WV^j)\frac{\partial}{\partial x^j}. \tag{4.6}$$

Proof. Because we know already that $[V, W]$ is a smooth vector field, its values are determined locally: $([V, W]f)|_U = [V, W](f|_U)$. Thus it suffices to compute in a single smooth chart, where we have:

$$\begin{aligned}
[V, W]f &= V^i \frac{\partial}{\partial x^i}\left(W^j \frac{\partial f}{\partial x^j}\right) - W^j \frac{\partial}{\partial x^j}\left(V^i \frac{\partial f}{\partial x^i}\right) \\
&= V^i \frac{\partial W^j}{\partial x^i}\frac{\partial f}{\partial x^j} + V^i W^j \frac{\partial^2 f}{\partial x^i \partial x^j} - W^j \frac{\partial V^i}{\partial x^j}\frac{\partial f}{\partial x^i} - W^j V^i \frac{\partial^2 f}{\partial x^j \partial x^i} \\
&= V^i \frac{\partial W^j}{\partial x^i}\frac{\partial f}{\partial x^j} - W^j \frac{\partial V^i}{\partial x^j}\frac{\partial f}{\partial x^i},
\end{aligned}$$

where in the last step we have used the fact that mixed partial derivatives of a smooth function can be taken in any order. Reversing the roles of the dummy indices i and j in the second term, we obtain (4.5). $\square$

One trivial application of formula (4.5) is to compute the Lie brackets of the coordinate vector fields $(\partial/\partial x^i)$ in any smooth chart: Because the component functions of the coordinate vector fields are all constants, it follows that $[\partial/\partial x^i, \partial/\partial x^j] = 0$ for any i and j. (This also follows from the definition of the Lie bracket, and is essentially a restatement of the fact that mixed partial derivatives of smooth functions commute.) Here is a slightly less trivial computation.

Example 4.14. Define smooth vector fields $V, W \in \mathfrak{T}(\mathbb{R}^3)$ by

$$\begin{aligned}
V &= x\frac{\partial}{\partial x} + \frac{\partial}{\partial y} + x(y+1)\frac{\partial}{\partial z}, \\
W &= \frac{\partial}{\partial x} + y\frac{\partial}{\partial z}.
\end{aligned}$$

Then formula (4.6) yields

$$\begin{aligned}
[V, W] &= V(1)\frac{\partial}{\partial x} + V(y)\frac{\partial}{\partial z} - W(x)\frac{\partial}{\partial x} - W(1)\frac{\partial}{\partial y} - W(x(y+1))\frac{\partial}{\partial z} \\
&= 0\frac{\partial}{\partial x} + 1\frac{\partial}{\partial z} - 1\frac{\partial}{\partial x} - 0\frac{\partial}{\partial y} - (y+1)\frac{\partial}{\partial z} \\
&= -\frac{\partial}{\partial x} - y\frac{\partial}{\partial z}.
\end{aligned}$$

Lemma 4.15 (Properties of the Lie Bracket). *The Lie bracket satisfies the following identities for all $V, W, X \in \mathfrak{T}(M)$:*

(*a*) Bilinearity: *For $a, b \in \mathbb{R}$,*

$$\begin{aligned}
[aV + bW, X] &= a[V, X] + b[W, X], \\
[X, aV + bW] &= a[X, V] + b[X, W].
\end{aligned}$$

(*b*) Antisymmetry:

$$[V, W] = -[W, V].$$

(*c*) Jacobi Identity:

$$[V, [W, X]] + [W, [X, V]] + [X, [V, W]] = 0.$$

(*d*) *For* $f, g \in C^\infty(M)$,

$$[fV, gW] = fg[V, W] + (fVg)W - (gWf)V. \tag{4.7}$$

Proof. Bilinearity and antisymmetry are obvious consequences of the definition. The proof of the Jacobi identity is just a computation:

$$\begin{aligned}
&[V,[W, X]]f + [W, [X, V]]f + [X, [V, W]]f \\
&\quad = V[W, X]f - [W, X]Vf + W[X, V]f \\
&\qquad - [X, V]Wf + X[V, W]f - [V, W]Xf \\
&\quad = VWXf - VXWf - WXVf + XWVf + WXVf - WVXf \\
&\qquad - XVWf + VXWf + XVWf - XWVf - VWXf + WVXf.
\end{aligned}$$

In this last expression, all the terms cancel in pairs. Part (d) is an easy consequence of the definition, and is left as an exercise. □

◇ **Exercise 4.6.** Prove part (d) of the preceding lemma.

Proposition 4.16 (Naturality of the Lie Bracket). *Let* $F\colon M \to N$ *be a smooth map, and let* $V_1, V_2 \in \mathcal{T}(M)$ *and* $W_1, W_2 \in \mathcal{T}(N)$ *be vector fields such that* V_i *is* F*-related to* W_i *for* $i = 1, 2$. *Then* $[V_1, V_2]$ *is* F*-related to* $[W_1, W_2]$.

Proof. Using Lemma 4.8 and the fact that V_i and W_i are F-related,

$$\begin{aligned}
V_1V_2(f \circ F) &= V_1((W_2 f) \circ F) \\
&= (W_1W_2 f) \circ F.
\end{aligned}$$

Similarly,

$$V_2V_1(f \circ F) = (W_2W_1 f) \circ F.$$

Therefore,

$$\begin{aligned}
[V_1, V_2](f \circ F) &= V_1V_2(f \circ F) - V_2V_1(f \circ F) \\
&= (W_1W_2 f) \circ F - (W_2W_1 f) \circ F \\
&= ([W_1, W_2]f) \circ F.
\end{aligned}$$

□

Corollary 4.17. *Suppose* $F\colon M \to N$ *is a diffeomorphism and* $V_1, V_2 \in \mathcal{T}(M)$. *Then* $F_*[V_1, V_2] = [F_*V_1, F_*V_2]$.

Proof. This is just the special case of Proposition 4.16 in which F is a diffeomorphism and $W_i = F_*V_i$. □

The Lie Algebra of a Lie Group

The most important application of Lie brackets occurs in the context of Lie groups. Suppose G is a Lie group. Any $g \in G$ defines maps $L_g, R_g \colon G \to G$, called *left translation* and *right translation*, respectively, by

$$L_g(h) = gh, \qquad R_g(h) = hg.$$

Because L_g can be written as the composition of smooth maps

$$G \xrightarrow{\iota_g} G \times G \xrightarrow{m} G,$$

where $\iota_g(h) = (g, h)$ and m is multiplication, it follows that L_g is smooth. It is actually a diffeomorphism of G, because $L_{g^{-1}}$ is a smooth inverse for it. Similarly, $R_g \colon G \to G$ is a diffeomorphism. Observe that, given any two points $g_1, g_2 \in G$, there is a unique left translation of G taking g_1 to g_2, namely left translation by $g_2 g_1^{-1}$. Many of the important properties of Lie groups follow, as you will see below and repeatedly in later chapters, from the fact that we can systematically map any point to any other by such a global diffeomorphism.

A vector field X on G is said to be *left-invariant* if it is invariant under all left translations, in the sense that it is L_g-related to itself for every $g \in G$. More explicitly, this means

$$(L_g)_* X_{g'} = X_{gg'}, \qquad \text{for all } g, g' \in G. \tag{4.8}$$

Since L_g is a diffeomorphism, this can be abbreviated by writing $(L_g)_* X = X$ for every $g \in G$.

Because $(L_g)_*(aX + bY) = a(L_g)_*X + b(L_g)_*Y$, the set of all smooth left-invariant vector fields on G is a linear subspace of $\mathcal{T}(M)$. But it is much more than that. The central fact is that it is closed under Lie brackets.

Lemma 4.18. *Let G be a Lie group, and suppose X and Y are smooth left-invariant vector fields on G. Then $[X, Y]$ is also left-invariant.*

Proof. Since $(L_g)_* X = X$ and $(L_g)_* Y = Y$ by definition of left-invariance, it follows from Corollary 4.17 that

$$(L_g)_*[X, Y] = [(L_g)_* X, (L_g)_* Y] = [X, Y].$$

Thus $[X, Y]$ is L_g-related to itself, i.e., is left-invariant. □

A *Lie algebra* is a real vector space $\mathfrak{g}$ endowed with a map called the *bracket* from $\mathfrak{g} \times \mathfrak{g}$ to $\mathfrak{g}$, usually denoted by $(X, Y) \mapsto [X, Y]$, that satisfies the following properties for all $X, Y, Z \in \mathfrak{g}$:

(i) Bilinearity: For $a, b \in \mathbb{R}$,

$$[aX + bY, Z] = a[X, Z] + b[Y, Z],$$
$$[Z, aX + bY] = a[Z, X] + b[Z, Y].$$

(ii) Antisymmetry:

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

(iii) Jacobi Identity:

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

Notice that the Jacobi identity is a substitute for associativity, which does not hold in general for brackets in a Lie algebra.

If $\mathfrak{g}$ is a Lie algebra, a linear subspace $\mathfrak{h} \subset \mathfrak{g}$ is called a *Lie subalgebra* of $\mathfrak{g}$ if it is closed under brackets. In this case $\mathfrak{h}$ is itself a Lie algebra with the restriction of the same bracket.

If $\mathfrak{g}$ and $\mathfrak{h}$ are Lie algebras, a linear map $A\colon \mathfrak{g} \to \mathfrak{h}$ is called a *Lie algebra homomorphism* if it preserves brackets: $A[X, Y] = [AX, AY]$. An invertible Lie algebra homomorphism is called a *Lie algebra isomorphism*. If there exists a Lie algebra isomorphism from $\mathfrak{g}$ to $\mathfrak{h}$, we say they are *isomorphic* as Lie algebras.

◇ **Exercise 4.7.** Verify that the kernel and image of a Lie algebra homomorphism are Lie subalgebras.

◇ **Exercise 4.8.** If $\mathfrak{g}$ and $\mathfrak{h}$ are finite-dimensional Lie algebras and $A\colon \mathfrak{g} \to \mathfrak{h}$ is a linear map, show that A is a Lie algebra homomorphism if and only if $A[E_i, E_j] = [AE_i, AE_j]$ for some basis $(E_1, \dots, E_n)$ of $\mathfrak{g}$.

Example 4.19 (Lie Algebras).

(a) The space $\mathcal{T}(M)$ of all smooth vector fields on a smooth manifold M is a Lie algebra under the Lie bracket by Lemma 4.12.

(b) If G is a Lie group, the set of all smooth left-invariant vector fields on G is a Lie subalgebra of $\mathcal{T}(G)$ and is therefore a Lie algebra.

(c) The vector space $\mathrm{M}(n, \mathbb{R})$ of $n \times n$ real matrices becomes an n^2-dimensional Lie algebra under the *commutator bracket*:

$$[A, B] = AB - BA.$$

Bilinearity and antisymmetry are obvious from the definition, and the Jacobi identity follows from a straightforward calculation. When we are regarding $\mathrm{M}(n, \mathbb{R})$ as a Lie algebra with this bracket, we will denote it by $\mathfrak{gl}(n, \mathbb{R})$.

(d) Similarly, $\mathfrak{gl}(n, \mathbb{C})$ is the $2n^2$-dimensional (real) Lie algebra obtained by endowing $\mathrm{M}(n, \mathbb{C})$ with the commutator bracket.

(e) If V is a vector space, the linear space $\mathfrak{gl}(V)$ of all linear maps from V to itself becomes a Lie algebra with the commutator bracket:

$$[A, B]x = A(Bx) - B(Ax).$$

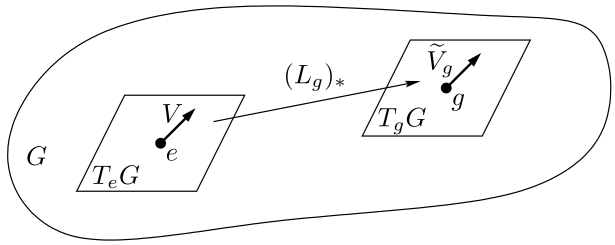


Figure 4.5. Defining a left-invariant vector field.

Under our usual identification of $n \times n$ matrices with linear maps from $\mathbb{R}^n$ to itself, $\mathfrak{gl}(\mathbb{R}^n)$ is the same as $\mathfrak{gl}(n, \mathbb{R})$.

(f) Any vector space V becomes a Lie algebra if we define all brackets to be zero. Such a Lie algebra is said to be *abelian*. (The name comes from the fact that brackets in most Lie algebras, as in the preceding examples, are defined as commutators in terms of underlying associative products; so "abelian" refers to the fact that all brackets are zero precisely when the underlying product is commutative.)

Example (b) is the most important one. The Lie algebra of all smooth left-invariant vector fields on a Lie group G is called the *Lie algebra of* G, and is denoted by $\mathrm{Lie}(G)$. (We will see below that the assumption of smoothness is redundant—see Corollary 4.21.) The fundamental fact is that $\mathrm{Lie}(G)$ is finite-dimensional, and in fact has the same dimension as G itself, as the following theorem shows.

Theorem 4.20. *Let G be a Lie group. The evaluation map $\varepsilon\colon \mathrm{Lie}(G) \to T_eG$, given by $\varepsilon(X) = X_e$, is a vector space isomorphism. Thus $\mathrm{Lie}(G)$ is finite-dimensional, with dimension equal to $\dim G$.*

Proof. We will prove the theorem by constructing an inverse for ε. For each $V \in T_eG$, define a (rough) vector field $\widetilde{V}$ on G by

$$\widetilde{V}_g = (L_g)_* V. \tag{4.9}$$

(See Figure 4.5.) If there is a left-invariant vector field on G whose value at the identity is V, clearly it has to be given by this formula.

First we need to check that $\widetilde{V}$ is smooth. By Lemma 4.6, it suffices to show that $\widetilde{V}f$ is smooth whenever f is a smooth real-valued function on an open set $U \subset G$. Choose a smooth curve $\gamma\colon (-\varepsilon, \varepsilon) \to G$ such that $\gamma(0) = e$

and $\gamma'(0) = V$. Then for $g \in U$,

$$\begin{aligned}(\widetilde{V} f)(g) &= \widetilde{V}_g f \\ &= ((L_g)_* V) f \\ &= V(f \circ L_g) \\ &= \gamma'(0)(f \circ L_g) \\ &= \left.\frac{d}{dt}\right|_{t=0} (f \circ L_g \circ \gamma)(t).\end{aligned}$$

If we define $\varphi\colon (-\varepsilon, \varepsilon) \times G \to \mathbb{R}$ by $\varphi(t, g) = f \circ L_g \circ \gamma(t) = f(g\gamma(t))$, the computation above shows that $(\widetilde{V} f)(g) = \partial\varphi/\partial t(0, g)$. Because φ is a composition of group multiplication, f, and γ, it is smooth. It follows that $\partial\varphi/\partial t(0, g)$ depends smoothly on g, so $\widetilde{V} f$ is smooth.

Next we need to verify that $\widetilde{V}$ is left-invariant, which is to say that $(L_h)_* \widetilde{V}_g = \widetilde{V}_{hg}$ for all $g, h \in G$. This follows from the definition of $\widetilde{V}$ and the fact that $L_h \circ L_g = L_{hg}$:

$$(L_h)_* \widetilde{V}_g = (L_h)_* (L_g)_* V = (L_{hg})_* V = \widetilde{V}_{hg}.$$

Thus $\widetilde{V} \in \operatorname{Lie}(G)$.

Finally, we check that the map $\tau\colon V \mapsto \widetilde{V}$ is an inverse for ε. On the one hand, given a vector $V \in T_e G$,

$$\varepsilon(\tau(V)) = \varepsilon(\widetilde{V}) = (\widetilde{V})_e = (L_e)_* V = V,$$

which shows that $\varepsilon \circ \tau$ is the identity on $T_e G$. On the other hand, given a vector field $X \in \operatorname{Lie}(G)$,

$$\tau(\varepsilon(X))_g = \tau(X_e)_g = \widetilde{X_e}|_g = (L_g)_* X_e = X_g,$$

which shows that $\tau \circ \varepsilon = \operatorname{Id}_{\operatorname{Lie}(G)}$. □

Given any vector $V \in T_e G$, we will consistently use the notation $\widetilde{V}$ to denote the smooth left-invariant vector field defined by (4.9).

It is worth observing that the preceding proof also shows that the assumption of smoothness in the definition of $\operatorname{Lie}(G)$ is unnecessary.

Corollary 4.21. *Every left-invariant rough vector field on a Lie group is smooth.*

Proof. Let V be a left-invariant rough vector field on a Lie group G. The fact that V is left-invariant implies that $V = \widetilde{V_e}$, which is smooth. □

Example 4.22. Let us determine the Lie algebras of some familiar Lie groups.

(a) *Euclidean space* $\mathbb{R}^n$: Left translation by an element $b \in \mathbb{R}^n$ is given by the affine map $L_b(x) = b + x$, whose push-forward $(L_b)_*$ is represented by the identity matrix in standard coordinates. Thus a vector field

$V^i \partial/\partial x^i$ is left-invariant if and only if its coefficients V^i are constants. Because the Lie bracket of two constant-coefficient vector fields is zero by (4.5), the Lie algebra of $\mathbb{R}^n$ is abelian, and is isomorphic to $\mathbb{R}^n$ itself with the trivial bracket. In brief, $\mathrm{Lie}(\mathbb{R}^n) \cong \mathbb{R}^n$.

(b) *The circle group* $\mathbb{S}^1$: In terms of appropriate angle coordinates, each left translation has a coordinate representation of the form $\theta \mapsto \theta + c$. Thus the vector field $d/d\theta$ defined in Example 4.4 is left-invariant, and is therefore a basis for the Lie algebra of $\mathbb{S}^1$. This Lie algebra is 1-dimensional and abelian, and therefore $\mathrm{Lie}(\mathbb{S}^1) \cong \mathbb{R}$.

(c) *The n-torus* $\mathbb{T}^n = \mathbb{S}^1 \times \cdots \times \mathbb{S}^1$: A similar analysis shows that $(\partial/\partial\theta^1, \ldots, \partial/\partial\theta^n)$ is a basis for $\mathrm{Lie}(\mathbb{T}^n)$, where $\partial/\partial\theta^i$ is the angle coordinate vector field on the ith $\mathbb{S}^1$ factor. Since the Lie brackets of these coordinate vector fields are all zero, $\mathrm{Lie}(\mathbb{T}^n) \cong \mathbb{R}^n$.

The Lie groups $\mathbb{R}^n$, $\mathbb{S}^1$, and $\mathbb{T}^n$ are abelian, and as the discussion above shows, their Lie algebras turn out also to be abelian. This is no accident—the Lie algebra of any abelian Lie group is abelian (see Problem 4-17). Just as we can view the tangent space as a "linear model" of a smooth manifold near a point, the Lie algebra of a Lie group provides a "linear model" of the group, which reflects many of the properties of the group. Because Lie groups have more structure than ordinary smooth manifolds, it should come as no surprise that their linear models have more structure than ordinary vector spaces. Since a finite-dimensional Lie algebra is a purely linear-algebraic object, it is in many ways simpler to understand than the group itself. Much of the progress in the theory of Lie groups has come from a careful analysis of Lie algebras.

We conclude this chapter by analyzing the Lie algebra of the most important nonabelian Lie group of all, the general linear group. Theorem 4.20 gives a vector space isomorphism between $\mathrm{Lie}(\mathrm{GL}(n,\mathbb{R}))$ and the tangent space to $\mathrm{GL}(n,\mathbb{R})$ at the identity matrix. Because $\mathrm{GL}(n,\mathbb{R})$ is an open subset of the vector space $\mathfrak{gl}(n,\mathbb{R})$, its tangent space is naturally isomorphic to $\mathfrak{gl}(n,\mathbb{R})$ itself. The composition of these two isomorphisms gives a vector space isomorphism $\mathrm{Lie}(\mathrm{GL}(n,\mathbb{R})) \cong \mathfrak{gl}(n,\mathbb{R})$.

Both $\mathrm{Lie}(\mathrm{GL}(n,\mathbb{R}))$ and $\mathfrak{gl}(n,\mathbb{R})$ have independently defined Lie algebra structures—the first coming from Lie brackets of vector fields and the second from commutator brackets of matrices. The next proposition shows that the natural vector space isomorphism between these spaces is in fact a Lie algebra isomorphism.

Proposition 4.23 (Lie Algebra of the General Linear Group). *The composition of the natural maps*

$$\mathfrak{gl}(n,\mathbb{R}) \to T_{I_n}\,\mathrm{GL}(n,\mathbb{R}) \to \mathrm{Lie}(\mathrm{GL}(n,\mathbb{R})) \tag{4.10}$$

gives a Lie algebra isomorphism between $\mathrm{Lie}(\mathrm{GL}(n,\mathbb{R}))$ *and the matrix algebra* $\mathfrak{gl}(n,\mathbb{R})$.

Proof. Using the matrix entries X^i_j as global coordinates on $\mathrm{GL}(n,\mathbb{R}) \subset \mathfrak{gl}(n,\mathbb{R})$, the natural isomorphism $\mathfrak{gl}(n,\mathbb{R}) \longleftrightarrow T_{I_n}\,\mathrm{GL}(n,\mathbb{R})$ takes the form

$$(A^i_j) \longleftrightarrow A^i_j \left.\frac{\partial}{\partial X^i_j}\right|_{I_n}.$$

(Because of the dual role of the indices i, j as coordinate indices and matrix row and column indices, in this case it is impossible to maintain our convention that all coordinates have upper indices. However, we will continue to observe the summation convention and the other index conventions associated with it. In particular, in the expression above, an upper index "in the denominator" is to be regarded as a lower index, and vice-versa.)

Let $\mathfrak{g}$ denote the Lie algebra of $\mathrm{GL}(n,\mathbb{R})$. Any matrix $A = (A^i_j) \in \mathfrak{gl}(n,\mathbb{R})$ determines a left-invariant vector field $\widetilde{A} \in \mathfrak{g}$ defined by (4.9), which in this case becomes

$$\widetilde{A}_X = (L_X)_* A = (L_X)_* \left(A^i_j \left.\frac{\partial}{\partial X^i_j}\right|_{I_n} \right).$$

Since L_X is the restriction to $\mathrm{GL}(n,\mathbb{R})$ of the linear map $A \mapsto XA$ on $\mathfrak{gl}(n,\mathbb{R})$, its push-forward is represented in coordinates by exactly the same linear map. In other words, the left-invariant vector field $\widetilde{A}$ determined by A is the one whose value at $X \in \mathrm{GL}(n,\mathbb{R})$ is

$$\widetilde{A}_X = X^i_j A^j_k \left.\frac{\partial}{\partial X^i_k}\right|_X. \tag{4.11}$$

Given two matrices $A, B \in \mathfrak{gl}(n,\mathbb{R})$, the Lie bracket of the corresponding left-invariant vector fields is given by

$$\begin{aligned}
[\widetilde{A}, \widetilde{B}] &= \left[X^i_j A^j_k \frac{\partial}{\partial X^i_k}, X^p_q B^q_r \frac{\partial}{\partial X^p_r} \right] \\
&= X^i_j A^j_k \frac{\partial}{\partial X^i_k}(X^p_q B^q_r) \frac{\partial}{\partial X^p_r} - X^p_q B^q_r \frac{\partial}{\partial X^p_r}(X^i_j A^j_k) \frac{\partial}{\partial X^i_k} \\
&= X^i_j A^j_k B^k_r \frac{\partial}{\partial X^i_r} - X^p_q B^q_r A^r_k \frac{\partial}{\partial X^p_k} \\
&= (X^i_j A^j_k B^k_r - X^i_j B^j_k A^k_r) \frac{\partial}{\partial X^i_r},
\end{aligned}$$

where we have used the fact that $\partial X^p_q / \partial X^i_k$ is equal to 1 if $p = i$ and $q = k$, and 0 otherwise, and A^i_j and B^i_j are constants. Evaluating this last expression when X is equal to the identity matrix, we get

$$[\widetilde{A}, \widetilde{B}]_{I_n} = (A^i_k B^k_r - B^i_k A^k_r) \left.\frac{\partial}{\partial X^i_r}\right|_{I_n}.$$

This is the vector corresponding to the matrix commutator bracket $[A, B]$. Since the left-invariant vector field $[\widetilde{A}, \widetilde{B}]$ is determined by its value at the identity, this implies that

$$[\widetilde{A}, \widetilde{B}] = \widetilde{[A, B]},$$

which is precisely the statement that the composite map (4.10) is a Lie algebra isomorphism. □

There is an analogue of this result for abstract vector spaces. If V is any finite-dimensional real vector space, recall that we have defined $\mathrm{GL}(V)$ as the Lie group of invertible linear transformations of V, and $\mathfrak{gl}(V)$ as the Lie algebra of all linear transformations. Just as in the case of $\mathrm{GL}(n, \mathbb{R})$, we can regard $\mathrm{GL}(V)$ as an open submanifold of $\mathfrak{gl}(V)$, and thus there are canonical vector space isomorphisms

$$\mathrm{Lie}(\mathrm{GL}(V)) \to T_{\mathrm{Id}} \, \mathrm{GL}(V) \to \mathfrak{gl}(V). \tag{4.12}$$

Corollary 4.24. *If V is any finite-dimensional real vector space, the composition of the canonical isomorphisms in* (4.12) *yields a Lie algebra isomorphism between* $\mathrm{Lie}(\mathrm{GL}(V))$ *and* $\mathfrak{gl}(V)$.

◇ **Exercise 4.9.** Prove the preceding corollary by choosing a basis for V and applying Proposition 4.23.

Induced Lie Algebra Homomorphisms

The importance of the Lie algebra of a Lie group stems, in large part, from the fact that each Lie group homomorphism induces a Lie algebra homomorphism, as the next theorem shows.

Theorem 4.25. *Let G and H be Lie groups, and let $\mathfrak{g}$ and $\mathfrak{h}$ be their Lie algebras. Suppose $F \colon G \to H$ is a Lie group homomorphism. For every $X \in \mathfrak{g}$, there is a unique vector field in $\mathfrak{h}$ that is F-related to X. Denoting this vector field by $F_* X$, the map $F_* \colon \mathfrak{g} \to \mathfrak{h}$ so defined is a Lie algebra homomorphism.*

Proof. If there is any vector field $Y \in \mathfrak{h}$ that is F-related to X, it must satisfy $Y_e = F_* X_e$, and thus it must be uniquely determined by

$$Y = \widetilde{F_* X_e}.$$

To show that this Y is F-related to X, we note that the fact that F is a homomorphism implies

$$\begin{aligned} F(gg') &= F(g)F(g') \\ &\implies F(L_g g') = L_{F(g)} F(g') \\ &\implies F \circ L_g = L_{F(g)} \circ F \\ &\implies F_* \circ (L_g)_* = (L_{F(g)})_* \circ F_*. \end{aligned}$$

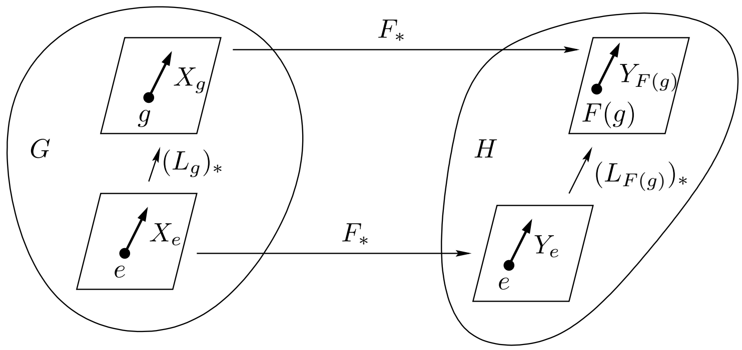


Figure 4.6. The induced Lie algebra homomorphism.

Thus

$$\begin{aligned} F_* X_g &= F_*(L_g)_* X_e \\ &= (L_{F(g)})_* F_* X_e \\ &= (L_{F(g)})_* Y_e \\ &= Y_{F(g)}. \end{aligned}$$

(See Figure 4.6.) This says precisely that X and Y are F-related.

Now, for each $X \in \mathfrak{g}$, let F_*X denote the unique vector field in $\mathfrak{h}$ that is F-related to X. It then follows immediately from the naturality of Lie brackets that $F_*[X, Y] = [F_*X, F_*Y]$, so F_* is a Lie algebra homomorphism. □

The map $F_*\colon \mathfrak{g} \to \mathfrak{h}$ whose existence is asserted in this theorem will be called the *induced Lie algebra homomorphism.* Note that the theorem implies that for any left-invariant vector field $X \in \mathfrak{g}$, F_*X is a well-defined smooth vector field on H, even though F may not be a diffeomorphism.

Proposition 4.26 (Properties of the Induced Homomorphism).

(*a*) *The homomorphism* $(\mathrm{Id}_G)*\colon \operatorname{Lie}(G) \to \operatorname{Lie}(G)$ *induced by the identity map of G is the identity of* $\operatorname{Lie}(G)$.

(*b*) *If $F_1\colon G \to H$ and $F_2\colon H \to K$ are Lie group homomorphisms, then* $(F_2 \circ F_1)_* = (F_2)_* \circ (F_1)_*\colon \operatorname{Lie}(G) \to \operatorname{Lie}(K)$.

(*c*) *Isomorphic Lie groups have isomorphic Lie algebras.*

Proof. Both of the relations $(\mathrm{Id}_G)* = \mathrm{Id}$ and $(F_2 \circ F_1)_* = (F_2)_* \circ (F_1)_*$ hold for push-forwards. Since the value of the induced homomorphism on a left-invariant vector field X is determined by the push-forward of X_e, this proves (a) and (b). If $F\colon G \to H$ is an isomorphism, (a) and (b) together imply that $F_* \circ (F^{-1})_* = (F \circ F^{-1})_* = \mathrm{Id} = (F^{-1})_* \circ F_*$, so $F_*\colon \operatorname{Lie}(G) \to \operatorname{Lie}(H)$ is an isomorphism. □

Problems

4-1. Extension Lemma for Vector Fields: Let M be a smooth manifold, and suppose Y is a smooth vector field defined on a closed subset $A \subset M$. (This means that $Y\colon A \to TM$ is a map satisfying $\pi \circ Y = \mathrm{Id}_A$, and for each $p \in A$, there is a neighborhood V_p of p in M and a smooth vector field $\widetilde{Y}$ on V_p that agrees with Y on $V_p \cap A$.) For any open set U containing A, show that there exists a smooth vector field $\widetilde{Y} \in \mathfrak{T}(M)$ such that $\widetilde{Y}|_A = Y$ and $\operatorname{supp} \widetilde{Y} \subset U$.

4-2. Let M be a nonempty manifold of dimension $n \geq 1$. Show that $\mathfrak{T}(M)$ is infinite-dimensional.

4-3. Show by finding a counterexample that Proposition 4.10 is false if we assume merely that F is smooth and bijective but not a diffeomorphism.

4-4. For each of the following vector fields on the plane, compute its coordinate representation in polar coordinates on the right half-plane $\{(x, y) : x > 0\}$.

(a) $V = x\dfrac{\partial}{\partial x} + y\dfrac{\partial}{\partial y}$.

(b) $W = x\dfrac{\partial}{\partial y} - y\dfrac{\partial}{\partial x}$.

(c) $X = (x^2 + y^2)\dfrac{\partial}{\partial x}$.

4-5. Show that there is a smooth vector field on $\mathbb{S}^2$ that vanishes at exactly one point. [Hint: Try using stereographic projection.]

4-6. Let M, N be smooth manifolds, and let $f\colon M \to N$ be a smooth map. Define $F\colon M \to M \times N$ by $F(x) = (x, f(x))$. For any $V \in \mathfrak{T}(M)$, show that there is a smooth vector field on $M \times N$ that is F-related to V.

4-7. Let $M_1, \dots, M_k$ be smooth manifolds, and for each $i = 1, \dots, k$, let $\pi_i\colon M_1 \times \cdots \times M_k \to M_i$ be the projection onto the ith factor. For any $X \in \mathfrak{T}(M_i)$, show that there is a smooth vector field on $M_1 \times \cdots \times M_k$ that is π_i-related to X.

4-8. Let $F\colon M \to N$ be a local diffeomorphism. For any $Y \in \mathfrak{T}(N)$, show there is a unique smooth vector field on M that is F-related to Y.

4-9. Define a (rough) vector field V on $\mathbb{S}^3$ by letting $V_z = \gamma_z'(0)$, where γ_z is the curve of Problem 3-5. Show that V is smooth and nowhere vanishing.

4-10. For each of the following pairs of vector fields V, W defined on $\mathbb{R}^3$, compute the Lie bracket $[V, W]$.

(a) $V = y\dfrac{\partial}{\partial z} - 2xy^2\dfrac{\partial}{\partial y}$; $\quad W = \dfrac{\partial}{\partial y}$.

(b) $V = x\dfrac{\partial}{\partial y} - y\dfrac{\partial}{\partial x}$; $\quad W = y\dfrac{\partial}{\partial z} - z\dfrac{\partial}{\partial y}$.

(c) $V = x\dfrac{\partial}{\partial y} - y\dfrac{\partial}{\partial x}$; $\quad W = x\dfrac{\partial}{\partial y} + y\dfrac{\partial}{\partial x}$.

4-11. Show that $\mathbb{R}^3$ with the cross product is a Lie algebra.

4-12. Let A be an associative algebra over $\mathbb{R}$. A *derivation* of A is a linear map $D\colon A \to A$ satisfying $D(xy) = xDy + yDx$ for all $x, y \in A$. If D_1 and D_2 are derivations of A, show that $[D_1, D_2] = D_1 \circ D_2 - D_2 \circ D_1$ is also a derivation. Show that the set of derivations of A is a Lie algebra.

4-13. Let $A \subset \mathcal{T}(\mathbb{R}^3)$ be the subspace with basis $\{X, Y, Z\}$, where

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

Show that A is a Lie subalgebra of $\mathcal{T}(\mathbb{R}^3)$, which is isomorphic to $\mathbb{R}^3$ with the cross product.

4-14. Prove that up to isomorphism, there are exactly one 1-dimensional Lie algebra and two 2-dimensional Lie algebras. Show that all three algebras are isomorphic to Lie subalgebras of $\mathfrak{gl}(2, \mathbb{R})$.

4-15. Let $\mathfrak{g}$ be a Lie algebra. A linear subspace $\mathfrak{h} \subset \mathfrak{g}$ is called an *ideal* in $\mathfrak{g}$ if for any $X \in \mathfrak{h}$ and $Y \in \mathfrak{g}$, $[X, Y] \in \mathfrak{h}$.

(a) If $\mathfrak{h}$ is an ideal in $\mathfrak{g}$, show that the quotient space $\mathfrak{g}/\mathfrak{h}$ has a unique Lie algebra structure such that the projection $\pi\colon \mathfrak{g} \to \mathfrak{g}/\mathfrak{h}$ is a Lie algebra homomorphism.

(b) Show that a subspace $\mathfrak{h} \subset \mathfrak{g}$ is an ideal if and only if it is the kernel of a Lie algebra homomorphism.

4-16. (a) If $\mathfrak{g}$ and $\mathfrak{h}$ are Lie algebras, show that $\mathfrak{g} \times \mathfrak{h}$ is a Lie algebra, called a *product Lie algebra*, with the bracket defined by

$$[(X, Y), (X', Y')] = ([X, X'], [Y, Y']).$$

(b) If G and H are Lie groups, prove that $\mathrm{Lie}(G \times H)$ is isomorphic to $\mathrm{Lie}(G) \times \mathrm{Lie}(H)$.

4-17. If G is an abelian Lie group, show that $\mathrm{Lie}(G)$ is abelian. [Hint: Show that the inversion map $i\colon G \to G$ is a group homomorphism, and use Problem 3-6.]

4-18. Let G and H be Lie groups, and suppose that $F\colon G \to H$ is a Lie group homomorphism that is also a smooth covering map. Show that the induced homomorphism $F_*\colon \mathrm{Lie}(G) \to \mathrm{Lie}(H)$ is an isomorphism of Lie algebras.

5
Vector Bundles

In the preceding chapter, we saw that the tangent bundle of a smooth manifold has a natural structure as a smooth manifold in its own right. The standard coordinates we constructed on TM make it look, locally, like the Cartesian product of an open subset of M with $\mathbb{R}^n$. As we will see later in the book, this kind of structure arises quite frequently—a collection of vector spaces, one for each point in M, glued together in a way that looks *locally* like the Cartesian product of M with $\mathbb{R}^n$, but globally may be "twisted." Such a structure is called a vector bundle.

The chapter begins with the definition of vector bundles and descriptions of a few examples. The most notable example, of course, is the tangent bundle to a smooth manifold. We then go on to discuss local and global sections of vector bundles (which correspond to vector fields in the case of the tangent bundle), and a natural notion of maps between bundles, called bundle maps.

At the end of the chapter, we give a brief introduction to the terminology of category theory, which puts the tangent bundle in a larger context.

Vector Bundles

Let M be a topological space. A (*real*) *vector bundle of rank k over M* is a topological space E together with a surjective continuous map $\pi\colon E \to M$ satisfying:

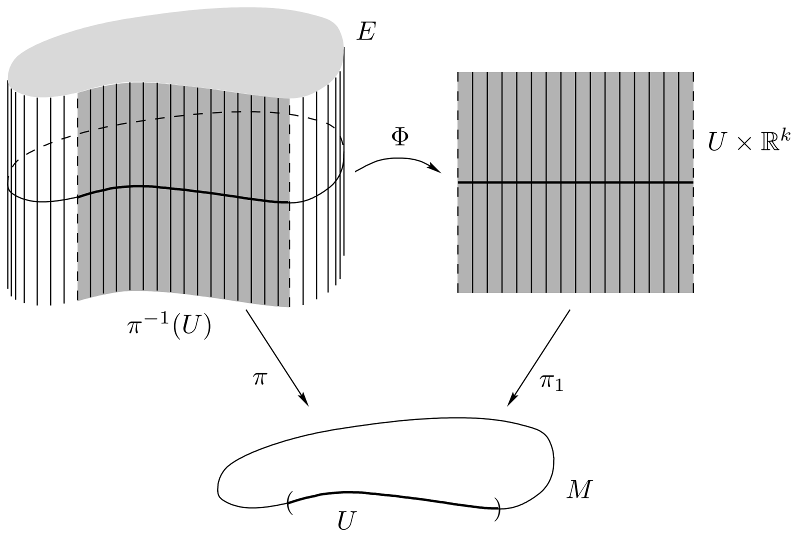


Figure 5.1. A local trivialization of a vector bundle.

(i) For each $p \in M$, the set $E_p = \pi^{-1}(p) \subset E$ (called the *fiber* of E over p) is endowed with the structure of a k-dimensional real vector space.

(ii) For each $p \in M$, there exists a neighborhood U of p in M and a homeomorphism $\Phi\colon \pi^{-1}(U) \to U \times \mathbb{R}^k$ (called a *local trivialization* of E over U), such that the following diagram commutes:

$$\begin{tikzcd} \pi^{-1}(U) \arrow[rr, "\Phi"] \arrow[dr, "\pi"'] & & U \times \mathbb{R}^k \arrow[dl, "\pi_1"] \\ & U & \end{tikzcd}$$

(where π_1 is the projection on the first factor); and such that for each $q \in U$, the restriction of Φ to E_q is a linear isomorphism from E_q to $\{q\} \times \mathbb{R}^k \cong \mathbb{R}^k$ (Figure 5.1).

If M and E are smooth manifolds, π is a smooth map, and the local trivializations can be chosen to be diffeomorphisms, then E is called a *smooth vector bundle*. In this case, we will call any local trivialization that is a diffeomorphism onto its image a *smooth local trivialization.*

A rank 1 vector bundle is often called a (*real*) *line bundle*. *Complex vector bundles* are defined similarly, with "real vector space" replaced by "complex vector space" and $\mathbb{R}^k$ replaced by $\mathbb{C}^k$ in the definition. We will not have occasion to treat complex vector bundles in this book, so all of our vector bundles will be understood without further comment to be real.

The space E is called the *total space* of the bundle, M is called its *base*, and π is its *projection*. Depending on what we wish to emphasize, we sometimes omit some or all of the ingredients from the notation, and write "E is a vector bundle over M," or "$E \to M$ is a vector bundle," or "$\pi\colon E \to M$ is a vector bundle." If $U \subset M$ is any open set, it is easy to verify that the subset $E|_U = \pi^{-1}(U)$ is again a vector bundle with the restriction of π as its projection map, called the *restriction of E to U*.

If there exists a local trivialization over all of M (called a *global trivialization* of E), then E is said to be a *trivial bundle*. In this case, E itself is homeomorphic to the product space $M \times \mathbb{R}^k$. If $E \to M$ is a smooth bundle that admits a smooth global trivialization, then we say E is *smoothly trivial*. In this case E is *diffeomorphic* to $M \times \mathbb{R}^k$, not just homeomorphic. For brevity, when we say that a smooth bundle is trivial, we will always understand this to mean smoothly trivial, not just trivial in the topological sense.

Example 5.1 (Product Bundles). One particularly simple example of a rank k vector bundle over any space M is the product manifold $E = M \times \mathbb{R}^k$ with $\pi = \pi_1\colon M \times \mathbb{R}^k \to M$ as its projection. This bundle is clearly trivial (with the identity map as a global trivialization). If M is a smooth manifold, then $M \times \mathbb{R}^k$ is smoothly trivial.

Although there are many vector bundles that are not trivial, the only one that is easy to visualize is the following.

Example 5.2 (The Möbius Bundle). Let $I = [0,1] \subset \mathbb{R}$ be the unit interval, and let $p\colon I \to \mathbb{S}^1$ be the quotient map $p(x) = e^{2\pi i x}$, which identifies the two endpoints of I. Consider the "infinite strip" $I \times \mathbb{R}$, and let $\pi_1\colon I \times \mathbb{R} \to I$ be the projection on the first factor. Let $\sim$ be the equivalence relation on $I \times \mathbb{R}$ that identifies each point $(0,y)$ in the fiber over 0 with the point $(1,-y)$ in the fiber over 1; in other words, the right-hand edge is given a half-twist to turn it upside-down, and then is glued to the left-hand edge. Let $E = (I \times \mathbb{R})/{\sim}$ denote the resulting quotient space, and let $q\colon I \times \mathbb{R} \to E$ be the quotient map (Figure 5.2).

Because $p \circ \pi_1$ is constant on each equivalence class, it descends to a continuous map $\pi\colon E \to \mathbb{S}^1$. A straightforward (if tedious) verification shows that this makes E into a smooth real line bundle over $\mathbb{S}^1$, called the *Möbius bundle*. (One local trivialization of E is obtained in an obvious way from the restriction of the identity map to $(0,1) \times \mathbb{R}$, which descends to the quotient to yield a homeomorphism from $\pi^{-1}(\mathbb{S}^1 \smallsetminus \{1\})$ to $(0,1) \times \mathbb{R}$. It takes a bit more work to construct a local trivialization whose domain includes the fiber where the gluing took place. Once this is done, the two local trivializations can be interpreted as coordinate charts defining the smooth structure on E. Problem 5-2 asks you to work out the details. Later in the book, Problem 9-16 will suggest a more powerful approach.) For any $r > 0$, the image under q of the rectangle $I \times [-r,r]$ is a smooth

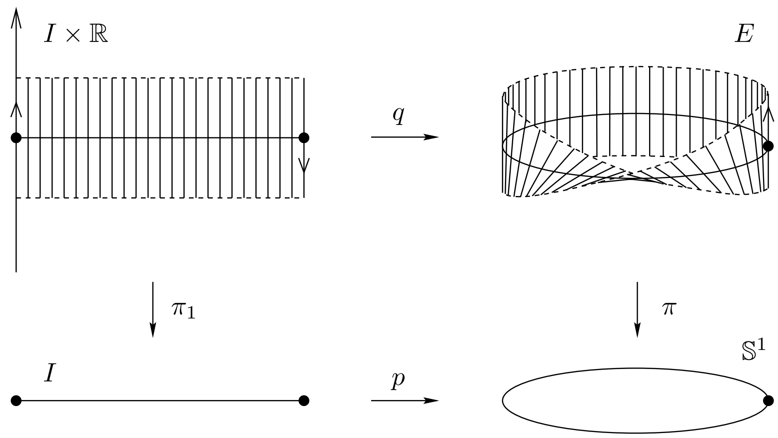


Figure 5.2. (Part of) the Möbius bundle.

compact manifold with boundary called the *Möbius band*; you can make a paper model of this space by gluing the ends of a strip of paper together with a half-twist.

The most important examples of vector bundles are tangent bundles of smooth manifolds.

Proposition 5.3 (The Tangent Bundle as a Vector Bundle). *Let M be a smooth n-manifold and let TM be its tangent bundle. With its standard projection map, its natural vector space structure on each fiber, and the smooth manifold structure constructed in Lemma 4.1, TM is a smooth vector bundle of rank n over M.*

Proof. Given any smooth chart (U, φ) for M with coordinate functions (x^i), define a map $\Phi\colon \pi^{-1}(U) \to U \times \mathbb{R}^n$ by

$$\Phi\left(v^i \frac{\partial}{\partial x^i}\bigg|_p \right) = (p, (v^1, \dots, v^n)). \tag{5.1}$$

This is obviously linear on fibers and satisfies $\pi_1 \circ \Phi = \pi$. The composite map

$$\pi^{-1}(U) \xrightarrow{\Phi} U \times \mathbb{R}^n \xrightarrow{\varphi \times \mathrm{Id}_{\mathbb{R}^n}} \varphi(U) \times \mathbb{R}^n$$

is equal to the coordinate map $\widetilde{\varphi}$ constructed in Lemma 4.1. Since both $\widetilde{\varphi}$ and $\varphi \times \mathrm{Id}_{\mathbb{R}^n}$ are diffeomorphisms, so is Φ. Thus Φ satisfies all the conditions for a smooth local trivialization. □

Any bundle that is not trivial, of course, will require more than one local trivialization. The next lemma shows that the composition of two smooth local trivializations has a simple form where they overlap.

Lemma 5.4. *Let $\pi\colon E \to M$ be a smooth vector bundle, and suppose $\Phi\colon \pi^{-1}(U) \to U \times \mathbb{R}^k$ and $\Psi\colon \pi^{-1}(V) \to V \times \mathbb{R}^k$ are two smooth local trivializations of E such that $U \cap V \neq \varnothing$. There exists a smooth map $\tau\colon U \cap V \to \mathrm{GL}(k,\mathbb{R})$ such that the composition $\Phi \circ \Psi^{-1}\colon (U \cap V) \times \mathbb{R}^k \to (U \cap V) \times \mathbb{R}^k$ has the form*

$$\Phi \circ \Psi^{-1}(p, v) = (p, \tau(p)v),$$

where $\tau(p)v$ denotes the usual action of the $k \times k$ matrix $\tau(p)$ on the vector $v \in \mathbb{R}^k$.

Proof. The following diagram commutes:

$$\begin{tikzcd}
(U \cap V) \times \mathbb{R}^k \arrow[dr, "\pi_1"'] & \pi^{-1}(U \cap V) \arrow[l, "\Psi"'] \arrow[r, "\Phi"] \arrow[d, "\pi"] & (U \cap V) \times \mathbb{R}^k \arrow[dl, "\pi_1"] \\
& U \cap V, &
\end{tikzcd} \tag{5.2}$$

where the maps on top are to be interpreted as the restrictions of Ψ and Φ to $\pi^{-1}(U \cap V)$. It follows that $\pi_1 \circ (\Phi \circ \Psi^{-1}) = \pi_1$, which means that $\Phi \circ \Psi^{-1}(p, v) = (p, \sigma(p, v))$ for some smooth map $\sigma\colon (U \cap V) \times \mathbb{R}^k \to \mathbb{R}^k$. Moreover, for each fixed $p \in U \cap V$, the map $v \mapsto \sigma(p, v)$ is a linear isomorphism of $\mathbb{R}^k$, so there is a nonsingular $k \times k$ matrix $\tau(p)$ such that $\sigma(p, v) = \tau(p)v$. It remains only to show that the map $\tau\colon U \cap V \to \mathrm{GL}(k,\mathbb{R})$ is smooth.

To see this, write the matrix entries of $\tau(p)$ as $\tau^i_j(p)$, so that $\tau(p)v = \tau^i_j(p)v^j e_i$. Note that $\tau^i_j(p) = \pi^i(\sigma(p, e_j))$, where e_j is the jth standard basis vector and $\pi^i\colon \mathbb{R}^k \to \mathbb{R}$ is projection onto the ith coordinate. This is smooth by composition. Since the matrix entries are (global) smooth coordinates on $\mathrm{GL}(k,\mathbb{R})$, this shows that τ is smooth. $\square$

The smooth map $\tau\colon U \cap V \to \mathrm{GL}(k,\mathbb{R})$ described in this lemma is called the *transition function* between the local trivializations Φ and Ψ. (This is one of the few situations in which it is traditional to use the word "function" even thought the range is not $\mathbb{R}$ or $\mathbb{R}^k$.) For example, if M is a smooth manifold and Φ and Ψ are the local trivializations of TM associated with two different smooth charts, then the transition function between them is just the Jacobian matrix of the coordinate transition map.

Like the tangent bundle, vector bundles are often most easily described by giving a collection of vector spaces, one for each point of the base manifold. In order to make such a set into a smooth vector bundle, we would first have to construct a manifold topology and a smooth structure on the disjoint union of all the vector spaces, and then construct the local trivializations and show that they have the requisite properties. The next lemma provides a shortcut, by showing that it is sufficient to construct the local trivializations, as long as they overlap with smooth transition functions. The proof is basically an adaptation of the proofs of Lemma 4.1 and Proposition 5.3. (See also Problem 5-4 for a stronger form of this result.)

Lemma 5.5 (Vector Bundle Construction Lemma). *Let M be a smooth manifold, and suppose that we are given*

- *for each $p \in M$, a real vector space E_p of some fixed dimension k.*

Let $E = \coprod_{p \in M} E_p$, and let $\pi\colon E \to M$ be the map that takes each element of E_p to the point p. Suppose furthermore that we are given

- *an open cover $\{U_\alpha\}_{\alpha \in A}$ of M;*
- *for each $\alpha \in A$, a bijective map $\Phi_\alpha\colon \pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{R}^k$ whose restriction to each E_p is a linear isomorphism from E_p to $\{p\} \times \mathbb{R}^k \cong \mathbb{R}^k$;*
- *for each $\alpha, \beta \in A$ such that $U_\alpha \cap U_\beta \neq \varnothing$, a smooth map $\tau_{\alpha\beta}\colon U_\alpha \cap U_\beta \to \mathrm{GL}(k, \mathbb{R})$ such that the composite map $\Phi_\alpha \circ \Phi_\beta^{-1}$ from $(U_\alpha \cap U_\beta) \times \mathbb{R}^k$ to itself has the form*

$$\Phi_\alpha \circ \Phi_\beta^{-1}(p, v) = (p, \tau_{\alpha\beta}(p)v). \tag{5.3}$$

Then E has a unique smooth manifold structure making it into a smooth vector bundle of rank k over M, with π as projection and the maps Φ_α as smooth local trivializations.

Proof. For each point $p \in M$, choose some U_α containing p; choose a smooth chart (V_p, φ_p) for M such that $p \in V_p \subset U_\alpha$; and let $\widetilde{V}_p = \varphi_p(V_p) \subset \mathbb{R}^n$ (where n is the dimension of M). Define a map $\widetilde{\varphi}_p\colon \pi^{-1}(V_p) \to \widetilde{V}_p \times \mathbb{R}^k$ by $\widetilde{\varphi}_p = (\varphi_p \times \mathrm{Id}_{\mathbb{R}^k}) \circ \Phi_\alpha$:

$$\pi^{-1}(V_p) \xrightarrow{\Phi_\alpha} V_p \times \mathbb{R}^k \xrightarrow{\varphi_p \times \mathrm{Id}_{\mathbb{R}^k}} \widetilde{V}_p \times \mathbb{R}^k.$$

We will show that the collection of all such charts $\{(\pi^{-1}(V_p), \widetilde{\varphi}_p) : p \in M\}$ satisfies the conditions of the smooth manifold construction lemma (Lemma 1.23), and therefore defines a smooth manifold structure on E.

As a composition of bijective maps, $\widetilde{\varphi}_p$ is bijective, and its image is an open subset of $\mathbb{R}^n \times \mathbb{R}^k = \mathbb{R}^{n+k}$. For any two points p and q, it is easy to check that

$$\widetilde{\varphi}_p(\pi^{-1}(V_p) \cap \pi^{-1}(V_q)) = \varphi_p(V_p \cap V_q) \times \mathbb{R}^k,$$

which is open because φ_p is a homeomorphism onto an open subset of $\mathbb{R}^n$. Wherever two such charts overlap, we have

$$\widetilde{\varphi}_p \circ \widetilde{\varphi}_q^{-1} = (\varphi_p \times \mathrm{Id}_{\mathbb{R}^k}) \circ \Phi_\alpha \circ \Phi_\beta^{-1} \circ (\varphi_q \times \mathrm{Id}_{\mathbb{R}^k})^{-1}.$$

Since $\varphi_p \times \mathrm{Id}_{\mathbb{R}^k}$, $\varphi_q \times \mathrm{Id}_{\mathbb{R}^k}$, and $\Phi_\alpha \circ \Phi_\beta^{-1}$ are all diffeomorphisms, this composition is a diffeomorphism. Thus conditions (i)–(iii) of Lemma 1.23 are satisfied. Because the open cover $\{V_p : p \in M\}$ has a countable subcover, (iv) is satisfied as well.

To check the Hausdorff condition (v), just note that any two points in the same space E_p lie in one of the charts we have constructed; while if $\xi \in E_p$ and $\eta \in E_q$ with $p \neq q$, we can choose V_p and V_q to be disjoint neighborhoods of p and q, so that the sets $\pi^{-1}(V_p)$ and $\pi^{-1}(V_q)$ are disjoint coordinate neighborhoods containing ξ and η, respectively. Thus we have defined a smooth manifold structure on E.

With respect to this structure, each of the maps Φ_α is a diffeomorphism, because in terms of the coordinate charts $(\pi^{-1}(V_p), \widetilde{\varphi}_p)$ for E and $(V_p \times \mathbb{R}^k, \varphi_p \times \mathrm{Id}_{\mathbb{R}^k})$ for $V_p \times \mathbb{R}^k$, the coordinate representation of Φ_α is the identity map. The coordinate representation of π, with respect to the same chart for E and the chart (V_p, φ_p) for M, is $\pi(x, v) = x$, so π is smooth as well. Because each Φ_α maps E_p to $\{p\} \times \mathbb{R}^k$, it is immediate that $\pi_1 \circ \Phi_\alpha = \pi$, and Φ_α is linear on fibers by hypothesis. Thus Φ_α satisfies all the conditions for a smooth local trivialization.

The fact that this is the unique such smooth structure follows easily from the requirement that the maps Φ_α be diffeomorphisms onto their images: If $\widetilde{E}$ represents the same set E with another topology and smooth structure satisfying the conclusions of the lemma, the identity map from E to $\widetilde{E}$ is locally equal to $\Phi_\alpha^{-1} \circ \Phi_\alpha$ and therefore is a local diffeomorphism between E and $\widetilde{E}$. Because it is a bijective local diffeomorphism, it is a diffeomorphism and thus the two smooth structures are identical. □

Local and Global Sections of Vector Bundles

Let $\pi : E \to M$ be a vector bundle over a manifold M. A *section of E* (also sometimes called a *cross section*) is a section of the map π, i.e., a continuous map $\sigma : M \to E$ satisfying $\pi \circ \sigma = \mathrm{Id}_M$. Specifically, this means that $\sigma(p)$ is an element of the fiber E_p for each $p \in U$. A *local section of E* is a section $\sigma : U \to E$ defined only on some open subset $U \subset M$ (see Figure 5.3). To emphasize the distinction, a section defined on all of M will sometimes be called a *global section*. Note that a local section of E over $U \subset M$ is the same as a global section of the restricted bundle $E|_U$. If M is a smooth manifold and E is a smooth vector bundle, a *smooth section* of E is a (local or global) section that is smooth as a map between manifolds.

Just as with vector fields, for some purposes it is useful also to consider maps that would be sections except that they might not be continuous. Thus we define a *rough section* of E over a set $U \subset M$ to be a map $\sigma : U \to E$ (not necessarily continuous) such that $\pi \circ \sigma = \mathrm{Id}_U$. A "section" without further qualification will always mean a continuous section.

The *zero section* of E is the global section $\zeta : M \to E$ defined by

$$\zeta(p) = 0 \in E_p \text{ for each } p \in M.$$

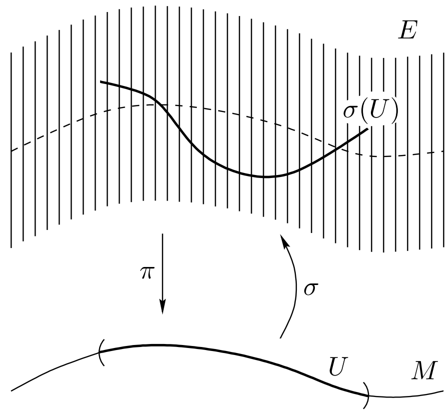


Figure 5.3. A local section of a vector bundle.

As in the case of vector fields, *support* of a section σ is the closure of the set $\{p \in M : \sigma(p) \neq 0\}$.

◇ **Exercise 5.1.** Show that the zero section of any smooth vector bundle is smooth. [Hint: Consider $\Phi \circ \zeta$, where Φ is any smooth local trivialization.]

If $E \to M$ is a smooth vector bundle, the set $\mathcal{E}(M)$ of all smooth global sections of E is a vector space under pointwise addition and scalar multiplication:

$$(c_1\sigma_1 + c_2\sigma_2)(p) = c_1\sigma_1(p) + c_2\sigma_2(p).$$

In addition, just like vector fields, smooth sections can be multiplied by smooth real-valued functions: If $f \in C^\infty(M)$ and $\sigma \in \mathcal{E}(M)$, we obtain a new section $f\sigma$ defined by

$$(f\sigma)(p) = f(p)\sigma(p).$$

◇ **Exercise 5.2.** If $\sigma, \tau \in \mathcal{E}(M)$ and $f, g \in C^\infty(M)$, show that $f\sigma + g\tau \in \mathcal{E}(M)$.

◇ **Exercise 5.3.** Show that $\mathcal{E}(M)$ is a module over the ring $C^\infty(M)$.

Lemma 5.6 (Extension Lemma for Vector Bundles). *Let $\pi\colon E \to M$ be a smooth vector bundle over a smooth manifold M, and suppose $\sigma\colon A \to E$ is a smooth section of E defined on a closed subset $A \subset M$ (in the sense that σ extends to a smooth section in a neighborhood of each point). For any open set U containing A, there exists a smooth section $\widetilde{\sigma} \in \mathcal{E}(M)$ such that $\widetilde{\sigma}|_A = \sigma$ and $\operatorname{supp}\widetilde{\sigma} \subset U$.*

◇ **Exercise 5.4.** Prove the preceding lemma.

For a smooth vector bundle $E \to M$, the notations $\Gamma(E)$, $C^\infty(M;E)$, and $\mathcal{E}(M)$ are all in common use for the space of smooth global sections of E. We will generally prefer the latter notation, using the script letter corresponding to the name of a bundle to denote its space of smooth sections. For example, we have already defined $\mathcal{T}(M)$ to mean the space of smooth sections of the tangent bundle TM. Besides being concise, this notation has the advantage that it adapts easily to spaces of local sections, with $\mathcal{E}(U)$ denoting the space of smooth local sections of E over a fixed open set $U \subset M$.

Example 5.7 (Sections of Vector Bundles).

(a) For a smooth manifold M, sections of TM are vector fields on M.

(b) If $E\colon M \times \mathbb{R}^k$ is a product bundle, there is a natural one-to-one correspondence between (smooth) sections of E and (smooth) functions from M to $\mathbb{R}^k$: A function $F\colon M \to \mathbb{R}^k$ determines a section $\widetilde{F}\colon M \to M \times \mathbb{R}^k$ by $\widetilde{F}(x) = (x, F(x))$, and vice versa. In particular, $C^\infty(M)$ can be naturally identified with the space of smooth sections of the trivial line bundle $M \times \mathbb{R}$.

Local and Global Frames

Let $E \to M$ be a vector bundle. If $U \subset M$ is an open set, local sections $\sigma_1, \dots, \sigma_k$ of E over U are said to be *independent* if their values $\sigma_1(p), \dots, \sigma_k(p)$ are linearly independent elements of E_p for each $p \in U$. Similarly, they are said to *span* E if their values span E_p for each $p \in U$. A *local frame* for E over U is an ordered k-tuple $(\sigma_1, \dots, \sigma_k)$ of independent local sections over U that span E; thus $(\sigma_1(p), \dots, \sigma_k(p))$ is a basis for the fiber E_p for each $p \in U$. It is called a *global frame* if $U = M$. A local or global frame is said to be smooth if each section σ_i is smooth. We will often use the shorthand notation (σ_i) to denote a frame $(\sigma_1, \dots, \sigma_k)$.

Local frames are intimately connected with local trivializations, as the next two examples show.

Example 5.8 (Global Frame for a Product Bundle). If $E = M \times \mathbb{R}^k$ is a product bundle, the standard basis $(e_1, \dots, e_k)$ for $\mathbb{R}^k$ yields a global frame $(\widetilde{e}_i)$ for E, defined by $\widetilde{e}_i(p) = (p, e_i)$. If M is a smooth manifold, this global frame is smooth.

Example 5.9 (Local Frames and Local Trivializations). Suppose $\pi\colon E \to M$ is a smooth vector bundle. If $\Phi\colon \pi^{-1}(U) \to U \times \mathbb{R}^k$ is a smooth local trivialization of E, we can use the same idea as the preceding example to construct a local frame for E. Define maps $\sigma_1, \dots, \sigma_k\colon U \to E$ by $\sigma_i(p) =$

$\Phi^{-1}(p, e_i) = \Phi^{-1} \circ \widetilde{e}_i(p)$:

$$\begin{tikzcd}
\pi^{-1}(U) \arrow[r, "\Phi"] \arrow[dr, "\pi"'] & U \times \mathbb{R}^k \arrow[d, "\pi_1"] \\
& U. \arrow[ul, bend left, "\sigma_i"] \arrow[u, bend right, "\widetilde{e}_i"']
\end{tikzcd}$$

Then σ_i is smooth because Φ is a diffeomorphism, and the fact that $\pi_1 \circ \Phi = \pi$ implies that

$$\begin{aligned}
\pi \circ \sigma_i(p) &= \pi \circ \Phi^{-1}(p, e_i) \\
&= \pi_1(p, e_i) \\
&= p,
\end{aligned}$$

so σ_i is a smooth section. To see that $(\sigma_i(p))$ forms a basis for E_p, just note that Φ restricts to a linear isomorphism from E_p to $\{p\} \times \mathbb{R}^k$, and $\Phi(\sigma_i(p)) = (p, e_i)$, which forms a basis for $\{p\} \times \mathbb{R}^k$. We say that this local frame (σ_i) is *associated with* Φ.

Proposition 5.10. *Every smooth local frame for a smooth vector bundle is associated with a smooth local trivialization as in Example 5.9.*

Proof. Suppose that (σ_i) is a smooth local frame for E over an open set $U \subset M$. We define a map $\Psi \colon U \times \mathbb{R}^k \to \pi^{-1}(U)$ by

$$\Psi(p, (v^1, \dots, v^k)) = v^i \sigma_i(p). \tag{5.4}$$

The fact that $(\sigma_i(p))$ forms a basis for E_p at each $p \in U$ implies that Ψ is bijective, and an easy computation shows that $\sigma_i = \Psi \circ \widetilde{e}_i$. Thus if we can show that Ψ is a diffeomorphism, then Ψ^{-1} will be a smooth local trivialization whose associated local frame is (σ_i).

Since Ψ is bijective, to show that it is a diffeomorphism it suffices to show that it is a local diffeomorphism. Given $q \in U$, we can choose a neighborhood $V \subset M$ of q over which there exists a smooth local trivialization $\Phi \colon \pi^{-1}(V) \to V \times \mathbb{R}^k$, and by shrinking V if necessary we may assume that $V \subset U$. Since Φ is a diffeomorphism, if we can show that $\Phi \circ \Psi|_{V \times \mathbb{R}^k}$ is a diffeomorphism from $V \times \mathbb{R}^k$ to itself, it follows that Ψ is a diffeomorphism from $V \times \mathbb{R}^k$ to $\pi^{-1}(V)$:

$$\begin{tikzcd}
V \times \mathbb{R}^k \arrow[r, "\Psi|_{V \times \mathbb{R}^k}"] \arrow[dr, "\pi_1"'] & \pi^{-1}(V) \arrow[r, "\Phi"] \arrow[d, "\pi"] & V \times \mathbb{R}^k \arrow[dl, "\pi_1"] \\
& V. &
\end{tikzcd}$$

For each smooth section σ_i, the composite map $\Phi \circ \sigma_i|_V : V \to V \times \mathbb{R}^k$ is smooth, and therefore there are smooth functions $\sigma_i^1, \dots, \sigma_i^k : V \to \mathbb{R}$ such that

$$\Phi \circ \sigma_i(p) = (p, (\sigma_i^1(p), \dots, \sigma_i^k(p))).$$

On $V \times \mathbb{R}^k$, therefore,

$$\Phi \circ \Psi(p, (v^1, \dots, v^k)) = (p, (v^i \sigma_i^1(p), \dots, v^i \sigma_i^k(p))),$$

which is clearly smooth.

To show that $(\Phi \circ \Psi)^{-1}$ is smooth, note that the matrix $(\sigma_i^j(p))$ is invertible for each p, because $(\sigma_i(p))$ is a basis for E_p. Let $(\tau_i^j(p))$ denote the inverse matrix. Because matrix inversion is a smooth map (see Example 2.7(a)), the functions τ_i^j are smooth. It follows from the computations in the preceding paragraph that

$$(\Phi \circ \Psi)^{-1}(p, (w^1, \dots, w^k)) = (p, (w^i \tau_i^1(p), \dots, w^i \tau_i^k(p))),$$

which is also smooth. □

Corollary 5.11. *A smooth vector bundle is trivial if and only if it admits a smooth global frame.*

Proof. Taken together, Example 5.9 and Proposition 5.10 show that there exists a smooth local trivialization over an open subset $U \subset M$ if and only if there exists a smooth local frame over U. The corollary is just the special case of this statement when $U = M$. □

Corollary 5.12. *Let $\pi : E \to M$ be a smooth vector bundle of rank k, let (V, φ) be a smooth chart on M, and suppose there exists a smooth local frame (σ_i) for E over V. Then the map $\widetilde{\varphi} : \pi^{-1}(V) \to \varphi(V) \times \mathbb{R}^k$ given by*

$$\widetilde{\varphi}(v^i \sigma_i(p)) = (x^1(p), \dots, x^n(p), v^1, \dots, v^k)$$

is a smooth coordinate map for E.

Proof. Just check that $\widetilde{\varphi}$ is equal to the composition $(\varphi \times \mathrm{Id}_{\mathbb{R}^k}) \circ \Phi$, where Φ is the local trivialization associated with (σ_i). As a composition of diffeomorphisms, it is a diffeomorphism. □

Just as smoothness of vector fields can be characterized in terms of their component functions in any smooth chart, smoothness of sections of vector bundles can be characterized in terms of local frames. Suppose (σ_i) is a smooth local frame for E over some open set $U \subset M$. If $\tau : M \to E$ is a rough section, the value of τ at any point $p \in U$ can be written

$$\tau(p) = \tau^i(p) \sigma_i(p)$$

for some uniquely determined numbers $(\tau^1(p), \dots, \tau^n(p))$. This defines k functions $\tau^i : U \to \mathbb{R}$, called the *component functions* of τ with respect to the given local frame.

Lemma 5.13 (Local Frame Criterion for Smoothness). *Let $\pi\colon E \to M$ be a smooth vector bundle, and let $\tau\colon M \to E$ be a rough section. If (σ_i) is a smooth local frame for E over some open subset $U \subset M$, then τ is smooth on U if and only if its component functions with respect to (σ_i) are smooth.*

Proof. Let $\Phi\colon \pi^{-1}(U) \to U \times \mathbb{R}^k$ be the local trivialization associated with the local frame (σ_i). Because Φ is a diffeomorphism, τ is smooth on U if and only if $\Phi \circ \tau$ has the same property. It is straightforward to check that $\Phi \circ \tau(p) = (p, (\tau^1(p), \dots, \tau^k(p)))$, where (τ^i) are the component functions of τ with respect to (σ_i), so $\Phi \circ \tau$ is smooth if and only if the component functions τ^i are smooth. □

◇ **Exercise 5.5.** If $E \to M$ is any vector bundle, show that a rough section of E is continuous if and only if its component functions in any local frame are continuous.

It is worth remarking that Lemma 5.13 applies equally well to local sections, since a local section of E over an open set $V \subset M$ is a global section of the restricted bundle $E|_V$.

The correspondence between local frames and local trivializations leads to the following uniqueness result characterizing the smooth structure on the tangent bundle of a smooth manifold.

Lemma 5.14. *Let M be a smooth n-manifold. The smooth manifold structure on TM constructed in Lemma 4.1 is the unique one with respect to which $\pi\colon TM \to M$ is a smooth vector bundle of rank n over M and all coordinate vector fields are smooth local sections.*

Proof. Suppose TM is endowed with some smooth structure making it into a smooth rank n vector bundle for which the coordinate vector fields are smooth sections. Over any smooth coordinate domain $U \subset M$, the coordinate frame $(\partial/\partial x^i)$ is a smooth local frame, so by Proposition 5.10 there is a smooth local trivialization $\Phi\colon \pi^{-1}(U) \to U \times \mathbb{R}^n$ associated with this local frame. Referring back to the construction of Example 5.9, it is easy to see that this local trivialization is none other than the map Φ constructed in Proposition 5.3. It follows from Corollary 5.12 that the standard coordinate chart $\widetilde{\varphi} = (\varphi \times \mathrm{Id}_U) \circ \Phi$ belongs to the given smooth structure. Thus the given smooth structure is equal to the one constructed in Lemma 4.1. □

If M is a smooth manifold, we will use the term *local frame for M* to mean a local frame for the tangent bundle TM, or in other words an n-tuple of independent vector fields that span TM over some open subset of M. A *global frame for M* is defined similarly. For example, the coordinate vector fields $(\partial/\partial x^i)$ form a smooth local frame over any smooth coordinate domain, called a *coordinate frame.*

A smooth manifold M is said to be *parallelizable* if it admits a smooth global frame. By Corollary 5.11, this is equivalent to TM being a trivial bundle.

Because it has a global coordinate frame, $\mathbb{R}^n$ is parallelizable. Many more examples are provided by the following proposition.

Proposition 5.15. *Every Lie group is parallelizable.*

Proof. If G is a Lie group, any basis for $\mathrm{Lie}(G)$ is a smooth global frame for G, as is easily verified. $\square$

The proof of this proposition shows that in fact every Lie group has a global frame consisting of left-invariant vector fields. We will call any such frame a *left-invariant frame*.

Because they are Lie groups, $\mathbb{S}^1$ and $\mathbb{T}^n$ are parallelizable. (We exhibited left-invariant frames for them in Example 4.22.) It turns out that $\mathbb{S}^3$ and $\mathbb{S}^7$ are parallelizable as well, as you will be asked to show in Problems 5-10 and 8-21. However, despite the evidence of these examples, most smooth manifolds are not parallelizable. The simplest example of a non-parallelizable manifold is $\mathbb{S}^2$, but the proof of this fact will have to wait until we have developed more machinery (see Problem 14-23). In fact, it was shown in 1958 by Raoul Bott and John Milnor [BM58] using more advanced methods from algebraic topology that $\mathbb{S}^1$, $\mathbb{S}^3$, and $\mathbb{S}^7$ are the *only* spheres that are parallelizable. Thus these are the only positive-dimensional spheres that can possibly admit Lie group structures. The first two do (see Example 2.7(g) and Problem 8-19); but it turns out that $\mathbb{S}^7$ has no Lie group structure (see [Bre93, page 301]). A remarkable theorem of Joseph Wolf [Wol71, Wol72] shows that the only compact, simply connected manifolds that are parallelizable are products of Lie groups and copies of $\mathbb{S}^7$.

Bundle Maps

If $\pi\colon E \to M$ and $\pi'\colon E' \to M'$ are two vector bundles, a *bundle map* from E to E' is a pair of continuous maps $F\colon E \to E'$ and $f\colon M \to M'$ such that $\pi' \circ F = f \circ \pi$:

$$\begin{tikzcd} E \arrow[r, "F"] \arrow[d, "\pi"'] & E' \arrow[d, "\pi'"] \\ M \arrow[r, "f"'] & M', \end{tikzcd}$$

and with the property that for each $p \in M$, the restricted map $F|_{E_p}\colon E_p \to E'_{f(p)}$ is linear. When the manifolds and vector bundles are smooth, it is called a *smooth bundle map* if both F and f are smooth. We often refer to F itself as the bundle map, and say that F *covers* f. A bijective bundle map $F\colon E \to E'$ whose inverse is also a bundle map is called a *bundle* *isomorphism*; if F is also a diffeomorphism, it is called a *smooth bundle isomorphism*. If there exists a (smooth) bundle isomorphism between E and E', the two bundles are said to be (*smoothly*) *isomorphic*.

In the special case in which both E and E' are vector bundles over the same base manifold M, a slightly more restrictive notion of bundle map is usually more useful. A *bundle map over* M is a bundle map covering the identity map of M; in other words, a continuous map $F\colon E \to E'$ such that $\pi' \circ F = \pi$:

$$\begin{tikzcd} E \arrow[rr, "F"] \arrow[dr, "\pi"'] & & E' \arrow[dl, "\pi'"] \\ & M, & \end{tikzcd}$$

and whose restriction to each fiber is linear. If $F\colon E \to E'$ is a bundle map over M that is also a (smooth) bundle isomorphism, then we say that E and E' are (*smoothly*) *isomorphic over* M.

◇ **Exercise 5.6.** Show that a smooth rank k vector bundle over M is trivial if and only if it is smoothly isomorphic over M to the product bundle $M \times \mathbb{R}^k$.

◇ **Exercise 5.7.** Suppose $F\colon M \to N$ is a smooth map. Show that $F_*\colon TM \to TN$ is a smooth bundle map.

Suppose $E \to M$ and $E' \to M$ are smooth vector bundles over M, and let $\mathcal{E}(M)$, $\mathcal{E}'(M)$ denote their spaces of smooth sections. If $F\colon E \to E'$ is any smooth bundle map over M, then composition with F induces a map from $\mathcal{E}(M)$ to $\mathcal{E}'(M)$, also denoted by F, as follows:

$$(F(\sigma))(p) = F(\sigma(p)). \tag{5.5}$$

In other words, $F(\sigma)$ is just the composition $F \circ \sigma$. It is easy to check that $F(\sigma)$ is a section of E', and it is smooth by composition.

Because a bundle map is linear on fibers, the resulting map on sections $F\colon \mathcal{E}(M) \to \mathcal{E}'(M)$ is linear over $\mathbb{R}$. In fact, it satisfies a stronger linearity property. A map $F\colon \mathcal{E}(M) \to \mathcal{E}'(M)$ is said to be *linear over* $C^\infty(M)$ if for any smooth functions $u_1, u_2 \in C^\infty(M)$ and smooth sections $\sigma_1, \sigma_2 \in \mathcal{E}(M)$,

$$F(u_1\sigma_1 + u_2\sigma_2) = u_1F(\sigma_1) + u_2F(\sigma_2).$$

It follows easily from the definition (5.5) that the map on sections induced by a bundle map is linear over $C^\infty(M)$. The next proposition shows that the converse is true as well.

Proposition 5.16. *Let $\pi\colon E \to M$ and $\pi'\colon E' \to M$ be smooth vector bundles over a smooth manifold M, and let $\mathcal{E}(M)$, $\mathcal{E}'(M)$ denote their spaces of sections. A map $\mathcal{F}\colon \mathcal{E}(M) \to \mathcal{E}'(M)$ is linear over $C^\infty(M)$ if and only if there is a smooth bundle map $F\colon E \to E'$ over M such that $\mathcal{F}(\sigma) = F \circ \sigma$ for all $\sigma \in \mathcal{E}(M)$.*

Proof. We noted above that the map on sections induced by a smooth bundle map is linear over $C^\infty(M)$. Conversely, suppose $\mathcal{F}\colon \mathcal{E}(M) \to \mathcal{E}'(M)$ is linear over $C^\infty(M)$. First we will show that $\mathcal{F}$ acts locally: If $\sigma_1 \equiv \sigma_2$ in some open set $U \subset M$, then $\mathcal{F}(\sigma_1) \equiv \mathcal{F}(\sigma_2)$ in U. Writing $\tau = \sigma_1 - \sigma_2$, it suffices by linearity of $\mathcal{F}$ to assume that τ vanishes in U and show that $\mathcal{F}(\tau)$ does too. For any $p \in U$, let $\psi \in C^\infty(M)$ be a smooth bump function that is supported in U and is equal to 1 at p. Because $\psi\tau$ is identically zero on M, the fact that $\mathcal{F}$ is linear over $C^\infty(M)$ implies

$$0 = \mathcal{F}(\psi\tau) = \psi\mathcal{F}(\tau).$$

Evaluating at p shows that $\mathcal{F}(\tau)(p) = \psi(p)\mathcal{F}(\tau)(p) = 0$; since the same is true for every $p \in U$, the claim follows.

Next we show that $\mathcal{F}$ actually acts pointwise: If $\sigma_1(p) = \sigma_2(p)$, then $\mathcal{F}(\sigma_1)(p) = \mathcal{F}(\sigma_2)(p)$. Once again, it suffices to assume that $\tau(p) = 0$ and show that $\mathcal{F}(\tau)(p) = 0$. Let $(\sigma_1, \dots, \sigma_k)$ be a smooth local frame for E in some neighborhood of p, and write τ in terms of this frame as $\tau = u^i\sigma_i$ for some smooth functions u^i defined near p. The fact that $\tau(p) = 0$ means that $u^1(p) = \cdots = u^k(p) = 0$. By the extension lemmas for vector bundles and for functions, there exist smooth global sections $\widetilde{\sigma}_i \in \mathcal{E}(M)$ that agree with σ_i in a neighborhood of p, and smooth functions $\widetilde{u}^i \in C^\infty(M)$ that agree with u^i in a neighborhood of p. Then since $\tau = \widetilde{u}^i\widetilde{\sigma}_i$ near p, we have

$$\mathcal{F}(\tau)(p) = \mathcal{F}(\widetilde{u}^i\widetilde{\sigma}_i)(p) = \widetilde{u}^i(p)\mathcal{F}(\widetilde{\sigma}_i)(p) = 0,$$

which proves the claim.

Define a bundle map $F\colon E \to E'$ as follows. For any $p \in M$ and $v \in E_p$, let $F(v) = \mathcal{F}(\widetilde{v})(p) \in E'_p$, where $\widetilde{v}$ is any global smooth section of E such that $\widetilde{v}(p) = v$. The discussion above shows that the resulting element of E'_p is independent of the choice of section. This map F clearly satisfies $\pi' \circ F = \pi$, and it is linear on each fiber because of the linearity of $\mathcal{F}$. It also satisfies $F \circ \sigma(p) = \mathcal{F}(\sigma)(p)$ for any $\sigma \in \mathcal{E}(M)$ by definition. It remains only to show that F is smooth. It suffices to show that it is smooth in a neighborhood of each point.

Given $p \in M$, let (σ_i) be a local frame for E on some neighborhood of p. By the extension lemma, there are global sections $\widetilde{\sigma}_i$ that agree with σ_i in a (smaller) neighborhood U of p. Shrinking U further if necessary, we may also assume that there exists a smooth local frame (σ'_j) for E' over U. Because $\mathcal{F}$ maps smooth global sections of E to smooth global sections of E', there are smooth functions $A_i^j \in C^\infty(U)$ such that $\mathcal{F}(\widetilde{\sigma}_i)|_U = A_i^j\sigma'_j$.

For any $q \in U$ and $v \in E_q$, we can write $v = v^i\sigma_i(q)$ for some real numbers $(v^1, \dots, v^k)$, and then

$$F(v^i\sigma_i(q)) = \mathcal{F}(v^i\widetilde{\sigma}_i)(q) = v^i\mathcal{F}(\widetilde{\sigma}_i)(q) = v^i\mathcal{A}_i^j(q)\sigma'_j(q),$$

because $v^i\widetilde{\sigma}_i$ is a global smooth section of E whose value at q is v. If Φ and Φ' denote the local trivializations of E associated with the frames (σ_i) and (σ'_i), respectively, it follows that the composite map $\Phi' \circ F \circ \Phi^{-1}\colon U \times \mathbb{R}^k \to$

$U \times \mathbb{R}^m$ has the form

$$\Phi' \circ F \circ \Phi^{-1}(q, (v^1, \dots, v^k)) = (q, (A_i^1(q)v^i, \dots, A_i^m(q)v^i)),$$

which is smooth. Because Φ and Φ' are diffeomorphisms, this shows that F is smooth on $\pi^{-1}(U)$. $\square$

Later, after we have developed more tools, we will see many examples of smooth bundle maps. For now, here are two elementary examples.

Example 5.17 (Bundle Maps).

(a) If X is any smooth vector field on $\mathbb{R}^3$, the cross product with X defines a map from $\mathcal{T}(\mathbb{R}^3)$ to itself: $Y \mapsto X \times Y$. Since it is linear over $C^\infty(\mathbb{R}^3)$ in Y, it corresponds to a bundle map from $T\mathbb{R}^3$ to $T\mathbb{R}^3$.

(b) Similarly, the Euclidean dot product with X defines a map from $\mathcal{T}(\mathbb{R}^3)$ to $C^\infty(\mathbb{R}^3)$, which is linear over $C^\infty(\mathbb{R}^3)$ and therefore corresponds to a bundle map from $T\mathbb{R}^3$ to the trivial line bundle $\mathbb{R}^3 \times \mathbb{R}$.

Because of Proposition 5.16, we will frequently use the same symbol for both a bundle map $F\colon E \to E'$ over M and the linear map $F\colon \mathcal{E}(M) \to \mathcal{E}'(M)$ that it induces on sections, and we will refer to a map of either of these types as a bundle map. Because the action on sections is obtained simply by applying the bundle map pointwise, this should cause no confusion. In fact, we have been doing the same thing all along—for example, we use the same notation $X \mapsto 2X$ to denote both the operation of multiplying vectors in each tangent space T_pM by 2, and the operation of multiplying vector fields by 2. Because multiplying by 2 is a bundle map from TM to itself, there is no ambiguity about what is meant.

It should be noted that most maps that involve differentiation are *not* bundle maps. For example, if X is a smooth vector field on M, the derivation $X\colon C^\infty(M) \to C^\infty(M)$ is not a bundle map from the trivial line bundle to itself, because it is not linear over C^∞. Similarly, for a fixed vector field X, the map $Y \mapsto [X, Y]$ is a linear map from $\mathcal{T}(M)$ to itself, but it is not a bundle map because it satisfies (4.7) instead of linearity over $C^\infty(M)$. As a rule of thumb, a linear map that takes smooth sections of one bundle to smooth sections of another is likely to be a bundle map if it acts pointwise, but not if it involves differentiation.

Categories and Functors

Another useful perspective on the tangent bundle is provided by the theory of categories. In this section, we summarize the basic definitions of category theory. We will not do much with the theory in this book, but we mention it because it provides a convenient and powerful language for talking about many of the mathematical structures we will meet.

A *category* C consists of three things:

- a class of *objects*;
- for each pair X, Y of objects a set $\mathrm{Hom}_{\mathsf{C}}(X, Y)$ whose elements are called *morphisms*;
- and for each triple X, Y, Z of objects a map called *composition*: $\mathrm{Hom}_{\mathsf{C}}(X, Y) \times \mathrm{Hom}_{\mathsf{C}}(Y, Z) \to \mathrm{Hom}_{\mathsf{C}}(X, Z)$, written $(f, g) \mapsto g \circ f$.

The morphisms are required to satisfy the following properties:

(i) Associativity: $(f \circ g) \circ h = f \circ (g \circ h)$.

(ii) Existence of Identities: For each object X in C, there exists an *identity morphism* $\mathrm{Id}_X \in \mathrm{Hom}_{\mathsf{C}}(X, X)$, satisfying $\mathrm{Id}_Y \circ f = f = f \circ \mathrm{Id}_X$ for all $f \in \mathrm{Hom}_{\mathsf{C}}(X, Y)$.

A morphism $f \in \mathrm{Hom}_{\mathsf{C}}(X, Y)$ is called an *isomorphism* in C if there exists a morphism $g \in \mathrm{Hom}_{\mathsf{C}}(Y, X)$ such that $f \circ g = \mathrm{Id}_Y$ and $g \circ f = \mathrm{Id}_X$.

Example 5.18 (Categories). In most of the categories that one meets "in nature," the objects are sets with some extra structure, the morphisms are maps that preserve that structure, and the composition laws and identity morphisms are the obvious ones. Some of the categories of this type that will appear in this book (implicitly or explicitly) are listed below. In each case, we describe the category by giving its objects and its morphisms.

- SET: Sets and maps.
- TOP: Topological spaces and continuous maps.
- TM: Topological manifolds and continuous maps.
- SM: Smooth manifolds and smooth maps.
- VB: Smooth vector bundles and smooth bundle maps.
- $\mathsf{VECT}_{\mathbb{R}}$: Real vector spaces and real-linear maps.
- $\mathsf{VECT}_{\mathbb{C}}$: Complex vector spaces and complex-linear maps.
- GROUP: Groups and group homomorphisms.
- AB: Abelian groups and group homomorphisms.
- LIE: Lie groups and Lie group homomorphisms.
- lie: Lie algebras and Lie algebra homomorphisms.

The reason we are careful to use the word "class" instead of "set" for the collection of objects in a category is that some categories are "too large" to be considered sets. For example, in the category SET, the class of objects is the class of all sets; any attempt to treat this class as a set in its own right leads to the well-known Russell paradox of set theory. (See the Appendix to [Lee00] or almost any book on set theory for more.)

The most important construction in category theory is the following. If C and D are categories, a *covariant functor* from C to D is a rule $\mathcal{F}$ that assigns to each object X in C an object $\mathcal{F}(X)$ in D, and to each morphism $f \in \mathrm{Hom}_{\mathsf{C}}(X,Y)$ a morphism $\mathcal{F}(f) \in \mathrm{Hom}_{\mathsf{D}}(\mathcal{F}(X), \mathcal{F}(Y))$, so that identities and composition are preserved:

$$\mathcal{F}(\mathrm{Id}_X) = \mathrm{Id}_{\mathcal{F}(X)}; \qquad \mathcal{F}(g \circ h) = \mathcal{F}(g) \circ \mathcal{F}(h).$$

We will also need to consider functors that reverse morphisms: A *contravariant functor* $\mathcal{F}$ from C to D assigns to each object X in C an object $\mathcal{F}(X)$ in D, and to each morphism $g \in \mathrm{Hom}_{\mathsf{C}}(X,Y)$ a morphism $\mathcal{F}(g) \in \mathrm{Hom}_{\mathsf{D}}(\mathcal{F}(Y), \mathcal{F}(X))$, such that

$$\mathcal{F}(\mathrm{Id}_X) = \mathrm{Id}_{\mathcal{F}(X)}; \qquad \mathcal{F}(g \circ h) = \mathcal{F}(h) \circ \mathcal{F}(g).$$

If the functor is understood, it is common for the morphism induced by a covariant functor to be denoted by g_* instead of $\mathcal{F}(g)$, and that induced by a contravariant functor by g^*.

◇ **Exercise 5.8.** Show that any (covariant or contravariant) functor from C to D takes isomorphisms in C to isomorphisms in D.

One trivial example of a covariant functor from any category to itself is the *identity functor*, which takes each object and each morphism to itself. Another example is the *forgetful functor*: If C is a category whose objects are sets with some additional structure and whose morphisms are maps preserving that structure (as are all the categories listed in Example 5.18 except the first), the forgetful functor $\mathcal{F}\colon \mathsf{C} \to \mathsf{SET}$ assigns to each object its underlying set, and to each morphism the same map thought of as a map between sets.

More interesting functors arise when we associate "invariants" to classes of mathematical objects. For example, Proposition 4.26 shows that the assignment $G \mapsto \mathrm{Lie}(G)$, $F \mapsto F_*$ is a covariant functor from the category of Lie groups to the category of Lie algebras. If we define TOP_* to be the category whose objects are *pointed topological spaces* (i.e., topological spaces together with a choice of base point in each) and whose morphisms are continuous maps taking base points to base points, then the fundamental group is a covariant functor from TOP_* to GROUP.

The discussion in this chapter has given us another important example of a functor: The *tangent functor* is a covariant functor from the category SM of smooth manifolds to the category VB of smooth vector bundles. To each smooth manifold M it assigns the tangent bundle $TM \to M$, and to each smooth map $F\colon M \to N$ it assigns the push-forward $F_*\colon TM \to TN$, which is a smooth bundle map by Exercise 5.7. The fact that this is a functor is the content of parts (b) and (c) of Lemma 3.5.

Problems

5-1. If E is a vector bundle over a topological space M, show that the projection map $\pi\colon E \to M$ is a homotopy equivalence.

5-2. Prove that the space E constructed in Example 5.2, together with the projection $\pi\colon E \to \mathbb{S}^1$, is a smooth rank 1 vector bundle over $\mathbb{S}^1$, and show that it is nontrivial.

5-3. Let $\pi\colon E \to M$ be a smooth vector bundle of rank k over a smooth manifold M. Suppose $\{U_\alpha\}_{\alpha\in A}$ is an open cover of M, and for each $\alpha \in A$ we are given a smooth local trivialization $\Phi_\alpha\colon \pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{R}^k$ of E. For each $\alpha, \beta \in A$ such that $U_\alpha \cap U_\beta \neq \varnothing$, let $\tau_{\alpha\beta}\colon U_\alpha \cap U_\beta \to \mathrm{GL}(k, \mathbb{R})$ be the transition function defined by (5.3). Show that the following identity is satisfied for all $\alpha, \beta, \gamma \in A$:

$$\tau_{\alpha\beta}(p)\tau_{\beta\gamma}(p) = \tau_{\alpha\gamma}(p), \qquad p \in U_\alpha \cap U_\beta \cap U_\gamma. \tag{5.6}$$

(Here juxtaposition of matrices represents matrix multiplication.)

5-4. Let M be a smooth manifold and let $\{U_\alpha\}_{\alpha\in A}$ be an open cover of M. Suppose for each $\alpha, \beta \in A$ we are given a smooth map $\tau_{\alpha\beta}\colon U_\alpha \cap U_\beta \to \mathrm{GL}(k, \mathbb{R})$ such that (5.6) is satisfied for all $\alpha, \beta, \gamma \in A$. Show that there is a smooth rank k vector bundle $E \to M$ with smooth local trivializations $\Phi_\alpha\colon \pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{R}^k$ whose transition functions are the given maps $\tau_{\alpha\beta}$. [Hint: Define an appropriate equivalence relation on $\coprod_{\alpha\in A}(U_\alpha \times \mathbb{R}^k)$, and use the bundle construction lemma.]

5-5. Let $\pi\colon E \to M$ and $\widetilde{\pi}\colon \widetilde{E} \to M$ be two smooth rank k vector bundles over a smooth manifold M. Suppose $\{U_\alpha\}_{\alpha\in A}$ is an open cover of M such that both E and $\widetilde{E}$ admit smooth local trivializations over each U_α. Let $\{\tau_{\alpha\beta}\}$ and $\{\widetilde{\tau}_{\alpha\beta}\}$ denote the transition functions determined by the given local trivializations of E and $\widetilde{E}$, respectively. Show that E and $\widetilde{E}$ are smoothly isomorphic over M if and only if for each $\alpha \in A$ there exists a smooth map $\sigma_\alpha\colon U_\alpha \to \mathrm{GL}(k, \mathbb{R})$ such that

$$\widetilde{\tau}_{\alpha\beta}(p) = \sigma_\alpha(p)^{-1}\tau_{\alpha\beta}(p)\sigma_\beta(p), \qquad p \in U_\alpha \cap U_\beta.$$

5-6. Let $U = \mathbb{S}^1 \smallsetminus \{1\}$ and $V = \mathbb{S}^1 \smallsetminus \{-1\}$, and define $\tau\colon U \cap V \to \mathrm{GL}(1, \mathbb{R})$ by

$$\tau(z) = \begin{cases} (1), & \operatorname{Im} z > 0, \\ (-1), & \operatorname{Im} z < 0. \end{cases}$$

By the result of Problem 5-4, there is a smooth real line bundle $F \to \mathbb{S}^1$ that is trivial over U and V, and has τ as transition function. Show that F is smoothly isomorphic to the Möbius bundle of Example 5.2.

5-7. Compute the transition function for $T\mathbb{S}^2$ associated with the two local trivializations determined by stereographic coordinates.

5-8. Let $\pi\colon E \to M$ be a smooth vector bundle of rank k, and suppose $\sigma_1, \dots, \sigma_m$ are independent smooth local sections over an open subset $U \subset M$. Show that for each $p \in U$ there are smooth sections $\sigma_{m+1}, \dots, \sigma_k$ defined on some neighborhood V of p such that $(\sigma_1, \dots, \sigma_k)$ is a smooth local frame for E over $U \cap V$.

5-9. Suppose E and E' are vector bundles over a smooth manifold M, and $F\colon E \to E'$ is a bijective bundle map over M. Show that F is a bundle isomorphism.

5-10. Consider the following vector fields on $\mathbb{R}^4$:

$$\begin{aligned} X_1 &= -x^2 \frac{\partial}{\partial x^1} + x^1 \frac{\partial}{\partial x^2} + x^4 \frac{\partial}{\partial x^3} - x^3 \frac{\partial}{\partial x^4}, \\ X_2 &= -x^3 \frac{\partial}{\partial x^1} - x^4 \frac{\partial}{\partial x^2} + x^1 \frac{\partial}{\partial x^3} + x^2 \frac{\partial}{\partial x^4}, \\ X_3 &= -x^4 \frac{\partial}{\partial x^1} + x^3 \frac{\partial}{\partial x^2} - x^2 \frac{\partial}{\partial x^3} + x^1 \frac{\partial}{\partial x^4}. \end{aligned}$$

Show that there are smooth vector fields V_1, V_2, V_3 on $\mathbb{S}^3$ such that V_j is ι-related to X_j for $j = 1, \dots, 3$, where $\iota\colon \mathbb{S}^3 \hookrightarrow \mathbb{R}^4$ is inclusion. Conclude that $\mathbb{S}^3$ is parallelizable.

5-11. Let V be a finite-dimensional vector space, and let $\mathrm{G}_k(V)$ be the Grassmannian of k-dimensional subspaces of V. Let T be the disjoint union of all these k-dimensional subspaces:

$$T = \coprod_{S \in \mathrm{G}_k(V)} S;$$

and let $\pi\colon T \to \mathrm{G}_k(V)$ be the natural map sending each point $x \in S$ to S. Show that T has a unique smooth manifold structure making it into a smooth rank k vector bundle over $\mathrm{G}_k(V)$, with π as projection and with the vector space structure on each fiber inherited from V. [Remark: T is sometimes called the *tautological vector bundle* over $\mathrm{G}_k(V)$, because the fiber over each point $S \in \mathrm{G}_k(V)$ is S itself.]

5-12. Show that the tautological vector bundle over $\mathrm{G}_1(\mathbb{R}^2)$ is isomorphic to the Möbius bundle. (See Problems 5-2, 5-6, and 5-11.)

5-13. Let V_0 be the category whose objects are finite-dimensional real vector spaces and whose morphisms are linear isomorphisms. If $\mathcal{F}$ is a covariant functor from V_0 to itself, for each finite-dimensional vector space V we get a map $\mathcal{F}\colon \mathrm{GL}(V) \to \mathrm{GL}(\mathcal{F}(V))$ sending each isomorphism $A\colon V \to V$ to the induced isomorphism $\mathcal{F}(A)\colon \mathcal{F}(V) \to \mathcal{F}(V)$. We say $\mathcal{F}$ is a *smooth functor* if this map is smooth for every V. If $\mathcal{F}\colon \mathsf{V}_0 \to \mathsf{V}_0$ is a smooth functor, show that for every smooth vector bundle $E \to M$ there is a smooth vector bundle $\mathcal{F}(E) \to M$ whose fiber at each point $p \in M$ is $\mathcal{F}(E_p)$.

6
The Cotangent Bundle

In this chapter, we introduce a construction that is not typically seen in elementary calculus: tangent covectors, which are linear functionals on the tangent space at a point $p \in M$. The space of all covectors at p is a vector space called the cotangent space at p; in linear-algebraic terms, it is the dual space to T_pM. The union of all cotangent spaces at all points of M is a vector bundle called the cotangent bundle.

Whereas tangent vectors give us a coordinate-free interpretation of derivatives of curves, it turns out that derivatives of real-valued functions on a manifold are most naturally interpreted as tangent covectors. Thus we define the differential of a real-valued function as a covector field (a smooth section of the cotangent bundle); it is a sort of coordinate-independent analogue of the classical gradient. We then explore the behavior of covector fields under smooth maps, and show that smooth covector fields on the range of a smooth map always pull back to smooth covector fields on the domain.

In the second half of the chapter, we define line integrals of covector fields. This allows us to generalize the classical notion of line integrals to manifolds. Then we explore the relationships among three closely related types of covector fields: exact (those that are the differentials of functions), conservative (those whose line integrals around closed curves are zero), and closed (those that satisfy a certain differential equation in coordinates).

Covectors

Let V be a finite-dimensional vector space. (As usual, all of our vector spaces are assumed to be real.) We define a *covector* on V to be a real-valued linear functional on V, that is, a linear map $\omega\colon V \to \mathbb{R}$. The space of all covectors on V is itself a real vector space under the obvious operations of pointwise addition and scalar multiplication. It is denoted by V^* and called the *dual space* to V.

The most important fact about V^* is expressed in the following proposition.

Proposition 6.1. *Let V be a finite-dimensional vector space. If $(E_1, \ldots, E_n)$ is any basis for V, then the covectors $(\varepsilon^1, \ldots, \varepsilon^n)$, defined by*

$$\varepsilon^i(E_j) = \delta^i_j = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases}$$

form a basis for V^*, called the dual basis to (E_i). Therefore $\dim V^* = \dim V$.

Remark. The symbol δ^i_j used in this proposition, meaning 1 if $i = j$ and 0 otherwise, is called the *Kronecker delta*.

◇ **Exercise 6.1.** Prove Proposition 6.1.

For example, if (e_i) denotes the standard basis for $\mathbb{R}^n$, we denote the dual basis by $(e^1, \ldots, e^n)$ (note the upper indices), and call it the *standard dual basis*. These basis covectors are the linear functions from $\mathbb{R}^n$ to $\mathbb{R}$ given by

$$e^j(v) = e^j(v^1, \ldots, v^n) = v^j.$$

In other words, e^j is just the linear functional that picks out the jth component of a vector. In matrix notation, a linear map from $\mathbb{R}^n$ to $\mathbb{R}$ is represented by a $1 \times n$ matrix, i.e., a row matrix. The basis covectors can therefore also be thought of as the linear functionals represented by the row matrices

$$e^1 = (1\ 0\ \ldots\ 0), \quad \ldots, \quad e^n = (0\ \ldots\ 0\ 1).$$

In general, if (E_i) is a basis for V and (ε^j) is its dual basis, then for any vector $X = X^i E_i \in V$, we have

$$\varepsilon^j(X) = X^i \varepsilon^j(E_i) = X^i \delta^j_i = X^j.$$

Thus, just as in the case of $\mathbb{R}^n$, ε^j picks out the jth component of a vector with respect to the basis (E_i). More generally, Proposition 6.1 shows that we can express an arbitrary covector $\omega \in V^*$ in terms of the dual basis as

$$\omega = \omega_j \varepsilon^j, \tag{6.1}$$

where the components ω_j are determined by

$$\omega_j = \omega(E_j).$$

The action of ω on a vector $X = X^i E_i$ is

$$\omega(X) = \omega_j X^j. \tag{6.2}$$

We will always write basis covectors with upper indices, and components of a covector with lower indices, because this helps to ensure that mathematically meaningful expressions such as (6.1) and (6.2) will always follow our index conventions: Any index that is to be summed over in a given term appears exactly twice, once as a subscript and once as a superscript.

Suppose V and W are vector spaces and $A\colon V \to W$ is a linear map. We define a linear map $A^*\colon W^* \to V^*$, called the *dual map* or *transpose* of A, by

$$(A^*\omega)(X) = \omega(AX) \qquad \text{for } \omega \in W^*,\ X \in V.$$

◇ **Exercise 6.2.** Show that $A^*\omega$ is actually a linear functional on V, and that A^* is a linear map.

Proposition 6.2. *The dual map satisfies the following properties.*

(*a*) $(A \circ B)^* = B^* \circ A^*$.

(*b*) $(\mathrm{Id}_V)^*\colon V^* \to V^*$ *is the identity map of* V^*.

◇ **Exercise 6.3.** Prove the preceding proposition.

Corollary 6.3. *The assignment that sends a vector space to its dual space and a linear map to its dual map is a contravariant functor from the category of real vector spaces to itself.*

Apart from the fact that the dimension of V^* is the same as that of V, the second most important fact about dual spaces is the following characterization of the *second dual space* $V^{**} = (V^*)^*$. For each vector space V there is a natural, basis-independent map $\xi\colon V \to V^{**}$, defined as follows. For each vector $X \in V$, define a linear functional $\xi(X)\colon V^* \to \mathbb{R}$ by

$$\xi(X)(\omega) = \omega(X) \qquad \text{for } \omega \in V^*. \tag{6.3}$$

◇ **Exercise 6.4.** Let V be a vector space.

(a) For any $X \in V$, show that $\xi(X)(\omega)$ depends linearly on ω, so that $\xi(X) \in V^{**}$.

(b) Show that the map $\xi\colon V \to V^{**}$ is linear.

Proposition 6.4. *For any finite-dimensional vector space* V*, the map* $\xi\colon V \to V^{**}$ *is an isomorphism.*

Proof. Because V and V^* have the same dimension, it suffices to verify that ξ is injective. Suppose $X \in V$ is not zero. Extend X to a basis $(X = E_1, \ldots, E_n)$ for V, and let $(\varepsilon^1, \ldots, \varepsilon^n)$ denote the dual basis for V^*. Then

$$\xi(X)(\varepsilon^1) = \varepsilon^1(X) = \varepsilon^1(E_1) = 1 \neq 0,$$

so $\xi(X) \neq 0$. □

The preceding proposition shows that when V is finite dimensional, we can unambiguously identify V^{**} with V itself, because the map ξ is canonically defined, without reference to any basis. It is important to observe that although V^* is also isomorphic to V (for the simple reason that any two vector spaces of the same dimension are isomorphic), there is no *canonical* isomorphism $V \cong V^*$. One way to make this statement precise is indicated in Problem 6-1.

Because of Proposition 6.4, the real number $\omega(X)$ obtained by applying a covector ω to a vector X is sometimes denoted by either of the more symmetric-looking notations $\langle \omega, X \rangle$ or $\langle X, \omega \rangle$; both expressions can be thought of either as the action of the covector $\omega \in V^*$ on the vector $X \in V$, or as the action of the covector $\xi(X) \in V^{**}$ on the element $\omega \in V^*$. There should be no cause for confusion with the use of the same angle bracket notation for inner products: Whenever one of the arguments is a vector and the other a covector, the notation $\langle \omega, X \rangle$ is always to be interpreted as the natural pairing between vectors and covectors, not as an inner product. We will typically omit any mention of the map ξ, and think of $X \in V$ either as a vector or as a linear functional on V^*, depending on the context.

There is also a symmetry between bases and dual bases for a finite-dimensional vector space V: Any basis for V determines a dual basis for V^*, and conversely any basis for V^* determines a dual basis for $V^{**} = V$. It is easy to check that if (ε^i) is the basis for V^* dual to a basis (E_i) for V, then (E_i) is the basis dual to (ε^i), because both statements are equivalent to the relation $\langle \varepsilon^i, E_j \rangle = \delta^i_j$.

Tangent Covectors on Manifolds

Now let M be a smooth manifold. For each $p \in M$, we define the *cotangent space* at p, denoted by T^*_pM, to be the dual space to T_pM:

$$T^*_pM = (T_pM)^*.$$

Elements of T^*_pM are called *tangent covectors* at p, or just covectors at p.

If (x^i) are smooth local coordinates on an open subset $U \subset M$, then for each $p \in U$, the coordinate basis $(\partial/\partial x^i|_p)$ gives rise to a dual basis for T^*_pM, which we denote for the moment by $(\lambda^i|_p)$. (In a short while, we will come up with a better notation.) Any covector $\omega \in T_p^*M$ can thus be written uniquely as $\omega = \omega_i \lambda^i|_p$, where

$$\omega_i = \omega\left(\frac{\partial}{\partial x^i}\bigg|_p \right).$$

Suppose now that $(\widetilde{x}^j)$ is another set of smooth coordinates whose domain overlaps U, and let $(\widetilde{\lambda}^j|_p)$ denote the basis for T_p^*M dual to $(\partial/\partial\widetilde{x}^j|_p)$. We can compute the components of the same covector ω with respect to the new coordinate system as follows. First observe that the computations in Chapter 3 show that the coordinate vector fields transform as follows:

$$\frac{\partial}{\partial x^i}\bigg|_p = \frac{\partial \widetilde{x}^j}{\partial x^i}(p)\, \frac{\partial}{\partial \widetilde{x}^j}\bigg|_p. \tag{6.4}$$

Writing ω in both systems as

$$\omega = \omega_i \lambda^i|_p = \widetilde{\omega}_j \widetilde{\lambda}^j|_p,$$

we can use (6.4) to compute the components ω_i in terms of $\widetilde{\omega}_j$:

$$\omega_i = \omega\left(\frac{\partial}{\partial x^i}\bigg|_p \right) = \omega\left(\frac{\partial \widetilde{x}^j}{\partial x^i}(p)\, \frac{\partial}{\partial \widetilde{x}^j}\bigg|_p \right) = \frac{\partial \widetilde{x}^j}{\partial x^i}(p)\widetilde{\omega}_j. \tag{6.5}$$

As we mentioned in Chapter 3, in the early days of smooth manifold theory, before most of the abstract coordinate-free definitions we are using were developed, mathematicians tended to think of a tangent vector at a point p as an assignment of an n-tuple of real numbers to each smooth coordinate system, with the property that the n-tuples $(X^1, \ldots, X^n)$ and $(\widetilde{X}^1, \ldots, \widetilde{X}^n)$ assigned to two different coordinate systems (x^i) and $(\widetilde{x}^j)$ were related by the transformation law that we derived in Chapter 3:

$$\widetilde{X}^j = \frac{\partial \widetilde{x}^j}{\partial x^i}(p) X^i. \tag{6.6}$$

Similarly, a tangent covector was thought of as an n-tuple $(\omega_1, \ldots, \omega_n)$ that transforms, by virtue of (6.5), according to the following slightly different rule:

$$\omega_i = \frac{\partial \widetilde{x}^j}{\partial x^i}(p)\widetilde{\omega}_j. \tag{6.7}$$

Since the transformation law (6.4) for the coordinate partial derivatives follows directly from the chain rule, it can be thought of as fundamental. Thus it became customary to call tangent covectors *covariant vectors* because their components transform in the same way as ("vary with") the coordinate partial derivatives, with the Jacobian matrix $(\partial\widetilde{x}^j/\partial x^i)$ multiplying the objects associated with the "new" coordinates $(\widetilde{x}^j)$ to obtain those associated with the "old" coordinates (x^i). Analogously, tangent vectors were called *contravariant vectors*, because their components transform in the opposite way. (Remember, it was the component n-tuples that were thought of as the objects of interest.) Admittedly, it does not make a lot of sense, but by now the terms are well entrenched, and we will see them again in Chapter 11. Note that this use of the terms covariant and contravariant has nothing to do with the covariant and contravariant functors of category theory!

The Cotangent Bundle

The disjoint union

$$T^*M = \coprod_{p \in M} T_p^*M$$

is called the *cotangent bundle* of M. It has a natural projection map $\pi\colon T^*M \to M$ sending $\omega \in T_p^*M$ to $p \in M$. As above, given any smooth local coordinates (x^i) on $U \subset M$, for each $p \in U$ we denote the basis for T_p^*M dual to $(\partial/\partial x^i|_p)$ by $(\lambda^i|_p)$. This defines n maps $\lambda^1, \dots, \lambda^n\colon U \to T^*M$, called *coordinate covector fields.*

Proposition 6.5. *Let M be a smooth manifold and let T^*M be its cotangent bundle. With its standard projection map and the natural vector space structure on each fiber, T^*M has a unique smooth manifold structure making it into a rank n vector bundle over M for which all coordinate covector fields are smooth local sections.*

Proof. The proof is just like the one we gave for the tangent bundle. Given a smooth chart (U, φ) on M, with coordinate functions (x^i), define $\Phi\colon \pi^{-1}(U) \to U \times \mathbb{R}^n$ by

$$\Phi\left(\xi_i \lambda^i|_p\right) = (p, (\xi_1, \dots, \xi_n)),$$

where λ^i is the ith coordinate covector field associated with (x^i). Suppose $(\widetilde{U}, \widetilde{\varphi})$ is another smooth chart with coordinate functions $(\widetilde{x}^j)$, and let $\widetilde{\Phi}\colon \pi^{-1}(\widetilde{U}) \to \widetilde{U} \times \mathbb{R}^n$ be defined analogously. On $\pi^{-1}(U \cap V)$, it follows from (6.5) that

$$\Phi \circ \widetilde{\Phi}^{-1}(p, (\widetilde{\xi}_1, \dots, \widetilde{\xi}_n)) = \left(p, \left(\frac{\partial \widetilde{x}^j}{\partial x^1}(p)\widetilde{\xi}_j, \dots, \frac{\partial \widetilde{x}^j}{\partial x^n}(p)\widetilde{\xi}_j\right)\right).$$

The $\mathrm{GL}(n, \mathbb{R})$-valued function $(\partial \widetilde{x}^j / \partial x^i)$ is smooth, so it follows from the bundle construction lemma that T^*M has a smooth structure making it into a smooth vector bundle for which the maps Φ are smooth local trivializations. Uniqueness follows as in the proof of Proposition 5.3. □

As in the case of the tangent bundle, smooth local coordinates for M yield smooth local coordinates for its cotangent bundle. If (x^i) are smooth coordinates on an open set $U \subset M$, Corollary 5.12 shows that the map from $\pi^{-1}(U)$ to $\mathbb{R}^{2n}$ given by

$$\xi_i \lambda^i|_p \mapsto (x^1(p), \dots, x^n(p), \xi_1, \dots, \xi_n)$$

is a smooth coordinate chart for T^*M. We will call (x^i, ξ_i) the *standard coordinates* for T^*M associated with (x^i). (In this situation, we must forego our insistence that coordinate functions have upper indices, because the fiber coordinates ξ_i are already required by our index conventions to have lower indices. Nonetheless, the convention still holds that each index to be summed over in a given term appears once as a superscript and once as a subscript.)

A section of T^*M is called a *covector field* or a (*differential*) 1*-form*. (The reason for the latter terminology will become clear in Chapter 12, when we define differential k-forms for $k > 1$.) As we did with vector fields, we will write the value of a covector field ω at a point $p \in M$ as ω_p instead of $\omega(p)$, to avoid conflict with the notation for the action of a covector on a vector. In any smooth local coordinates on an open set $U \subset M$, a covector field ω can be written in terms of the coordinate covector fields (λ^i) as $\omega = \omega_i \lambda^i$ for n functions $\omega_i \colon U \to \mathbb{R}$ called the *component functions* of ω. They are characterized by

$$\omega_i(p) = \omega_p\left(\left.\frac{\partial}{\partial x^i}\right|_p \right).$$

Just as in the case of vector fields, there are several ways to check for smoothness of a covector field.

Lemma 6.6 (Smoothness Criteria for Covector Fields). *Let M be a smooth manifold, and let $\omega \colon M \to T^*M$ be a rough section.*

(*a*) *If $\omega = \omega_i \lambda^i$ is the coordinate representation for ω in any smooth chart $(U, (x^i))$ for M, then ω is smooth on U if and only if its component functions ω_i are smooth.*

(*b*) *ω is smooth if and only if for every smooth vector field X on an open subset $U \subset M$, the function $\langle \omega, X \rangle \colon U \to \mathbb{R}$, defined by*

$$\langle \omega, X \rangle(p) = \langle \omega_p, X_p \rangle = \omega_p(X_p)$$

is smooth.

◇ **Exercise 6.5.** Prove Lemma 6.6.

An ordered n-tuple of covector fields $(\varepsilon^1, \dots, \varepsilon^n)$ defined on some open set $U \subset M$ is called a *local coframe for M over U* if $(\varepsilon^i|_p)$ forms a basis for T_p^*M at each point $p \in U$. If $U = M$, it is called a *global coframe*. (A local coframe is just a local frame for T^*M, in the terminology introduced in Chapter 5.) Given a local frame $(E_1, \dots, E_n)$ for TM over an open set U, there is a uniquely determined local coframe $(\varepsilon^1, \dots, \varepsilon^n)$ satisfying

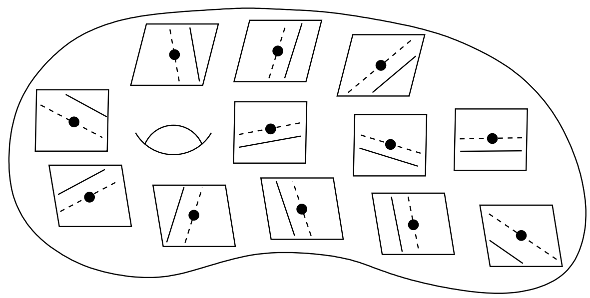


Figure 6.1. A covector field.

$\varepsilon^i(E_j) = \delta^i_j$. This coframe is called the *dual coframe* to the given frame. If (E_i) is a smooth frame, then its dual coframe (ε_i) is smooth by part (b) of the preceding lemma. For example, in any smooth chart, the coordinate vector fields (λ^i) constitute a smooth local coframe, called a *coordinate coframe*. It is the coframe dual to the coordinate frame.

We denote the real vector space of all smooth covector fields on M by $\mathcal{T}^*(M)$. As smooth sections of a vector bundle, elements of $\mathcal{T}^*(M)$ can be multiplied by smooth real-valued functions: If $f \in C^\infty(M)$ and $\omega \in \mathcal{T}^*(M)$, the covector field $f\omega$ is defined by

$$(f\omega)_p = f(p)\omega_p. \tag{6.8}$$

Like the space of smooth vector fields, $\mathcal{T}^*(M)$ is a module over $C^\infty(M)$.

Geometrically, we think of a vector field on M as a rule that attaches an arrow to each point of M. What kind of geometric picture can we form of a covector field? The key idea is that a nonzero linear functional $\omega_p \in T^*_pM$ is completely determined by two pieces of data: its kernel, which is a codimension-1 linear subspace of T_pM (a *hyperplane*), and the set of vectors X for which $\omega_p(X) = 1$, which is an affine hyperplane parallel to the kernel. (Actually, the set where $\omega_p(X) = 1$ alone suffices, but it is useful to visualize the two parallel hyperplanes.) The value of $\omega_p(X)$ for any other vector X is then obtained by linear interpolation or extrapolation.

Thus you can visualize a covector field as defining a pair of affine hyperplanes in each tangent space, one through the origin and another parallel to it, and varying continuously from point to point (Figure 6.1). At points where the covector field takes on the value zero, one of the hyperplanes goes off to infinity.

The Differential of a Function

In elementary calculus, the gradient of a smooth real-valued function f on $\mathbb{R}^n$ is defined as the vector field whose components are the partial derivatives of f. In our notation, this would read:

$$\operatorname{grad} f = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i} \frac{\partial}{\partial x^i}.$$

Unfortunately, in this form, the gradient does not make coordinate-independent sense. (The fact that it violates our index conventions should be taken as a strong clue.)

◇ **Exercise 6.6.** Let $f(x, y) = x^2$ on $\mathbb{R}^2$, and let X be the vector field

$$X = \operatorname{grad} f = 2x \frac{\partial}{\partial x}.$$

Compute the coordinate expression of X in polar coordinates (on some open set on which they are defined) using (6.4) and show that it is *not* equal to

$$\frac{\partial f}{\partial r} \frac{\partial}{\partial r} + \frac{\partial f}{\partial \theta} \frac{\partial}{\partial \theta}.$$

Although the first partial derivatives of a smooth function cannot be interpreted in a coordinate-independent way as the components of a vector field, it turns out that they *can* be interpreted as the components of a covector field. This is the most important application of covector fields.

Let f be a smooth real-valued function on a smooth manifold M. (As usual, all of this discussion applies to functions defined on an open subset $U \subset M$, simply by replacing M by U throughout.) We define a covector field df, called the *differential* of f, by

$$df_p(X_p) = X_p f \qquad \text{for } X_p \in T_pM.$$

Lemma 6.7. *The differential of a smooth function is a smooth covector field.*

Proof. It is straightforward to verify that at each point $p \in M$, $df_p(X_p)$ depends linearly on X_p, so that df_p is indeed a covector at p. To see that df is smooth, we use Lemma 6.6(b): For any smooth vector field X on an open subset $U \subset M$, the function $\langle df, X \rangle$ is smooth because it is equal to Xf. □

To see what df looks like more concretely, we need to compute its coordinate representation. Let (x^i) be smooth coordinates on an open subset $U \subset M$, and let (λ^i) be the corresponding coordinate coframe on U. Writing df in coordinates as $df_p = A_i(p)\lambda^i|_p$ for some functions $A_i \colon U \to \mathbb{R}$, the definition of df implies

$$A_i(p) = df_p\left(\left.\frac{\partial}{\partial x^i}\right|_p \right) = \left.\frac{\partial}{\partial x^i}\right|_p f = \frac{\partial f}{\partial x^i}(p).$$

This yields the following formula for the coordinate representation of df:

$$df_p = \frac{\partial f}{\partial x^i}(p)\lambda^i|_p. \tag{6.9}$$

Thus the component functions of df in any smooth coordinate chart are the partial derivatives of f with respect to those coordinates. Because of this, we can think of df as an analogue of the classical gradient, reinterpreted in a way that makes coordinate-independent sense on a manifold.

If we apply (6.9) to the special case in which f is one of the coordinate functions $x^j : U \to \mathbb{R}$, we find

$$dx^j|_p = \frac{\partial x^j}{\partial x^i}(p)\lambda^i|_p = \delta^j_i \lambda^i|_p = \lambda^j|_p.$$

In other words, *the coordinate covector field λ^j is none other than dx^j!* Therefore, the formula (6.9) for df_p can be rewritten as

$$df_p = \frac{\partial f}{\partial x^i}(p)dx^i|_p,$$

or as an equation between covector *fields* instead of covectors:

$$df = \frac{\partial f}{\partial x^i}dx^i. \tag{6.10}$$

In particular, in the 1-dimensional case, this reduces to

$$df = \frac{df}{dx}dx.$$

Thus we have recovered the familiar classical expression for the differential of a function f in coordinates. Henceforth, we will abandon the notation λ^i for the coordinate coframe, and use dx^i instead.

Example 6.8. If $f(x,y) = x^2 y \cos x$ on $\mathbb{R}^2$, then df is given by the formula

$$\begin{aligned} df &= \frac{\partial(x^2 y\cos x)}{\partial x}dx + \frac{\partial(x^2 y \cos x)}{\partial y}dy \\ &= (2xy\cos x - x^2 y \sin x)dx + x^2\cos x\, dy. \end{aligned}$$

Proposition 6.9 (Properties of the Differential). *Let M be a smooth manifold, and let $f, g \in C^\infty(M)$.*

(*a*) *For any constants a, b, $d(af + bg) = a\,df + b\,dg$.*

(*b*) *$d(fg) = f\,dg + g\,df$.*

(*c*) *$d(f/g) = (g\,df - f\,dg)/g^2$ on the set where $g \neq 0$.*

(*d*) *If $J \subset \mathbb{R}$ is an interval containing the image of f, and $h\colon J \to \mathbb{R}$ is a smooth function, then $d(h \circ f) = (h' \circ f)\,df$.*

(*e*) *If f is constant, then $df = 0$.*

◇ **Exercise 6.7.** Prove Proposition 6.9.

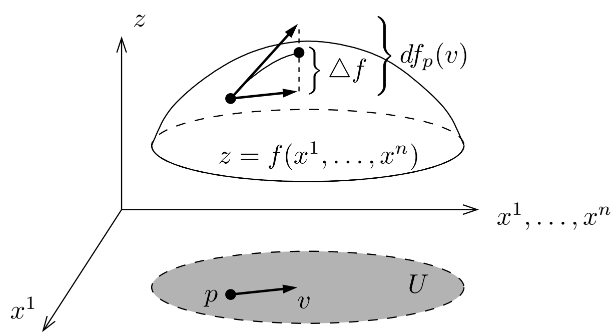


Figure 6.2. The differential as an approximation to $\triangle f$.

One very important property of the differential is the following characterization of smooth functions with vanishing differentials.

Proposition 6.10 (Functions with Vanishing Differentials). *If f is a smooth real-valued function on a smooth manifold M, then $df = 0$ if and only if f is constant on each component of M.*

Proof. It suffices to assume that M is connected and show that $df = 0$ if and only if f is constant. One direction is immediate: If f is constant, then $df = 0$ by Proposition 6.9(e). Conversely, suppose $df = 0$, let $p \in M$, and let $\mathcal{C} = \{q \in M : f(q) = f(p)\}$. If q is any point in $\mathcal{C}$, let U be a smooth coordinate ball centered at q. From (6.10) we see that $\partial f / \partial x^i \equiv 0$ in U for each i, so by elementary calculus f is constant on U. This shows that $\mathcal{C}$ is open, and since it is closed by continuity it must be all of M. Thus f is everywhere equal to the constant $f(p)$. $\square$

In elementary calculus, one thinks of df as an approximation for the small change in the value of f caused by small changes in the independent variables x^i. In our present context, df has the same meaning, provided we interpret everything appropriately. Suppose M is a smooth manifold and $f \in C^\infty(M)$, and let p be a point in M. By choosing smooth coordinates on a neighborhood of p, we can think of f as a function on an open subset $U \subset \mathbb{R}^n$. Recall that $dx^i|_p$ is the linear functional that picks out the ith component of a tangent vector at p. Writing $\Delta f = f(p+v) - f(p)$ for $v \in \mathbb{R}^n$, Taylor's theorem shows that Δf is well approximated when v is small by

$$\Delta f \approx \frac{\partial f}{\partial x^i}(p) v^i = \frac{\partial f}{\partial x^i}(p) dx^i|_p(v) = df_p(v)$$

(where now we are considering v as an element of $T_p\mathbb{R}^n$ via our usual identification $T_p\mathbb{R}^n \leftrightarrow \mathbb{R}^n$). In other words, df_p is the linear functional that best approximates Δf near p (Figure 6.2). The great power of the

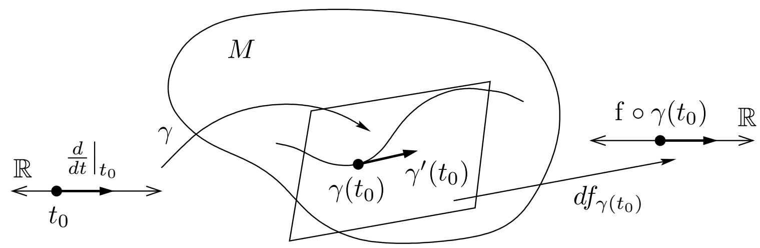


Figure 6.3. Derivative of a function along a curve.

concept of the differential comes from the fact that we can define df invariantly on any manifold, without resorting to any vague arguments involving infinitesimals.

The next result is an analogue of Proposition 3.12 for the differential.

Proposition 6.11 (Derivative of a Function Along a Curve). *Suppose M is a smooth manifold, $\gamma\colon J \to M$ is a smooth curve, and $f\colon M \to \mathbb{R}$ is a smooth function. Then the derivative of the real-valued function $f \circ \gamma\colon \mathbb{R} \to \mathbb{R}$ is given by*

$$(f \circ \gamma)'(t) = df_{\gamma(t)}(\gamma'(t)). \tag{6.11}$$

Proof. See Figure 6.3. Directly from the definitions, for any $t_0 \in J$,

$$\begin{aligned}
df_{\gamma(t_0)}(\gamma'(t_0)) &= \gamma'(t_0) f && \text{(definition of } df\text{)}\\
&= \left(\gamma_* \frac{d}{dt}\bigg|_{t_0}\right) f && \text{(definition of } \gamma'(t)\text{)}\\
&= \frac{d}{dt}\bigg|_{t_0} (f \circ \gamma) && \text{(definition of } \gamma_*\text{)}\\
&= (f \circ \gamma)'(t_0) && \text{(definition of } d/dt|_{t_0}\text{).} \qquad \square
\end{aligned}$$

It is important to observe that for a smooth real-valued function $f\colon M \to \mathbb{R}$, we have now defined two different kinds of derivative of f at a point $p \in M$. In Chapter 3, we defined the push-forward f_* as a linear map from T_pM to $T_{f(p)}\mathbb{R}$. In this chapter, we defined the differential df_p as a covector at p, which is to say a linear map from T_pM to $\mathbb{R}$. These are really the same object, once we take into account the canonical identification between $\mathbb{R}$ and its tangent space at any point; one easy way to see this is to note that both are represented in coordinates by the row matrix whose components are the partial derivatives of f.

Similarly, if γ is a smooth curve in M, we have two different meanings for the expression $(f \circ \gamma)'(t)$. On the one hand, $f \circ \gamma$ can be interpreted as a smooth curve in $\mathbb{R}$, and thus $(f \circ \gamma)'(t)$ is its tangent vector at the point $f \circ \gamma(t)$, an element of the tangent space $T_{f\circ\gamma(t)}\mathbb{R}$. Proposition 3.12 shows that this tangent vector is equal to $f_*(\gamma'(t))$. On the other hand, $f \circ \gamma$ can also be considered simply as a real-valued function of one real variable, and then $(f \circ \gamma)'(t)$ is just its ordinary derivative. Proposition 6.11 shows that this derivative is equal to the real number $df_{\gamma(t)}(\gamma'(t))$. Which of these interpretations we choose will depend on the purpose we have in mind.

Pullbacks

As we have seen, a smooth map yields a linear map on tangent vectors called the push-forward. Dualizing this leads to a linear map on covectors going in the opposite direction.

Let $F\colon M \to N$ be a smooth map, and let $p \in M$ be arbitrary. The push-forward map

$$F_*\colon T_pM \to T_{F(p)}N$$

yields a dual linear map

$$(F_*)^*\colon T^*_{F(p)}N \to T^*_pM.$$

To avoid a proliferation of stars, we write this map, called the *pullback* associated with F, as

$$F^*\colon T^*_{F(p)}N \to T^*_pM.$$

Unraveling the definitions, F^* is characterized by

$$(F^*\omega)(X) = \omega(F_*X), \qquad \text{for } \omega \in T^*_{F(p)}N,\ X \in T_pM.$$

When we introduced the push-forward map, we made a point of noting that vector *fields* do not push forward to vector fields, except in the special case of a diffeomorphism. The surprising thing about pullbacks is that smooth covector fields always pull back to smooth covector fields. Given a smooth map $G\colon M \to N$ and a smooth covector field ω on N, define a covector field $G^*\omega$ on M by

$$(G^*\omega)_p = G^*(\omega_{G(p)}). \tag{6.12}$$

Observe that there is no ambiguity here about what point to pull back from, in contrast to the vector field case. We will prove in Proposition 6.13 below that $G^*\omega$ is smooth. Before doing so, let us examine two important special cases.

Lemma 6.12. *Let $G\colon M \to N$ be a smooth map, and suppose $f \in C^\infty(N)$ and $\omega \in \mathcal{T}^*(N)$. Then*

$$G^*df = d(f \circ G); \tag{6.13}$$
$$G^*(f\omega) = (f \circ G)G^*\omega. \tag{6.14}$$

Proof. To prove (6.13), we let $X_p \in T_pM$ be arbitrary, and compute

$$\begin{aligned}
(G^*df)_p(X_p) &= (G^*(df_{G(p)}))(X_p) && \text{(by (6.12))}\\
&= df_{G(p)}(G_*X_p) && \text{(by definition of } G^*)\\
&= (G_*X_p)f && \text{(by definition of } df)\\
&= X_p(f \circ G) && \text{(by definition of } G_*)\\
&= d(f \circ G)_p(X_p) && \text{(by definition of } d(f \circ G)).
\end{aligned}$$

Similarly, for (6.14), we compute

$$\begin{aligned}
(G^*(f\omega))_p &= G^*((f\omega)_{G(p)}) && \text{(by (6.12))}\\
&= G^*(f(G(p))\omega_{G(p)}) && \text{(by (6.8))}\\
&= f(G(p))G^*(\omega_{G(p)}) && \text{(by linearity of } G^*)\\
&= f(G(p))(G^*\omega)_p && \text{(by (6.12))}\\
&= ((f \circ G)G^*\omega)_p && \text{(by (6.8)).}
\end{aligned}$$

□

Proposition 6.13. *Suppose $G\colon M \to N$ is smooth, and let ω be a smooth covector field on N. Then $G^*\omega$ is a smooth covector field on M.*

Proof. Let $p \in M$ be arbitrary, and choose smooth coordinates (x^i) for M near p and (y^j) for N near $G(p)$. Writing ω in coordinates as $\omega = \omega_j dy^j$ for smooth functions ω_j defined near $G(p)$ and using Lemma 6.12 twice, we have the following computation in a neighborhood of p:

$$\begin{aligned}
G^*\omega &= G^*(\omega_j dy^j)\\
&= (\omega_j \circ G)G^*dy^j\\
&= (\omega_j \circ G)d(y^j \circ G).
\end{aligned}$$

Because this expression is smooth, it follows that $G^*\omega$ is smooth. □

In the course of the preceding proof, we derived the following formula for the pullback of a covector field with respect to smooth coordinates (x^i) on the domain and (y^j) on the range:

$$G^*\omega = G^*(\omega_j dy^j) = (\omega_j \circ G)d(y^j \circ G) = (\omega_j \circ G)dG^j, \tag{6.15}$$

where G^j is the jth component function of G in these coordinates. This formula makes the computation of pullbacks in coordinates exceedingly simple, as the next example shows.

Example 6.14. Let $G\colon \mathbb{R}^3 \to \mathbb{R}^2$ be the map given by

$$(u, v) = G(x, y, z) = (x^2y, y \sin z),$$

and let $\omega \in \mathcal{T}^*(\mathbb{R}^2)$ be the covector field

$$\omega = u\,dv + v\,du.$$

According to (6.15), the pullback $G^*\omega$ is given by

$$\begin{aligned} G^*\omega &= (u\circ G)d(v\circ G) + (v\circ G)d(u\circ G) \\ &= (x^2y)d(y\sin z) + (y\sin z)d(x^2y) \\ &= x^2y(\sin z\,dy + y\cos z\,dz) + y\sin z(2xy\,dx + x^2\,dy) \\ &= 2xy^2\sin z\,dx + 2x^2y\sin z\,dy + x^2y^2\cos z\,dz. \end{aligned}$$

In other words, to compute $G^*\omega$, all you need to do is substitute the component functions of G for the coordinate functions of N everywhere they appear in ω!

This also yields an easy way to remember the transformation law for a covector field under a change of coordinates. Again, an example will convey the idea better than a general formula.

Example 6.15. Let (r,θ) be polar coordinates on, say, the right half-plane $H = \{(x,y) : x > 0\}$. We can think of the change of coordinates $(x,y) = (r\cos\theta, r\sin\theta)$ as the coordinate expression for the identity map of H, but using (r,θ) as coordinates for the domain and (x,y) for the range. Then the pullback formula (6.15) tells us that we can compute the polar coordinate expression for a covector field simply by substituting $x = r\cos\theta$, $y = r\sin\theta$ and expanding. For example,

$$\begin{aligned} x\,dy - y\,dx &= \mathrm{Id}^*(x\,dy - y\,dx) \\ &= (r\cos\theta)d(r\sin\theta) - (r\sin\theta)d(r\cos\theta) \\ &= (r\cos\theta)(\sin\theta\,dr + r\cos\theta\,d\theta) - (r\sin\theta)(\cos\theta\,dr - r\sin\theta\,d\theta) \\ &= (r\cos\theta\sin\theta - r\sin\theta\cos\theta)dr + (r^2\cos^2\theta + r^2\sin^2\theta)d\theta \\ &= r^2\,dr. \end{aligned}$$

Line Integrals

Another important application of covector fields is to make coordinate-independent sense of the notion of a line integral.

We begin with the simplest case: an interval in the real line. Suppose $[a,b]\subset\mathbb{R}$ is a compact interval, and ω is a smooth covector field on $[a,b]$. (This means that the component function of ω admits a smooth extension to some neighborhood of $[a,b]$.) If we let t denote the standard coordinate on $\mathbb{R}$, ω can be written $\omega_t = f(t)\,dt$ for some smooth function $f\colon [a,b]\to\mathbb{R}$. The similarity between this and the usual notation $\int f(t)\,dt$ for an integral suggests that there might be a connection between covector fields and integrals, and indeed there is. We define the *integral of ω over $[a,b]$* to be

$$\int_{[a,b]} \omega = \int_a^b f(t)\,dt.$$

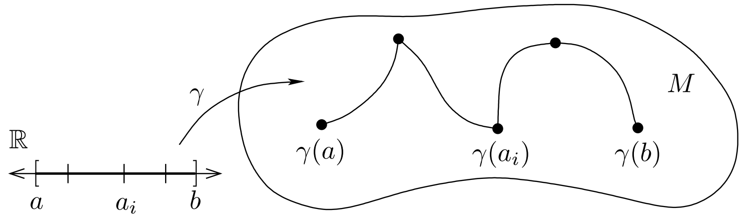


Figure 6.4. A piecewise smooth curve segment.

The next proposition should convince you that this is more than just a trick of notation.

Proposition 6.16 (Diffeomorphism Invariance of the Integral). *Let ω be a smooth covector field on the compact interval $[a,b] \subset \mathbb{R}$. If $\varphi\colon [c,d] \to [a,b]$ is an increasing diffeomorphism (meaning that $t_1 < t_2$ implies $\varphi(t_1) < \varphi(t_2)$), then*

$$\int_{[c,d]} \varphi^*\omega = \int_{[a,b]} \omega.$$

Proof. If we let s denote the standard coordinate on $[c,d]$ and t that on $[a,b]$, then (6.15) shows that the pullback $\varphi^*\omega$ has the coordinate expression $(\varphi^*\omega)_s = f(\varphi(s))\varphi'(s)\,ds$. Inserting this into the definition of the line integral and using the change of variables formula for ordinary integrals, we obtain

$$\int_{[c,d]} \varphi^*\omega = \int_c^d f(\varphi(s))\varphi'(s)\,ds = \int_a^b f(t)\,dt = \int_{[a,b]} \omega. \qquad \square$$

◇ **Exercise 6.8.** If $\varphi\colon [c,d] \to [a,b]$ is a decreasing diffeomorphism, show that $\int_{[c,d]} \varphi^*\omega = -\int_{[a,b]} \omega$.

Now let M be a smooth manifold. By a *curve segment* in M we mean a continuous curve $\gamma\colon [a,b] \to M$ whose domain is a compact interval. It is a *smooth curve segment* if it is has a smooth extension to an open interval containing $[a,b]$. A *piecewise smooth curve segment* is a curve segment $\gamma\colon [a,b] \to M$ with the property that there exists a finite partition $a = a_0 < a_1 < \cdots < a_k = b$ of $[a,b]$ such that $\gamma|_{[a_{i-1},a_i]}$ is smooth for each i (Figure 6.4). Continuity of γ means that $\gamma(t)$ approaches the same value as t approaches any of the points a_i (other than a_0 or a_k) from the left or the right. Smoothness of γ on each subinterval means that γ has one-sided tangent vectors at each such a_i when approaching from the left or the right, but these one-sided tangent vectors need not be equal.

Lemma 6.17. *If M is a connected smooth manifold, any two points of M can be joined by a piecewise smooth curve segment.*

Proof. Let p be an arbitrary point of M, and define a subset $\mathcal{C} \subset M$ by $\mathcal{C} = \{q \in M :$ there is a piecewise smooth curve in M from p to $q\}$. Clearly $p \in \mathcal{C}$, so $\mathcal{C}$ is nonempty. To show $\mathcal{C} = M$, we need to show it is open and closed.

Let $q \in \mathcal{C}$ be arbitrary, which means that there is a piecewise smooth curve segment γ going from p to q. Let U be a smooth coordinate ball centered at q. If q' is any point in U, then it is easy to construct a piecewise smooth curve segment from p to q' by first following γ from p to q, and then following a straight-line path in coordinates from q to q'. Thus $U \subset \mathcal{C}$, which shows that $\mathcal{C}$ is open. On the other hand, if $q \in \partial\mathcal{C}$, let U be a smooth coordinate ball around q as above. The fact that q is a boundary point of $\mathcal{C}$ means that there is some point $q' \in \mathcal{C} \cap U$. In this case, we can construct a piecewise smooth curve from p to q by first following one from p to q' and then following a straight-line path in coordinates from q' to q. This shows that $q \in \mathcal{C}$, so $\mathcal{C}$ is also closed. □

If $\gamma\colon [a,b] \to M$ is a smooth curve segment and ω is a smooth covector field on M, we define the *line integral* of ω over γ to be the real number

$$\int_\gamma \omega = \int_{[a,b]} \gamma^*\omega.$$

Because $\gamma^*\omega$ is a smooth covector field on $[a,b]$, this definition makes sense. More generally, if γ is piecewise smooth, we define

$$\int_\gamma \omega = \sum_{i=1}^{k} \int_{[a_{i-1},a_i]} \gamma^*\omega,$$

where $[a_{i-1},a_i]$, $i = 1,\dots,k$, are the intervals on which γ is smooth. This definition gives a rigorous meaning to classical line integrals such as $\int_\gamma P\,dx + Q\,dy$ in the plane or $\int_\gamma P\,dx + Q\,dy + R\,dz$ in $\mathbb{R}^3$.

Proposition 6.18 (Properties of Line Integrals). *Let M be a smooth manifold. Suppose $\gamma\colon [a,b] \to M$ is a piecewise smooth curve segment and $\omega, \omega_1, \omega_2 \in \mathcal{T}^*(M)$.*

(*a*) *For any $c_1, c_2 \in \mathbb{R}$,*

$$\int_\gamma (c_1\omega_1 + c_2\omega_2) = c_1 \int_\gamma \omega_1 + c_2 \int_\gamma \omega_2.$$

(*b*) *If γ is a constant map, then $\int_\gamma \omega = 0$.*

(*c*) *If $a < c < b$, then*

$$\int_\gamma \omega = \int_{\gamma_1} \omega + \int_{\gamma_2} \omega,$$

where $\gamma_1 = \gamma|_{[a,c]}$ and $\gamma_2 = \gamma|_{[c,b]}$.

◇ **Exercise 6.9.** Prove Proposition 6.18.

The next lemma gives a useful alternative expression for the line integral of a covector field.

Proposition 6.19. *If $\gamma\colon [a,b] \to M$ is a piecewise smooth curve segment, the line integral of ω over γ can also be expressed as the ordinary integral*

$$\int_\gamma \omega = \int_a^b \omega_{\gamma(t)}(\gamma'(t))\,dt. \tag{6.16}$$

Proof. First suppose that γ is smooth and that its image is contained in the domain of a single smooth chart. Writing the coordinate representations of γ and ω as $(\gamma^1(t), \dots, \gamma^n(t))$ and $\omega_i\,dx^i$, respectively, we have

$$\omega_{\gamma(t)}(\gamma'(t)) = \omega_i(\gamma(t))dx^i(\gamma'(t)) = \omega_i(\gamma(t))(\gamma^i)'(t).$$

Combining this with the coordinate formula (6.15) for the pullback, we obtain

$$\begin{aligned}(\gamma^*\omega)_t &= (\omega_i \circ \gamma)(t)\,d(\gamma^i)_t\\ &= \omega_i(\gamma(t))\,(\gamma^i)'(t)\,dt\\ &= \omega_{\gamma(t)}(\gamma'(t))\,dt.\end{aligned}$$

Therefore, by definition of the line integral,

$$\begin{aligned}\int_\gamma \omega &= \int_{[a,b]} \gamma^*\omega\\ &= \int_a^b \omega_{\gamma(t)}(\gamma'(t))\,dt.\end{aligned}$$

Now, if γ is an arbitrary smooth curve segment, by compactness there exists a finite partition $a = a_0 < a_1 < \cdots < a_k = b$ of the interval $[a,b]$ such that $\gamma[a_{i-1}, a_i]$ is contained in the domain of a single smooth chart for each $i = 1, \dots, k$, so we can apply the computation above on each such subinterval. Finally, if γ is only piecewise smooth, we simply apply the same argument on each subinterval on which γ is smooth. □

Example 6.20. Let $M = \mathbb{R}^2 \smallsetminus \{0\}$, let ω be the covector field on M given by

$$\omega = \frac{x\,dy - y\,dx}{x^2 + y^2},$$

and let $\gamma\colon [0, 2\pi] \to M$ be the curve segment defined by

$$\gamma(t) = (\cos t, \sin t).$$

Since $\gamma^*\omega$ can be computed by substituting $x = \cos t$ and $y = \sin t$ everywhere in the formula for ω, we find that

$$\int_\gamma \omega = \int_{[0,2\pi]} \frac{\cos t(\cos t\,dt) - \sin t(-\sin t\,dt)}{\sin^2 t + \cos^2 t} = \int_0^{2\pi} dt = 2\pi.$$

One of the most significant features of line integrals is that they are independent of parametrization, in a sense we now make precise. If $\gamma\colon [a,b] \to M$ and $\widetilde{\gamma}\colon [c,d] \to M$ are smooth curve segments, we say that $\widetilde{\gamma}$ is a *reparametrization* of γ if $\widetilde{\gamma} = \gamma \circ \varphi$ for some diffeomorphism $\varphi\colon [c,d] \to [a,b]$. If φ is an increasing function, we say $\widetilde{\gamma}$ is a *forward reparametrization*, and if φ is decreasing, it is a *backward reparametrization*. (More generally, one can allow φ to be piecewise smooth, but we will have no need for this extra generality.)

Proposition 6.21 (Parameter Independence of Line Integrals). *Suppose M is a smooth manifold, ω is a smooth covector field on M, and γ is a piecewise smooth curve segment in M. For any reparametrization $\widetilde{\gamma}$ of γ, we have*

$$\int_{\widetilde{\gamma}} \omega = \begin{cases} \displaystyle\int_\gamma \omega & \text{if } \widetilde{\gamma} \text{ is a forward reparametrization,} \\ \displaystyle -\int_\gamma \omega & \text{if } \widetilde{\gamma} \text{ is a backward reparametrization.} \end{cases}$$

Proof. First assume that $\gamma\colon [a,b] \to M$ is smooth, and suppose $\varphi\colon [c,d] \to [a,b]$ is an increasing diffeomorphism. Then Proposition 6.16 implies

$$\begin{aligned} \int_{\widetilde{\gamma}} \omega &= \int_{[c,d]} (\gamma \circ \varphi)^*\omega \\ &= \int_{[c,d]} \varphi^*\gamma^*\omega \\ &= \int_{[a,b]} \gamma^*\omega \\ &= \int_\gamma \omega. \end{aligned}$$

When φ is decreasing, the analogous result follows from Exercise 6.8. If γ is only piecewise smooth, the result follows simply by applying the preceding argument on each subinterval where γ is smooth. □

There is one special case in which a line integral is trivial to compute: the line integral of a differential.

Theorem 6.22 (Fundamental Theorem for Line Integrals). *Let M be a smooth manifold. Suppose f is a smooth real-valued function on M*

and $\gamma\colon [a,b] \to M$ *is a piecewise smooth curve segment in* M*. Then*

$$\int_\gamma df = f(\gamma(b)) - f(\gamma(a)).$$

Proof. Suppose first that γ is smooth. By Propositions 6.11 and 6.19,

$$\int_\gamma df = \int_a^b df_{\gamma(t)}(\gamma'(t))\,dt = \int_a^b (f\circ\gamma)'(t)\,dt.$$

By the one-variable version of the fundamental theorem of calculus, this is equal to $f\circ\gamma(b) - f\circ\gamma(a)$.

If γ is merely piecewise smooth, let $a = a_0 < \cdots < a_k = b$ be the endpoints of the subintervals on which γ is smooth. Applying the above argument on each subinterval and summing, we find that

$$\int_\gamma df = \sum_{i=1}^k \bigl(f(\gamma(a_i)) - f(\gamma(a_{i-1}))\bigr) = f(\gamma(b)) - f(\gamma(a)),$$

because the contributions from all the interior points cancel. □

Conservative Covector Fields

Theorem 6.22 shows that the line integral of any covector field ω that can be written as the differential of a smooth function can be computed extremely easily once the smooth function is known. For this reason, there is a special term for covector fields with this property. We say a smooth covector field ω on a manifold M is *exact* (or an *exact differential*) on M if there is a function $f \in C^\infty(M)$ such that $\omega = df$. In this case, the function f is called a *potential* for ω. The potential is not uniquely determined, but by Lemma 6.10, the difference between any two potentials for ω must be constant on each component of M.

Because exact differentials are so easy to integrate, it is important to develop criteria for deciding whether a covector field is exact. Theorem 6.22 provides an important clue. It shows that the line integral of an exact covector field depends only on the endpoints $p = \gamma(a)$ and $q = \gamma(b)$: Any other curve segment from p to q would give the same value for the line integral. In particular, if γ is a *closed* curve segment, meaning that $\gamma(a) = \gamma(b)$, then the integral of df over γ is zero.

We say a smooth covector field ω is *conservative* if the line integral of ω over any closed piecewise smooth curve segment is zero. This terminology comes from physics, where a force field is called conservative if the change in energy caused by the force acting along any closed path is zero ("energy is conserved"). (In elementary physics, force fields are usually thought of as vector fields rather than covector fields; see Problem 6-12 for the connection.)

The following lemma gives a useful alternative characterization of conservative covector fields.

Lemma 6.23. *A smooth covector field ω is conservative if and only if the line integral of ω depends only on the endpoints of the curve, i.e., $\int_\gamma \omega = \int_{\widetilde{\gamma}} \omega$ whenever γ and $\widetilde{\gamma}$ are piecewise smooth curve segments with the same starting and ending points.*

◇ **Exercise 6.10.** Prove Lemma 6.23. [Observe that this would be much harder to prove if we defined conservative fields in terms of smooth curves instead of piecewise smooth ones.]

Theorem 6.24. *A smooth covector field is conservative if and only if it is exact.*

Proof. If $\omega \in \mathcal{T}^*(M)$ is exact, Theorem 6.22 shows that it is conservative, so we need only prove the converse. Suppose therefore that ω is conservative, and assume for the moment that M is connected. Because the line integrals of ω are path independent, we can adopt the following notation: For any points $p, q \in M$, we will use the notation $\int_p^q \omega$ to denote the value of any line integral of the form $\int_\gamma \omega$, where γ is a piecewise smooth curve segment from p to q. Observe that Proposition 6.18(c) implies that

$$\int_{p_1}^{p_2} \omega + \int_{p_2}^{p_3} \omega = \int_{p_1}^{p_3} \omega \tag{6.17}$$

for any three points $p_1, p_2, p_3 \in M$.

Now choose any base point $p_0 \in M$, and define a function $f\colon M \to \mathbb{R}$ by

$$f(q) = \int_{p_0}^{q} \omega.$$

We will show that $df = \omega$. To accomplish this, let $q_0 \in M$ be an arbitrary point, let $(U, (x^i))$ be a smooth chart centered at q_0, and write the coordinate representation of ω in U as $\omega = \omega_i dx^i$. We will show that

$$\frac{\partial f}{\partial x^j}(q_0) = \omega_j(q_0)$$

for $j = 1, \dots, n$, which implies that $df_{q_0} = \omega_{q_0}$.

Fix j, and let $\gamma\colon [-\varepsilon, \varepsilon] \to U$ be the smooth curve segment defined in coordinates by $\gamma(t) = (0, \dots, t, \dots, 0)$, with t in the jth place, and with ε chosen small enough that $\gamma[-\varepsilon, \varepsilon] \subset U$ (Figure 6.5). Let $p_1 = \gamma(-\varepsilon)$, and define a new function $\widetilde{f}\colon M \to \mathbb{R}$ by $\widetilde{f}(q) = \int_{p_1}^{q} \omega$. Note that (6.17) implies

$$f(q) - \widetilde{f}(q) = \int_{p_0}^{q} \omega - \int_{p_1}^{q} \omega = \int_{p_0}^{p_1} \omega,$$

which does not depend on q. Thus $\widetilde{f}$ and f differ by a constant, so it suffices to show that $\partial \widetilde{f}/\partial x^j(q_0) = \omega_j(q_0)$.

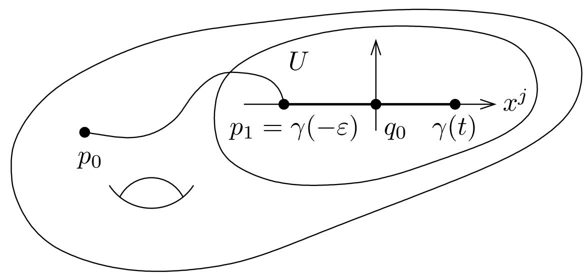


Figure 6.5. Proof that a conservative covector field is exact.

Now $\gamma'(t) = \partial/\partial x^j|_{\gamma(t)}$ by construction, so

$$\omega_{\gamma(t)}(\gamma'(t)) = \omega_i(\gamma(t))dx^i\left(\left.\frac{\partial}{\partial x^j}\right|_{\gamma(t)}\right) = \omega_j(\gamma(t)).$$

Since the restriction of γ to $[-\varepsilon, t]$ is a smooth curve from p_1 to $\gamma(t)$, we have

$$\begin{aligned}\widetilde{f}\circ\gamma(t) &= \int_{p_1}^{\gamma(t)} \omega \\ &= \int_{-\varepsilon}^{t} \omega_{\gamma(s)}(\gamma'(s))\, ds \\ &= \int_{-\varepsilon}^{t} \omega_j(\gamma(s))\, ds.\end{aligned}$$

Thus by the fundamental theorem of calculus,

$$\begin{aligned}\frac{\partial\widetilde{f}}{\partial x^j}(q_0) &= \gamma'(0)\widetilde{f} \\ &= \left.\frac{d}{dt}\right|_{t=0} \widetilde{f}\circ\gamma(t) \\ &= \left.\frac{d}{dt}\right|_{t=0} \int_{-\varepsilon}^{t} \omega_j(\gamma(s))\, ds \\ &= \omega_j(\gamma(0)) = \omega_j(q_0).\end{aligned}$$

This completes the proof that $df = \omega$.

Finally, if M is not connected, let $\{M_i\}$ be the components of M. The argument above shows that for each i there is a smooth function $f_i \in C^\infty(M_i)$ such that $df_i = \omega$ on M_i. Letting $f\colon M \to \mathbb{R}$ be the function that is equal to f_i on M_i, we have $df = \omega$, thus completing the proof. □

It would be nice if every smooth covector field were exact, for then the evaluation of any line integral would just be a matter of finding a potential function and evaluating it at the endpoints, a process analogous to evaluating an ordinary integral by finding an indefinite integral or primitive. However, this is too much to hope for.

Example 6.25. The covector field ω of Example 6.20 cannot be exact on $\mathbb{R}^2 \smallsetminus \{0\}$, because it is not conservative: The computation in that example showed that $\int_\gamma \omega = 2\pi \neq 0$, where γ is the unit circle traversed counterclockwise.

Because exactness has such important consequences for the evaluation of line integrals, we would like to have an easy way to check whether a given covector field is exact. Fortunately, there is a very simple necessary condition, which follows from the fact that partial derivatives of smooth functions can be taken in any order.

To see what this condition is, suppose that ω is exact. Let f be any potential function for ω, and let $(U, (x^i))$ be any smooth chart on M. Because f is smooth, it satisfies the following identity on U:

$$\frac{\partial^2 f}{\partial x^i \partial x^j} = \frac{\partial^2 f}{\partial x^j \partial x^i}. \tag{6.18}$$

Writing $\omega = \omega_i dx^i$ in coordinates, the fact that $\omega = df$ is equivalent to $\omega_i = \partial f / \partial x^i$. Substituting this into (6.18), we find that the component functions of ω satisfy

$$\frac{\partial \omega_j}{\partial x^i} = \frac{\partial \omega_i}{\partial x^j}. \tag{6.19}$$

We say that a smooth covector field ω is *closed* if its components in every smooth chart satisfy (6.19). The following lemma summarizes the computation above.

Lemma 6.26. *Every exact covector field is closed.*

The significance of this result is that the property of being closed is one that can be easily checked. First we need the following result, which says that it is not necessary to check the closedness condition in *every* smooth chart, just in a collection of smooth charts that cover the manifold. (The proof of this lemma is a tedious computation. After the proof of Proposition 6.30 below, we will be able to give a more conceptual proof, so you are free to skip this proof if you wish. See also Chapter 12, in which we will prove a much more general result.)

Proposition 6.27. *Let ω be a smooth covector field. If ω satisfies* (6.19) *in some smooth chart around every point, then it is closed.*

Proof. Let $(U, (x^i))$ be an arbitrary smooth chart. For each $p \in U$, the hypothesis guarantees that there are some smooth coordinates $(\widetilde{x}^j)$ defined near p in which the analogue of (6.19) holds. Using formula (6.7) for the transformation of the components of ω together with the chain rule, we find

$$\begin{aligned}
\frac{\partial \omega_i}{\partial x^j} - \frac{\partial \omega_j}{\partial x^i} &= \frac{\partial}{\partial x^j}\left(\frac{\partial \widetilde{x}^k}{\partial x^i}\widetilde{\omega}_k\right) - \frac{\partial}{\partial x^i}\left(\frac{\partial \widetilde{x}^k}{\partial x^j}\widetilde{\omega}_k\right) \\
&= \left(\frac{\partial^2 \widetilde{x}^k}{\partial x^j \partial x^i}\widetilde{\omega}_k + \frac{\partial \widetilde{x}^k}{\partial x^i}\frac{\partial \widetilde{\omega}_k}{\partial x^j}\right) - \left(\frac{\partial^2 \widetilde{x}^k}{\partial x^i \partial x^j}\widetilde{\omega}_k + \frac{\partial \widetilde{x}^k}{\partial x^j}\frac{\partial \widetilde{\omega}_k}{\partial x^i}\right) \\
&= \frac{\partial^2 \widetilde{x}^k}{\partial x^j \partial x^i}\widetilde{\omega}_k + \frac{\partial \widetilde{x}^k}{\partial x^i}\frac{\partial \widetilde{x}^l}{\partial x^j}\frac{\partial \widetilde{\omega}_k}{\partial \widetilde{x}^l} - \frac{\partial^2 \widetilde{x}^k}{\partial x^i \partial x^j}\widetilde{\omega}_k - \frac{\partial \widetilde{x}^k}{\partial x^j}\frac{\partial \widetilde{x}^l}{\partial x^i}\frac{\partial \widetilde{\omega}_k}{\partial \widetilde{x}^l} \\
&= \left(\frac{\partial^2 \widetilde{x}^k}{\partial x^j \partial x^i} - \frac{\partial^2 \widetilde{x}^k}{\partial x^i \partial x^j}\right)\widetilde{\omega}_k + \frac{\partial \widetilde{x}^k}{\partial x^i}\frac{\partial \widetilde{x}^l}{\partial x^j}\left(\frac{\partial \widetilde{\omega}_k}{\partial \widetilde{x}^l} - \frac{\partial \widetilde{\omega}_l}{\partial \widetilde{x}^k}\right) \\
&= 0 + 0,
\end{aligned}$$

where the fourth equation follows from the third by interchanging the roles of k and l in the last term. □

Corollary 6.28. *If $G\colon M \to N$ is a local diffeomorphism, then the pullback $G^*\colon \mathcal{T}^*(N) \to \mathcal{T}^*(M)$ takes closed covector fields to closed covector fields, and exact ones to exact ones.*

Proof. The result for exact covector fields follows immediately from (6.13). For closed covector fields, if (U, φ) is any smooth chart for N, then $\varphi \circ G$ is a smooth chart for M in a neighborhood of each point of $G^{-1}(U)$. In these coordinates, the coordinate representation of G is the identity, so if ω satisfies (6.19) in U, then $G^*\omega$ satisfies (6.19) in $G^{-1}(U)$. □

Consider the following covector field on $\mathbb{R}^2$:

$$\omega = y \cos xy \, dx + x \cos xy \, dy.$$

It is easy to check that

$$\frac{\partial(y \cos xy)}{\partial y} = \frac{\partial(x \cos xy)}{\partial x} = \cos xy - xy \sin xy,$$

so ω is closed. In fact, you might guess that $\omega = d(\sin xy)$. On the other hand, the covector field

$$\omega = x \cos xy \, dx + y \cos xy \, dy$$

is not closed, because

$$\frac{\partial(x \cos xy)}{\partial y} = -x^2 \sin xy, \qquad \frac{\partial(y \cos xy)}{\partial x} = -y^2 \sin xy.$$

Thus ω is not exact.

The question then naturally arises whether the converse of Lemma 6.26 is true: Is every closed covector field exact? The answer is *almost* yes, but there is an important restriction. It turns out that the answer to the question depends in a subtle way on the shape of the domain, as the next example illustrates.

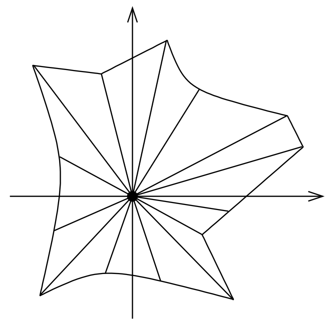


Figure 6.6. A star-shaped subset of $\mathbb{R}^2$.

Example 6.29. Look once again at the covector field ω of Example 6.20. A straightforward computation shows that ω is closed; but as we observed above, it is not exact on $\mathbb{R}^2 \smallsetminus \{0\}$. On the other hand, if we restrict the domain to the right half-plane $U = \{(x, y) : x > 0\}$, a computation shows that $\omega = d(\tan^{-1} y/x)$ there. This can be seen more clearly in polar coordinates, where $\omega = d\theta$. The problem, of course, is that there is no smooth (or even continuous) angle function on all of $\mathbb{R}^2 \smallsetminus \{0\}$, which is a consequence of the "hole" in the center.

This last example illustrates a key fact: The question of whether a particular closed covector field is exact is a global one, depending on the shape of the domain in question. This observation is the starting point for *de Rham cohomology*, which expresses a deep relationship between smooth structures and topology. We will pursue this relationship in more depth in Chapter 15, but for now we can prove the following result. A subset $V \subset \mathbb{R}^n$ is said to be *star-shaped* with respect to a point $c \in V$ if for every $x \in V$, the line segment from c to x is entirely contained in V (Figure 6.6). For example, a convex subset is star-shaped with respect to each of its points.

Proposition 6.30. *If U is a star-shaped open subset of $\mathbb{R}^n$, then every closed covector field on U is exact.*

Proof. Suppose U is star-shaped with respect to $c \in U$, and let $\omega = \omega_i dx^i$ be a closed covector field on U. As in the proof of Theorem 6.24, we will construct a potential function for ω by integrating along smooth curve segments from c. However, in this case we do not know a priori that the line integrals are path-independent, so we must integrate along specific paths.

For any point $x \in U$, let $\gamma_x \colon [0,1] \to U$ denote the line segment from c to x, parametrized as follows:

$$\gamma_x(t) = c + t(x - c).$$

The hypothesis guarantees that the image of γ_x lies entirely in U for each $x \in U$. Define a function $f\colon U \to \mathbb{R}$ by

$$f(x) = \int_{\gamma_x} \omega. \tag{6.20}$$

We will show that f is a potential for ω, or equivalently that $\partial f/\partial x^i = \omega_i$ for $i = 1, \dots, n$. To begin, we compute

$$\begin{aligned} f(x) &= \int_0^1 \omega_{\gamma_x(t)}(\gamma_x'(t))\,dt \\ &= \int_0^1 \sum_{i=1}^n \omega_i(c + t(x-c))(x^i - c^i)\,dt. \end{aligned}$$

To compute the partial derivatives of f, we note that the integrand is smooth in all variables, so it is permissible to differentiate under the integral sign to obtain

$$\frac{\partial f}{\partial x^j}(x) = \int_0^1 \left(\sum_{i=1}^n t \frac{\partial \omega_i}{\partial x^j}(c + t(x-c))(x^i - c^i) + \omega_j(c + t(x-c)) \right) dt.$$

Because ω is closed, this reduces to

$$\begin{aligned} \frac{\partial f}{\partial x^j}(x) &= \int_0^1 \left(\sum_{i=1}^n t \frac{\partial \omega_j}{\partial x^i}(c + t(x-c))(x^i - c^i) + \omega_j(c + t(x-c)) \right) dt \\ &= \int_0^1 \frac{d}{dt}\bigl(t\omega_j(c + t(x-c))\bigr)\,dt \\ &= \Bigl[t\omega_j(c + t(x-c)) \Bigr]_{t=0}^{t=1} \\ &= \omega_j(x). \end{aligned}$$

□

Corollary 6.31 (Local Exactness of Closed Covector Fields). *Let ω be a closed covector field on a smooth manifold M. Then every $p \in M$ has a neighborhood on which ω is exact.*

Proof. Let $p \in M$ be arbitrary. The hypothesis implies that ω satisfies (6.19) in some smooth coordinate ball (U, φ) containing p. Because a ball is star-shaped, we can apply Proposition 6.30 to the coordinate representation of ω and conclude that there is a function $f \in C^\infty(U)$ such that $\omega|_U = df$. □

Note that the proof of this corollary used only the fact that ω satisfies (6.19) in *some* smooth chart about each point, not in every smooth chart. This fact leads to the following simpler and more conceptual proof of the fact that closedness is coordinate-independent.

Another proof of Proposition 6.27. Let $p \in M$, and let $(U, (x^i))$ be a smooth chart containing p with respect to which ω satisfies (6.19). Shrinking U if necessary, we may assume that it is a coordinate ball, and then the proof of Corollary 6.31 shows that $\omega|_U$ is exact. Therefore, by Lemma 6.26, $\omega|_U$ is closed. Since the same is true in a neighborhood of each point of M, it follows that ω is closed. □

The key to the construction of a potential function in Proposition 6.30 is that we can reach every point $x \in M$ by a definite path γ_x from c to x, chosen in such a way that γ_x varies smoothly as x varies. That is what fails in the case of the closed covector field ω on the punctured plane (Example 6.25): Because of the hole, it is impossible to choose a smoothly-varying family of paths starting at a fixed base point and reaching every point of the domain. In Chapter 15, we will generalize Proposition 6.30 to show that every closed covector field is exact on any simply connected manifold.

When you actually have to *compute* a potential function for a given covector field that is known to be exact, there is a much simpler procedure that almost always works. Rather than describe it in complete generality, we illustrate it with an example.

Example 6.32. Let ω be a smooth covector field on $\mathbb{R}^3$, say

$$\omega = e^{y^2}dx + 2xye^{y^2}dy - 2z\,dz.$$

You can check that ω is closed. For f to be a potential for ω, we must have

$$\frac{\partial f}{\partial x} = e^{y^2}, \quad \frac{\partial f}{\partial y} = 2xye^{y^2}, \quad \frac{\partial f}{\partial z} = -2z.$$

Holding y and z fixed and integrating the first equation with respect to x, we obtain

$$f(x,y,z) = \int e^{y^2}dx = xe^{y^2} + C_1(y,z),$$

where the "constant" of integration $C_1(y,z)$ may depend on the choice of (y,z). Now the second equation implies

$$\begin{aligned} 2xye^{y^2} &= \frac{\partial}{\partial y}\big(xe^{y^2} + C_1(y,z)\big) \\ &= 2xye^{y^2} + \frac{\partial C_1}{\partial y}, \end{aligned}$$

which forces $\partial C_1/\partial y = 0$, so C_1 is actually a function of z only. Finally, the third equation implies

$$\begin{aligned} -2z &= \frac{\partial}{\partial z}\big(xe^{y^2} + C_1(z)\big) \\ &= \frac{\partial C_1}{\partial z}, \end{aligned}$$

from which we conclude that $C_1(z) = -z^2 + C$, where C is an arbitrary constant. Thus a potential function for ω is given by $f(x,y,z) = xe^{y^2} - z^2$. Any other potential differs from this one by a constant.

You should convince yourself that the formal procedure we followed in this example is equivalent to choosing an arbitrary base point $c \in \mathbb{R}^3$, and defining $f(x, y, z)$ by integrating ω along a path from c to (x, y, z) consisting of three straight line segments parallel to the axes. This works for any closed covector field defined on an open rectangle in $\mathbb{R}^n$ (which we know must be exact, because a rectangle is convex). In practice, once a formula is found for f on some open rectangle, the same formula typically works for the entire domain. (This is because most of the covector fields for which one can explicitly compute the integrals as we did above are real-analytic, and real-analytic functions are determined by their behavior in any open set.)

Problems

6-1. (a) If V and W are finite-dimensional vector spaces and $A\colon V \to W$ is any linear map, show that the following diagram commutes:

$$\begin{tikzcd} V \arrow[r, "A"] \arrow[d, "\xi_V"'] & W \arrow[d, "\xi_W"] \\ V^{**} \arrow[r, "(A^*)^*"'] & W^{**}, \end{tikzcd}$$

where ξ_V and ξ_W denote the isomorphisms defined by (6.3) for V and W, respectively.

(b) Show that there does not exist a rule that assigns to each finite-dimensional vector space V an isomorphism $\beta_V\colon V \to V^*$ in such a way that for every linear map $A\colon V \to W$, the following diagram commutes:

$$\begin{tikzcd} V \arrow[r, "A"] \arrow[d, "\beta_V"'] & W \arrow[d, "\beta_W"] \\ V^* & W^*. \arrow[l, "A^*"] \end{tikzcd}$$

6-2. (a) If $F\colon M \to N$ is a smooth map, show that $F^*\colon T^*N \to T^*M$ is a smooth bundle map.

(b) Show that the assignment $M \mapsto T^*M$, $F \mapsto F^*$ defines a contravariant functor from the category of smooth manifolds to the category of smooth vector bundles.

6-3. If M is a smooth manifold, show that T^*M is a trivial bundle if and only if TM is trivial.

6-4. Let $f\colon \mathbb{R}^3 \to \mathbb{R}$ be the function $f(x, y, z) = x^2 + y^2 + z^2$, and let $F\colon \mathbb{R}^2 \to \mathbb{R}^3$ be the following map (the inverse of stereographic projection):

$$F(u,v) = \left(\frac{2u}{u^2+v^2+1}, \frac{2v}{u^2+v^2+1}, \frac{u^2+v^2-1}{u^2+v^2+1} \right).$$

Compute F^*df and $d(f \circ F)$ separately, and verify that they are equal.

6-5. In each of the cases below, M is a smooth manifold and $f\colon M \to \mathbb{R}$ is a smooth function. Compute the coordinate representation for df, and determine the set of all points $p \in M$ at which $df_p = 0$.

(a) $M = \{(x,y) \in \mathbb{R}^2 : x > 0\}$; $f(x,y) = x/(x^2+y^2)$. Use standard coordinates (x,y).
(b) M and f are as in part (a); this time use polar coordinates (r,θ).
(c) $M = S^2 \subset \mathbb{R}^3$; $f(p) = z(p)$ (the z-coordinate of p, thought of as a point in $\mathbb{R}^3$). Use stereographic coordinates.
(d) $M = \mathbb{R}^n$; $f(x) = |x|^2$. Use standard coordinates.

6-6. Let $\det\colon \mathrm{GL}(n,\mathbb{R}) \to \mathbb{R}$ denote the determinant function.

(a) Using matrix entries (X_i^j) as global coordinates on $\mathrm{GL}(n,\mathbb{R})$, show that the partial derivatives of det are given by

$$\frac{\partial}{\partial X_i^j}(\det X) = (\det X)(X^{-1})_j^i.$$

[Hint: Expand $\det X$ by minors along the ith column and use Cramer's rule.]

(b) Conclude that the differential of the determinant function is

$$d(\det)_X(B) = (\det X)\operatorname{tr}(X^{-1}B)$$

for $X \in \mathrm{GL}(n,\mathbb{R})$ and $B \in T_X \mathrm{GL}(n,\mathbb{R}) \cong \mathrm{M}(n,\mathbb{R})$, where $\operatorname{tr} X = \sum_i X_i^i$ is the trace of X.

(c) Considering $\det\colon \mathrm{GL}(n,\mathbb{R}) \to \mathbb{R}^*$ as a Lie group homomorphism, show that its induced Lie algebra homomorphism is $\operatorname{tr}\colon \mathfrak{gl}(n,\mathbb{R}) \to \mathbb{R}$.

6-7. Let M be a smooth manifold and $p \in M$. Problem 3-8 showed that tangent vectors at p can be viewed as equivalence classes of smooth curves, which are smooth maps from (open subsets of) $\mathbb{R}$ to M. This problem shows that covectors at p can be viewed dually as equivalence classes of smooth functions from M to $\mathbb{R}$. Let $\mathcal{F}_p$ denote the subspace of $C^\infty(M)$ consisting of smooth functions that vanish at p, and let $\mathcal{F}_p^2$ be the subspace of $\mathcal{F}_p$ spanned by functions of the form fg for some $f, g \in \mathcal{F}_p$. Define a map $\Phi\colon \mathcal{F}_p \to T_p^*M$ by setting $\Phi(f) = df_p$. Show that the restriction of Φ to $\mathcal{F}_p^2$ is zero, and that Φ descends to an isomorphism from $\mathcal{F}_p/\mathcal{F}_p^2$ to T_p^*M.

6-8. Let M be a smooth manifold.

(a) Given a smooth covector field ω on M, show that the map $\widetilde{\omega}\colon \mathfrak{T}(M) \to C^\infty(M)$ defined by

$$\widetilde{\omega}(X)(p) = \omega_p(X_p)$$

is linear over $C^\infty(M)$ (see page 117).

(b) Show that a map

$$\widetilde{\omega}\colon \mathfrak{T}(M) \to C^\infty(M)$$

is induced by a smooth covector field as above if and only if it is linear over $C^\infty(M)$.

6-9. Suppose $F\colon M \to N$ is any smooth map, $\omega \in \mathfrak{T}^*(N)$, and γ is a piecewise smooth curve segment in M. Show that

$$\int_\gamma F^*\omega = \int_{F\circ\gamma} \omega.$$

6-10. Consider the following two covector fields on $\mathbb{R}^3$:

$$\omega = -\frac{4z\,dx}{(x^2+1)^2} + \frac{2y\,dy}{y^2+1} + \frac{2x\,dz}{x^2+1},$$
$$\eta = -\frac{4xz\,dx}{(x^2+1)^2} + \frac{2y\,dy}{y^2+1} + \frac{2\,dz}{x^2+1}.$$

(a) Set up and evaluate the line integral of each covector field along the straight line segment from $(0,0,0)$ to $(1,1,1)$.

(b) Determine whether either of these covector fields is exact.

(c) For each one that is exact, find a potential function and use it to recompute the line integral.

6-11. The *length* of a smooth curve $\gamma\colon [a,b] \to \mathbb{R}^n$ is defined to be the value of the (ordinary) integral

$$L(\gamma) = \int_a^b |\gamma'(t)|\,dt.$$

Show that there is no smooth covector field $\omega \in \mathfrak{T}^*(\mathbb{R}^n)$ with the property that $\int_\gamma \omega = L(\gamma)$ for every smooth curve γ.

6-12. Line Integrals of Vector Fields: Suppose X is a smooth vector field on an open set $U \subset \mathbb{R}^n$. For any piecewise smooth curve segment $\gamma\colon [a,b] \to U$, define the line integral of X over γ, denoted by $\int_\gamma X \cdot ds$, as

$$\int_\gamma X \cdot ds = \int_a^b X(\gamma(t)) \cdot \gamma'(t)\,dt,$$

where the dot on the right-hand side denotes the Euclidean dot product between tangent vectors at $\gamma(t)$, identified with elements of $\mathbb{R}^n$. We say a vector field is conservative if its line integral around any piecewise smooth closed curve is zero.

(a) Show that X is conservative if and only if there exists a smooth function $f \in C^\infty(U)$ such that $X = \operatorname{grad} f$. [Hint: Consider the covector field ω defined by $\omega_x(Y) = X_x \cdot Y$.]

(b) If $n = 3$ and X is conservative, show that $\operatorname{curl} X = 0$, where

$$\begin{aligned}\operatorname{curl} X = \left(\frac{\partial X^3}{\partial x^2} - \frac{\partial X^2}{\partial x^3}\right)\frac{\partial}{\partial x^1} + \left(\frac{\partial X^1}{\partial x^3} - \frac{\partial X^3}{\partial x^1}\right)\frac{\partial}{\partial x^2} \\ + \left(\frac{\partial X^2}{\partial x^1} - \frac{\partial X^1}{\partial x^2}\right)\frac{\partial}{\partial x^3}.\end{aligned}$$

(c) If $U \subset \mathbb{R}^3$ is star-shaped, show that X is conservative on U if and only if $\operatorname{curl} X = 0$.

6-13. (a) If M is a compact manifold, show that every exact covector field on M vanishes at least at two points.

(b) Is there a smooth covector field on $\mathbb{S}^2$ that vanishes at exactly one point?

6-14. Let $\mathbb{T}^n = \mathbb{S}^1 \times \cdots \times \mathbb{S}^1 \subset \mathbb{C}^n$ denote the n-torus. For each $i = 1, \ldots, n$, let $\gamma_i \colon [0,1] \to \mathbb{T}^n$ be the curve segment

$$\gamma_i(t) = (1, \ldots, e^{2\pi i t}, \ldots, 1) \qquad \text{(with } e^{2\pi i t} \text{ in the } i\text{th place).}$$

Show that a closed covector field ω on $\mathbb{T}^n$ is exact if and only if $\int_{\gamma_i} \omega = 0$ for $i = 1, \ldots, n$. [Hint: Consider first $E^*\omega$, where $E \colon \mathbb{R}^n \to \mathbb{T}^n$ is the smooth covering map $E(x^1, \ldots, x^n) = (e^{2\pi i x^1}, \ldots, e^{2\pi i x^n})$.]

6-15. Let V_0 be the category of finite-dimensional vector spaces and linear isomorphisms as in Problem 5-13. Define a functor $\mathcal{F} \colon \mathsf{V}_0 \to \mathsf{V}_0$ by setting $\mathcal{F}(V) = V^*$ for a vector space V, and $\mathcal{F}(A) = (A^{-1})^*$ for an isomorphism A. Show that $\mathcal{F}$ is a smooth covariant functor, and show that $\mathcal{F}(TM)$ and T^*M are canonically smoothly isomorphic vector bundles for any smooth manifold M.

7
Submersions, Immersions, and Embeddings

Because the push-forward of a smooth map represents the "best linear approximation" to the map near a given point, we can learn a great deal about the map itself by studying linear-algebraic properties of its push-forward at each point. The most important such property is its rank (the dimension of its image).

In this chapter, we will undertake a detailed study of the ways in which geometric properties of maps can be detected from their push-forwards. The maps for which push-forwards give good local models turn out to be the ones whose push-forwards have constant rank. Three special categories of such maps will play particularly important roles: submersions (smooth maps whose push-forwards are surjective), immersions (smooth maps whose push-forwards are injective), and smooth embeddings (injective immersions that are also homeomorphisms onto their images).

The engine that powers this discussion is an analytic result that will prove indispensable in the theory of smooth manifolds: the inverse function theorem. This theorem and its corollaries (the rank theorem and the implicit function theorem) show that, under appropriate hypotheses on its rank, a smooth map behaves locally like its push-forward.

After discussing some general applications of these results to manifolds, we use them to study an important category of maps—surjective submersions—which play a role in smooth manifold theory closely analogous to the role played by quotient maps in topology.

Maps of Constant Rank

If $F\colon M \to N$ is a smooth map, we define the *rank* of F at $p \in M$ to be the rank of the linear map $F_*\colon T_pM \to T_{F(p)}N$; it is of course just the rank of the matrix of partial derivatives of F in any smooth chart, or the dimension of $\operatorname{Im} F_* \subset T_{F(p)}N$. If F has the same rank k at every point, we say it has *constant rank*, and write $\operatorname{rank} F = k$.

A smooth map $F\colon M \to N$ is called a *submersion* if F_* is surjective at each point (or equivalently if $\operatorname{rank} F = \dim N$). It is called an *immersion* if F_* is injective at each point (equivalently, $\operatorname{rank} F = \dim M$). As we will see in this chapter, submersions and immersions behave locally like surjective and injective linear maps, respectively.

One special kind of immersion is particularly important. A *smooth embedding* is an immersion $F\colon M \to N$ that is also a topological embedding, i.e., a homeomorphism onto its image $F(M) \subset N$ in the subspace topology. Notice that although submersions and immersions are smooth maps by definition, there are two types of embeddings—topological and smooth. A smooth embedding is a map that is both a topological embedding and an immersion, not just a topological embedding that happens to be smooth.

Example 7.1 (Submersions, Immersions, and Embeddings).

(a) If $M_1, \dots, M_k$ are smooth manifolds, each of the projections $\pi_i\colon M_1 \times \cdots \times M_k \to M_i$ is a submersion. In particular, the projection $\pi\colon \mathbb{R}^{n+k} \to \mathbb{R}^n$ onto the first n coordinates is a submersion.

(b) Similarly, with $M_1, \dots, M_k$ as above, if $p_i \in M_i$ are arbitrarily chosen points, each of the maps $\iota_j\colon M_j \to M_1 \times \cdots \times M_k$ given by

$$\iota_j(q) = (p_1, \dots, p_{j-1}, q, p_{j+1}, \dots, p_k)$$

is a smooth embedding. In particular, the inclusion map $\mathbb{R}^n \hookrightarrow \mathbb{R}^{n+k}$ given by sending $(x^1, \dots, x^n)$ to $(x^1, \dots, x^n, 0, \dots, 0)$ is a smooth embedding.

(c) If $\gamma\colon J \to M$ is a smooth curve in a smooth manifold M, then γ is an immersion if and only if $\gamma'(t) \neq 0$ for all $t \in J$.

(d) If $F\colon M \to N$ is a local diffeomorphism, then F is both an immersion and a submersion by Exercise 3.3. In particular, any smooth covering map is both an immersion and a submersion.

(e) If E is a smooth vector bundle over a smooth manifold M, the projection map $\pi\colon E \to M$ is a submersion.

(f) The smooth map $X\colon \mathbb{R}^2 \to \mathbb{R}^3$ given by

$$X(\varphi, \theta) = ((2 + \cos\varphi)\cos\theta, (2 + \cos\varphi)\sin\theta, \sin\varphi),$$

is an immersion of $\mathbb{R}^2$ into $\mathbb{R}^3$ whose image is the doughnut-shaped surface obtained by revolving the circle $(y-2)^2 + z^2 = 1$ in the

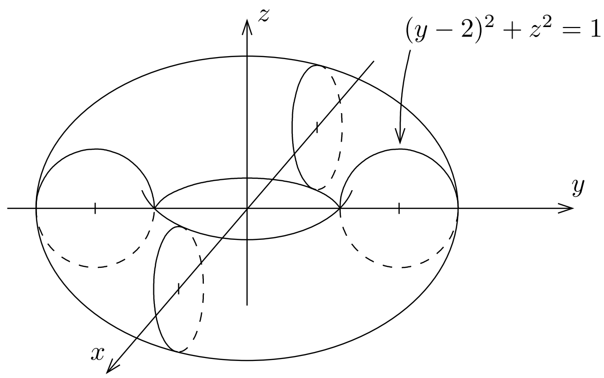


Figure 7.1. An embedding of the torus in $\mathbb{R}^3$.

(y,z)-plane about the z-axis (Figure 7.1). Problem 7-1 shows that X descends to an embedding of the torus into $\mathbb{R}^3$.

◇ **Exercise 7.1.** Verify the claims made in parts (a)–(e) of the preceding example.

◇ **Exercise 7.2.** Show that a composition of submersions is a submersion, a composition of immersions is an immersion, and a composition of smooth embeddings is a smooth embedding.

To understand more fully what it means to be an embedding, it is useful to bear in mind some examples of injective immersions that are *not* embeddings. The next two examples illustrate two rather different ways in which an injective immersion can fail to be an embedding.

Example 7.2. Consider the map $\gamma\colon(-\pi/2,3\pi/2)\to\mathbb{R}^2$ given by

$$\gamma(t)=(\sin 2t,\cos t).$$

Its image is a curve that looks like a figure eight in the plane (Figure 7.2). (It is the locus of points (x,y) where $x^2=4y^2(1-y^2)$, as you can check.) It is easy to see that γ is an injective immersion because $\gamma'(t)$ never vanishes; but it is not a topological embedding, because its image is compact in the subspace topology while its domain is not.

Example 7.3. Let $\mathbb{T}^2=\mathbb{S}^1\times\mathbb{S}^1\subset\mathbb{C}^2$ denote the torus, and let c be any irrational number. The map $\gamma\colon\mathbb{R}\to\mathbb{T}^2$ given by

$$\gamma(t)=(e^{2\pi it},e^{2\pi ict})$$

is an immersion because $\gamma'(t)$ never vanishes. It is also injective, because $\gamma(t_1)=\gamma(t_2)$ implies that both t_1-t_2 and ct_1-ct_2 are integers, which is impossible unless $t_1=t_2$.

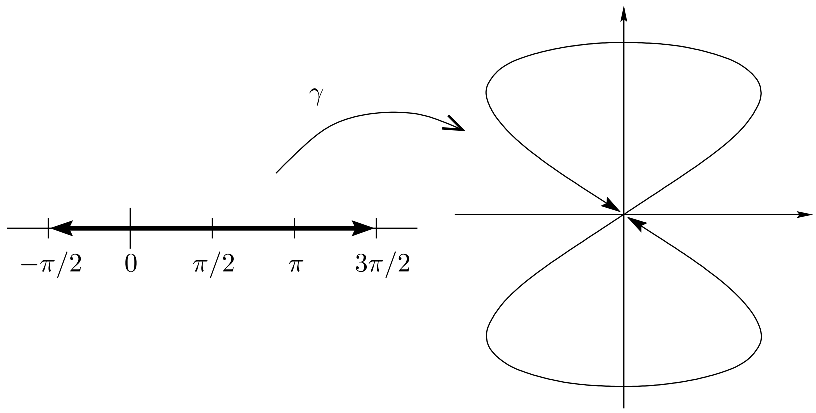


Figure 7.2. The figure eight curve of Example 7.2.

Consider the set $\gamma(\mathbb{Z}) = \{\gamma(n) : n \in \mathbb{Z}\}$. If γ were a homeomorphism onto its image, this set would have no limit point in $\gamma(\mathbb{R})$, because $\mathbb{Z}$ has no limit point in $\mathbb{R}$. However, we will show that $\gamma(0)$ is a limit point of $\gamma(\mathbb{Z})$. To prove this claim, we need to show that given any $\varepsilon > 0$, there is a nonzero integer k such that $|\gamma(k) - \gamma(0)| < \varepsilon$.

Since $\mathbb{S}^1$ is compact, the infinite set $\{e^{2\pi icn} : n \in \mathbb{Z}\}$ has a limit point, say $z_0 \in \mathbb{S}^1$. Given $\varepsilon > 0$, we can choose distinct integers n_1 and n_2 such that $|e^{2\pi icn_j} - z_0| < \varepsilon/2$, and therefore $|e^{2\pi icn_1} - e^{2\pi icn_2}| < \varepsilon$. Taking $k = n_1 - n_2$, this implies that

$$|e^{2\pi ick} - 1| = |e^{-2\pi in_2}(e^{2\pi icn_1} - e^{2\pi icn_2})| = |e^{2\pi icn_1} - e^{2\pi icn_2}| < \varepsilon,$$

and so

$$|\gamma(k) - \gamma(0)| = |(1, e^{2\pi ick}) - (1, 1)| < \varepsilon.$$

In fact, it is not hard to show that the image set $\gamma(\mathbb{R})$ is actually dense in $\mathbb{T}^2$ (see Problem 7-3).

Since any closed, injective, continuous map is a topological embedding (see Lemma A.13 in the Appendix), one criterion that rules out such cases is that F be a closed map. This is always the case when the domain is compact or the map is proper.

Proposition 7.4. *Suppose $F\colon M \to N$ is an injective immersion. If either of the following conditions holds, then F is a smooth embedding with closed image.*

(*a*) *M is compact.*

(*b*) *F is a proper map.*

Proof. Either hypothesis implies that F is a closed map—when F is proper, it follows from Proposition 2.18, and when M is compact, it follows either from that same proposition or from the closed map lemma (Lemma A.19 in the Appendix). This immediately implies that $F(M)$ is closed in N, and Lemma A.13 shows that F is a topological embedding. □

The Inverse Function Theorem and Its Friends

In this section, we will prove several analytic theorems that will provide the keys to understanding how the local behavior of a smooth map is modeled by the behavior of its push-forward. The simplest of these is the inverse function theorem.

Theorem 7.5 (Inverse Function Theorem). *Suppose U and V are open subsets of $\mathbb{R}^n$, and $F\colon U \to V$ is a smooth map. If $DF(p)$ is nonsingular at some point $p \in U$, then there exist connected neighborhoods $U_0 \subset U$ of p and $V_0 \subset V$ of $F(p)$ such that $F|_{U_0}\colon U_0 \to V_0$ is a diffeomorphism.*

The proof of this theorem is based on an elementary result about metric spaces, which we describe first.

Let X be a metric space. A map $G\colon X \to X$ is said to be a *contraction* if there is a constant $\lambda < 1$ such that $d(G(x), G(y)) \le \lambda d(x, y)$ for all $x, y \in X$. Clearly any contraction is continuous.

Lemma 7.6 (Contraction Lemma). *Let X be a complete metric space. Every contraction $G\colon X \to X$ has a unique fixed point, i.e., a point $x \in X$ such that $G(x) = x$.*

Proof. Uniqueness is immediate, for if x and x' are both fixed points of G, the contraction property implies $d(x, x') = d(G(x), G(x')) \le \lambda d(x, x')$, which is possible only if $x = x'$.

To prove the existence of a fixed point, let x_0 be an arbitrary point in X, and define a sequence $\{x_n\}$ inductively by $x_{n+1} = G(x_n)$. For any $i \ge 1$ we have $d(x_i, x_{i+1}) = d(G(x_{i-1}), G(x_i)) \le \lambda d(x_{i-1}, x_i)$, and therefore by induction

$$d(x_i, x_{i+1}) \le \lambda^i d(x_0, x_1).$$

If $j \ge i \ge N$,

$$\begin{aligned}
d(x_i, x_j) &\le d(x_i, x_{i+1}) + d(x_{i+1}, x_{i+2}) + \cdots + d(x_{j-1}, x_j) \\
&\le (\lambda^i + \cdots + \lambda^{j-1}) d(x_0, x_1) \\
&\le \lambda^i \left(\sum_{n=0}^{\infty} \lambda^n \right) d(x_0, x_1) \\
&\le \lambda^N \frac{1}{1-\lambda} d(x_0, x_1).
\end{aligned}$$

Since this last expression can be made as small as desired by choosing N large, the sequence $\{x_n\}$ is Cauchy and therefore converges to a limit $x \in X$. Because G is continuous,

$$G(x) = G\left(\lim_{n\to\infty} x_n\right) = \lim_{n\to\infty} G(x_n) = \lim_{n\to\infty} x_{n+1} = x,$$

so x is the desired fixed point. □

Proof of the inverse function theorem. We begin by making some simple modifications to the map F to streamline the proof. First, the map

$$F_1(x) = F(x+p) - F(p)$$

is smooth on a neighborhood of 0 and satisfies $F_1(0) = 0$ and $DF_1(0) = DF(p)$; clearly F is a diffeomorphism on a connected neighborhood of p if and only if F_1 is a diffeomorphism on a connected neighborhood of 0. Second, the map $F_2 = DF_1(0)^{-1} \circ F_1$ is smooth on the same neighborhood of 0 and satisfies $F_2(0) = 0$ and $DF_2(0) = \mathrm{Id}$; and if F_2 is a diffeomorphism near 0, then so is F_1 and therefore also F. Henceforth, replacing F by F_2, we will assume that F is defined in a neighborhood U of 0, $F(0) = 0$, and $DF(0) = \mathrm{Id}$.

Let $H(x) = x - F(x)$ for $x \in U$. Then $DH(0) = \mathrm{Id} - \mathrm{Id} = 0$. Because the matrix entries of $DH(x)$ are continuous functions of x, there is a number $\delta > 0$ such that $|DH(x)| \le \frac{1}{2}$ for all $x \in \overline{B}_\delta(0)$. If $x, x' \in \overline{B}_\delta(0)$, the Lipschitz estimate for smooth functions (Proposition A.69 in the Appendix) implies

$$|H(x') - H(x)| \le \tfrac{1}{2}|x' - x|. \tag{7.1}$$

In particular, taking $x' = 0$, this implies

$$|H(x)| \le \tfrac{1}{2}|x|. \tag{7.2}$$

Since $x' - x = F(x') - F(x) + H(x') - H(x)$, it follows that

$$|x' - x| \le |F(x') - F(x)| + |H(x') - H(x)| \le |F(x') - F(x)| + \tfrac{1}{2}|x' - x|.$$

Subtracting $\frac{1}{2}|x' - x|$ from both sides, we conclude that

$$|x' - x| \le 2|F(x') - F(x)| \tag{7.3}$$

for all $x, x' \in \overline{B}_\delta(0)$. In particular, this shows that F is injective on $\overline{B}_\delta(0)$.

Now let $y \in B_{\delta/2}(0)$ be arbitrary. We will show that there exists a unique point $x \in B_\delta(0)$ such that $F(x) = y$. Let $G(x) = y + H(x) = y + x - F(x)$, so that $G(x) = x$ if and only if $F(x) = y$. If $|x| \le \delta$, (7.2) implies

$$|G(x)| \le |y| + |H(x)| < \frac{\delta}{2} + \frac{1}{2}|x| \le \delta, \tag{7.4}$$

so G maps $\overline{B}_\delta(0)$ to itself. It follows from (7.1) that $|G(x) - G(x')| = |H(x) - H(x')| \le \frac{1}{2}|x - x'|$, so G is a contraction. Since $\overline{B}_\delta(0)$ is compact and therefore complete, the contraction lemma implies that G has a unique fixed point $x \in \overline{B}_\delta(0)$. From (7.4), $|x| = |G(x)| < \delta$, so in fact $x \in B_\delta(0)$, thus proving the claim.

Let $U_1 = B_\delta(0) \cap F^{-1}(B_{\delta/2}(0))$. Then U_1 is open in $\mathbb{R}^n$, and the argument above shows that $F\colon U_1 \to B_{\delta/2}(0)$ is bijective, so $F^{-1}\colon B_{\delta/2}(0) \to U_1$ exists. Substituting $x = F^{-1}(y)$ and $x' = F^{-1}(y')$ into (7.3) shows that F^{-1} is continuous. Let U_0 be the connected component of U_1 containing 0, and $V_0 = F(U_0)$. Then $F\colon U_0 \to V_0$ is a homeomorphism.

To show that it is a diffeomorphism, the only thing that needs to be proved is that F^{-1} is smooth. If F^{-1} were differentiable at $b \in V_0$, the chain rule would imply

$$\begin{aligned}\mathrm{Id} &= D(F \circ F^{-1})(b)\\ &= DF(F^{-1}(b)) \circ DF^{-1}(b),\end{aligned}$$

from which it would follow that $DF^{-1}(b) = DF(F^{-1}(b))^{-1}$. We will begin by showing that F^{-1} is differentiable at each point of V_0, with total derivative given by this formula.

Let $b \in V_0$ and set $a = F^{-1}(b) \in U_0$. For $v, w \in \mathbb{R}^n$ small enough that $a + v \in U_0$ and $b + w \in V_0$, define $R(v)$ and $S(w)$ by

$$\begin{aligned}R(v) &= F(a+v) - F(a) - DF(a)v,\\ S(w) &= F^{-1}(b+w) - F^{-1}(b) - DF(a)^{-1}w.\end{aligned}$$

Because F is smooth, it is differentiable at a, which means that $\lim_{v\to 0} R(v)/|v| = 0$. We need to show that $\lim_{w\to 0} S(w)/|w| = 0$.

For sufficiently small $w \in \mathbb{R}^n$, define

$$v(w) = F^{-1}(b+w) - F^{-1}(b) = F^{-1}(b+w) - a.$$

It follows that

$$\begin{aligned}F^{-1}(b+w) &= F^{-1}(b) + v(w) = a + v(w),\\ w &= (b+w) - b = F(a + v(w)) - F(a),\end{aligned} \tag{7.5}$$

and therefore

$$\begin{aligned}S(w) &= F^{-1}(b+w) - F^{-1}(b) - DF(a)^{-1}w\\ &= v(w) - DF(a)^{-1}w\\ &= DF(a)^{-1}(DF(a)v(w) - w)\\ &= DF(a)^{-1}(DF(a)v(w) + F(a) - F(a+v(w)))\\ &= -DF(a)^{-1}R(v(w)).\end{aligned}$$

We will show below that there are positive constants c and C such that

$$c|w| \le |v(w)| \le C|w| \tag{7.6}$$

for all sufficiently small w. In particular, this implies that $v(w) \neq 0$ when w is sufficiently small and nonzero. From this together with the result of

Exercise A.56 in the Appendix, we conclude that

$$\begin{aligned}\frac{|S(w)|}{|w|} &\le |DF(a)^{-1}|\,\frac{|R(v(w))|}{|w|}\\ &= |DF(a)^{-1}|\,\frac{|R(v(w))|}{|v(w)|}\,\frac{|v(w)|}{|w|}\\ &\le C|DF(a)^{-1}|\,\frac{|R(v(w))|}{|v(w)|},\end{aligned}$$

which approaches zero as $w \to 0$ because $v(w) \to 0$ and F is differentiable.

To complete the proof that F^{-1} is differentiable, it remains only to prove (7.6). From the definition of $R(v)$ and (7.5),

$$\begin{aligned}v(w) &= DF(a)^{-1}DF(a)v(w)\\ &= DF(a)^{-1}(F(a+v(w)) - F(a) - R(v(w)))\\ &= DF(a)^{-1}(w - R(v(w))),\end{aligned}$$

which implies

$$|v(w)| \le |DF(a)^{-1}|\,|w| + |DF(a)^{-1}|\,|R(v(w))|.$$

Because $|R(v)|/|v| \to 0$ as $v \to 0$, there exists $\delta_1 > 0$ such that $|v| < \delta_1$ implies $|R(v)| \le |v|/(2|DF(a)^{-1}|)$. By continuity of F^{-1}, there exists $\delta_2 > 0$ such that $|w| < \delta_2$ implies $|v(w)| < \delta_1$, and therefore

$$|v(w)| \le |DF(a)^{-1}|\,|w| + \tfrac{1}{2}|v(w)|.$$

Subtracting $\frac{1}{2}|v(w)|$ from both sides, we obtain

$$|v(w)| \le 2|DF(a)^{-1}|\,|w|$$

whenever $|w| < \delta_2$. This is the second inequality of (7.6). To prove the first, we use (7.5) again to get

$$w = F(a+v(w)) - F(a) = DF(a)v(w) + R(v(w)).$$

Therefore, when $|w| < \delta_2$,

$$|w| \le |DF(a)|\,|v(w)| + |R(v(w))| \le \left(|DF(a)| + \frac{1}{2|DF(a)^{-1}|}\right)|v(w)|.$$

This completes the proof that F^{-1} is differentiable.

By Lemma A.54, the partial derivatives of F^{-1} are defined at each point $y \in V_0$. Observe that the formula $DF^{-1}(y) = DF(F^{-1}(y))^{-1}$ implies that the map $DF^{-1}\colon V_0 \to \mathrm{GL}(n,\mathbb{R})$ can be written as the composition

$$V_0 \xrightarrow{F^{-1}} U_0 \xrightarrow{DF} GL(n,\mathbb{R}) \xrightarrow{i} GL(n,\mathbb{R}), \tag{7.7}$$

where $i(A) = A^{-1}$. Matrix inversion is a smooth map as we observed in Example 2.7(a). Also, DF is smooth because its component functions are the partial derivatives of F, which are assumed to be smooth. Because

DF^{-1} is a composition of continuous maps, it is continuous, and thus the partial derivatives of F^{-1} are continuous, which means that F^{-1} is of class C^1.

Now assume by induction that we have shown F^{-1} is of class C^k. This means that each of the maps in (7.7) is of class C^k. Because DF^{-1} is a composition of C^k functions, it is itself C^k; this implies that the partial derivatives of F^{-1} are of class C^k, so F^{-1} itself is of class C^{k+1}. Continuing by induction, we conclude that F^{-1} is smooth. □

For our purposes, the most important consequence of the inverse function theorem is the following, which says that a constant-rank smooth map can be placed locally into a particularly simple canonical form by a change of coordinates. It is a nonlinear version of the canonical form theorem for linear maps given in the Appendix (Theorem A.33).

Theorem 7.7 (Rank Theorem). *Suppose $U \subset \mathbb{R}^m$ and $V \subset \mathbb{R}^n$ are open sets and $F\colon U \to V$ is a smooth map with constant rank k. For any $p \in U$, there exist smooth coordinate charts (U_0, φ) for $\mathbb{R}^m$ and (V_0, ψ) for $\mathbb{R}^m$, with $p \in U_0 \subset U$ and $F(U_0) \subset V_0 \subset V$, such that*

$$\psi \circ F \circ \varphi^{-1}(x^1, \ldots, x^k, x^{k+1}, \ldots, x^m) = (x^1, \ldots, x^k, 0, \ldots, 0).$$

Proof. The fact that $DF(p)$ has rank k implies that its matrix has some $k \times k$ minor with nonzero determinant. By reordering the coordinates, we may assume that it is the upper left minor, $(\partial F^i/\partial x^j)$ for $i, j = 1, \ldots, k$. Let us relabel the standard coordinates as $(x, y) = (x^1, \ldots, x^k, y^1, \ldots, y^{m-k})$ in $\mathbb{R}^m$ and $(v, w) = (v^1, \ldots, v^k, w^1, \ldots, w^{n-k})$ in $\mathbb{R}^n$. By an initial translation of the coordinates, we may assume without loss of generality that $p = (0, 0)$ and $F(p) = (0, 0)$. If we write $F(x, y) = (Q(x, y), R(x, y))$ for some smooth maps $Q\colon U \to \mathbb{R}^k$ and $R\colon U \to \mathbb{R}^{n-k}$, then our hypothesis is that $(\partial Q^i/\partial x^j)$ is nonsingular at $(0, 0)$.

Define $\varphi\colon U \to \mathbb{R}^m$ by

$$\varphi(x, y) = (Q(x, y), y).$$

Its total derivative at $(0, 0)$ is

$$D\varphi(0,0) = \begin{pmatrix} \dfrac{\partial Q^i}{\partial x^j}(0,0) & \dfrac{\partial Q^i}{\partial y^j}(0,0) \\ 0 & I_{m-k} \end{pmatrix},$$

which is nonsingular because its columns are independent. Therefore, by the inverse function theorem, there are connected neighborhoods U_0 of $(0, 0)$ and $\widetilde{U}_0$ of $\varphi(0, 0) = (0, 0)$ such that $\varphi\colon U_0 \to \widetilde{U}_0$ is a diffeomorphism. Writing the inverse map as $\varphi^{-1}(x, y) = (A(x, y), B(x, y))$ for some smooth functions $A\colon \widetilde{U}_0 \to \mathbb{R}^k$ and $B\colon \widetilde{U}_0 \to \mathbb{R}^{m-k}$, we compute

$$\begin{aligned}(x, y) &= \varphi(A(x, y), B(x, y)) \\ &= (Q(A(x, y), B(x, y)), B(x, y)).\end{aligned} \tag{7.8}$$

Comparing y components, it follows that $B(x,y) = y$, and therefore φ^{-1} has the form

$$\varphi^{-1}(x,y) = (A(x,y),y).$$

Observe that $\varphi \circ \varphi^{-1} = \mathrm{Id}$ implies $Q(A(x,y),y) = x$, and therefore $F \circ \varphi^{-1}$ has the form

$$F \circ \varphi^{-1}(x,y) = (x, \widetilde{R}(x,y)),$$

where $\widetilde{R}\colon \widetilde{U}_0 \to \mathbb{R}^k$ is defined by $\widetilde{R}(x,y) = R(A(x,y),y)$. The Jacobian matrix of this map at an arbitrary point $(x,y) \in \widetilde{U}_0$ is

$$D(F \circ \varphi^{-1})(x,y) = \begin{pmatrix} I_k & 0 \\ \dfrac{\partial \widetilde{R}^i}{\partial x^j} & \dfrac{\partial \widetilde{R}^i}{\partial y^j} \end{pmatrix}.$$

Since composing with a diffeomorphism does not change the rank of a map, this matrix has rank equal to k everywhere in $\widetilde{U}_0$. Since the first k columns are obviously independent, the rank can be k only if the partial derivatives $\partial \widetilde{R}^i / \partial y^j$ vanish identically on $\widetilde{U}_0$, which implies that $\widetilde{R}$ is actually independent of $(y^1, \dots, y^{m-k})$. Thus if we let $S(x) = \widetilde{R}(x,0)$, we have

$$F \circ \varphi^{-1}(x,y) = (x, S(x)). \tag{7.9}$$

To complete the proof, we need to define a smooth chart for $\mathbb{R}^n$ near $(0,0)$. Let $V_0 \subset V$ be the open set

$$V_0 = \{(v,w) \in V : (v,0) \in \widetilde{U}_0\},$$

which is a neighborhood of $(0,0)$ because $(0,0) \in \widetilde{U}_0$, and define $\psi\colon V_0 \to \mathbb{R}^n$ by

$$\psi(v,w) = (v, w - S(v)).$$

This is a diffeomorphism onto its image, because its inverse is given explicitly by $\psi^{-1}(s,t) = (s, t + S(s))$; thus (V_0, ψ) is a smooth chart. It follows from (7.9) that

$$\psi \circ F \circ \varphi^{-1}(x,y) = \psi(x, S(x)) = (x, S(x) - S(x)) = (x,0),$$

which was to be proved. $\square$

Another useful consequence of the inverse function theorem is the implicit function theorem, which gives conditions under which a level set of a smooth map is locally the graph of a smooth function.

Theorem 7.8 (Implicit Function Theorem). *Let $U \subset \mathbb{R}^n \times \mathbb{R}^k$ be an open set, and let $(x,y) = (x^1, \dots, x^n, y^1, \dots, y^k)$ denote the standard coordinates on U. Suppose $\Phi\colon U \to \mathbb{R}^k$ is a smooth map, $(a,b) \in U$, and*

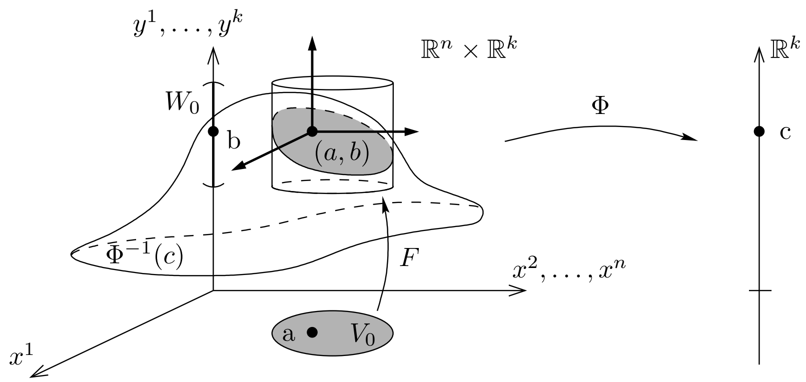


Figure 7.3. The implicit function theorem.

$c = \Phi(a, b)$. *If the* $k \times k$ *matrix*

$$\left(\frac{\partial \Phi^i}{\partial y^j}(a, b) \right)$$

is nonsingular, then there exist neighborhoods $V_0 \subset \mathbb{R}^n$ *of* a *and* $W_0 \subset \mathbb{R}^k$ *of* b *and a smooth map* $F \colon V_0 \to W_0$ *such that* $\Phi^{-1}(c) \cap V_0 \times W_0$ *is the graph of* F, *i.e.*, $\Phi(x, y) = c$ *for* $(x, y) \in V_0 \times W_0$ *if and only if* $y = F(x)$ (*Figure 7.3*).

Proof. Under the hypotheses of the theorem, consider the smooth map $\Psi \colon U \to \mathbb{R}^n \times \mathbb{R}^k$ defined by $\Psi(x, y) = (x, \Phi(x, y))$. Its total derivative at (a, b) is

$$D\Psi(a, b) = \begin{pmatrix} I_n & 0 \\ \dfrac{\partial \Phi^i}{\partial x^j}(a, b) & \dfrac{\partial \Phi^i}{\partial y^j}(a, b) \end{pmatrix},$$

which is nonsingular by hypothesis. Thus by the inverse function theorem there exist connected neighborhoods U_0 of (a, b) and Y_0 of (a, c) such that $\Psi \colon U_0 \to Y_0$ is a diffeomorphism. Shrinking U_0 and Y_0 if necessary, we may assume $U_0 = V \times W$ is a product neighborhood. Arguing exactly as in the proof of the rank theorem (but with the roles of x and y reversed), we see that the inverse map has the form

$$\Psi^{-1}(x, y) = (x, B(x, y))$$

for some smooth map $B \colon Y_0 \to W$.

Now let $V_0 = \{x \in V : (x, c) \in Y_0\}$ and $W_0 = W$, and define $F \colon V_0 \to W_0$ by $F(x) = B(x, c)$. Comparing y components in the relation $(x, c) = \Psi \circ$

$\Psi^{-1}(x, c)$ yields

$$c = \Phi(x, B(x, c)) = \Phi(x, F(x))$$

whenever $x \in V_0$, so the graph of F is contained in $\Phi^{-1}(c)$. Conversely, suppose $(x, y) \in V_0 \times W_0$ and $\Phi(x, y) = c$. Then $\Psi(x, y) = (x, \Phi(x, y)) = (x, c)$, so

$$(x, y) = \Psi^{-1}(x, c) = (x, B(x, c)) = (x, F(x)),$$

which implies that $y = F(x)$. This completes the proof. □

Constant Rank Maps Between Manifolds

The inverse function theorem and the rank theorem are, on the face of it, results about maps between open subsets of Euclidean spaces. However, we will usually apply them to maps between manifolds, so it is useful to restate them in this context.

Theorem 7.9 (Inverse Function Theorem for Manifolds). *Suppose M and N are smooth manifolds, $p \in M$, and $F \colon M \to N$ is a smooth map such that $F_* \colon T_pM \to T_pN$ is bijective. Then there exist connected neighborhoods U_0 of p and V_0 of $F(p)$ such that $F|_{U_0} \colon U_0 \to V_0$ is a diffeomorphism.*

Proof. The fact that F_* is bijective implies that M and N have the same dimension, and then the result follows from the Euclidean inverse function theorem applied to the coordinate representation of F. □

Corollary 7.10. *Suppose M and N are smooth manifolds of the same dimension, and $F \colon M \to N$ is an immersion or submersion. Then F is a local diffeomorphism. If F is bijective, it is a diffeomorphism.*

Proof. The fact that F is a local diffeomorphism is an immediate consequence of the inverse function theorem. If F is bijective, then it is a diffeomorphism by Exercise 2.9. □

Example 7.11 (Spherical Coordinates). As you probably know from calculus, *spherical coordinates* (ρ, φ, θ) on $\mathbb{R}^3$ are defined by the relations

$$\begin{aligned} x &= \rho \sin\varphi \cos\theta, \\ y &= \rho \sin\varphi \sin\theta, \\ z &= \rho \cos\varphi. \end{aligned} \tag{7.10}$$

Geometrically, ρ is the distance from the origin, φ is the angle from the positive z-axis, and θ is the angle from the $x > 0$ half of the (x, z)-plane. (Figure 7.4).

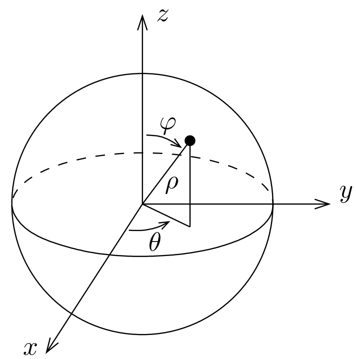


Figure 7.4. Spherical coordinates.

If we set $F(\rho,\varphi,\theta) = (\rho\sin\varphi\cos\theta, \rho\sin\varphi\sin\theta, \rho\cos\varphi)$, this defines F as a smooth map from $\mathbb{R}^3$ to $\mathbb{R}^3$, called the *spherical coordinate parametrization*. A computation shows that the Jacobian determinant of F is $\rho^2\sin\varphi$. Letting $U \subset \mathbb{R}^3$ be the open subset

$$U = \{(\rho,\varphi,\theta) \in \mathbb{R}^3 : \rho > 0,\ 0 < \varphi < \pi\},$$

it follows from Corollary 7.10 that F is a local diffeomorphism from U to V. Therefore, if we set $V_0 = F(U_0)$, where $U_0 \subset U$ is any open subset on which F is injective, the inverse map $(F|_{U_0})^{-1} \colon V_0 \to U_0$ is a smooth coordinate map. For example, we could take

$$\begin{aligned} U_0 &= \{(\rho,\varphi,\theta) \in \mathbb{R}^3 : \rho > 0,\ 0 < \varphi < \pi,\ 0 < \theta < 2\pi\}, \\ V_0 &= \{(x,y,z) \in \mathbb{R}^3 : y \neq 0 \text{ or } x < 0\}. \end{aligned}$$

Notice how much easier it is to argue this way than to construct an explicit inverse for F.

◇ **Exercise 7.3.** Verify the claims in the preceding example.

◇ **Exercise 7.4.** Carry out a similar analysis for polar coordinates in the plane.

Theorem 7.12 (Rank Theorem for Manifolds). *Suppose M and N are smooth manifolds of dimensions m and n, respectively, and $F\colon M \to N$ is a smooth map with constant rank k. For each $p \in M$ there exist smooth coordinates $(x^1,\dots,x^m)$ centered at p and $(v^1,\dots,v^n)$ centered at $F(p)$ in which F has the coordinate representation*

$$F(x^1,\dots,x^k,x^{k+1},\dots,x^m) = (x^1,\dots,x^k,0,\dots,0). \tag{7.11}$$

Proof. Replacing M and N by smooth coordinate domains $U \subset M$ near p and $V \subset N$ near $F(p)$ and replacing F by its coordinate representation, the theorem is reduced to the rank theorem in Euclidean space. □

The following corollary can be viewed as a somewhat more invariant statement of the rank theorem. It says that constant-rank maps are precisely the ones whose local behavior is the same as that of their push-forwards.

Corollary 7.13. *Let $F \colon M \to N$ be a smooth map, and suppose M is connected. Then the following are equivalent:*

(*a*) *For each $p \in M$ there exist smooth charts near p and $F(p)$ in which the coordinate representation of F is linear.*

(*b*) *F has constant rank.*

Proof. First suppose F has a linear coordinate representation in a neighborhood of each point. Since any linear map has constant rank, it follows that the rank of F is constant in a neighborhood of each point, and thus by connectedness it is constant on all of M. Conversely, if F has constant rank, the rank theorem shows that it has the linear coordinate representation (7.11) in a neighborhood of each point. □

The next theorem is a powerful consequence of the rank theorem. (Part of the proof needs to be delayed until Chapter 10, but we state the entire theorem here because it fits nicely with our study of constant-rank maps.)

Theorem 7.14. *Let $F \colon M \to N$ be a smooth map of constant rank.*

(*a*) *If F is surjective, then it is a submersion.*

(*b*) *If F is injective, then it is an immersion.*

(*c*) *If F is bijective, then it is a diffeomorphism.*

Proof. First note that (c) follows from (a) and (b), because a bijective smooth map of constant rank is a submersion by part (a) and an immersion by part (b), so M and N have the same dimension; and then Corollary 7.10 implies that F is a diffeomorphism.

To prove (b), let $m = \dim M$, $n = \dim N$, and suppose F has constant rank k. If F is not an immersion, then $k < m$. By the rank theorem, in a neighborhood of any point there is a smooth chart in which F has the coordinate representation

$$F(x^1, \dots, x^k, x^{k+1}, \dots, x^m) = (x^1, \dots, x^k, 0, \dots, 0). \tag{7.12}$$

It follows that $F(0, \dots, 0, \varepsilon) = F(0, \dots, 0, 0)$ for any sufficiently small ε, so F is not injective.

For the proof of (a), see Chapter 10 (page 244). □

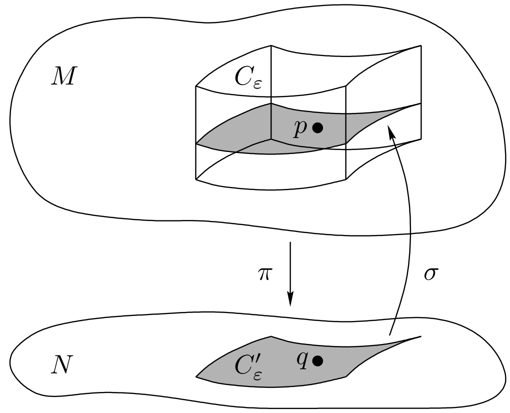


Figure 7.5. Local section of a submersion.

Submersions

Another important application of the rank theorem is to vastly expand our understanding of the properties of submersions.

Proposition 7.15 (Properties of Submersions). *Let $\pi\colon M \to N$ be a submersion.*

(*a*) *π is an open map.*

(*b*) *Every point of M is in the image of a smooth local section of π.*

(*c*) *If π is surjective, it is a quotient map.*

Proof. Given $p \in M$, let $q = \pi(p) \in N$. Because a submersion has constant rank, by the rank theorem we can choose smooth coordinates $(x^1, \dots, x^m)$ centered at p and $(y^1, \dots, y^k)$ centered at q in which π has the coordinate representation $\pi(x^1, \dots, x^k, x^{k+1}, \dots, x^m) = (x^1, \dots, x^k)$. If ε is a sufficiently small positive number, the coordinate cube

$$C_\varepsilon = \{x : |x^i| < \varepsilon \text{ for } i = 1, \dots, m\}$$

is a neighborhood of p whose image under π is the cube

$$C'_\varepsilon = \{y : |y^i| < \varepsilon \text{ for } i = 1, \dots, k\}.$$

The map $\sigma\colon C'_\varepsilon \to C_\varepsilon$ whose coordinate representation is $\sigma(x^1, \dots, x^k) = (x^1, \dots, x^k, 0, \dots, 0)$ is a smooth local section of π satisfying $\sigma(q) = p$ (Figure 7.5). This proves (b).

Suppose W is any open subset of M and $q \in \pi(W)$. For any $p \in W$ such that $\pi(p) = q$, W contains an open coordinate cube C_ε centered at p as above, and thus $\pi(W)$ contains an open coordinate cube centered at $\pi(p)$.

This proves that $\pi(W)$ is open, so (a) holds. Because a surjective open map is automatically a quotient map, (c) follows from (a). ☐

The next three propositions provide important tools that we will use frequently when studying submersions. The general philosophy of the proofs is this: To "push" a smooth object (such as a smooth map) down via a submersion, pull it back via local sections.

Proposition 7.16. *Suppose M, N, and P are smooth manifolds, $\pi\colon M \to N$ is a surjective submersion, and $F\colon N \to P$ is any map. Then F is smooth if and only if $F \circ \pi$ is smooth:*

$$\begin{tikzcd} M \arrow[d, "\pi"'] \arrow[dr, "F \circ \pi"] & \\ N \arrow[r, "F"'] & P. \end{tikzcd}$$

Proof. If F is smooth, then $F \circ \pi$ is smooth by composition. Conversely, suppose that $F \circ \pi$ is smooth, and let $q \in N$ be arbitrary. For any $p \in \pi^{-1}(q)$, Proposition 7.15(b) guarantees the existence of a neighborhood U of q and a smooth local section $\sigma\colon U \to M$ of π such that $\sigma(q) = p$. Then $\pi \circ \sigma = \mathrm{Id}_U$ implies

$$F|_U = F|_U \circ \mathrm{Id}_U = F|_U \circ (\pi \circ \sigma) = (F \circ \pi) \circ \sigma,$$

which is a composition of smooth maps. This shows that F is smooth in a neighborhood of each point, so it is smooth. ☐

The next proposition gives a very general sufficient condition under which a smooth map can be "pushed down" by a submersion.

Proposition 7.17 (Passing Smoothly to the Quotient). *Suppose $\pi\colon M \to N$ is a surjective submersion. If $F\colon M \to P$ is a smooth map that is constant on the fibers of π, then there is a unique smooth map $\widetilde{F}\colon N \to P$ such that $\widetilde{F} \circ \pi = F$:*

$$\begin{tikzcd} M \arrow[d, "\pi"'] \arrow[dr, "F"] & \\ N \arrow[r, "\widetilde{F}"'] & P. \end{tikzcd}$$

Proof. Clearly, if $\widetilde{F}$ exists, it will have to satisfy $\widetilde{F}(q) = F(p)$ whenever $p \in \pi^{-1}(q)$. We use this to *define* $\widetilde{F}$: Given $q \in N$, let $\widetilde{F}(q) = F(p)$, where $p \in M$ is any point in the fiber over q. (Such a point exists because we are assuming that π is surjective.) This is well-defined because F is constant on the fibers of π, and it satisfies $\widetilde{F} \circ \pi = F$ by construction. Thus $\widetilde{F}$ is smooth by Proposition 7.16. ☐

Our third proposition can be interpreted as a uniqueness result for smooth manifolds defined as quotients of other smooth manifolds by submersions.

Proposition 7.18 (Uniqueness of Smooth Quotients). *Suppose $\pi_1 \colon M \to N_1$ and $\pi_2 \colon M \to N_2$ are surjective submersions that are constant on each other's fibers. Then there exists a unique diffeomorphism $F \colon N_1 \to N_2$ such that $F \circ \pi_1 = \pi_2$:*

$$\begin{tikzcd} & M \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_2"] & \\ N_1 \arrow[rr, "F"'] & & N_2. \end{tikzcd}$$

◇ **Exercise 7.5.** Prove Proposition 7.18.

Problems

7-1. Using the covering map $\pi \colon \mathbb{R}^2 \to \mathbb{T}^2$ defined by $\pi(\varphi, \theta) = (e^{i\varphi}, e^{i\theta})$, show that the immersion $X \colon \mathbb{R}^2 \to \mathbb{R}^3$ defined in Example 7.1(f) descends to an embedding of $\mathbb{T}^2$ into $\mathbb{R}^3$.

7-2. Define a map $F \colon \mathbb{S}^2 \to \mathbb{R}^4$ by

$$F(x, y, z) = (x^2 - y^2, xy, xz, yz).$$

Using the smooth covering map $p \colon \mathbb{S}^2 \to \mathbb{RP}^2$ described in Example 2.5(d) and Problem 2-9, show that F descends to a smooth embedding of $\mathbb{RP}^2$ into $\mathbb{R}^4$.

7-3. Let $\gamma \colon \mathbb{R} \to \mathbb{T}^2$ be the curve of Example 7.3. Show that the image set $\gamma(\mathbb{R})$ is dense in $\mathbb{T}^2$.

7-4. Suppose M is a smooth manifold, $p \in M$, and $y^1, \dots, y^n$ are smooth real-valued functions defined on a neighborhood of p in M.

(a) If $dy^1|_p, \dots, dy^n|_p$ form a basis for T_p^*M, show that $(y^1, \dots, y^n)$ are smooth coordinates for M in some neighborhood of p.

(b) If $dy^1|_p, \dots, dy^n|_p$ are independent, show that there are real-valued functions $y^{n+1}, \dots, y^m$ such that $(y^1, \dots, y^m)$ are smooth coordinates for M in some neighborhood of p.

(c) If $dy^1|_p, \dots, dy^n|_p$ span T_p^*M, show that there are indices $i_1, \dots, i_k$ such that $(y^{i_1}, \dots, y^{i_k})$ are smooth coordinates for M in some neighborhood of p.

7-5. Let M be a smooth compact manifold. Show that there is no submersion $F \colon M \to \mathbb{R}^k$ for any $k > 0$.

7-6. Suppose $\pi \colon M \to N$ is a smooth map such that every point of M is in the image of a smooth local section of π. Show that π is a submersion.

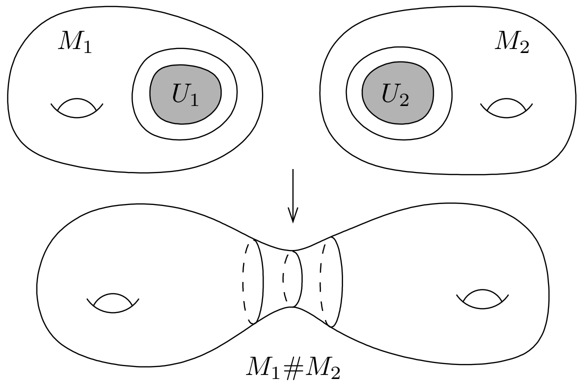


Figure 7.6. A connected sum.

7-7. Let M be a smooth n-manifold with boundary. Recall from Chapter 1 that a point $p \in M$ called a *boundary point* of M if $\varphi(p) \in \partial\mathbb{H}^n$ for some smooth chart (U, φ), and an *interior point* if $\varphi(p) \in \operatorname{Int}\mathbb{H}^n$ for some smooth chart. Show that the set of boundary points and the set of interior points are disjoint. [Hint: If $\varphi(p) \in \partial\mathbb{H}^n$ and $\psi(p) \in \operatorname{Int}\mathbb{H}^n$, consider $\varphi \circ \psi^{-1}$ as a map into $\mathbb{R}^n$.]

7-8. If $\pi\colon M \to N$ is a submersion and $X \in \mathfrak{T}(N)$, show that there is a smooth vector field on M (called a *lift of* X) that is π-related to X. Is it unique?

7-9. Suppose $\pi\colon M \to N$ is a surjective submersion. If X is a smooth vector field on M such that $\pi_* X_p = \pi_* X_q$ whenever $\pi(p) = \pi(q)$, show that there exists a unique smooth vector field on N that is π-related to X.

7-10. Let M_1, M_2 be connected smooth manifolds of dimension n. For $i = 1, 2$, let (W_i, φ_i) be a smooth coordinate domain centered at some point $p_i \in M_i$ such that $\varphi_i(W_i) = B_2(0) \subset \mathbb{R}^n$. Define $U_i = \varphi_i^{-1}(B_1(0)) \subset W_i$ and $M_i' = M_i \smallsetminus U_i$. The *connected sum* of M_1 and M_2, denoted by $M_1 \# M_2$, is the quotient space of $M_1' \amalg M_2'$ obtained by identifying each $q \in \partial U_1$ with $\varphi_2^{-1} \circ \varphi_1(q) \in \partial U_2$ (Figure 7.6). Show that $M_1 \# M_2$ is a connected topological n-manifold, and has a unique smooth structure such that the restriction of the quotient map to each M_i' is a smooth embedding (where M_i' is thought of as a smooth manifold with boundary). Show that there are open subsets $\widetilde{M}_1, \widetilde{M}_2 \subset M_1 \# M_2$ that are diffeomorphic to $M_1 \smallsetminus \{p_1\}$ and $M_2 \smallsetminus \{p_2\}$, respectively, and such that $\widetilde{M}_1 \cap \widetilde{M}_2$ is diffeomorphic to $B_2(0) \smallsetminus \{0\}$.

This is page 173
Printer: Opaque this

8
Submanifolds

Many of the most familiar examples of manifolds arise naturally as subsets of other manifolds—for example, the n-sphere is a subset of $\mathbb{R}^{n+1}$, and the n-torus $\mathbb{T}^n = \mathbb{S}^1 \times \cdots \times \mathbb{S}^1$ is a subset of $\mathbb{C} \times \cdots \times \mathbb{C} = \mathbb{C}^n$. In this chapter, we will explore conditions under which a subset of a smooth manifold can be considered as a smooth manifold in its own right. As you will soon discover, the situation is quite a bit more subtle than the analogous theory of topological subspaces.

We begin by defining the most important type of smooth submanifolds, called embedded submanifolds. These are modeled locally on linear subspaces of Euclidean space, and turn out to be exactly the images of smooth embeddings. Then, because submanifolds are most often presented as level sets of smooth maps, we devote some time to analyzing the conditions under which such sets are smooth manifolds. We will see that level sets of constant-rank maps (in particular, submersions) are always embedded submanifolds.

Next, we introduce a more general kind of submanifolds, called immersed submanifolds, which turn out to be the images of injective immersions. They look locally like embedded submanifolds, but may not be globally embedded. We also explore conditions under which smooth maps, vector fields, and covector fields can be restricted to submanifolds.

At the end of the chapter, we apply the theory of submanifolds to study Lie subgroups of Lie groups and subbundles of vector bundles.

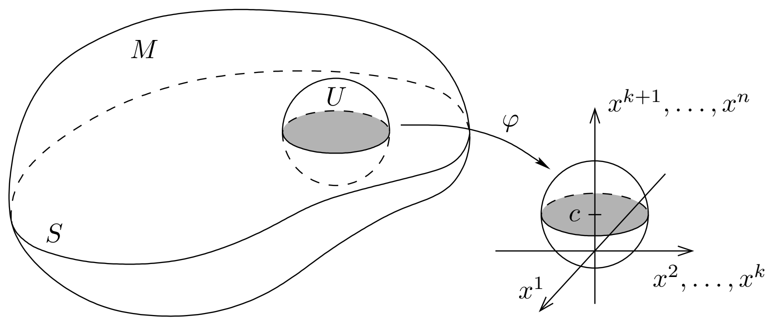


Figure 8.1. A slice chart.

Embedded Submanifolds

Smooth submanifolds are modeled locally on the standard embedding of $\mathbb{R}^k$ into $\mathbb{R}^n$, identifying $\mathbb{R}^k$ with the subspace

$$\{(x^1,\dots,x^k,x^{k+1},\dots,x^n) : x^{k+1} = \cdots = x^n = 0\}$$

of $\mathbb{R}^n$. Somewhat more generally, if U is an open subset of $\mathbb{R}^n$, a *k-slice* of U is any subset of the form

$$S = \{(x^1,\dots,x^k,x^{k+1},\dots,x^n) \in U : x^{k+1} = c^{k+1},\dots,x^n = c^n\}$$

for some constants $c^{k+1},\dots,c^n$. Clearly any k-slice is homeomorphic to an open subset of $\mathbb{R}^k$. (Sometimes it is convenient to consider slices defined by setting some other subset of the coordinates equal to constants instead of the last ones. The meaning should be clear from the context.)

Let M be a smooth n-manifold, and let (U,φ) be a smooth chart on M. If S is a subset of U such that $\varphi(S)$ is a k-slice of $\varphi(U)$, then we say simply that S is a k-slice of U. A subset $S \subset M$ is called an *embedded submanifold of dimension k* (or *embedded k-submanifold* for short) if for each point $p \in S$, there exists a smooth chart (U,φ) for M such that $p \in U$ and $U \cap S$ is a k-slice of U (Figure 8.1). In this situation, we call the chart (U,φ) a *slice chart for S in M*, and the corresponding coordinates $(x^1,\dots,x^n)$ are called *slice coordinates*. (Although in general we allow our slices to be defined by arbitrary constants $c^{k+1},\dots,c^n$, it is sometimes useful to have slice coordinates for which the constants are all zero, which can easily be achieved by subtracting a constant from each coordinate function.) If the dimension of S is understood or irrelevant, we will just call it an *embedded submanifold*. Embedded submanifolds are also called *regular submanifolds* by some authors.

If S is an embedded submanifold of M, the difference $\dim M - \dim S$ is called the *codimension of S in M*. An *embedded hypersurface* is an embedded submanifold of codimension 1. By convention, we consider an open submanifold to be an embedded submanifold of codimension zero.

The definition of an embedded submanifold is a local one. It is useful to express this formally as a lemma.

Lemma 8.1. *Let M be a smooth manifold and let S be a subset of M. Suppose that for some k, every point $p \in S$ has a neighborhood $U \subset M$ such that $U \cap S$ is an embedded k-submanifold of U. Then S is an embedded k-submanifold of M.*

◇ **Exercise 8.1.** Prove Lemma 8.1.

The next theorem explains the reason for the name "embedded submanifold."

Theorem 8.2. *Let $S \subset M$ be an embedded k-dimensional submanifold. With the subspace topology, S is a topological manifold of dimension k, and it has a unique smooth structure such that the inclusion map $S \hookrightarrow M$ is a smooth embedding.*

Proof. It is automatic that S is Hausdorff and second countable, because M is and both properties are inherited by subspaces. To see that it is locally Euclidean, we will construct an atlas. The basic idea of the construction is that if $(x^1, \ldots, x^n)$ are slice coordinates for S in M, we can use $(x^1, \ldots, x^k)$ as local coordinates for S.

For this proof, let $\pi\colon \mathbb{R}^n \to \mathbb{R}^k$ denote the projection onto the first k coordinates. For any slice chart (U, φ), let

$$\begin{aligned} V &= U \cap S, \\ \widetilde{V} &= \pi \circ \varphi(V), \\ \psi &= \pi \circ \varphi|_V \colon V \to \widetilde{V}. \end{aligned}$$

Then $\widetilde{V}$ is open in $\mathbb{R}^k$ because both φ and π are open maps, and ψ is a homeomorphism because it has a continuous inverse given by $\varphi^{-1} \circ j|_{\widetilde{V}}$, where $j\colon \mathbb{R}^k \to \mathbb{R}^n$ is an affine map of the form

$$j(x^1, \ldots, x^k) = (x^1, \ldots, x^k, c^{k+1}, \ldots, c^n).$$

Thus S is a topological k-manifold, and the inclusion map $\iota\colon S \hookrightarrow M$ is a topological embedding.

To put a smooth structure on S, we will verify that the charts constructed above are smoothly compatible. Suppose (U, φ) and (U', φ') are two slice charts for S in M, and let (V, ψ), (V', ψ') be the corresponding charts for S. The transition map is given by $\psi' \circ \psi^{-1} = \pi \circ \varphi' \circ \varphi^{-1} \circ j$, which is a composition of the smooth maps π, $\varphi' \circ \varphi^{-1}$, and j (Figure 8.2). Thus the atlas we have constructed is in fact a smooth atlas, and defines a smooth structure on S. Using a slice chart (U, φ) for M and the corresponding chart

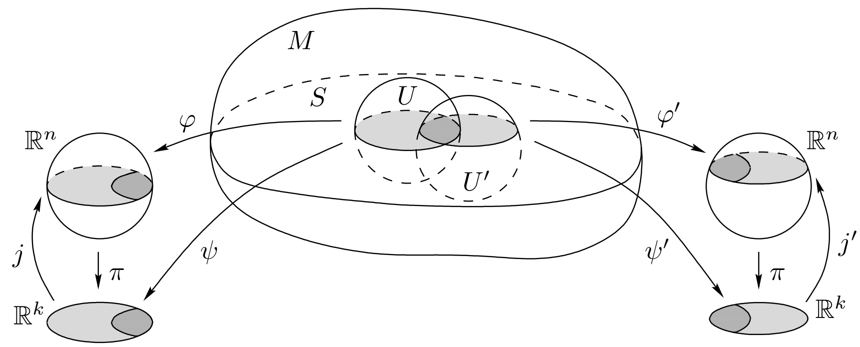


Figure 8.2. Smooth compatibility of slice charts.

(V, ψ) for S, the inclusion map $S \hookrightarrow M$ has a coordinate representation of the form

$$(x^1, \dots, x^k) \mapsto (x^1, \dots, x^k, c^{k+1}, \dots, c^n),$$

which is obviously an immersion. Since the inclusion is an injective immersion and a topological embedding, it is a smooth embedding as claimed.

The last thing we have to prove is that this is the unique smooth structure making the inclusion map a smooth embedding. Suppose that $\mathcal{A}$ is a (possibly different) smooth structure on S with the property that the inclusion map $\iota\colon (S, \mathcal{A}) \hookrightarrow M$ is a smooth embedding. To show that $\mathcal{A}$ is equal to the smooth structure we have constructed, it suffices to show that each of the charts we constructed above is compatible with every chart in $\mathcal{A}$. Thus let (U, φ) be a slice chart for S in M, let (V, ψ) be the corresponding chart for S constructed above, and let (W, θ) be an arbitrary chart in $\mathcal{A}$. We need to show that $\psi \circ \theta^{-1}\colon \theta(W \cap V) \to \psi(W \cap V)$ is a diffeomorphism (Figure 8.3).

Observe first that $\psi \circ \theta^{-1}$ is a composition of homeomorphisms and therefore is itself a homeomorphism. It is smooth because it can be written as the following composition of smooth maps:

$$\theta(W \cap V) \xrightarrow{\theta^{-1}} W \cap V \overset{\iota}{\hookrightarrow} U \xrightarrow{\varphi} \mathbb{R}^n \xrightarrow{\pi} \mathbb{R}^k,$$

where we think of $W \cap V$ as an open subset of S (with the smooth structure $\mathcal{A}$) and U as an open subset of M. By Corollary 7.10, to prove that it is a diffeomorphism, it suffices to show that it is an immersion.

By the argument above, $(\psi \circ \theta^{-1})_* = \pi_* \circ \varphi_* \circ \iota_* \circ (\theta^{-1})_*$. Each of the linear maps φ_*, ι_*, and $(\theta^{-1})_*$ is injective—in fact φ_* and $(\theta^{-1})_*$ are bijective—and thus their composition is injective. Although π_* is not injective, the composition will be injective provided $\operatorname{Im}(\varphi \circ \iota \circ \theta^{-1})_* \cap \operatorname{Ker} \pi_* = \{0\}$ (see Exercise A.41(d) in the Appendix). Since ι takes its values in S, $\varphi \circ \iota \circ \theta^{-1}$

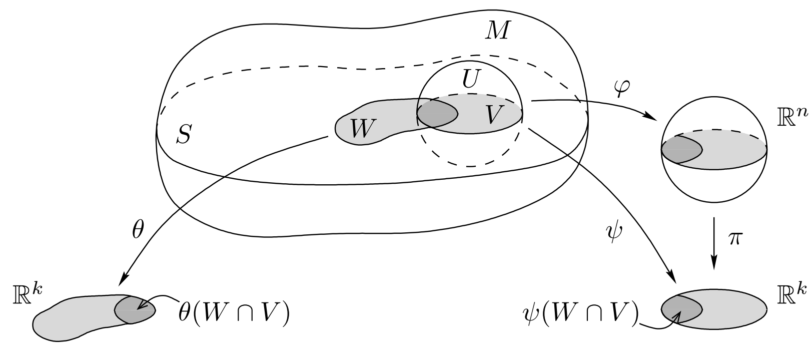


Figure 8.3. Uniqueness of the smooth structure on S.

takes its values in the slice where the coordinates $x^{k+1}, \dots, x^n$ are constant:

$$\varphi \circ \iota \circ \theta^{-1}(y^1, \dots, y^k) = (x^1(y), \dots, x^k(y), c^{k+1}, \dots, c^n).$$

It follows easily that the push-forward of this map at any point takes its values in the span of $(e_1, \dots, e_k)$, while $\operatorname{Ker} \pi_* = \operatorname{span}(e_{k+1}, \dots, e_n)$. □

This theorem has the following important converse.

Theorem 8.3. *The image of a smooth embedding is an embedded submanifold.*

Proof. Let $F\colon N \to M$ be a smooth embedding. We need to show that each point of $F(N)$ has a coordinate neighborhood $U \subset M$ in which $F(N) \cap U$ is a slice.

Let $p \in N$ be arbitrary. Since a smooth embedding has constant rank, by the rank theorem there are smooth charts (U, φ), (V, ψ) centered at p and $F(p)$ in which $F|_U\colon U \to V$ has the coordinate representation

$$F(x^1, \dots, x^k) = (x^1, \dots, x^k, 0, \dots, 0).$$

In particular (shrinking V if necessary), this implies that $F(U)$ is a slice in V. Because an embedding is a homeomorphism onto its image with the subspace topology, the fact that $F(U)$ is open in $F(N)$ means that there is an open set $W \subset M$ such that $F(U) = W \cap F(N)$. Replacing V by $\widetilde{V} = V \cap W$, we obtain a slice chart $(\widetilde{V}, \psi|_{\widetilde{V}})$ containing $F(p)$ such that $\widetilde{V} \cap F(N) = \widetilde{V} \cap F(U)$ is a slice of $\widetilde{V}$. □

The preceding two theorems can be summarized by the following corollary.

Corollary 8.4. *Embedded submanifolds are precisely the images of smooth embeddings.*

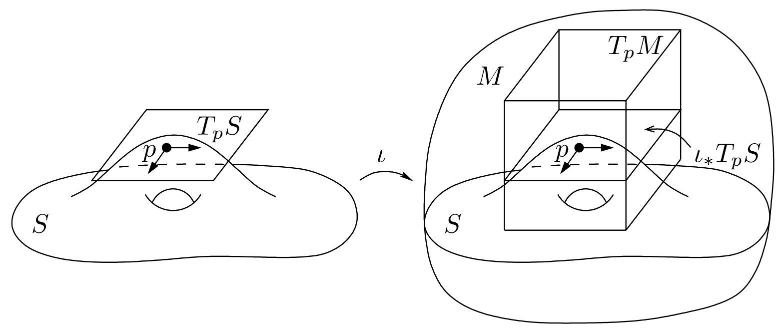


Figure 8.4. The tangent space to an embedded submanifold.

The Tangent Space to an Embedded Submanifold

If S is an embedded submanifold of $\mathbb{R}^n$, we intuitively think of the tangent space T_pS at a point of S as a subspace of the tangent space $T_p\mathbb{R}^n$. Similarly, the tangent space to a submanifold of an abstract smooth manifold can be viewed as a subspace of the tangent space to the ambient manifold, once we make appropriate identifications.

Let M be a smooth manifold, and let $S \subset M$ be an embedded submanifold. Since the inclusion map $\iota\colon S \hookrightarrow M$ is a smooth embedding, at each point $p \in S$ we have an injective linear map $\iota_*\colon T_pS \to T_pM$. In terms of derivations, this injection works in the following way: For any vector $X \in T_pS$, the image vector $\widetilde{X} = \iota_* X \in T_pM$ acts on smooth functions on M by

$$\widetilde{X} f = (\iota_* X) f = X(f \circ \iota) = X\left(f|_S\right).$$

We will adopt the convention of *identifying* T_pS with its image under this map, thereby thinking of T_pS as a certain linear subspace of T_pM (Figure 8.4).

The next proposition gives a useful way to characterize T_pS as a subspace of T_pM.

Proposition 8.5. *Suppose $S \subset M$ is an embedded submanifold and $p \in S$. As a subspace of T_pM, the tangent space T_pS is given by*

$$T_pS = \{X \in T_pM : Xf = 0 \text{ whenever } f \in C^\infty(M) \text{ and } f|_S \equiv 0\}.$$

Proof. First suppose $X \in T_pS \subset T_pM$. This means, more precisely, that $X = \iota_* Y$ for some $Y \in T_pS$. If f is any smooth real-valued function on M that vanishes on S, then $f \circ \iota \equiv 0$, so

$$Xf = (\iota_* Y) f = Y(f \circ \iota) \equiv 0.$$

Conversely, if $X \in T_pM$ satisfies $Xf = 0$ whenever f vanishes on S, we need to show that there is a vector $Y \in T_pS$ such that $X = \iota_* Y$. Let

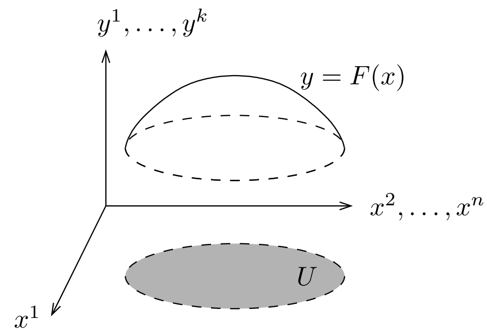


Figure 8.5. A graph is an embedded submanifold.

$(x^1, \dots, x^n)$ be slice coordinates for S in some neighborhood U of p, so that $U \cap S$ is the subset of U where $x^{k+1} = \cdots = x^n = 0$, and $(x^1, \dots, x^k)$ are coordinates for $U \cap S$. Because the inclusion map $\iota \colon S \cap U \hookrightarrow M$ has the coordinate representation

$$\iota(x^1, \dots, x^k) = (x^1, \dots, x^k, 0, \dots, 0)$$

in these coordinates, it follows that T_pS (that is, $\iota_* T_pS$) is exactly the subspace of T_pM spanned by $\partial/\partial x^1|_p, \dots, \partial/\partial x^k|_p$. If we write the coordinate representation of X as

$$X = \sum_{i=1}^{n} X^i \left.\frac{\partial}{\partial x^i}\right|_p,$$

we see that $X \in T_pS$ if and only if $X^i = 0$ for $i > k$.

Let φ be a smooth bump function supported in U that is equal to 1 in a neighborhood of p. Choose an index $j > k$, and consider the function $f(x) = \varphi(x)x^j$, extended to be zero on $M \smallsetminus U$. Then f vanishes identically on S, so

$$0 = Xf = \sum_{i=1}^{n} X^i \frac{\partial(\varphi(x)x^j)}{\partial x^i}(p) = X^j.$$

Thus $X \in T_pS$ as desired. □

Examples of Embedded Submanifolds

One straightforward way to construct embedded submanifolds is by using the graphs of smooth functions.

Lemma 8.6 (Graphs as Submanifolds). *If $U \subset \mathbb{R}^n$ is open and $F \colon U \to \mathbb{R}^k$ is smooth, then the graph of F is an embedded n-dimensional submanifold of $\mathbb{R}^{n+k}$ (see Figure 8.5).*

Proof. Define a map $\varphi\colon U \times \mathbb{R}^k \to U \times \mathbb{R}^k$ by

$$\varphi(x, y) = (x, y - F(x)).$$

It is clearly smooth, and in fact it is a diffeomorphism because its inverse can be written explicitly:

$$\varphi^{-1}(u, v) = (u, v + F(u)).$$

Because $\varphi(\Gamma(F))$ is the slice $\{(u, v) : v = 0\}$ of $U \times \mathbb{R}^k$, this shows that $\Gamma(F)$ is an embedded submanifold. □

Example 8.7 (Spheres). For any $n \geq 0$, $\mathbb{S}^n$ is an embedded submanifold of $\mathbb{R}^{n+1}$, because it is locally the graph of a smooth function: As we showed in Example 1.2, the intersection of $\mathbb{S}^n$ with the open set $\{x : x^i > 0\}$ is the graph of the smooth function

$$x^i = f(x^1, \dots, x^{i-1}, x^{i+1}, \dots, x^{n+1}),$$

where $f\colon \mathbb{B}^n \to \mathbb{R}$ is given by $f(u) = \sqrt{1 - |u|^2}$. Similarly, the intersection of $\mathbb{S}^n$ with $\{x : x^i < 0\}$ is the graph of $-f$. Since every point in $\mathbb{S}^n$ is in one of these sets, Lemma 8.1 shows that $\mathbb{S}^n$ is an embedded submanifold of $\mathbb{R}^{n+1}$. The smooth structure thus induced on $\mathbb{S}^n$ is the same as the one we defined in Chapter 1: In fact, the coordinates for $\mathbb{S}^n$ determined by these slice charts are exactly the graph coordinates we defined in Example 1.20.

◇ **Exercise 8.2.** Show that spherical coordinates (Example 7.11) form a slice chart for $\mathbb{S}^2$ in $\mathbb{R}^3$ on any open set where they are defined.

Level Sets

As Example 8.7 illustrates, showing directly from the definition that a subset of a manifold is an embedded submanifold can be somewhat cumbersome. We know that images of smooth embeddings are always embedded submanifolds (Theorem 8.3), so one way to show that a subset is an embedded submanifold is to exhibit it as the image of a smooth embedding. However, in practice, a submanifold is most often presented as the set of points where some map takes on a fixed value. If $\Phi\colon M \to N$ is any map and c is any point of N, the set $\Phi^{-1}(c)$ is called a *level set* of Φ (Figure 8.6). (In the special case $N = \mathbb{R}^k$ and $c = 0$, the level set $\Phi^{-1}(0)$ is usually called the *zero set* of Φ.) For example, the n-sphere $\mathbb{S}^n \subset \mathbb{R}^{n+1}$ is the level set $\Phi^{-1}(1)$, where $\Phi\colon \mathbb{R}^{n+1} \to \mathbb{R}$ is the function $\Phi(x) = |x|^2$.

We do not yet have an effective criterion for deciding when a level set of a smooth map is an embedded submanifold. It is easy to construct examples that are not: For instance, consider the two smooth maps $\Phi, \Psi\colon \mathbb{R}^2 \to \mathbb{R}$ defined by $\Phi(x, y) = x^3 - y^2$ and $\Psi(x, y) = x^2 - y^2$ (Figure 8.7). The zero set of Φ is a curve that has a "cusp" or "kink" at the origin, while the

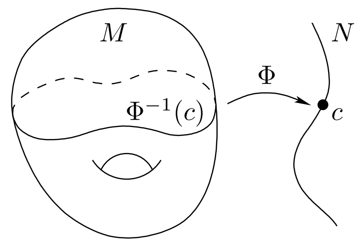


Figure 8.6. A level set.

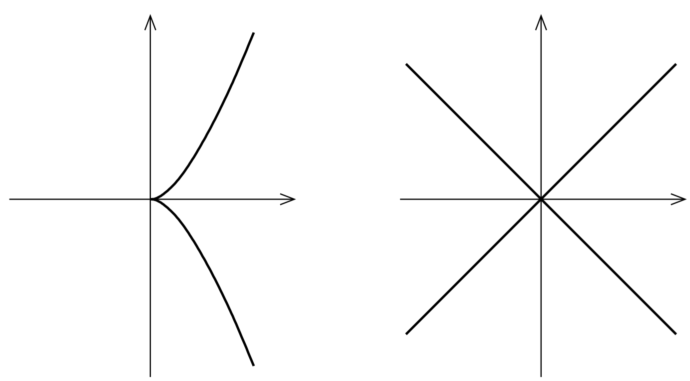


Figure 8.7. Level sets that are not embedded submanifolds.

zero set of Ψ is the union of the lines $x = y$ and $x = -y$. As Problem 8-13 shows, neither of these sets is an embedded submanifold of $\mathbb{R}^2$.

In this section, we will use the tools we developed in Chapter 7 to give some very general criteria for level sets to be submanifolds. To set the stage, consider first a linear version of the problem. Any k-dimensional linear subspace $S \subset \mathbb{R}^n$ is the kernel of some linear map. (Such a linear map is easily constructed by choosing a basis for S and extending it to a basis for $\mathbb{R}^n$.) By the rank-nullity law, if $S = \operatorname{Ker} L$, then $\operatorname{Im} L$ must have dimension $n-k$. Therefore, a natural way to specify a k-dimensional subspace $S \subset \mathbb{R}^n$ is to give a surjective linear map $L\colon \mathbb{R}^n \to \mathbb{R}^{n-k}$ whose kernel is S. The vector equation $Lx = 0$ is equivalent to $n-k$ independent scalar equations, each of which can be thought of as cutting down one of the degrees of freedom in $\mathbb{R}^n$, leaving a subspace S of dimension k.

In the context of smooth manifolds, the analogue of a surjective linear map is a submersion. Thus we might expect that a level set of a submersion from an n-manifold to an $(n-k)$-manifold should be an embedded k-dimensional submanifold. We will see below that this is the case. In fact, thanks to the rank theorem, something even stronger is true.

Theorem 8.8 (Constant Rank Level Set Theorem). *Let M and N be smooth manifolds, and let $\Phi\colon M \to N$ be a smooth map with constant rank equal to k. Each level set of Φ is a closed embedded submanifold of codimension k in M.*

Proof. Let $c \in N$ be arbitrary, and let S denote the level set $\Phi^{-1}(c) \subset M$. Clearly S is closed in M by continuity. To show it is an embedded submanifold, we need to show that for each $p \in S$ there is a slice chart for S in M near p. From the rank theorem, there are smooth charts (U, φ) centered at p and (V, ψ) centered at $c = \Phi(p)$ in which Φ has a coordinate representation of the form (7.11), and therefore $S \cap U$ is the slice $\{(x^1, \dots, x^m) \in U : x^1 = \cdots = x^k = 0\}$. □

Corollary 8.9 (Submersion Level Set Theorem). *If $\Phi\colon M \to N$ is a submersion, then each level set of Φ is a closed embedded submanifold whose codimension is equal to the dimension of N.*

Proof. A submersion has constant rank equal to the dimension of N. □

This result can be strengthened considerably, because we need only check the rank condition on the level set we are interested in. If $\Phi\colon M \to N$ is a smooth map, a point $p \in M$ is said to be a *regular point* of Φ if $\Phi_*\colon T_pM \to T_{\Phi(p)}N$ is surjective; it is a *critical point* otherwise. (This means, in particular, that every point is critical if $\dim M < \dim N$.) A point $c \in N$ is said to be a *regular value* of Φ if every point of the level set $\Phi^{-1}(c)$ is a regular point, and a *critical value* otherwise. In particular, if $\Phi^{-1}(c) = \varnothing$, c is regular. Finally, a level set $\Phi^{-1}(c)$ is called a *regular level set* if c is a regular value; in other words, a regular level set is a level set consisting entirely of regular points.

Corollary 8.10 (Regular Level Set Theorem). *Every regular level set of a smooth map is a closed embedded submanifold whose codimension is equal to the dimension of the range.*

Proof. Let $\Phi\colon M \to N$ be a smooth map and let $c \in N$ be a regular value such that $\Phi^{-1}(c) \neq \varnothing$. The fact that c is a regular value means that Φ_* has rank equal to the dimension of N at every point of $\Phi^{-1}(c)$. To prove the corollary, it suffices to show that the set U of points where $\operatorname{rank}\Phi_* = \dim N$ is open in M, for then $\Phi|_U\colon U \to N$ is a submersion, and we can apply the preceding corollary with M replaced by U, noting that an embedded submanifold of U is also an embedded submanifold of M.

To see that U is open, let $m = \dim M$, $n = \dim N$, and suppose $p \in U$. Choosing smooth coordinates near p and $\Phi(p)$, the assumption

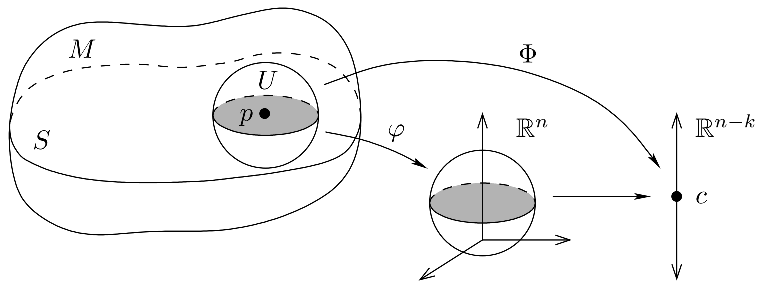


Figure 8.8. An embedded submanifold is locally a level set.

that rank $\Phi_* = n$ at p means that the $n \times m$ matrix representing Φ_* in coordinates has an $n \times n$ minor whose determinant is nonzero. By continuity, this determinant will be nonzero in some neighborhood of p, which means that Φ has rank n in this whole neighborhood. $\square$

$\diamond$ **Exercise 8.3.** If $f\colon M \to \mathbb{R}$ is a smooth real-valued function, show that $p \in M$ is a regular point of f if and only if $df_p \neq 0$.

Example 8.11 (Spheres). Now we can give a much easier proof that $\mathbb{S}^n$ is an embedded submanifold of $\mathbb{R}^{n+1}$. The sphere is easily seen to be a regular level set of the function $f\colon \mathbb{R}^{n+1} \to \mathbb{R}$ given by $f(x) = |x|^2$, since $df = 2\sum_i x^i\, dx^i$ vanishes only at the origin. Thus $\mathbb{S}^n$ is an embedded n-dimensional submanifold of $\mathbb{R}^{n+1}$.

Not all embedded submanifolds are naturally given as level sets of submersions as in these examples. However, the next proposition shows that every embedded submanifold is at least locally of this form.

Proposition 8.12. *Let S be a subset of a smooth n-manifold M. Then S is an embedded k-submanifold of M if and only if every point $p \in S$ has a neighborhood U in M such that $U \cap S$ is a level set of a submersion $\Phi\colon U \to \mathbb{R}^{n-k}$.*

Proof. First suppose S is an embedded k-submanifold. If $(x^1, \dots, x^n)$ are slice coordinates for S on an open set $U \subset M$, the map $\Phi\colon U \to \mathbb{R}^{n-k}$ given in coordinates by $\Phi(x) = (x^{k+1}, \dots, x^n)$ is easily seen to be a submersion one of whose level sets is $S \cap U$ (Figure 8.8). Conversely, suppose that around every point $p \in S$ there is a neighborhood U and a submersion $\Phi\colon U \to \mathbb{R}^{n-k}$ such that $S \cap U = \Phi^{-1}(c)$ for some $c \in \mathbb{R}^{n-k}$. By the submersion level set theorem, $S \cap U$ is an embedded submanifold of U, and so by Lemma 8.1, S is itself an embedded submanifold. $\square$

If $S \subset M$ is an embedded submanifold, a smooth map $\Phi\colon M \to N$ such that S is a regular level set of Φ is called a *defining map* for S. In the special case $N = \mathbb{R}^{n-k}$ (so that Φ is a real-valued or vector-valued function), it is usually called a *defining function*. Example 8.11 shows that $f(x) = |x|^2$ is a defining function for the sphere. More generally, if U is an open subset of M and $\Phi\colon U \to N$ is a smooth map such that $S \cap U$ is a regular level set of Φ, then Φ is called a *local defining map* (or *local defining function*) for S. Proposition 8.12 says that every embedded submanifold admits a local defining function in a neighborhood of each of its points. If $S \subset M$ is the zero set of a defining function $\Phi\colon M \to \mathbb{R}^m$, then S is the set of common zeros of the component functions of Φ; just as in the linear case, these m functions can be thought of as cutting down the number of degrees of freedom from n to $n - m$.

In specific examples, finding a (local or global) defining function for a submanifold is usually just a matter of using geometric information about how the submanifold is defined together with some computational ingenuity. Here is an example.

Example 8.13. Let D be the doughnut-shaped torus of revolution described in Example 7.1(f). Because D is obtained by revolving the circle $(y-2)^2 + z^2 = 1$ around the z-axis, a point (x, y, z) is in D if and only if it satisfies $(r-2)^2 + z^2 = 1$, where $r = \sqrt{x^2 + y^2}$ is the distance from the z axis. Thus D is the zero set of the function $\Phi(x, y, z) = (r-2)^2 + z^2 - 1$, which is smooth on $\mathbb{R}^3$ minus the z-axis. A straightforward computation shows that $d\Phi$ does not vanish on D, so Φ is a global defining function for D.

Our next example is a bit more complicated. It will be of use to us in the next chapter.

Example 8.14 (Matrices of Fixed Rank). As in Chapter 1, let $\mathrm{M}(m \times n, \mathbb{R})$ denote the mn-dimensional vector space of $m \times n$ real matrices. For any k, let $\mathrm{M}_k(m \times n, \mathbb{R})$ denote the subset of $\mathrm{M}(m \times n, \mathbb{R})$ consisting of matrices of rank k. We showed in Example 1.19 that $\mathrm{M}_k(m \times n, \mathbb{R})$ is an open submanifold of $\mathrm{M}(m \times n, \mathbb{R})$ when $k = \min(m, n)$. Now we will show that when $0 \le k \le \min(m, n)$, $\mathrm{M}_k(m \times n, \mathbb{R})$ is an embedded submanifold of codimension $(m-k)(n-k)$ in $\mathrm{M}(m \times n, \mathbb{R})$.

Let E_0 be an arbitrary $m \times n$ matrix of rank k. This implies that E_0 has some $k \times k$ minor with nonzero determinant. For the time being, let us assume that it is the upper left minor. Writing E_0 in block form as

$$E_0 = \begin{pmatrix} A_0 & B_0 \\ C_0 & D_0 \end{pmatrix},$$

where A_0 is a $k \times k$ matrix and D_0 is of size $(m-k) \times (n-k)$, our assumption is that A_0 is nonsingular.

Let U be the set

$$U = \left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{M}(m \times n, \mathbb{R}) : \det A \ne 0 \right\}.$$

By continuity of the determinant function, U is an open subset of $\mathrm{M}(m \times n, \mathbb{R})$ containing E_0. Given $E = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right) \in U$, consider the invertible $n \times n$ matrix

$$P = \begin{pmatrix} A^{-1} & -A^{-1}B \\ 0 & I_{n-k} \end{pmatrix}.$$

Since multiplication by an invertible matrix does not change the rank of a matrix, the rank of E is the same as that of

$$EP = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} A^{-1} & -A^{-1}B \\ 0 & I_{n-k} \end{pmatrix} = \begin{pmatrix} I_k & 0 \\ CA^{-1} & D - CA^{-1}B \end{pmatrix}. \tag{8.1}$$

Clearly EP has rank k if and only if $D - CA^{-1}B$ is the zero matrix. (To understand where P came from, observe that E has rank k if and only if it can be reduced by elementary column operations to a matrix whose first k columns are independent and whose last $n - k$ columns are zero. Since elementary column operations correspond to right multiplication by invertible matrices, it is natural, at least in retrospect, to look for a matrix P satisfying (8.1).)

Thus we are led to define $\Phi\colon U \to \mathrm{M}((m-k) \times (n-k), \mathbb{R})$ by

$$\Phi \begin{pmatrix} A & B \\ C & D \end{pmatrix} = D - CA^{-1}B.$$

Clearly Φ is smooth. To show that it is a submersion, we need to show that $D\Phi(E)$ is surjective for each $E \in U$. Since $\mathrm{M}((m-k) \times (n-k), \mathbb{R})$ is a vector space, tangent vectors at $\Phi(E)$ can be naturally identified with $(m-k) \times (n-k)$ matrices. Given $E = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ and any matrix $X \in \mathrm{M}((m-k) \times (n-k), \mathbb{R})$, define a curve $\gamma\colon \mathbb{R} \to U$ by

$$\gamma(t) = \begin{pmatrix} A & B \\ C & D + tX \end{pmatrix}.$$

Then

$$\Phi_*\gamma'(0) = (\Phi \circ \gamma)'(t) = \left.\frac{d}{dt}\right|_{t=0} (D + tX - CA^{-1}B) = X.$$

Thus Φ is a submersion and so $\mathrm{M}_k(m \times n, \mathbb{R}) \cap U$ is an embedded submanifold of U.

Now if E_0' is an arbitrary matrix of rank k, just note that it can be transformed to one in U by a rearrangement of its rows and columns. Such a rearrangement is a linear isomorphism $R\colon \mathrm{M}(m \times n, \mathbb{R}) \to \mathrm{M}(m \times n, \mathbb{R})$ that preserves rank, so $U' = R^{-1}(U)$ is a neighborhood of E_0' and $\Phi \circ R\colon U' \to \mathrm{M}((m-k) \times (n-k), \mathbb{R})$ is a submersion whose zero set is $\mathrm{M}_k(m \times n, \mathbb{R}) \cap U'$. Thus every point in $\mathrm{M}_k(m \times n, \mathbb{R})$ has a neighborhood U' in $\mathrm{M}(m \times n, \mathbb{R})$ such that $U' \cap \mathrm{M}_k(m \times n, \mathbb{R})$ is an embedded submanifold of U', so $\mathrm{M}_k(m \times n, \mathbb{R})$ is an embedded submanifold by Lemma 8.1.

The next lemma shows that defining maps give a concise characterization of the tangent space to an embedded submanifold.

Lemma 8.15. *Suppose $S \subset M$ is an embedded submanifold. If $\Phi\colon U \to N$ is any local defining map for S, then $T_pS = \operatorname{Ker}\Phi_*\colon T_pM \to T_{\Phi(p)}N$ for each $p \in S \cap U$.*

Proof. Recall that we identify T_pS with the subspace $\iota_*(T_pS) \subset T_pM$, where $\iota\colon S \hookrightarrow M$ is the inclusion map. Because $\Phi \circ \iota$ is constant on $S \cap U$, it follows that $\Phi_* \circ \iota_*$ is the zero map from T_pS to $T_{\Phi(p)}N$, and therefore $\operatorname{Im}\iota_* \subset \operatorname{Ker}\Phi_*$. On the other hand, Φ_* is surjective by the definition of a defining map, so the rank-nullity law implies that

$$\dim \operatorname{Ker}\Phi_* = \dim T_pM - \dim T_{\Phi(p)}N = \dim T_pS = \dim \operatorname{Im}\iota_*,$$

which implies that $\operatorname{Im}\iota_* = \operatorname{Ker}\Phi_*$. □

Immersed Submanifolds

Although embedded submanifolds are the most natural and common submanifolds and suffice for most purposes, it is sometimes important to consider a more general notion of submanifold. In particular, when we study Lie subgroups later in this chapter and foliations in Chapter 19, we will encounter subsets of smooth manifolds that are images of injective immersions, but not necessarily of embeddings. To see some of the kinds of phenomena that occur, look back again at the two examples we introduced earlier of injective immersions that are *not* embeddings. The "figure eight curve" of Example 7.2 and the dense curve on the torus of Example 7.3 are both images of injective immersions that are not embeddings. In fact, their image sets are not embedded submanifolds (see Problems 8-3 and 8-5).

So as to have a convenient language for talking about examples like these, we make the following definition. Let M be a smooth manifold. An *immersed submanifold of dimension k* (or *immersed k-submanifold*) of M is a subset $S \subset M$ endowed with a k-manifold topology (not necessarily the subspace topology) together with a smooth structure such that the inclusion map $\iota\colon S \hookrightarrow M$ is a smooth immersion (Figure 8.9). As for embedded submanifolds, we define the *codimension of N in M* to be $\dim M - \dim N$.

Clearly every embedded submanifold is also an immersed submanifold. More generally, immersed submanifolds usually arise in the following way. Given an injective immersion $F\colon N \to M$, we can give the image set $F(N) \subset M$ a unique manifold topology and smooth structure such that $F\colon N \to F(N)$ is a diffeomorphism: We simply declare a set $U \subset F(N)$ to be open if and only if $F^{-1}(U) \subset N$ is open, and take the smooth coordinate maps on $F(N)$ to be the maps of the form $\varphi \circ F^{-1}$, where φ is a smooth coordinate map for N. With this smooth manifold structure, $\iota\colon F(N) \hookrightarrow M$ is an injective immersion, because it is equal to the

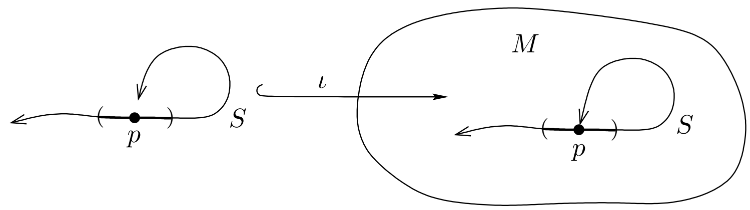


Figure 8.9. An immersed submanifold.

composition of a diffeomorphism followed by an injective immersion:

$$F(N) \xrightarrow{F^{-1}} N \xrightarrow{F} M.$$

The following proposition is an analogue for immersed submanifolds of Corollary 8.4.

Proposition 8.16. *Immersed submanifolds are precisely the images of injective immersions.*

Proof. If $N \subset M$ is an immersed submanifold, the inclusion map $\iota\colon N \hookrightarrow M$ is by definition an injective immersion. Conversely, the discussion above showed that the image of any injective immersion has a unique topology and smooth structure making it into an immersed submanifold such that the given immersion is a diffeomorphism onto its image. $\square$

Example 8.17 (Immersed Submanifolds). Because the figure eight of Example 7.2 and the dense curve of Example 7.3 are images of injective immersions, they are immersed submanifolds when given appropriate topologies and smooth structures. As smooth manifolds, they are diffeomorphic to $\mathbb{R}$. They are *not* embedded, because neither one has the subspace topology.

Because immersed submanifolds are the more general of the two types of submanifolds, the term "submanifold" without further qualification means an immersed submanifold, which includes an embedded submanifold as a special case. (Since we are not defining any submanifolds other than smooth ones, both types will be understood to be smooth.) Similarly, we will use the term "hypersurface" without qualification to mean an (immersed or embedded) submanifold of codimension 1. If there is room for confusion, it is usually better to specify explicitly which type of submanifold is meant, particularly because some authors do not follow this convention, but instead reserve the unqualified term "submanifold" to mean what we call an embedded submanifold.

Even though an immersed submanifold $S \subset M$ may not be a topological subspace of M, its tangent space at any point $p \in S$ can nonetheless be viewed as a linear subspace of T_pM, as for an embedded submanifold. If

$\iota\colon S \hookrightarrow M$ is the inclusion map, then ι is a smooth immersion, so $\iota_*\colon T_pS \to T_pM$ is injective. Just as in the case of embedded submanifolds, we will routinely identify T_pS with the subspace $\iota_* T_pS \subset T_pM$.

When deciding whether a given subset S of a smooth manifold M is a submanifold, it is important to bear in mind that there are two very different questions one can ask. The simplest question is whether S is an embedded submanifold. This is simply a question about the subset S itself, with its subspace topology—either it is an embedded submanifold or it is not, and if so it has a unique smooth structure making it into an embedded submanifold. (In fact, an even stronger form of uniqueness holds; see Problem 8-12.) A more subtle question is whether S can be an immersed submanifold. In this case, neither the topology nor the smooth structure is known in advance, so one needs to ask whether there exist *any* topology and smooth structure on S making it into an immersed submanifold. This question is not always straightforward to answer (see Problem 8-13).

Although many immersed submanifolds are not embedded, the following lemma shows that the *local* structure of an immersed submanifold is the same as that of an embedded one.

Lemma 8.18. *Let $F\colon N \to M$ be an immersion. Then F is locally an embedding: For any $p \in N$, there exists a neighborhood U of p in N such that $F|_U\colon U \to M$ is a smooth embedding.*

◇ **Exercise 8.4.** Prove Lemma 8.18.

It is important to be clear about what this lemma does and does not say. Given an immersed submanifold $N \subset M$ and a point $p \in N$, it is possible to find a neighborhood U of p (in N) such that U is embedded; but it may not be possible to find a neighborhood V of p *in M* such that $V \cap N$ is embedded.

Recall that a smooth local parametrization of a manifold N is a smooth map $F\colon U \to N$ whose domain is an open subset $U \subset \mathbb{R}^n$, whose image $F(U)$ is an open subset of N, and whose inverse is a smooth coordinate map. Local parametrizations are often especially useful for describing submanifolds, because a local parametrization of a submanifold $S \subset M$ can be thought of either as a smooth map into S or as a smooth map into M whose image lies in S (because the inclusion $S \hookrightarrow M$ is smooth).

Example 8.19. The map $F\colon \mathbb{B}^2 \to \mathbb{S}^2$ given by

$$F(u,v) = (u, v, \sqrt{1 - u^2 - v^2})$$

is a smooth local parametrization of $\mathbb{S}^2$ whose image is the open upper hemisphere. The map $\gamma\colon (-\pi/2, \pi/2) \to \mathbb{R}^2$ given by $\gamma(t) = (\sin 2t, \cos t)$ is a smooth local parametrization of the immersed figure eight curve of Example 7.2, whose image is the portion of the curve in the open upper half-plane.

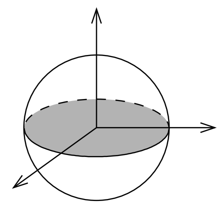


Figure 8.10. A submanifold (with boundary) of a manifold with boundary.

Submanifolds of Manifolds with Boundary

The definitions of this section extend easily to manifolds with boundary. First, if M and N are manifolds with boundary, a smooth map $F\colon M \to N$ is said to be a immersion if F_* is injective at each point, a submersion if F_* is surjective at each point, and a smooth embedding if it is an immersion and a topological embedding. A subset $S \subset M$ is said to be an *immersed submanifold* of M if S is endowed with a smooth manifold structure such that the inclusion map is an immersion, and an *embedded submanifold* if in addition S has the subspace topology. (We do not necessarily require the existence of slice coordinates for embedded submanifolds, because such coordinates can be problematic if S contains boundary points of M.)

More generally, an immersed or embedded *submanifold with boundary* in M is defined in exactly the same way, except that now S itself is allowed to be a manifold with boundary. For example, for any $k \le n$, the closed unit k-dimensional ball $\overline{\mathbb{B}^k}$ is an embedded submanifold with boundary in $\overline{\mathbb{B}^n}$, because the inclusion map $\overline{\mathbb{B}^k} \hookrightarrow \overline{\mathbb{B}^n}$ is easily seen to be a smooth embedding (Figure 8.10).

◇ **Exercise 8.5.** If M is a smooth manifold with boundary, show that ∂M is an embedded submanifold of M (without boundary).

Restricting Maps to Submanifolds

Given a smooth map $F\colon M \to N$, it is important to know whether F is still smooth when its domain or range is restricted to a submanifold. In the case of restricting the domain, the answer is easy.

Proposition 8.20 (Restricting the Domain of a Smooth Map). *If $F\colon M \to N$ is a smooth map and $S \subset M$ is an* (*immersed or embedded*) *submanifold* (*Figure 8.11*), *then $F|_S\colon S \to N$ is smooth.*

Proof. The inclusion map $\iota\colon S \hookrightarrow M$ is smooth by definition of an immersed submanifold. Since $F|_S = F \circ \iota$, the result follows. $\square$

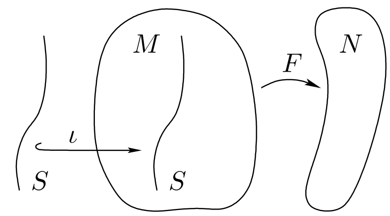


Figure 8.11. Restricting the domain.

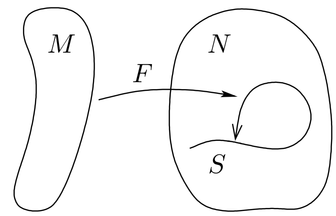


Figure 8.12. Restricting the range.

When the range is restricted, however, the resulting map may not be smooth, as the following example shows.

Example 8.21. Let $S \subset \mathbb{R}^2$ be the figure eight submanifold, with the topology and smooth structure induced by the immersion γ of Example 7.2. Define a smooth map $G\colon \mathbb{R} \to \mathbb{R}^2$ by

$$G(t) = (\sin 2t, \cos t).$$

(This is the same formula that we used to define γ, but now the domain is extended to the whole real line instead of being just a subinterval.) It is easy to check that the image of G lies in S. However, as a map from $\mathbb{R}$ to S, G is not even continuous, because $\gamma^{-1} \circ G$ is not continuous at $t = -\pi/2$.

The next proposition gives sufficient conditions for a map to be smooth when its range is restricted to an immersed submanifold. It shows that the failure of continuity is the only thing that can go wrong.

Proposition 8.22 (Restricting the Range of a Smooth Map). *Let $S \subset N$ be an immersed submanifold, and let $F\colon M \to N$ be a smooth map whose image is contained in S (Figure 8.12). If F is continuous as a map from M to S, then $F\colon M \to S$ is smooth.*

Proof. Let $p \in M$ be arbitrary and let $q = F(p) \in S$. Because the inclusion map $\iota\colon S \hookrightarrow N$ is an immersion, Lemma 8.18 guarantees that there is a neighborhood V of q in S such that $\iota|_V\colon V \hookrightarrow N$ is a smooth embedding. Thus there exists a slice chart (W, ψ) for V in N centered at q (Figure 8.13). (Of course, it might not be a slice chart for S in N.) The fact that (W, ψ) is a slice chart means that $(V_0, \widetilde{\psi})$ is a smooth chart for V, where $V_0 = W \cap V$ and $\widetilde{\psi} = \pi \circ \psi$, with $\pi\colon \mathbb{R}^n \to \mathbb{R}^k$ the projection onto the first $k = \dim S$ coordinates. Since $V_0 = (\iota|_V)^{-1}(W)$ is open in V, it is open in S in its given topology, and so $(V_0, \widetilde{\psi})$ is also a smooth chart for S.

Let $U = F^{-1}(V_0) \subset M$, which is an open set containing p. (Here is where we use the hypothesis that F is continuous.) Choose a smooth chart (U_0, φ) for M such that $p \in U_0 \subset U$. Then the coordinate representation

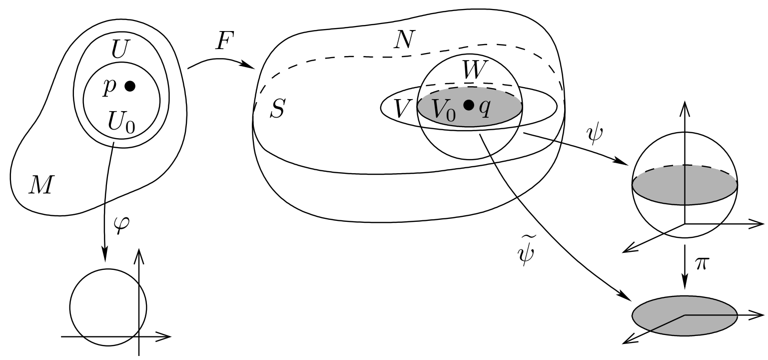


Figure 8.13. Proof of Proposition 8.22.

of $F\colon M \to S$ with respect to the charts (U_0, φ) and $(V_0, \widetilde{\psi})$ is

$$\widetilde{\psi} \circ F \circ \varphi^{-1} = \pi \circ (\psi \circ F \circ \varphi^{-1}),$$

which is smooth because $F\colon M \to N$ is smooth. □

In the special case in which the submanifold S is embedded, the continuity hypothesis is always satisfied.

Corollary 8.23 (Embedded Case). *Let $S \subset N$ be an embedded submanifold. Then any smooth map $F\colon M \to N$ whose image is contained in S is also smooth as a map from M to S.*

Proof. Since $S \subset N$ has the subspace topology, a continuous map $F\colon M \to N$ whose image is contained in S is automatically continuous into S, by the characteristic property of the subspace topology (Lemma A.5(a) in the Appendix). □

Vector Fields and Covector Fields on Submanifolds

If $S \subset M$ is an immersed or embedded submanifold, a vector field X on M does not necessarily restrict to a vector field on S, because X_p may not lie in the subspace $T_pS \subset T_pM$ at a point $p \in S$. A vector field X on M is said to be *tangent to S* if $X_p \in T_pS \subset T_pM$ for each $p \in S$ (Figure 8.14).

Lemma 8.24. *Let M be a smooth manifold, let $S \subset M$ be an embedded submanifold, and let Y be a smooth vector field on M. Then Y is tangent to S if and only if Yf vanishes on S for every $f \in C^\infty(M)$ such that $f|_S \equiv 0$.*

Proof. This is an immediate consequence of Proposition 8.5. □

Suppose $S \subset M$ is an immersed submanifold, and let Y be a smooth vector field on M. If there is a vector field $X \in \mathfrak{T}(S)$ that is ι-related to

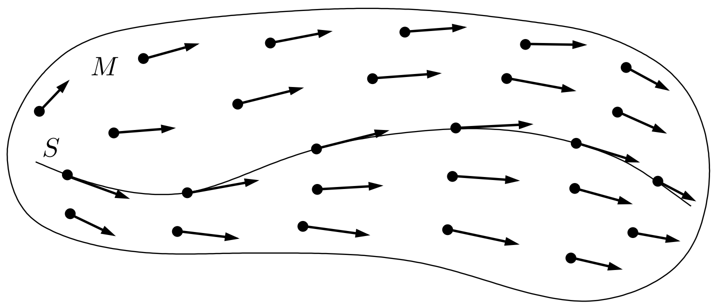


Figure 8.14. A vector field tangent to a submanifold.

Y, where $\iota\colon S \hookrightarrow M$ is the inclusion map, then clearly Y is tangent to S, because $Y_p = \iota_* X_p$ is in the image of ι_* for each $p \in S$. The next proposition shows that the converse is true.

Proposition 8.25 (Restricting Vector Fields to Submanifolds). *Let $S \subset M$ be an immersed submanifold, and let $\iota\colon S \hookrightarrow M$ denote the inclusion map. If $Y \in \mathcal{T}(M)$ is tangent to S, then there is a unique smooth vector field on S, denoted by $Y|_S$, that is ι-related to Y.*

Proof. The fact that Y is tangent to S means by definition that Y_p is in the image of ι_* for each p. Thus for each p there is a vector $X_p \in T_pS$ such that $Y_p = \iota_* X_p$. Since ι_* is injective, X_p is unique, so this defines X as a rough vector field on S. If we can show that X is smooth, it is the unique vector field that is ι-related to Y. It suffices to show that it is smooth in a neighborhood of each point.

Let p be any point in S. Since an immersed submanifold is locally embedded, there is a neighborhood V of p in S that is embedded in M. Let U be the domain of a slice chart for V in M near p, and let $S = V \cap U$, which is a slice in U and thus a closed subset of U. If $f \in C^\infty(S)$ is arbitrary, let $\widetilde{f} \in C^\infty(U)$ be an extension of f to a smooth function on U. (Such an extension exists by the extension lemma, Lemma 2.27.) Then for each $p \in S$,

$$Xf(p) = X_p(f) = X_p(\widetilde{f}|_S) = X_p(\widetilde{f} \circ \iota) = \iota_* X_p(\widetilde{f}) = Y_p(\widetilde{f}) = Y\widetilde{f}(p).$$

It follows that $Xf = (Y\widetilde{f})|_S$, which is smooth on S. Since the same is true in a neighborhood of each point of S, X is smooth by Lemma 4.2. □

This result has the following important consequence for Lie brackets, which will play a role in Chapter 19.

Corollary 8.26. *Let M be a smooth manifold and let S be an immersed submanifold. If Y_1 and Y_2 are smooth vector fields on M that are tangent to S, then $[Y_1, Y_2]$ is tangent to S as well.*

Proof. By the preceding lemma, there exist smooth vector fields X_1 and X_2 on S such that X_i is ι-related to Y_i for $i = 1, 2$ (where $\iota\colon S \to M$ is the inclusion). By the naturality of Lie brackets (Proposition 4.16), $[X_1, X_2]$ is ι-related to $[Y_1, Y_2]$, and is therefore tangent to S. □

The restriction of covector fields to submanifolds is much simpler. Suppose $S \subset M$ is an immersed submanifold, and let $\iota\colon S \hookrightarrow M$ denote the inclusion map. If ω is any smooth covector field on M, the pullback by ι yields a smooth covector field $\iota^*\omega$ on S. To see what this means, let $X_p \in T_pS$ be arbitrary, and compute

$$\begin{aligned}(\iota^*\omega)_p(X_p) &= \omega_p(\iota_* X_p)\\ &= \omega_p(X_p),\end{aligned}$$

since $\iota_*\colon T_pS \to T_pM$ is just the inclusion map, under our usual identification of T_pS with a subspace of T_pM. Thus $\iota^*\omega$ is just the restriction of ω to vectors tangent to S. For this reason we often write $\omega|_S$ in place of $\iota^*\omega$, and call it the *restriction* of ω to S. Be warned, however, that $\omega|_S$ might equal zero at a given point of S, even though *considered as a covector field on* M, ω might not vanish there. An example will help to clarify this distinction.

Example 8.27. Let $\omega = dy$ on $\mathbb{R}^2$, and let S be the x-axis, considered as a submanifold of $\mathbb{R}^2$. As a covector field on $\mathbb{R}^2$, ω does not vanish at any point, because one of its components is always 1. However, the restriction $\omega|_S$ is identically zero:

$$\omega|_S = \iota^* dy = d(y \circ \iota) = 0,$$

because y vanishes identically on S.

To distinguish the two ways in which we might interpret the statement "ω vanishes on S," one usually says that ω *vanishes along* S or *vanishes at points of* S if $\omega_p = 0$ for every point $p \in S$. The weaker condition that $\omega|_S = 0$ is expressed by saying that the restriction of ω to S vanishes.

◇ **Exercise 8.6.** Suppose M is a smooth manifold and $S \subset M$ is an immersed submanifold. If $f \in C^\infty(M)$, show that $d(f|_S) = (df)|_S$. Conclude that the restriction of df to S is zero if and only if f is constant on each component of S.

Lie Subgroups

A *Lie subgroup* of a Lie group G is a subgroup of G endowed with a topology and smooth structure making it into a Lie group and an immersed submanifold of G. The following proposition shows that embedded subgroups are automatically Lie subgroups.

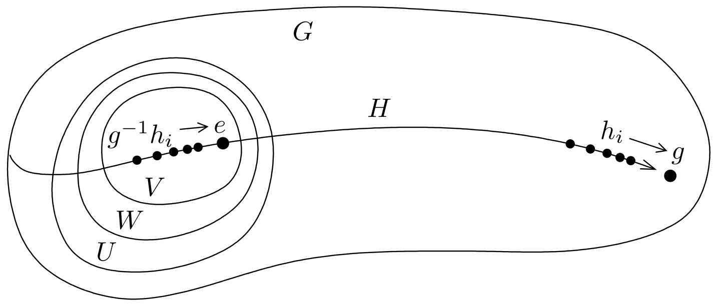


Figure 8.15. An embedded Lie subgroup.

Proposition 8.28. *Let G be a Lie group, and suppose $H \subset G$ is a subgroup that is also an embedded submanifold. Then H is a closed Lie subgroup of G.*

Proof. To prove that H is a Lie subgroup, we need only check that multiplication $H \times H \to H$ and inversion $H \to H$ are smooth maps. Because multiplication is a smooth map from $G \times G$ into G, its restriction is clearly smooth from $H \times H$ into G (this is true even if H is merely immersed). Because H is a subgroup, multiplication takes $H \times H$ into H, and since H is embedded, this is a smooth map into H by Corollary 8.23. A similar argument applies to inversion. This proves that H is a Lie subgroup.

To prove that H is closed, let g be an arbitrary point of $\overline{H}$. Then there is a sequence $\{h_i\}$ of points in H converging to g. Let U be the domain of a slice chart for H containing the identity, and let W be a smaller neighborhood of e such that $\overline{W} \subset U$. Since the map $\mu\colon G \times G \to G$ given by $\mu(g_1, g_2) = g_1^{-1}g_2$ is continuous, there is a neighborhood V of the identity with the property that $V \times V \subset \mu^{-1}(W)$, which means that $g_1^{-1}g_2 \in W$ whenever $g_1, g_2 \in V$ (Figure 8.15).

Because $g^{-1}h_i \to e$, by discarding finitely many terms of the sequence we may assume that $g^{-1}h_i \in V$ for all i. This implies that

$$h_j^{-1}h_i = (g^{-1}h_j)^{-1}(g^{-1}h_i) \in W$$

for all i and j. Fixing j and letting $i \to \infty$, we find $h_j^{-1}h_i \to h_j^{-1}g \in \overline{W} \subset U$. Since $H \cap U$ is a slice, it is closed in U, and therefore $h_j^{-1}g \in H$, which implies $g \in H$. Thus H is closed. □

We will see in Chapter 20 that this proposition has an important converse, called the closed subgroup theorem—it asserts that any subgroup of a Lie group that is topologically closed is automatically an embedded Lie subgroup.

Here are some important examples of Lie subgroups.

Example 8.29 (Matrices with Positive Determinant). The subset $\mathrm{GL}^+(n,\mathbb{R}) \subset \mathrm{GL}(n,R)$ consisting of real $n \times n$ matrices with positive determinant is a subgroup because $\det(AB) = (\det A)(\det B)$. It is an open subset of $\mathrm{GL}(n,\mathbb{R})$ by continuity of the determinant function, and therefore it is an embedded Lie subgroup of dimension n^2.

Example 8.30 (The Circle Group). The circle $\mathbb{S}^1$ is a Lie subgroup of $\mathbb{C}^*$ because it is a subgroup and an embedded submanifold.

Example 8.31 (The Orthogonal Group). A real $n \times n$ matrix A is said to be *orthogonal* if as a linear map $A\colon \mathbb{R}^n \to \mathbb{R}^n$ it preserves the Euclidean dot product:

$$(Ax) \cdot (Ay) = x \cdot y \quad \text{for all } x, y \in \mathbb{R}^n.$$

The set $\mathrm{O}(n)$ of all orthogonal $n \times n$ matrices is clearly a subgroup of $\mathrm{GL}(n,\mathbb{R})$, called the n-dimensional *orthogonal group*. It is easy to see that a matrix A is orthogonal if and only if it takes the standard basis of $\mathbb{R}^n$ to an orthonormal basis, which is equivalent to the columns of A being orthonormal. Since the (i,j)-entry of the matrix A^TA is the dot product of the ith and jth columns of A, this condition is also equivalent to the requirement that $A^TA = I_n$.

Let $\mathrm{S}(n,\mathbb{R})$ denote the set of symmetric $n \times n$ matrices, which is easily seen to be a linear subspace of $\mathrm{M}(n,\mathbb{R})$ of dimension $n(n+1)/2$ because each symmetric matrix is uniquely determined by its values on and above the main diagonal. Define $\Phi\colon\ \mathrm{GL}(n,\mathbb{R}) \to \mathrm{S}(n,\mathbb{R})$ by

$$\Phi(A) = A^TA.$$

We will show that the identity matrix I_n is a regular value of Φ, from which it follows that $\mathrm{O}(n) = \Phi^{-1}(I_n)$ is an embedded submanifold of codimension $n(n+1)/2$, that is, of dimension $n^2 - n(n+1)/2 = n(n-1)/2$, and thus an embedded Lie subgroup of $\mathrm{GL}(n,\mathbb{R})$. It is a compact group because it is a closed and bounded subset of $\mathrm{M}(n,\mathbb{R})$: closed because it is a level set of the continuous map $\Phi(A) = A^TA$, and bounded because each column of an orthogonal matrix has norm 1, which implies that the Euclidean norm of $A \in \mathrm{O}(n)$ is $(\sum_{ij}(A_i^j)^2)^{1/2} = \sqrt{n}$.

Let $A \in \mathrm{O}(n)$ be arbitrary, and let us compute the push-forward $\Phi_*\colon T_A\,\mathrm{GL}(n,\mathbb{R}) \to T_{\Phi(A)}\,\mathrm{S}(n,\mathbb{R})$. We can identify the tangent spaces $T_A\,\mathrm{GL}(n,\mathbb{R})$ and $T_{\Phi(A)}\,\mathrm{S}(n,\mathbb{R})$ with $\mathrm{M}(n,\mathbb{R})$ and $\mathrm{S}(n,\mathbb{R})$, respectively, because $\mathrm{GL}(n,\mathbb{R})$ is an open subset of the vector space $\mathrm{M}(n,\mathbb{R})$ and $\mathrm{S}(n,\mathbb{R})$ is itself a vector space. For any $B \in \mathrm{M}(n,\mathbb{R})$, the curve $\gamma(t) = A + tB$ satisfies $\gamma(0) = A$ and $\gamma'(0) = B$. We compute

$$\begin{aligned}\Phi_*B = \Phi_*\gamma'(0) = (\Phi \circ \gamma)'(0) &= \left.\frac{d}{dt}\right|_{t=0} \Phi(A+tB) \\ &= \left.\frac{d}{dt}\right|_{t=0} (A+tB)^T(A+tB) = B^TA + A^TB.\end{aligned}$$

If $C \in \mathrm{S}(n,\mathbb{R})$ is arbitrary,

$$\Phi_*(\tfrac{1}{2}AC) = \tfrac{1}{2}C^T A^T A + \tfrac{1}{2}A^T AC = \tfrac{1}{2}C + \tfrac{1}{2}C = C.$$

Thus Φ_* is surjective, which proves the claim.

Example 8.32 (The Special Linear Group). The set of $n \times n$ matrices with determinant equal to 1 is called the *special linear group*. Because $\mathrm{SL}(n,\mathbb{R})$ is the kernel of the group homomorphism $\det\colon \mathrm{GL}(n,\mathbb{R}) \to \mathbb{R}^*$, $\mathrm{SL}(n,\mathbb{R})$ is a subgroup of $\mathrm{GL}(n,\mathbb{R})$. We will show that det is a submersion, from which it follows that $\mathrm{SL}(n,\mathbb{R}) = \det^{-1}(1)$ is an embedded submanifold of codimension 1 in $\mathrm{GL}(n,\mathbb{R})$, and therefore a Lie subgroup.

Let $A \in \mathrm{GL}(n,\mathbb{R})$ be arbitrary. Problem 6-6 shows that the differential of det is given by $d(\det)_A(B) = (\det A)\operatorname{tr}(A^{-1}B)$ for $B \in T_A \mathrm{GL}(n,\mathbb{R}) \cong \mathrm{M}(n,\mathbb{R})$. Choosing $B = A$ yields

$$d(\det)_A(A) = (\det A)\operatorname{tr}(A^{-1}A) = (\det A)\operatorname{tr}(I_n) = n \det A \neq 0.$$

This shows that $d(\det)$ never vanishes on $\mathrm{GL}(n,\mathbb{R})$, so det is a submersion and $\mathrm{SL}(n,\mathbb{R})$ is an embedded (n^2-1)-dimensional Lie subgroup of $\mathrm{GL}(n,\mathbb{R})$.

Example 8.33 (The Special Orthogonal Group). The *special orthogonal group* is defined as $\mathrm{SO}(n) = \mathrm{O}(n) \cap \mathrm{SL}(n,\mathbb{R}) \subset \mathrm{GL}(n,\mathbb{R})$. Because every matrix $A \in \mathrm{O}(n)$ satisfies

$$1 = \det I_n = \det(A^T A) = (\det A)(\det A^T) = (\det A)^2,$$

it follows that $\det A = \pm 1$ for all $A \in \mathrm{O}(n)$. Therefore, $\mathrm{SO}(n)$ is the open subgroup of $\mathrm{O}(n)$ consisting of matrices of positive determinant, and is therefore also an embedded Lie subgroup of dimension $n(n-1)/2$ in $\mathrm{GL}(n,\mathbb{R})$. It is a compact group because it is a closed subset of $\mathrm{O}(n)$.

Example 8.34 (A Dense Lie Subgroup of the Torus). Let $H \subset \mathbb{T}^2$ be the dense immersed submanifold of the torus that is the image of the immersion $\gamma\colon \mathbb{R} \to \mathbb{T}^2$ defined in Example 7.3. It is easy to check that γ is a group homomorphism and therefore H is a subgroup of $\mathbb{T}^2$. Because the smooth structure on H is defined so that $\gamma\colon \mathbb{R} \to H$ is a diffeomorphism, H is a Lie group (in fact, isomorphic to $\mathbb{R}$) and is therefore a Lie subgroup of $\mathbb{T}^2$.

Example 8.35 (The Complex General Linear Group). Define a map $\beta\colon \mathrm{GL}(n,\mathbb{C}) \to \mathrm{GL}(2n,\mathbb{R})$ by replacing each complex matrix entry $a + ib$ with the 2×2 block $\left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right)$:

$$\beta \begin{pmatrix} a_1^1 + ib_1^1 & \dots & a_1^n + ib_1^n \\ \vdots & & \vdots \\ a_n^1 + ib_n^1 & \dots & a_n^n + ib_n^n \end{pmatrix} = \begin{pmatrix} a_1^1 & -b_1^1 & \multirow{2}{*}{$\cdots$} & a_1^n & -b_1^n \\ b_1^1 & a_1^1 & & b_1^n & a_1^n \\ \vdots & & & \vdots & \\ a_n^1 & -b_n^1 & \multirow{2}{*}{$\cdots$} & a_n^n & -b_n^n \\ b_n^1 & a_n^1 & & b_n^n & a_n^n \end{pmatrix}.$$

It is straightforward to verify that β is an injective Lie group homomorphism whose image is a closed Lie subgroup of $\mathrm{GL}(2n,\mathbb{R})$. Thus $\mathrm{GL}(n,\mathbb{C})$ is isomorphic to this Lie subgroup of $\mathrm{GL}(2n,\mathbb{R})$. (You can check that β arises naturally from the identification of $(x^1+iy^1,\dots,x^n+iy^n)\in\mathbb{C}^n$ with $(x^1,y^1,\dots,x^n,y^n)\in\mathbb{R}^{2n}$.)

The Lie Algebra of a Lie Subgroup

If G is a Lie group and $H\subset G$ is a Lie subgroup, we might hope that the Lie algebra of H would be a Lie subalgebra of that of G. However, elements of $\mathrm{Lie}(H)$ are vector fields on H, not G, and so, strictly speaking, are not elements of $\mathrm{Lie}(G)$. Nonetheless, the next proposition gives us a way to view $\mathrm{Lie}(H)$ as a subalgebra of $\mathrm{Lie}(G)$.

Proposition 8.36 (The Lie Algebra of a Lie Subgroup). *Suppose $H\subset G$ is a Lie subgroup. The subset $\widetilde{\mathfrak{h}}\subset\mathrm{Lie}(G)$ defined by*

$$\widetilde{\mathfrak{h}}=\{X\in\mathrm{Lie}(G): X_e\in T_eH\}$$

is a Lie subalgebra of $\mathrm{Lie}(G)$ canonically isomorphic to $\mathrm{Lie}(H)$.

Proof. Let $\mathfrak{g}=\mathrm{Lie}(G)$ and $\mathfrak{h}=\mathrm{Lie}(H)$. Because the inclusion map $\iota\colon H\hookrightarrow G$ is a Lie group homomorphism, $\iota_*(\mathfrak{h})$ is a Lie subalgebra of $\mathfrak{g}$. By the way we defined the induced Lie algebra homomorphism, this subalgebra is precisely the set of left-invariant vector fields on G whose values at the identity are of the form ι_*V for some $V\in T_eH$. Since the push-forward map $\iota_*\colon T_eH\to T_eG$ is the inclusion of T_eH as a subspace in T_eG, it follows that $\iota_*(\mathfrak{h})=\widetilde{\mathfrak{h}}$, and ι_* is injective on $\mathfrak{h}$ because it is injective on T_eH. □

Using this proposition, whenever H is a Lie subgroup of G, we will typically identify $\mathrm{Lie}(H)$ as a subalgebra of $\mathrm{Lie}(G)$. As we mentioned above, elements of $\mathrm{Lie}(H)$ are not themselves left-invariant vector fields on G. But the preceding proposition shows that every element of $\mathrm{Lie}(H)$ corresponds to a unique element of $\mathrm{Lie}(G)$, determined by its value at the identity, and the injection of $\mathrm{Lie}(H)$ into $\mathrm{Lie}(G)$ thus determined respects Lie brackets; so by thinking of $\mathrm{Lie}(H)$ as a subalgebra of $\mathrm{Lie}(G)$ we are not committing a grave error.

This identification is especially illuminating in the case of Lie subgroups of $\mathrm{GL}(n,\mathbb{R})$.

Example 8.37 (The Lie Algebra of O(n)). Consider $\mathrm{O}(n)$ as a Lie subgroup of $\mathrm{GL}(n,\mathbb{R})$. By Example 8.31, it is equal to the level set $\Phi^{-1}(I_n)$, where $\Phi\colon\mathrm{GL}(n,\mathbb{R})\to\mathrm{S}(n,\mathbb{R})$ is the map $\Phi(A)=A^TA$. By Lemma 8.15, $T_{I_n}\mathrm{O}(n)=\operatorname{Ker}\Phi_*\colon T_{I_n}\mathrm{GL}(n,\mathbb{R})\to T_{I_n}\mathrm{S}(n,\mathbb{R})$. By the computation in Example 8.31, this push-forward is $\Phi_*B=B^T+B$, so

$$\begin{aligned}T_{I_n}\mathrm{O}(n)&=\{B\in\mathfrak{gl}(n,\mathbb{R}):B^T+B=0\}\\&=\{\text{skew-symmetric } n\times n \text{ matrices}\}.\end{aligned}$$

We denote this subspace of $\mathfrak{gl}(n,\mathbb{R})$ by $\mathfrak{o}(n)$. Proposition 8.36 then implies that $\mathfrak{o}(n)$ is a Lie subalgebra of $\mathfrak{gl}(n,\mathbb{R})$ that is canonically isomorphic to $\mathrm{Lie}(\mathrm{O}(n))$. Notice that we did not even have to verify directly that $\mathfrak{o}(n)$ is a subalgebra.

In Chapter 4, we showed that the Lie algebra of $\mathrm{GL}(n,\mathbb{R})$ is naturally isomorphic to the matrix algebra $\mathfrak{gl}(n,\mathbb{R})$. Using the tools developed in this chapter, we can prove a similar result for $\mathrm{GL}(n,\mathbb{C})$. Just as in the real case, our usual identification of $\mathrm{GL}(n,\mathbb{C})$ as an open subset of $\mathfrak{gl}(n,\mathbb{C})$ yields a sequence of vector space isomorphisms

$$\mathrm{Lie}(\mathrm{GL}(n,\mathbb{C})) \xrightarrow{\varepsilon} T_{I_n}\mathrm{GL}(n,\mathbb{C}) \xrightarrow{\varphi} \mathfrak{gl}(n,\mathbb{C}), \tag{8.2}$$

where ε is the evaluation map and φ is the usual identification between the tangent space to an open subset of a vector space and the vector space itself. (Note that we are considering these as real vector spaces, not complex ones.)

Proposition 8.38 (The Lie Algebra of $\mathbf{GL}(n,\mathbb{C})$). *The composition of the maps in* (8.2) *yields a Lie algebra isomorphism between* $\mathrm{Lie}(\mathrm{GL}(n,\mathbb{C}))$ *and the matrix algebra* $\mathfrak{gl}(n,\mathbb{C})$.

Proof. The injective Lie group homomorphism $\beta\colon \mathrm{GL}(n,\mathbb{C}) \to \mathrm{GL}(2n,\mathbb{R})$ constructed in Example 8.35 induces a Lie algebra homomorphism $\beta_*\colon \mathrm{Lie}(\mathrm{GL}(n,\mathbb{C})) \to \mathrm{Lie}(\mathrm{GL}(2n,\mathbb{R}))$. Composing β_* with our canonical isomorphisms yields a commutative diagram

$$\begin{tikzcd}
\mathrm{Lie}(\mathrm{GL}(n,\mathbb{C})) \arrow[r, "\varepsilon"] \arrow[d, "\beta_*"] & T_{I_n}\mathrm{GL}(n,\mathbb{C}) \arrow[r, "\varphi"] \arrow[d, "\beta_*"] & \mathfrak{gl}(n,\mathbb{C}) \arrow[d, "\alpha"] \\
\mathrm{Lie}(\mathrm{GL}(2n,\mathbb{R})) \arrow[r, "\varepsilon"] & T_{I_n}\mathrm{GL}(2n,\mathbb{R}) \arrow[r, "\varphi"] & \mathfrak{gl}(2n,\mathbb{R}),
\end{tikzcd} \tag{8.3}$$

in which the β_* in the middle is the push-forward at the identity, and $\alpha = \varphi \circ \beta_* \circ \varphi^{-1}$. Proposition 4.23 showed that the composition of the isomorphisms in the bottom row is a Lie algebra isomorphism; we need to show the same thing for the top row.

It is easy to see from the formula in Example 8.35 that β is (the restriction of) a linear map. It follows that $\beta_*\colon T_{I_n}\mathrm{GL}(n,\mathbb{C}) \to T_{I_n}\mathrm{GL}(2n,\mathbb{R})$ is given by exactly the same formula, as is $\alpha\colon \mathfrak{gl}(n,\mathbb{C}) \to \mathfrak{gl}(2n,\mathbb{R})$. Because $\beta(AB) = \beta(A)\beta(B)$, it follows that α preserves matrix commutators:

$$\alpha[A,B] = \alpha(AB - BA) = \alpha(A)\alpha(B) - \alpha(B)\alpha(A) = [\alpha(A), \alpha(B)].$$

Thus α is an injective Lie algebra homomorphism from $\mathfrak{gl}(n,\mathbb{C})$ to $\mathfrak{gl}(2n,\mathbb{R})$ (considering both as matrix algebras). Replacing the bottom row in (8.3) by the images of the vertical maps, we obtain a commutative diagram of

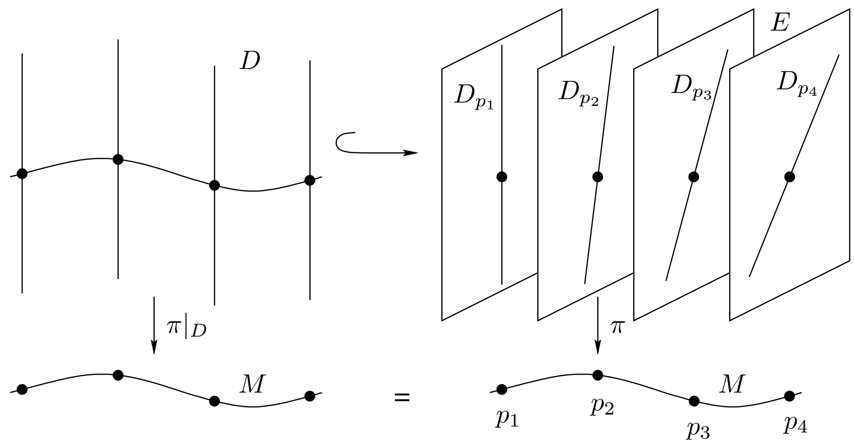


Figure 8.16. A subbundle of a vector bundle.

vector space isomorphisms

$$\begin{tikzcd}
\operatorname{Lie}(\mathrm{GL}(n,\mathbb{C})) \arrow[r, "\cong"] \arrow[d, "\beta_*"'] & \mathfrak{gl}(n,\mathbb{C}) \arrow[d, "\alpha"] \\
\beta_*(\operatorname{Lie}(\mathrm{GL}(n,\mathbb{C}))) \arrow[r, "\cong"'] & \alpha(\mathfrak{gl}(n,\mathbb{C})),
\end{tikzcd}$$

in which the bottom map and the two vertical maps are Lie algebra isomorphisms; it follows that the top map is also a Lie algebra isomorphism. □

Vector Subbundles

Given a smooth vector bundle $\pi\colon E \to M$, a *smooth subbundle* of E (see Figure 8.16) is a subset $D \subset E$ with the following properties:

(i) D is an embedded submanifold of E.

(ii) For each $p \in M$, the fiber $D_p = D \cap \pi^{-1}(p)$ is a linear subspace of $E_p = \pi^{-1}(p)$.

(iii) With the vector space structure on each D_p inherited from E_p and the projection $\pi|_D\colon D \to M$, D is a smooth vector bundle over M.

Note that the condition that D be a vector bundle over M implies that the projection $\pi|_D\colon D \to M$ must be surjective, and that all the fibers D_p must have the same dimension.

◇ **Exercise 8.7.** If $D \subset E$ is a smooth subbundle, show that the inclusion map $\iota\colon D \hookrightarrow E$ is a smooth bundle map over M.

The following lemma gives a convenient condition for checking that a collection of subspaces $\{D_p \subset E_p : p \in M\}$ is a smooth subbundle.

Lemma 8.39 (Local Frame Criterion for Subbundles). *Let $\pi\colon E \to M$ be a smooth vector bundle, and suppose for each $p \in M$ we are given an m-dimensional linear subspace $D_p \subset E_p$. Then $D = \coprod_{p\in M} D_p \subset E$ is a smooth subbundle if and only if the following condition is satisfied:*

Each point $p \in M$ has a neighborhood U on which there are smooth local sections $\sigma_1, \dots, \sigma_m\colon U \to E$ such that $\sigma_1|_q, \dots, \sigma_m|_q$ form a basis for D_q at each $q \in U$. (8.4)

Proof. If D is a smooth subbundle, then by definition any $p \in M$ has a neighborhood U over which there exists a smooth local trivialization of D, and Example 5.9 shows that there exists a smooth local frame for D over any such set U. Such a local frame is by definition a collection of smooth sections $\tau_1, \dots, \tau_m\colon U \to D$ whose images form a basis for D_p at each point $p \in U$. The smooth sections of E that we seek are obtained simply by composing with the inclusion map $\iota\colon D \hookrightarrow E$: $\sigma_j = \iota \circ \tau_j$.

Conversely, suppose that D satisfies (8.4). Condition (ii) in the definition of a subbundle is true by hypothesis, so we need to show that D satisfies conditions (i) and (iii).

To prove that D is an embedded submanifold, it suffices to show that each point $p \in M$ has a neighborhood U such that $D \cap \pi^{-1}(U)$ is an embedded submanifold of $\pi^{-1}(U) \subset TM$. Given $p \in M$, let $\sigma_1, \dots, \sigma_m$ be smooth local sections of E defined on a neighborhood of p and satisfying (8.4). By the result of Problem 5-8, we can extend these to a smooth local frame $(\sigma_1, \dots, \sigma_k)$ for E over some neighborhood U of p. By Proposition 5.10, this local frame is associated with a smooth local trivialization $\pi^{-1}(U) \to U \times \mathbb{R}^k$, defined by

$$s^i \sigma_i|_q \mapsto (q, (s^1, \dots, s^k)).$$

In terms of this trivialization, $D \cap \pi^{-1}(U)$ corresponds to $U \times \mathbb{R}^m = \{(q, (s^1, \dots, s^m, 0, \dots, 0))\} \subset U \times \mathbb{R}^k$, which is obviously an embedded submanifold. Moreover, the map $\Phi|_{D\cap\pi^{-1}(U)}\colon D \cap \pi^{-1}(U) \to U \times \mathbb{R}^m$ is obviously a smooth local trivialization of D, showing that D is itself a smooth vector bundle. □

Example 8.40. Suppose M is any parallelizable manifold, and let $(E_1, \dots, E_n)$ be a smooth global frame for M. If $1 \le k \le n$, the subset $D \subset TM$ defined by $D_p = \operatorname{span}(E_1, \dots, E_k)$ is a smooth subbundle of TM.

Problems

8-1. Consider the map $F\colon \mathbb{R}^4 \to \mathbb{R}^2$ defined by
$$F(x,y,s,t) = (x^2+y, x^2+y^2+s^2+t^2+y).$$
Show that $(0,1)$ is a regular value of F, and that the level set $F^{-1}(0,1)$ is diffeomorphic to $\mathbb{S}^2$.

8-2. Let $F\colon \mathbb{R}^2 \to \mathbb{R}$ be defined by
$$F(x,y) = x^3 + xy + y^3.$$
Which level sets of F are embedded submanifolds of $\mathbb{R}^2$?

8-3. Show that the image of the curve $\gamma\colon (-\pi/2, 3\pi/2) \to \mathbb{R}^2$ of Example 7.2 is not an embedded submanifold of $\mathbb{R}^2$. [Be careful: This is not the same as showing that γ is not an embedding.]

8-4. Let $S \subset \mathbb{R}^2$ be the boundary of the square of side 2 centered at the origin (see Problem 3-4). Show that S does not have a topology and smooth structure in which it is an immersed submanifold of $\mathbb{R}^2$.

8-5. Let $\gamma\colon \mathbb{R} \to \mathbb{T}^2$ be the curve of Example 7.3. Show that $\gamma(\mathbb{R})$ is not an embedded submanifold of the torus.

8-6. Our definition of Lie groups included the requirement that both the multiplication map and the inversion map are smooth. Show that smoothness of the inversion map is redundant: If G is a smooth manifold with a group structure such that the multiplication map $m\colon G \times G \to G$ is smooth, then G is a Lie group.

8-7. This problem generalizes the result of Problem (4-9) to higher dimensions. For any integer $n \geq 1$, define a vector field on $\mathbb{S}^{2n-1} \subset \mathbb{C}^n$ by $V_z = \gamma_z'(0)$, where $\gamma_z\colon \mathbb{R} \to \mathbb{S}^{2n-1}$ is the curve $\gamma_z(t) = e^{it}z$. Show that V is smooth and nowhere vanishing.

8-8. Show that an embedded submanifold is closed if and only if the inclusion map is proper.

8-9. For each $a \in \mathbb{R}$, let M_a be the subset of $\mathbb{R}^2$ defined by
$$M_a = \{(x,y) : y^2 = x(x-1)(x-a)\}.$$
For which values of a is M_a an embedded submanifold of $\mathbb{R}^2$? For which values can M_a be given a topology and smooth structure making it into an immersed submanifold?

8-10. If $F\colon M \to N$ is a smooth map, $c \in N$, and $F^{-1}(c)$ is an embedded submanifold of M whose codimension is equal to the dimension of N, must c be a regular value of F?

8-11. Let $S \subset N$ be a closed embedded submanifold.

(a) Suppose $f \in C^\infty(S)$. (This means that f is smooth when considered as a function on S, not as a function on a closed subset of N.) Show that f is the restriction of a smooth function on N.
(b) If $X \in \mathfrak{T}(S)$, show that there is a smooth vector field Y on N such that $X = Y|_S$.
(c) Find counterexamples to both results if the hypothesis that S is closed is omitted.

8-12. Suppose M is a smooth manifold and $S \subset M$ is a submanifold.

(a) If S is embedded, show that the subspace topology and the smooth structure of Theorem 8.2 are the unique such structures on S for which it is an immersed submanifold (of any dimension).
(b) If S is immersed, show that for the given topology on S, there is only one smooth structure making S into an immersed submanifold (of any dimension).
(c) If S is immersed, show by example that there may be more than one topology and smooth structure with respect to which S is an immersed submanifold.

[Hint: Use Proposition 8.22 and Theorem 7.14.]

8-13. Let $\Phi, \Psi \colon \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$\Phi(x, y) = x^3 - y^2,$$
$$\Psi(x, y) = x^2 - y^2.$$

(a) Show that neither $\Phi^{-1}(0)$ nor $\Psi^{-1}(0)$ is an embedded submanifold of $\mathbb{R}^2$.
(b) Can either set be given a topology and smooth structure making it into an immersed submanifold of $\mathbb{R}^2$? [Hint: Consider tangent vectors to the submanifold at the origin.]

8-14. If $S \subset M$ is an embedded submanifold and $\gamma \colon J \to M$ is a smooth curve whose image happens to lie in S, show that $\gamma'(t)$ is in the subspace $T_{\gamma(t)}S$ of $T_{\gamma(t)}M$ for all $t \in J$. Give a counterexample if S is not embedded.

8-15. Show by giving a counterexample that Proposition 8.5 is false if S is merely immersed.

8-16. This problem gives a generalization of the regular level set theorem. Suppose $\Phi \colon M \to N$ is a smooth map and $S \subset N$ is an embedded submanifold. We say that Φ is *transverse to* Σ if for every $p \in \Phi^{-1}(S)$, the spaces $T_{\Phi(p)}S$ and $\Phi_* T_p M$ together span $T_{\Phi(p)}N$. If Φ is transverse to S, show that $\Phi^{-1}(S)$ is an embedded submanifold of M whose codimension is equal to $\dim N - \dim S$.

8-17. Let M be a smooth manifold. Two embedded submanifolds $S_1, S_2 \subset M$ are said to be *transverse* (or to *intersect transversely*) if for each

$p \in S_1 \cap S_2$, the tangent spaces T_pS_1 and T_pS_2 together span T_pM. If S_1 and S_2 are transverse, show that $S_1 \cap S_2$ is an embedded submanifold of M of dimension $\dim S_1 + \dim S_2 - \dim M$. Give a counterexample when S_1 and S_2 are not transverse.

8-18. Let M be a smooth manifold, and let $C \subset M$ be an embedded submanifold that admits a global defining function $\Phi\colon M \to \mathbb{R}^k$. Let $f \in C^\infty(M)$, and suppose $p \in C$ is a point at which f attains its maximum or minimum value among points in C. Show that there are real numbers $\lambda_1, \dots, \lambda_k$ (called *Lagrange multipliers*) such that

$$df_p = \lambda_1 d\Phi^1|_p + \cdots + \lambda_k d\Phi^k|_p.$$

8-19. Let $\mathbb{H} = \mathbb{C} \times \mathbb{C}$ (considered as a real vector space), and define a bilinear product $\mathbb{H} \times \mathbb{H} \to \mathbb{H}$ by

$$(a,b)(c,d) = (ac - \overline{d}b, da + b\overline{c}), \qquad \text{for } a,b,c,d \in \mathbb{C}.$$

With this product, $\mathbb{H}$ is a 4-dimensional algebra over $\mathbb{R}$, called the algebra of *quaternions*. For any $p = (a,b) \in \mathbb{H}$, define $p^* = (\overline{a}, -b)$. It is useful to work with the basis $(\mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k})$ for $\mathbb{H}$ defined by

$$\mathbf{1} = (1,0), \quad \mathbf{i} = (i,0), \quad \mathbf{j} = (0,1), \quad \mathbf{k} = (0,i).$$

It is straightforward to verify that this basis satisfies

$$\begin{aligned}
\mathbf{1}q &= q\mathbf{1} = q \qquad \text{for all } q \in \mathbb{H},\\
\mathbf{ij} &= -\mathbf{ji} = \mathbf{k},\\
\mathbf{jk} &= -\mathbf{kj} = \mathbf{i},\\
\mathbf{ki} &= -\mathbf{ik} = \mathbf{j},\\
\mathbf{i}^2 &= \mathbf{j}^2 = \mathbf{k}^2 = -\mathbf{1},
\end{aligned}$$

and

$$(x^1\mathbf{1} + x^2\mathbf{i} + x^3\mathbf{j} + x^4\mathbf{k})^* = x^1\mathbf{1} - x^2\mathbf{i} - x^3\mathbf{j} - x^4\mathbf{k}.$$

(a) Show that quaternionic multiplication is associative but not commutative.
(b) Show that $(pq)^* = q^*p^*$ for all $p, q \in \mathbb{H}$.
(c) Show that $\langle p, q\rangle = \frac{1}{2}(p^*q + q^*p)$ is an inner product on $\mathbb{H}$, whose associated norm satisfies $|pq| = |p|\,|q|$.
(d) Show that every nonzero quaternion has a two-sided multiplicative inverse given by $p^{-1} = |p|^{-2}p^*$.
(e) Show that the set $\mathcal{S}$ of unit quaternions is a Lie group under quaternionic multiplication, diffeomorphic to $\mathbb{S}^3$.

8-20. Let $\mathbb{H}$ be the algebra of quaternions (Problem 8-19), and let $\mathcal{S} \subset \mathbb{H}$ be the group of unit quaternions. A quaternion p is said to be *real* if $p^* = p$, and *imaginary* if $p^* = -p$.

(a) If $p \in \mathbb{H}$ is imaginary, show that qp is tangent to $\mathbb{S}$ at each $q \in \mathbb{S}$. (Here we are identifying each tangent space to $\mathbb{H}$ with $\mathbb{H}$ itself in the usual way.)

(b) Define vector fields X_1, X_2, X_3 on $\mathbb{H}$ by

$$X_1|_q = q\mathbf{i},$$
$$X_2|_q = q\mathbf{j},$$
$$X_3|_q = q\mathbf{k}.$$

Show that these vector fields restrict to a smooth left-invariant global frame on $\mathbb{S}$.

(c) Under the isomorphism $(x^1, x^2, x^3, x^4) \leftrightarrow x^1\mathbf{1} + x^2\mathbf{i} + x^3\mathbf{j} + x^4\mathbf{k}$ between $\mathbb{R}^4$ and $\mathbb{H}$, show that these vector fields are the same as the ones defined in Problem 5-10.

8-21. The algebra of *octonions* (also called the *Cayley numbers*) is the 8-dimensional real vector space $\mathbb{O} = \mathbb{H} \times \mathbb{H}$ (where $\mathbb{H}$ is the space of quaternions—see Problem 8-19) with the following bilinear product:

$$(p, q)(r, s) = (pr - s^* q, sp + qr^*), \qquad \text{for } p, q, r, s \in \mathbb{H}. \tag{8.5}$$

Show that $\mathbb{O}$ is a noncommutative, nonassociative algebra over $\mathbb{R}$, and prove that $\mathbb{S}^7$ is parallelizable by imitating as much of Problem 8-20 as you can. [Hint: It might be helpful to prove that $(PQ^*)Q = P(Q^*Q)$ for all $P, Q \in \mathbb{O}$, where $(p, q)^* = (p^*, -q)$.]

8-22. The algebra of *sedenions* is the 16-dimensional real vector space $\mathbb{S} = \mathbb{O} \times \mathbb{O}$ with the product defined by (8.5), but with p, q, r, and s interpreted as elements of $\mathbb{O}$. (The name comes Latin *sedecim*, meaning sixteen.) Why does sedenionic multiplication not yield a global frame for $\mathbb{S}^{15}$? [Remark: A *division algebra* is an algebra with a multiplicative identity element and no zero divisors (i.e., $ab = 0$ if and only if $a = 0$ or $b = 0$). It follows from the work of Bott and Milnor [BM58] on parallelizability of spheres that a finite-dimensional division algebra over $\mathbb{R}$ must have dimension 1, 2, 4, or 8.]

8-23. Proposition 8.36 implies that the Lie algebra of any Lie subgroup of $\mathrm{GL}(n, \mathbb{R})$ is canonically isomorphic to a subalgebra of $\mathfrak{gl}(n, \mathbb{R})$. Under this isomorphism, show that $\mathrm{Lie}(\mathrm{SL}(n, \mathbb{R})) \cong \mathfrak{sl}(n, \mathbb{R})$ and $\mathrm{Lie}(\mathrm{SO}(n)) \cong \mathfrak{o}(n)$, where

$$\mathfrak{sl}(n, \mathbb{R}) = \{A \in \mathfrak{gl}(n, \mathbb{R}) : \operatorname{tr} A = 0\},$$
$$\mathfrak{o}(n) = \{A \in \mathfrak{gl}(n, \mathbb{R}) : A^T + A = 0\}.$$

8-24. Let M be a smooth n-manifold with boundary. Show that ∂M is a topological $(n-1)$-manifold (without boundary), and has a unique smooth structure such that the inclusion $\partial M \hookrightarrow M$ is a smooth embedding.

This is page 205
Printer: Opaque this

9
Lie Group Actions

The most important applications of Lie groups involve actions by Lie groups on other manifolds. These typically arise in situations involving some kind of symmetry. For example, if M is a vector space or smooth manifold endowed with a certain geometric structure (such as an inner product, a norm, a metric, or a distinguished vector or covector field), the set of diffeomorphisms of M that preserve the structure (called the *symmetry group* of the structure) frequently turns out to be a Lie group acting smoothly on M. The properties of the group action can shed considerable light on the properties of the structure. This chapter is devoted to studying Lie group actions on manifolds.

We begin by defining group actions and describing a number of examples. Our first application of the theory is the equivariant rank theorem, which gives an easily verified condition under which a map between spaces with group actions has constant rank.

Next we introduce proper actions, which are the ones whose quotients have nice properties. The main result of the chapter is the quotient manifold theorem, which describes conditions under which the quotient of a smooth manifold by a proper group action is again a smooth manifold. At the end of the chapter, we explore two significant applications of the quotient manifold theorem: to homogeneous spaces, which are manifolds admitting a transitive smooth Lie group action; and to smooth covering maps, which can be realized as quotients of manifolds by discrete group actions.

Group Actions

If G is a group and M is a set, a *left action* of G on M is a map $G \times M \to M$, often written as $(g,p) \mapsto g \cdot p$, that satisfies

$$\begin{aligned} g_1 \cdot (g_2 \cdot p) &= (g_1 g_2) \cdot p \qquad && \text{for all } g_1, g_2 \in G \text{ and } p \in M; \\ e \cdot p &= p && \text{for all } p \in M. \end{aligned} \tag{9.1}$$

A *right action* is defined analogously as a map $M \times G \to M$ with the appropriate composition law:

$$\begin{aligned} (p \cdot g_1) \cdot g_2 &= p \cdot (g_1 g_2) \qquad && \text{for all } g_1, g_2 \in G \text{ and } p \in M; \\ p \cdot e &= p && \text{for all } p \in M. \end{aligned}$$

Now suppose G is a Lie group and M is a manifold. An action of G on M is said to be *continuous* if the map $G \times M \to M$ or $M \times G \to M$ defining the action is continuous. A manifold M endowed with a continuous G-action is called a (left or right) *G-space*. If M is a smooth manifold and the action is smooth, M is called a *smooth G-space*.

Sometimes it is useful to give a name to an action, such as $\theta\colon G \times M \to M$, with the action of a group element g on a point p usually written $\theta_g(p)$. In terms of this notation, the conditions (9.1) for a left action read

$$\begin{aligned} \theta_{g_1} \circ \theta_{g_2} &= \theta_{g_1 g_2}, \\ \theta_e &= \mathrm{Id}_M, \end{aligned} \tag{9.2}$$

while for a right action the first equation is replaced by

$$\theta_{g_2} \circ \theta_{g_1} = \theta_{g_1 g_2}.$$

For a continuous action, each map $\theta_g\colon M \to M$ is a homeomorphism, because $\theta_{g^{-1}}$ is a continuous inverse for it. If the action is smooth, then each θ_g is a diffeomorphism.

For left actions, we will generally use the notations $g \cdot p$ and $\theta_g(p)$ interchangeably. The latter notation contains a bit more information, and is useful when it is important to specify the specific action under consideration, while the former is often more convenient when the action is understood. For right actions, the notation $p \cdot g$ is generally preferred because of the way composition works.

A right action can always be converted to a left action by the trick of defining $g \cdot p$ to be $p \cdot g^{-1}$, and a left action can similarly be converted to a right action. Thus any results about left actions can be translated into results about right actions, and vice versa. We will usually focus our attention on left actions, because their group law (9.2) has the property that multiplication of group elements corresponds to composition of maps. However, there are some circumstances in which right actions arise naturally; we will see several such actions later in this chapter.

Let us introduce some basic terminology regarding group actions. Suppose $\theta\colon G \times M \to M$ is a left action of a group G on a set M. (The definitions for right actions are analogous.)

- For any $p \in M$, the *orbit* of p under the action is the set
$$G \cdot p = \{g \cdot p : g \in G\},$$
the set of all images of p under the action by elements of G.
- The action is *transitive* if for any two points $p, q \in M$, there is a group element g such that $g \cdot p = q$, or equivalently if the orbit of any point is all of M.
- Given $p \in M$, the *isotropy group* of p, denoted by G_p, is the set of elements $g \in G$ that fix p:
$$G_p = \{g \in G : g \cdot p = p\}.$$
- The action is said to be *free* if the only element of G that fixes any element of M is the identity: $g \cdot p = p$ for some $p \in M$ implies $g = e$. This is equivalent to the requirement that $G_p = \{e\}$ for every $p \in M$.

Example 9.1 (Lie Group Actions).

(a) If G is any Lie group and M is any smooth manifold, the *trivial action* of G on M is defined by $g \cdot p = p$ for all $g \in G$. It is easy to see that it is a smooth action, and the isotropy group of each point is all of G.

(b) The natural action of $\mathrm{GL}(n, \mathbb{R})$ on $\mathbb{R}^n$ is the left action given by matrix multiplication: $(A, x) \mapsto Ax$, considering $x \in \mathbb{R}^n$ as a column matrix. This is an action because $I_n x = x$ and matrix multiplication is associative: $(AB)x = A(Bx)$. It is smooth because the components of Ax depend polynomially on the matrix entries of A and the components of x. Because any nonzero vector can be taken to any other by some linear transformation, there are exactly two orbits: $\{0\}$ and $\mathbb{R}^n \smallsetminus \{0\}$.

(c) The restriction of the natural action to $\mathrm{O}(n) \times \mathbb{R}^n \to \mathbb{R}^n$ defines a smooth left action of $\mathrm{O}(n)$ on $\mathbb{R}^n$. In this case, the orbits are the origin and the spheres centered at the origin. To see why, note that any orthogonal linear transformation preserves norms, so $\mathrm{O}(n)$ takes the sphere of radius R to itself; on the other hand, any nonzero vector of length R can be taken to any other by an orthogonal matrix. (If v and v' are such vectors, complete $v/|v|$ and $v'/|v'|$ to orthonormal bases and let A and A' be the orthogonal matrices whose columns are these orthonormal bases; then it is easy to check that $A'A^{-1}$ takes v to v'.)

(d) Further restricting the natural action to $\mathrm{O}(n) \times \mathbb{S}^{n-1} \to \mathbb{S}^{n-1}$, we obtain a transitive action of $\mathrm{O}(n)$ on $\mathbb{S}^{n-1}$. It is smooth by Corollary 8.23, because $\mathbb{S}^{n-1}$ is an embedded submanifold of $\mathbb{R}^n$.

(e) The natural action of $\mathrm{O}(n)$ restricts to a smooth action of $\mathrm{SO}(n)$ on $\mathbb{S}^{n-1}$. When $n = 1$, this action is trivial because $\mathrm{SO}(1)$ is the trivial group consisting of the matrix (1) alone. But when $n > 1$, $\mathrm{SO}(n)$ acts transitively on $\mathbb{S}^{n-1}$. To see this, it suffices to show that for any $v \in \mathbb{S}^n$, there is a matrix $A \in \mathrm{SO}(n)$ taking the first standard basis vector e_1 to v. Since $\mathrm{O}(n)$ acts transitively, there is a matrix $A \in \mathrm{O}(n)$ taking e_1 to v. Either $\det A = 1$, in which case $A \in \mathrm{SO}(n)$, or $\det A = -1$, in which case the matrix obtained by multiplying the last column of A by -1 is in $\mathrm{SO}(n)$ and still takes e_1 to v.

(f) Any Lie group G acts smoothly, freely, and transitively on itself by left or right translation. More generally, if H is a Lie subgroup of G, then the restriction of the multiplication map to $H \times G \to G$ defines a smooth, free (but generally not transitive) left action of H on G; similarly, restriction to $G \times H \to G$ defines a free right action of H on G.

(g) Similarly, any Lie group acts smoothly on itself by conjugation: $g \cdot h = ghg^{-1}$. For any $h \in G$, the isotropy group G_h is the set of all elements of G that commute with h.

(h) An action of a discrete group Γ on a manifold M is smooth if and only if for each $g \in \Gamma$, the map $p \mapsto g \cdot p$ is a smooth map from M to itself. Thus, for example, $\mathbb{Z}^n$ acts smoothly on the left on $\mathbb{R}^n$ by translation:

$$(m^1, \ldots, m^n) \cdot (x^1, \ldots, x^n) = (m^1 + x^1, \ldots, m^n + x^n).$$

Representations

One kind of Lie group action plays a fundamental role in many branches of mathematics and science. If G is a Lie group, a (*finite-dimensional*) *representation* of G is a Lie group homomorphism $\rho\colon G \to \mathrm{GL}(V)$ for some finite-dimensional real or complex vector space V. (Although it is useful for some applications to consider also the case in which V is infinite-dimensional, in this book we will consider only finite-dimensional representations.) Any representation ρ yields a smooth left action of G on V, defined by

$$g \cdot v = \rho(g)v, \qquad \text{for } g \in G,\ v \in V. \tag{9.3}$$

A seemingly more general notion is the following: If G is a Lie group, an action of G on a finite-dimensional vector space V is said to be *linear* if for each $g \in G$, the map from V to itself given by $v \mapsto g \cdot v$ is linear.

The next exercise shows that smooth linear actions correspond precisely to representations.

◇ **Exercise 9.1.** Let G be a Lie group and let V be a finite-dimensional vector space. Show that a smooth action of G on V is linear if and only if it is of the form (9.3) for some representation ρ of G. [Hint: Use the same idea as the proof of Lemma 5.4.]

As we will see later (Problem 9-15), the image of a representation $\rho\colon G \to \mathrm{GL}(V)$ is a Lie subgroup of $\mathrm{GL}(V)$. If ρ is injective, it is said to be a *faithful representation*, in which case it gives a Lie group isomorphism between G and $\rho(G) \subset \mathrm{GL}(V)$. By choosing a basis of V, we obtain a Lie group isomorphism $\mathrm{GL}(V) \cong \mathrm{GL}(n,\mathbb{R})$ or $\mathrm{GL}(n,\mathbb{C})$, so a Lie group admits a faithful representation if and only if it is isomorphic to a Lie subgroup of $\mathrm{GL}(n,\mathbb{R})$ or $\mathrm{GL}(n,\mathbb{C})$ for some n.

The study of representations of Lie groups is a vast subject in its own right, with applications to fields as diverse as differential geometry, differential equations, harmonic analysis, number theory, quantum physics, and engineering; we can do no more than touch on it here.

Example 9.2 (Lie Group Representations).

(a) If G is any Lie subgroup of $\mathrm{GL}(n,\mathbb{R})$, the inclusion map $G \hookrightarrow \mathrm{GL}(n,\mathbb{R}) = \mathrm{GL}(\mathbb{R}^n)$ is a faithful representation, called the *defining representation* of G. The defining representation of a subgroup of $\mathrm{GL}(n,\mathbb{C})$ is defined similarly.

(b) The inclusion map $\mathbb{S}^1 \hookrightarrow \mathbb{C}^* \cong \mathrm{GL}(1,\mathbb{C})$ is a faithful representation of the circle group. More generally, the map $\rho\colon \mathbb{T}^n \to \mathrm{GL}(n,\mathbb{C})$ given by

$$\rho(z^1,\dots,z^n) = \begin{pmatrix} z^1 & 0 & \dots & 0 \\ 0 & z^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & z^n \end{pmatrix}$$

is a faithful representation of $\mathbb{T}^n$.

(c) Let $\sigma\colon \mathbb{R}^n \to \mathrm{GL}(n+1,\mathbb{R})$ be the map that sends $x \in \mathbb{R}^n$ to the matrix $\sigma(x)$ defined in block form by

$$\sigma(x) = \begin{pmatrix} I_n & x \\ 0 & 1 \end{pmatrix},$$

where I_n is the $n \times n$ identity matrix and x is regarded as an $n \times 1$ column matrix. A straightforward computation shows that σ is a faithful representation of the additive Lie group $\mathbb{R}^n$.

(d) Another faithful representation of $\mathbb{R}^n$ is the map $\mathbb{R}^n \to \mathrm{GL}(n,\mathbb{R})$ that sends $x = (x^1, \dots, x^n) \in \mathbb{R}^n$ to the diagonal matrix whose diagonal entries are $(e^{x^1}, \dots, e^{x^n})$.

(e) Yet another representation of $\mathbb{R}^n$ is the map $\mathbb{R}^n \to \mathrm{GL}(n,\mathbb{C})$ sending x to the diagonal matrix with diagonal entries $(e^{2\pi i x^1}, \dots, e^{2\pi i x^n})$. This one is not faithful, because its kernel is the subgroup $\mathbb{Z}^n \subset \mathbb{R}^n$.

(f) For positive integers n and d, let $\mathcal{P}^n_d$ denote the vector space of real-valued polynomial functions $p\colon \mathbb{R}^n \to \mathbb{R}$ of degree d. For any matrix $A \in \mathrm{GL}(n,\mathbb{R})$, define a linear map $\tau(A)\colon \mathcal{P}^n_d \to \mathcal{P}^n_d$ by

$$\tau(A)p = p \circ A^{-1}.$$

A straightforward computation shows that $\tau\colon \mathrm{GL}(n,\mathbb{R}) \to \mathrm{GL}(\mathcal{P}^n_d)$ is a faithful representation.

Example 9.3 (The Adjoint Representation). Let G be a Lie group. For any $g \in G$, the conjugation map $C_g\colon G \to G$ given by $C_g(h) = ghg^{-1}$ is a Lie group homomorphism (see Example 2.8(f)). We let $\mathrm{Ad}(g) = (C_g)_*\colon \mathfrak{g} \to \mathfrak{g}$ denote its induced Lie algebra homomorphism. Because $C_{g_1g_2} = C_{g_1} \circ C_{g_2}$ for any $g_1, g_2 \in G$, it follows immediately that $\mathrm{Ad}(g_1g_2) = \mathrm{Ad}(g_1) \circ \mathrm{Ad}(g_2)$, and $\mathrm{Ad}(g)$ is invertible with inverse $\mathrm{Ad}(g^{-1})$. Once we show that $\mathrm{Ad}\colon G \to \mathrm{GL}(\mathfrak{g})$ is smooth, it will follow that it is a representation, called the *adjoint representation* of G.

To see that Ad is smooth, let $C\colon G \times G \to G$ be the smooth map defined by $C(g,h) = ghg^{-1}$. Let $X \in \mathfrak{g}$ and $g \in G$ be arbitrary, and let $\gamma\colon (-\varepsilon, \varepsilon) \to G$ be a smooth curve satisfying $\gamma(0) = e$ and $\gamma'(0) = X_e$. Then $\mathrm{Ad}(g)X$ is the left-invariant vector field whose value at $e \in G$ is

$$\begin{aligned}(C_g)_* X_e &= \left.\frac{d}{dt}\right|_{t=0} C_g(\gamma(t)) \\ &= \left.\frac{d}{dt}\right|_{t=0} C(g, \gamma(t)) \\ &= C_*(0_g, X_e),\end{aligned}$$

where we are regarding $(0_g, X_e)$ as an element of $T_{(g,e)}(G \times G)$ under the canonical isomorphism $T_{(g,e)}(G \times G) \cong T_gG \oplus T_eG$. Because $C_*\colon T(G \times G) \to TG$ is a smooth bundle map by Exercise 5.7, this expression depends smoothly on g. Smooth coordinates on $\mathrm{GL}(\mathfrak{g})$ are obtained by choosing a basis (E_i) for $\mathfrak{g}$ and using matrix entries with respect to this basis as coordinates. If (ε^j) is the dual basis, the matrix entries of $\mathrm{Ad}(g)\colon \mathfrak{g} \to \mathfrak{g}$ are given by $(\mathrm{Ad}(g))^j_i = \varepsilon^j(\mathrm{Ad}(g)E_i)$. The computation above with $X = E_i$ shows that these are smooth functions of g.

While we are on the subject, it is worth noting that there is an analogous notion of representations of Lie algebras. If $\mathfrak{g}$ is a finite-dimensional Lie algebra, a (*finite-dimensional*) *representation* of $\mathfrak{g}$ is a Lie algebra homomorphism $\varphi\colon \mathfrak{g} \to \mathfrak{gl}(V)$ for some finite-dimensional vector space V. If φ is injective, it is said to be a faithful representation, in which case $\mathfrak{g}$ is isomorphic to the Lie subalgebra $\varphi(\mathfrak{g}) \subset \mathfrak{gl}(V) \cong \mathfrak{gl}(n, \mathbb{R})$.

There is an intimate connection between representations of Lie groups and representations of their Lie algebras. Suppose G is a Lie group and $\mathfrak{g}$ is its Lie algebra. If $\rho\colon G \to \mathrm{GL}(V)$ is any representation of G, then $\rho_*\colon \mathfrak{g} \to \mathfrak{gl}(V)$ is easily seen to be a representation of $\mathfrak{g}$. (In Chapter 20, we will discover that that the converse is true when G is simply connected—see Exercise 20.1.)

The following deep algebraic theorem shows that every finite-dimensional Lie algebra is isomorphic to a Lie subalgebra of $\mathfrak{gl}(n, \mathbb{R})$ or $\mathfrak{gl}(n, \mathbb{C})$ for some n. Its proof requires far more algebra than we have at our disposal, so we simply refer the reader to the proof in [Var84]. (We will use this result only in the last section of the book, in the proof of Theorem 20.21.)

Theorem 9.4 (Ado's Theorem). *Every finite-dimensional Lie algebra admits a faithful finite-dimensional representation.*

It is important to note that the analogous result for Lie groups is false—there are Lie groups that are not isomorphic to any Lie subgroup of $\mathrm{GL}(n, \mathbb{R})$. One example of such a group is given in Problem 20-10.

Equivariant Maps

Suppose M and N are both (left or right) G-spaces. A map $F\colon M \to N$ is said to be *equivariant* with respect to the given G-actions if for each $g \in G$,

$$F(g \cdot p) = g \cdot F(p) \qquad \text{(for left actions)},$$

$$F(p \cdot g) = F(p) \cdot g \qquad \text{(for right actions)}.$$

Equivalently, if θ and φ are the given actions on M and N, respectively, F is equivariant if the following diagram commutes for each $g \in G$:

$$\begin{tikzcd} M \arrow[r, "F"] \arrow[d, "\theta_g"'] & N \arrow[d, "\varphi_g"] \\ M \arrow[r, "F"'] & N. \end{tikzcd}$$

This condition is also expressed by saying that F *intertwines* the two G-actions.

Example 9.5. Let $v = (v^1, \ldots, v^n) \in \mathbb{R}^n$ be any fixed nonzero vector. Define smooth left actions of $\mathbb{R}$ on $\mathbb{R}^n$ and $\mathbb{T}^n$ by

$$t \cdot (x^1, \ldots, x^n) = (x^1 + tv^1, \ldots, x^n + tv^n),$$
$$t \cdot (z^1, \ldots, z^n) = (e^{2\pi i t v^1} z^1, \ldots, e^{2\pi i t v^n} z^n)$$

for $t \in \mathbb{R}$, $(x^1, \ldots, x^n) \in \mathbb{R}^n$, and $(z^1, \ldots, z^n) \in \mathbb{T}^n$. The smooth covering map $E\colon \mathbb{R}^n \to \mathbb{T}^n$ given by $E(x^1, \ldots, x^n) = (e^{2\pi i x^1}, \ldots, e^{2\pi i x^n})$ is equivariant with respect to these actions.

Example 9.6. Let G and H be Lie groups, and let $F\colon G \to H$ be a Lie group homomorphism. There is a natural left action of G on itself by left translation. Define a left action θ of G on H by

$$\theta_g(h) = F(g)h.$$

To check that this is an action, we just observe that $\theta_e(h) = F(e)h = h$, and

$$\theta_{g_1} \circ \theta_{g_2}(h) = F(g_1)(F(g_2)h) = (F(g_1)F(g_2))h = F(g_1g_2)h = \theta_{g_1g_2}(h)$$

because F is a homomorphism. With respect to these G-actions, F is equivariant because

$$\theta_g \circ F(g') = F(g)F(g') = F(gg') = F \circ L_g(g').$$

The following theorem is an extremely useful tool for proving that certain sets are embedded submanifolds.

Theorem 9.7 (Equivariant Rank Theorem). *Let M and N be smooth manifolds and let G be a Lie group. Suppose $F\colon M \to N$ is a smooth map that is equivariant with respect to a transitive smooth G-action on M and any smooth G-action on N. Then F has constant rank. In particular, its level sets are closed embedded submanifolds of M.*

Proof. Let θ and φ denote the G-actions on M and N, respectively, and let p_0 be any point in M. For any other point $p \in M$, choose $g \in G$ such that $\theta_g(p_0) = p$. (Such a g exists because we are assuming G acts transitively on M.) Because $\varphi_g \circ F = F \circ \theta_g$, the following diagram commutes (see Figure 9.1):

$$\begin{tikzcd} T_{p_0}M \arrow[r, "F_*"] \arrow[d, "\theta_{g*}"'] & T_{F(p_0)}N \arrow[d, "\varphi_{g*}"] \\ T_pM \arrow[r, "F_*"'] & T_{F(p)}N. \end{tikzcd}$$

Because the vertical linear maps in this diagram are isomorphisms, the horizontal ones have the same rank. In other words, the rank of F_* at an arbitrary point p is the same as its rank at p_0, so F has constant rank. □

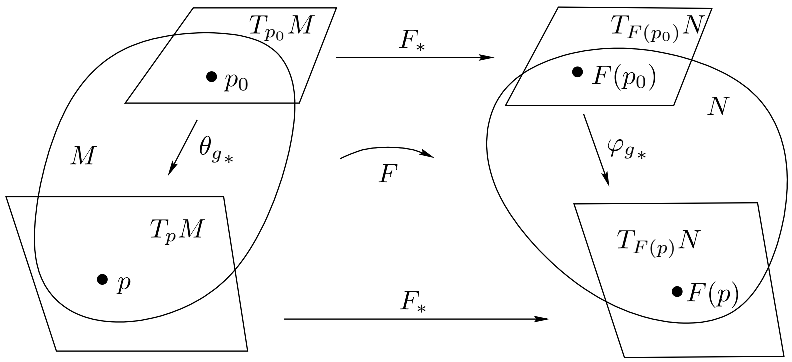


Figure 9.1. The equivariant rank theorem.

Here are some applications of the equivariant rank theorem.

Proposition 9.8. *Let $F\colon G \to H$ be a Lie group homomorphism. The kernel of F is an embedded Lie subgroup of G, whose codimension is equal to the rank of F.*

Proof. As in Example 9.6, F is equivariant with respect to suitable G-actions on G and H. Since the action of G on itself by left translation is transitive, it follows that F has constant rank, so its kernel $F^{-1}(e)$ is an embedded submanifold. It is thus a Lie subgroup by Proposition 8.28. □

As another application, we describe a few more Lie subgroups of $\mathrm{GL}(n,\mathbb{C})$. For any complex matrix A, let A^* denote the *adjoint* or conjugate transpose of A: $A^* = \overline{A}^T$. Observe that $(AB)^* = (\overline{AB})^T = \overline{B}^T\overline{A}^T = B^*A^*$. Consider the following subgroups of $\mathrm{GL}(n,\mathbb{C})$:

- The Complex Special Linear Group:

$$\mathrm{SL}(n,\mathbb{C}) = \{A \in \mathrm{GL}(n,\mathbb{C}) : \det A = 1\}.$$

- The Unitary Group:

$$\mathrm{U}(n) = \{A \in \mathrm{GL}(n,\mathbb{C}) : A^*A = I_n\}.$$

- The Special Unitary Group:

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

◇ **Exercise 9.2.** Show that $\mathrm{SL}(n,\mathbb{C})$, $\mathrm{U}(n)$, and $\mathrm{SU}(n)$ are subgroups of $\mathrm{GL}(n,\mathbb{C})$ (in the algebraic sense).

◇ **Exercise 9.3.** Show that a matrix is in $\mathrm{U}(n)$ if and only if its columns form an orthonormal basis for $\mathbb{C}^n$ with respect to the Hermitian dot product $z \cdot w = \sum_i z^i\overline{w^i}$.

Proposition 9.9. *The unitary group* $\mathrm{U}(n)$ *is an embedded* n^2*-dimensional Lie subgroup of* $\mathrm{GL}(n,\mathbb{C})$.

Proof. Clearly $\mathrm{U}(n)$ is a level set of the map $\Phi\colon \mathrm{GL}(n,\mathbb{C}) \to \mathrm{M}(n,\mathbb{C})$ defined by

$$\Phi(A) = A^*A.$$

To show that Φ has constant rank and therefore that $\mathrm{U}(n)$ is an embedded Lie subgroup, we will show that Φ is equivariant with respect to suitable right actions of $\mathrm{GL}(n,\mathbb{C})$. Let $\mathrm{GL}(n,\mathbb{C})$ act on itself by right multiplication, and define a right action of $\mathrm{GL}(n,\mathbb{C})$ on $\mathrm{M}(n,\mathbb{C})$ by

$$X \cdot B = B^*XB \qquad \text{for } X \in \mathrm{M}(n,\mathbb{C}),\ B \in \mathrm{GL}(n,\mathbb{C}).$$

It is easy to check that this is a smooth action, and Φ is equivariant because

$$\Phi(AB) = (AB)^*(AB) = B^*A^*AB = B^*\Phi(A)B = \Phi(A) \cdot B.$$

Thus $\mathrm{U}(n)$ is an embedded Lie subgroup of $\mathrm{GL}(n,\mathbb{C})$.

To determine its dimension, we need to compute the rank of Φ. Because the rank is constant, it suffices to compute it at the identity $I_n \in \mathrm{GL}(n,\mathbb{C})$. Thus for any $B \in T_{I_n}\mathrm{GL}(n,\mathbb{C}) = \mathrm{M}(n,\mathbb{C})$, let $\gamma\colon (-\varepsilon,\varepsilon) \to \mathrm{GL}(n,\mathbb{C})$ be the curve $\gamma(t) = I_n + tB$, and compute

$$\begin{aligned}\Phi_*B &= \left.\frac{d}{dt}\right|_{t=0} \Phi\circ\gamma(t)\\ &= \left.\frac{d}{dt}\right|_{t=0} (I_n + tB)^*(I_n + tB)\\ &= B^* + B.\end{aligned}$$

The image of this linear map is the set of all *Hermitian* $n \times n$ matrices, i.e., the set of $A \in \mathrm{M}(n,\mathbb{C})$ satisfying $A = A^*$. This is a (real) vector space of dimension n^2, as you can check. Therefore $\mathrm{U}(n)$ is an embedded Lie subgroup of dimension $2n^2 - n^2 = n^2$. □

Proposition 9.10. *The complex special linear group* $\mathrm{SL}(n,\mathbb{C})$ *is an embedded* $(2n^2-2)$*-dimensional Lie subgroup of* $\mathrm{GL}(n,\mathbb{C})$.

Proof. Just note that $\mathrm{SL}(n,\mathbb{C})$ is the kernel of the Lie group homomorphism $\det\colon \mathrm{GL}(n,\mathbb{C}) \to \mathbb{C}^*$. It is easy to check that this homomorphism is surjective, so it is a submersion by Theorem 7.14(a). Therefore $\mathrm{SL}(n,\mathbb{C}) = \mathrm{Ker}(\det)$ is an embedded Lie subgroup whose codimension is equal to $\dim \mathbb{C}^* = 2$. □

Proposition 9.11. *The special unitary group* $\mathrm{SU}(n)$ *is an embedded* (n^2-1)*-dimensional Lie subgroup of* $\mathrm{GL}(n,\mathbb{C})$.

Proof. We will show first that $\mathrm{SU}(n)$ is an embedded submanifold of $\mathrm{U}(n)$. Since the composition of smooth embeddings $\mathrm{SU}(n) \hookrightarrow \mathrm{U}(n) \hookrightarrow \mathrm{GL}(n,\mathbb{C})$ is again a smooth embedding, $\mathrm{SU}(n)$ is also embedded in $\mathrm{GL}(n,\mathbb{C})$.

If $A \in \mathrm{U}(n)$, then

$$1 = \det I_n = \det(A^*A) = (\det A^*)(\det A) = (\overline{\det A})(\det A) = |\det A|^2.$$

Thus $\det\colon \mathrm{U}(n) \to \mathbb{C}^*$ actually takes its values in $\mathbb{S}^1$. It is easy to check that it is surjective onto $\mathbb{S}^1$, so it is a submersion by Theorem 7.14(a). Therefore its kernel $\mathrm{SU}(n)$ is an embedded Lie subgroup of codimension 1 in $\mathrm{U}(n)$. $\square$

◇ **Exercise 9.4.** Use the techniques developed in this section to give simpler proofs than the ones in Chapter 8 that $\mathrm{O}(n)$, $\mathrm{SO}(n)$, and $\mathrm{SL}(n, \mathbb{R})$ are Lie subgroups of $\mathrm{GL}(n, \mathbb{R})$.

Proper Actions

Suppose we are given a continuous action of a Lie group G on a manifold M. The action is said to be *proper* if the map $G \times M \to M \times M$ given by $(g, p) \mapsto (g \cdot p, p)$ is a proper map. (Note that this is *not* the same as requiring that the map $G \times M \to M$ defining the action be a proper map.)

It is not always easy to tell whether a given action is proper. The next two propositions give alternative characterizations of proper actions that are often useful. Given an action of G on M, for any $g \in G$ and any subset $K \subset M$, we will use the notation $g \cdot K$ to denote the set $\{g \cdot x : x \in K\}$.

Proposition 9.12. *Suppose a Lie group G acts continously on a manifold M. The action is proper if and only if for every compact subset $K \subset M$, the set $G_K = \{g \in G\colon (g \cdot K) \cap K \neq \varnothing\}$ is compact.*

Proof. Let $\Theta\colon G \times M \to M \times M$ denote the map $\Theta(g, p) = (g \cdot p, p)$. Suppose first that Θ is proper. Then for any compact set $K \subset M$, it is easy to check that

$$\begin{aligned} G_K &= \{g \in G : \text{there exists } p \in K \text{ such that } g \cdot p \in K\} \\ &= \{g \in G : \text{there exists } p \in M \text{ such that } \Theta(g, p) \in K \times K\} \\ &= \pi_G\big(\Theta^{-1}(K \times K)\big), \end{aligned} \tag{9.4}$$

where $\pi_G\colon G \times M \to G$ is the projection (Figure 9.2). Thus G_K is compact. Conversely, suppose G_K is compact for every compact set $K \subset M$. If $L \subset M \times M$ is compact, let $K = \pi_1(L) \cup \pi_2(L) \subset M$, where $\pi_1, \pi_2\colon M \times M \to M$ are the projections onto the first and second factors, respectively. Then

$$\Theta^{-1}(L) \subset \Theta^{-1}(K \times K) \subset \{(g, p) : g \cdot p \in K \text{ and } p \in K\} \subset G_K \times K.$$

Since $\Theta^{-1}(L)$ is closed by continuity, it is a closed subset of the compact set $G_K \times K$ and is therefore compact. $\square$

Proposition 9.13. *Let M be a manifold, and let G be a Lie group acting continuously on M. The action is proper if and only if the following*

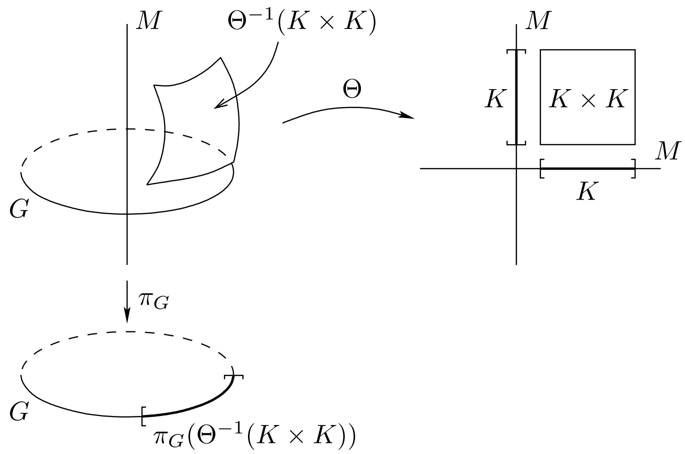


Figure 9.2. Characterizing proper actions.

condition is satisfied:

> *If $\{p_i\}$ is a convergent sequence in M and $\{g_i\}$ is a sequence in G such that $\{g_i \cdot p_i\}$ converges, then a subsequence of $\{g_i\}$ converges.* (9.5)

Proof. Let $\Theta(g, p) = (g \cdot p, p)$ as in the proof of the previous proposition. If Θ is proper and $\{p_i\}$, $\{g_i\}$ are sequences satisfying the hypotheses of (9.5), let U and V be precompact neighborhoods of the points $p = \lim_i p_i$ and $q = \lim_i (g_i \cdot q_i)$, respectively. The assumption means that the points $\Theta(g_i, p_i)$ all lie in the compact set $\overline{U} \times \overline{V}$ when i is large enough, so a subsequence of $\{(g_i, p_i)\}$ converges. In particular, this means that a subsequence of $\{g_i\}$ converges in G.

Conversely, suppose (9.5) holds, and let $L \subset M \times M$ be a compact set. If $\{(g_i, p_i)\}$ is any sequence in $\Theta^{-1}(L)$, then $\Theta(g_i, p_i) = (g_i \cdot p_i, p_i)$ lies in L, so passing to a subsequence, we obtain sequences $\{p_i\}$ and $\{g_i\}$ satisfying the hypotheses of (9.5). The corresponding subsequence of $\{(g_i, p_i)\}$ converges in $G \times M$, and since $\Theta^{-1}(L)$ is closed in $G \times M$ by continuity, the limit lies in $\Theta^{-1}(L)$. □

One case in which this condition is automatic is when the group is compact.

Corollary 9.14. *Any continuous action by a compact Lie group on a manifold is proper.*

Proof. If $\{p_i\}$ and $\{g_i\}$ are any sequences satisfying the hypotheses of (9.5) (or not!), a subsequence of $\{g_i\}$ converges. □

In the special case in which K is a one-point set $\{p\}$, the set G_K is just the isotropy group of p, as you can easily check. Thus a simple necessary condition for a Lie group action to be proper is that each isotropy group must be compact. For example, the action of $\mathbb{R}^*$ on $\mathbb{R}^n$ given by

$$t \cdot (x^1, \ldots, x^n) = (tx^1, \ldots, tx^n)$$

is not proper, because the isotropy group of the origin is all of $\mathbb{R}^*$, which is not compact.

Quotients of Manifolds by Group Actions

Suppose a Lie group G acts on a manifold M (on the left, say). Define a relation on M by setting $p \sim q$ if and only if there exists $g \in G$ such that $g \cdot p = q$. This is an equivalence relation, whose equivalence classes are exactly the orbits of G in M. The set of orbits is denoted by M/G; with the quotient topology, it is called the *orbit space* of the action. It is of great importance to determine conditions under which an orbit space is a smooth manifold.

One simple but important example to keep in mind is the action of $\mathbb{R}^k$ on $\mathbb{R}^k \times \mathbb{R}^n$ by translation in the $\mathbb{R}^k$ factor: $\theta_v(x, y) = (v + x, y)$. The orbits are the affine subspaces parallel to $\mathbb{R}^k$, and the orbit space $(\mathbb{R}^k \times \mathbb{R}^n)/\mathbb{R}^k$ is homeomorphic to $\mathbb{R}^n$. The quotient map $\pi\colon \mathbb{R}^k \times \mathbb{R}^n \to \mathbb{R}^n$ is a smooth submersion.

It is worth noting that some authors use distinctive notations such as M/G and $G\backslash M$ to distinguish between orbit spaces determined by left actions and right actions. We will rely on the context, not the notation, to distinguish between the two cases.

One simple property of quotients by Lie group actions that will prove useful is given in the following lemma.

Lemma 9.15. *For any continuous action of a Lie group G on a manifold M, the quotient map $\pi\colon M \to M/G$ is open.*

Proof. For any open set $U \subset M$, $\pi^{-1}(\pi(U))$ is equal to the union of all sets of the form $\theta_g(U)$ as g ranges over G. Since θ_g is a homeomorphism, each such set is open, and therefore $\pi^{-1}(\pi(U))$ is open in M. Because π is a quotient map, this implies that $\pi(U)$ is open in M/G, and therefore π is an open map. □

The following theorem gives a very general sufficient condition for the quotient of a smooth manifold by a group action to be a smooth manifold. It is one of the most important applications of the inverse function theorem that we will see.

Theorem 9.16 (Quotient Manifold Theorem). *Suppose a Lie group G acts smoothly, freely, and properly on a smooth manifold M. Then* *the orbit space M/G is a topological manifold of dimension equal to* $\dim M - \dim G$, *and has a unique smooth structure with the property that the quotient map $\pi\colon M \to M/G$ is a smooth submersion.*

Proof. First we prove the uniqueness of the smooth structure. Suppose M/G has two different smooth structures such that $\pi\colon M \to M/G$ is a smooth submersion. Let $(M/G)_1$ and $(M/G)_2$ denote M/G with the first and second smooth structures, respectively. By Proposition 7.16, the identity map is smooth from $(M/G)_1$ to $(M/G)_2$:

$$\begin{tikzcd} M \arrow[d, "\pi"'] \arrow[dr, "\pi"] & \\ (M/G)_1 \arrow[r, "\mathrm{Id}"'] & (M/G)_2. \end{tikzcd}$$

The same argument shows that it is also smooth in the opposite direction, so the two smooth structures are identical.

Next we prove that M/G is a topological manifold. Assume for definiteness that G acts on the left, and let $\theta\colon G \times M \to M$ denote the action and $\Theta\colon G \times M \to M \times M$ the proper map $\Theta(g,p) = (g \cdot p, p)$.

If $\{U_i\}$ is a countable basis for the topology of M, then $\{\pi(U_i)\}$ is a countable collection of open subsets of M/G (because π is an open map), and it is easy to check that it is a basis for the topology of M/G. Thus M/G is second countable.

To show that M/G is Hausdorff, define the *orbit relation* $\mathcal{O} \subset M \times M$ by

$$\mathcal{O} = \Theta(G \times M) = \{(g \cdot p, p) \in M \times M : p \in M,\ g \in G\}.$$

(It is called the orbit relation because $(q,p) \in \mathcal{O}$ if and only if p and q are in the same G-orbit.) Since proper continuous maps are closed (Proposition 2.18), it follows that $\mathcal{O}$ is a closed subset of $M \times M$. If $\pi(p)$ and $\pi(q)$ are distinct points in M/G, then p and q lie in distinct orbits, so $(p,q) \notin \mathcal{O}$. If $U \times V$ is a product neighborhood of (p,q) in $M \times M$ that is disjoint from $\mathcal{O}$, then $\pi(U)$ and $\pi(V)$ are disjoint open subsets of M/G containing $\pi(p)$ and $\pi(q)$, respectively. Thus M/G is Hausdorff.

Before proving that M/G is locally Euclidean, we will show that the G-orbits are embedded submanifolds of M diffeomorphic to G. For any $p \in M$, define the *orbit map* $\theta^{(p)}\colon G \to M$ by

$$\theta^{(p)}(g) = g \cdot p. \tag{9.6}$$

This is a smooth map whose image is exactly the G-orbit of p. We will show that $\theta^{(p)}$ is a smooth embedding. First, if $\theta^{(p)}(g') = \theta^{(p)}(g)$, then $g' \cdot p = g \cdot p$, which implies $(g^{-1}g') \cdot p = p$. Since we are assuming G acts freely on M, this can only happen if $g^{-1}g' = e$, which means $g = g'$; thus

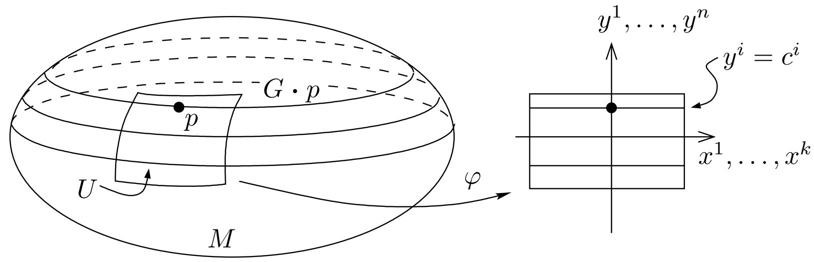


Figure 9.3. An adapted coordinate chart.

$\theta^{(p)}$ is injective. Observe that

$$\theta^{(p)}(g'g) = (g'g) \cdot p = g' \cdot (g \cdot p) = g' \cdot \theta^{(p)}(g), \tag{9.7}$$

so $\theta^{(p)}$ is equivariant with respect to left translation on G and the given action on M. Since G acts transitively on itself, the equivariant rank theorem implies that $\theta^{(p)}$ has constant rank. Since it is also injective, it is an immersion by Theorem 7.14.

If $K \subset M$ is a compact set, then $(\theta^{(p)})^{-1}(K)$ is closed in G by continuity, and since it is contained in $G_{K \cup \{p\}}$, it is compact by Proposition 9.12. Therefore, $\theta^{(p)}$ is a proper map. We have shown that $\theta^{(p)}$ is a proper injective immersion, so it is a smooth embedding by Proposition 7.4.

Let $k = \dim G$ and $n = \dim M - \dim G$. Let us say that a smooth chart (U, φ) on M, with coordinate functions $(x, y) = (x^1, \dots, x^k, y^1, \dots, y^n)$, is *adapted* to the G-action if

(i) $\varphi(U)$ is a product open set $U_1 \times U_2 \subset \mathbb{R}^k \times \mathbb{R}^n$, and

(ii) each orbit intersects U either in the empty set or in a single slice of the form $\{y^1 = c^1, \dots, y^n = c^n\}$.

(See Figure 9.3.)

We will show that for any $p \in M$, there exists an adapted coordinate chart centered at p. To prove this, we begin by choosing any slice chart (W, φ_0) centered at p for the orbit $G \cdot p$ in M. Write the coordinate functions of φ_0 as $(u^1, \dots, u^k, v^1, \dots, v^n)$, so that $(G \cdot p) \cap W$ is the slice $\{v^1 = \cdots = v^n = 0\}$. Let S be the submanifold of W defined by $u^1 = \cdots = u^k = 0$. (This is the slice "perpendicular" to the orbit in these coordinates.) Thus T_pM decomposes as the following direct sum:

$$T_pM = T_p(G \cdot p) \oplus T_pS,$$

where $T_p(G \cdot p)$ is the span of $(\partial/\partial u^i)$ and T_pS is the span of $(\partial/\partial v^i)$.

Let $\psi\colon G \times S \to M$ denote the restriction of the action θ to $G \times S \subset G \times M$. We will use the inverse function theorem to show that ψ is a diffeomorphism in a neighborhood of $(e, p) \in G \times S$. Let $i_p\colon G \to G \times S$ be

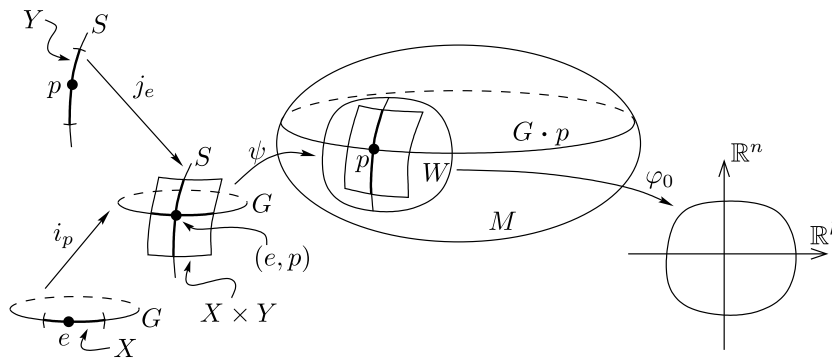


Figure 9.4. Finding an adapted chart.

the smooth embedding given by $i_p(g) = (g, p)$. The orbit map $\theta^{(p)}\colon G \to M$ is equal to the composition

$$G \xrightarrow{i_p} G \times S \xrightarrow{\psi} M.$$

(See Figure 9.4.) Since $\theta^{(p)}$ is a smooth embedding whose image is the orbit $G \cdot p$, it follows that $\theta^{(p)}_*(T_eG)$ is equal to the subspace $T_p(G \cdot p) \subset T_pM$, and thus the image of $\psi_*\colon T_{(e,p)}(G \times S) \to T_pM$ contains $T_p(G \cdot p)$. Similarly, if $j_e\colon S \to G \times S$ is the smooth embedding $j_e(q) = (e, q)$, then the inclusion $\iota\colon S \hookrightarrow M$ is equal to the composition

$$S \xrightarrow{j_e} G \times S \xrightarrow{\psi} M.$$

Therefore, the image of ψ_* also includes $T_pS \subset T_pM$. Since $T_p(G \cdot p)$ and T_pS together span T_pM, $\psi_*\colon T_{(e,p)}(G \times S) \to T_pM$ is surjective, and for dimensional reasons, it is bijective. By the inverse function theorem, there exist a neighborhood (which we may assume to be a product neighborhood) $X \times Y$ of (e, p) in $G \times S$ and a neighborhood U of p in M such that $\psi\colon X \times Y \to U$ is a diffeomorphism. Shrinking X and Y if necessary, we may assume that X and Y are precompact sets that are diffeomorphic to Euclidean balls in $\mathbb{R}^k$ and $\mathbb{R}^n$, respectively.

We need to show that $Y \subset S$ can be chosen small enough that each G-orbit intersects Y in at most a single point. Suppose this is not true. Then if $\{Y_i\}$ is a countable neighborhood basis for Y at p (e.g., a sequence of coordinate balls whose diameters decrease to 0), for each i there exist distinct points $p_i, p_i' \in Y_i$ that are in the same orbit, which is to say that $g_i \cdot p_i = p_i'$ for some $g_i \in G$. Since $\{Y_i\}$ is a neighborhood basis, both sequences $\{p_i\}$ and $\{p_i' = g_i \cdot p_i\}$ converge to p. By Proposition 9.13, we may pass to a subsequence and assume that $g_i \to g \in G$. By continuity, therefore,

$$g \cdot p = \lim_{i\to\infty} g_i \cdot p_i = \lim_{i\to\infty} p_i' = p.$$

Since G acts freely, this implies $g = e$. When i gets large enough, therefore, $g_i \in X$. But this contradicts the fact that $\psi = \theta|_{X\times Y}$ is injective, because

$$\theta_{g_i}(p_i) = p_i' = \theta_e(p_i'),$$

and we are assuming $p_i \neq p_i'$.

Choose diffeomorphisms $\alpha\colon \mathbb{B}^k \to X$ and $\beta\colon \mathbb{B}^n \to Y$ (where $\mathbb{B}^k$ and $\mathbb{B}^n$ are the open unit balls in $\mathbb{R}^k$ and $\mathbb{R}^n$, respectively), and define $\gamma\colon \mathbb{B}^k \times \mathbb{B}^n \to U$ by $\gamma(x,y) = \theta_{\alpha(x)}(\beta(y))$. Because γ is equal to the composition of diffeomorphisms

$$\mathbb{B}^k \times \mathbb{B}^n \xrightarrow{\alpha\times\beta} X \times Y \xrightarrow{\psi} U,$$

γ is a diffeomorphism. The map $\varphi = \gamma^{-1}$ is therefore a smooth coordinate map on U. We will show that φ is adapted to the G-action. Condition (i) is obvious by construction. Observe that each $y =$ constant slice is contained in a single orbit, because it is of the form $\theta(X\times\{p_0\}) \subset \theta(G\times\{p_0\}) = G\cdot p_0$, where $p_0 \in Y$ is the point whose y-coordinate is the given constant. Thus if an arbitrary orbit intersects U, it does so in a union of $y =$ constant slices. However, since an orbit can intersect Y at most once, and each $y =$ constant slice has a point in Y, it follows that each orbit intersects U in precisely one slice if at all. This completes the proof that adapted coordinate charts exist.

To finish the proof that M/G is locally Euclidean, let $q = \pi(p)$ be an arbitrary point of M/G, and let (U,φ) be an adapted coordinate chart for M centered at p, with $\varphi(U) = U_1 \times U_2 \subset \mathbb{R}^k \times \mathbb{R}^n$. Let $V = \pi(U)$ (Figure 9.5), which is an open subset of M/G because π is an open map. Writing the coordinate functions of φ as $(x^1,\dots,x^k,y^1,\dots,y^n)$ as before, let $Y \subset U$ be the slice $\{x^1 = \cdots = x^k = 0\}$. Note that $\pi\colon Y \to V$ is bijective by the definition of an adapted chart. Moreover, if W is an open subset of Y, then

$$\pi(W) = \pi\bigl(\{(x,y) : (0,y) \in W\}\bigr)$$

is open in M/G, and thus $\pi|_Y$ is a homeomorphism. Let $\sigma = (\pi|_Y)^{-1}\colon V \to Y \subset U$, which is a local section of π.

Define a map $\eta\colon V \to U_2$ by sending the equivalence class of a point (x,y) to y; this is well-defined by the definition of an adapted chart. Formally, $\eta = \pi_2 \circ \varphi \circ \sigma$, where $\pi_2\colon U_1 \times U_2 \to U_2 \subset \mathbb{R}^n$ is the projection onto the second factor. Because σ is a homeomorphism from V to Y and $\pi_2 \circ \varphi$ is a homeomorphism from Y to U_2, it follows that η is a homeomorphism. This completes the proof that M/G is a topological n-manifold.

Finally, we need to show that M/G has a smooth structure such that π is a submersion. We will use the atlas consisting of all charts (V,η) as constructed in the preceding paragraph. With respect to any such chart

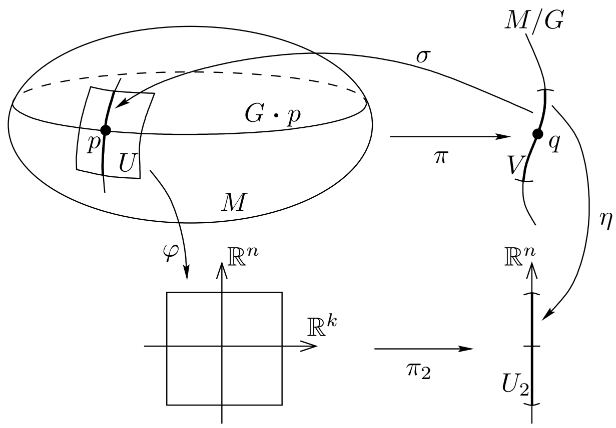


Figure 9.5. A coordinate chart for M/G.

for M/G and the corresponding adapted chart for M, π has the coordinate representation $\pi(x,y) = y$, which is certainly a submersion. Thus we need only show that any two such charts for M/G are smoothly compatible.

Let (U,φ) and $(\widetilde{U},\widetilde{\varphi})$ be two adapted charts for M, and let (V,η) and $(\widetilde{V},\widetilde{\eta})$ be the corresponding charts for M/G. First consider the case in which the two adapted charts are both centered at the same point $p \in M$. Writing the adapted coordinates as (x,y) and $(\widetilde{x},\widetilde{y})$, the fact that the coordinates are adapted to the G-action means that two points with the same y-coordinate are in the same orbit, and therefore also have the same $\widetilde{y}$-coordinate. This means that the transition map between these coordinates can be written $(\widetilde{x},\widetilde{y}) = (A(x,y), B(y))$, where A and B are smooth maps defined on some neighborhood of the origin. The transition map $\widetilde{\eta}\circ\eta^{-1}$ is just $\widetilde{y} = B(y)$, which is clearly smooth.

In the general case, suppose (U,φ) and $(\widetilde{U},\widetilde{\varphi})$ are adapted charts for M, and $p \in U$, $\widetilde{p} \in \widetilde{U}$ are points such that $\pi(p) = \pi(\widetilde{p}) = q$. Modifying both charts by adding constant vectors, we can assume that they are centered at p and $\widetilde{p}$, respectively. Since p and $\widetilde{p}$ are in the same orbit, there is a group element g such that $g\cdot p = \widetilde{p}$. Because θ_g is a diffeomorphism taking orbits to orbits, it follows that $\widetilde{\varphi}' = \widetilde{\varphi}\circ\theta_g$ is another adapted chart centered at p. Moreover, $\widetilde{\sigma}' = \theta_g^{-1}\circ\widetilde{\sigma}$ is the local section corresponding to $\widetilde{\varphi}'$, and therefore $\widetilde{\eta}' = \pi_2\circ\widetilde{\varphi}'\circ\widetilde{\sigma}' = \pi_2\circ\widetilde{\varphi}\circ\theta_g\circ\theta_g^{-1}\circ\widetilde{\sigma} = \pi_2\circ\widetilde{\varphi}\circ\widetilde{\sigma} = \widetilde{\eta}$. Thus we are back in the situation of the preceding paragraph, and the two charts are smoothly compatible. $\square$

Homogeneous Spaces

One of the most interesting kinds of group action is that in which a group acts transitively. A smooth manifold endowed with a transitive smooth action by a Lie group G is called a *homogeneous G-space*, or a *homogeneous space* or *homogeneous manifold* if it is not important to specify the group.

In most examples of homogeneous spaces, the group action preserves some property of the manifold (such as distances in some metric, or a class of curves such as straight lines in the plane); then the fact that the action is transitive means that the manifold "looks the same" everywhere from the point of view of this property. Often, homogeneous spaces are models for various kinds of geometric structures, and as such they play a central role in many areas of differential geometry.

Here are some important examples of homogeneous spaces.

Example 9.17 (Homogeneous Spaces).

(a) The natural action of $\mathrm{O}(n)$ on $\mathbb{S}^{n-1}$ is transitive, as we observed in Example 9.1. So is the natural action of $\mathrm{SO}(n)$ on $\mathbb{S}^{n-1}$ when $n \geq 2$. Thus for $n \geq 2$, $\mathbb{S}^{n-1}$ is a homogeneous space of either $\mathrm{O}(n)$ or $\mathrm{SO}(n)$.

(b) Let $E(n)$ denote the subgroup of $\mathrm{GL}(n+1,\mathbb{R})$ consisting of matrices of the form

$$\left\{ \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix} : A \in \mathrm{O}(n),\ b \in \mathbb{R}^n \right\},$$

where b is considered as an $n \times 1$ column matrix. It is straightforward to check that $E(n)$ is an embedded Lie subgroup. If $S \subset \mathbb{R}^{n+1}$ denotes the affine subspace defined by $x^{n+1} = 1$, then a simple computation shows that $E(n)$ takes S to itself. Identifying S with $\mathbb{R}^n$ in the obvious way, this induces a smooth action of $E(n)$ on $\mathbb{R}^n$, in which the matrix $\left(\begin{smallmatrix} A & b \\ 0 & 1 \end{smallmatrix}\right)$ sends x to $Ax + b$. It is not hard to prove that these are precisely the diffeomorphisms of $\mathbb{R}^n$ that preserve the Euclidean distance function (see Problem 9-1). For this reason, $E(n)$ is called the *Euclidean group*. Because any point in $\mathbb{R}^n$ can be taken to any other by a translation, $E(n)$ acts transitively on $\mathbb{R}^n$, so $\mathbb{R}^n$ is a homogeneous $E(n)$-space.

(c) The group $\mathrm{SL}(2,\mathbb{R})$ acts smoothly and transitively on the upper half-plane $\mathbb{H} = \{z \in \mathbb{C} : \operatorname{Im} z > 0\}$ by the formula

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{az+b}{cz+d}.$$

The resulting complex-analytic transformations of $\mathbb{H}$ are called *Möbius transformations.*

(d) The natural action of $\mathrm{GL}(n,\mathbb{C})$ on $\mathbb{C}^n$ restricts to natural actions of both $\mathrm{U}(n)$ and $\mathrm{SU}(n)$ on $\mathbb{S}^{2n-1}$, thought of as the set of unit vectors in $\mathbb{C}^n$. The next exercise shows that these actions are smooth and transitive.

◇ **Exercise 9.5.** Show that the natural actions of $\mathrm{U}(n)$ and $\mathrm{SU}(n)$ on $\mathbb{S}^{2n-1}$ are smooth and transitive.

Next we will describe a very general construction that can be used to generate a great number of homogeneous spaces, as quotients of Lie groups by closed Lie subgroups.

Let G be a Lie group and let $H \subset G$ be a Lie subgroup. For each $g \in G$, the *left coset* of g modulo H is the set

$$gH = \{gh : h \in H\}.$$

The set of left cosets modulo H is denoted by G/H; with the quotient topology determined by the natural map $\pi\colon G \to G/H$ sending each element $g \in G$ to its coset, it is called the *left coset space* of G modulo H. Two elements $g_1, g_2 \in G$ are in the same coset modulo H if and only if $g_1^{-1}g_2 \in H$; in this case we write $g_1 \equiv g_2 \pmod{H}$.

Theorem 9.18 (Homogeneous Space Construction Theorem). *Let G be a Lie group and let H be a closed Lie subgroup of G. The left coset space G/H has a unique smooth manifold structure such that the quotient map $\pi\colon G \to G/H$ is a smooth submersion. The left action of G on G/H given by*

$$g_1 \cdot (g_2H) = (g_1g_2)H \tag{9.8}$$

turns G/H into a homogeneous G-space.

Proof. If we let H act on G by right translation, then $g_1, g_2 \in G$ are in the same H-orbit if and only if $g_1h = g_2$ for some $h \in H$, which is the same as saying that g_1 and g_2 are in the same coset modulo H. In other words, the orbit space determined by the *right* action of H on G is precisely the *left* coset space G/H.

We already observed in Example 9.1(f) that H acts smoothly and freely on G. To see that the action is proper, we will use Proposition 9.13. Suppose $\{g_i\}$ is a convergent sequence in G and $\{h_i\}$ is a sequence in H such that $\{g_ih_i\}$ converges. By continuity, $h_i = g_i^{-1}(g_ih_i)$ converges to a point in G, and since H is closed in G it follows that $\{h_i\}$ converges in H.

The quotient manifold theorem now implies that G/H has a unique smooth manifold structure such that the quotient map $\pi\colon G \to G/H$ is a submersion. Since a product of submersions is a submersion, it follows that $\mathrm{Id}_G \times \pi\colon G \times G \to G \times G/H$ is also a submersion. Consider the following diagram:

$$\begin{tikzcd} G \times G \arrow[r, "m"] \arrow[d, "\mathrm{Id}_G \times \pi"'] & G \arrow[d, "\pi"] \\ G \times G/H \arrow[r, "\theta"'] & G/H, \end{tikzcd}$$

where m is group multiplication and θ is the action of G on G/H given by (9.8). It is straightforward to check that $\pi \circ m$ is constant on the fibers of $\mathrm{Id}_G \times \pi$, and therefore θ is well-defined and smooth by Proposition 7.17. Finally, given any two points $g_1H, g_2H \in G/H$, the element $g_2g_1^{-1} \in G$ satisfies $(g_2g_1^{-1}) \cdot g_1H = g_2H$, so the action is transitive. □

The homogeneous spaces constructed in this theorem turn out to be of central importance because, as the next theorem shows, *every* homogeneous space is equivalent to one of this type. First we note that the isotropy group of any smooth Lie group action is a closed Lie subgroup.

Lemma 9.19. *If M is a smooth G-space, then for each $p \in M$, the isotropy group G_p is a closed, embedded Lie subgroup of G.*

Proof. For each $p \in M$, $G_p = (\theta^{(p)})^{-1}(p)$, where $\theta^{(p)} \colon G \to M$ is the orbit map defined by (9.6). As we observed in the proof of the quotient manifold theorem, $\theta^{(p)}$ is equivariant with respect to the action by G on itself by left multiplication and the given G-action on M (see (9.7)). The equivariant rank theorem implies that G_p is an embedded submanifold of G, and therefore is a closed Lie subgroup by Proposition 8.28. □

Theorem 9.20 (Characterization of Homogeneous Spaces). *Let M be a homogeneous G-space, and let p be any point of M. Then the map $F \colon G/G_p \to M$ defined by $F(gG_p) = g \cdot p$ is an equivariant diffeomorphism.*

Proof. For simplicity, let us write $H = G_p$. To see that F is well-defined, assume that $g_1H = g_2H$, which means that $g_1^{-1}g_2 \in H$. Writing $g_1^{-1}g_2 = h$, we see that

$$F(g_2H) = g_2 \cdot p = g_1h \cdot p = g_1 \cdot p = F(g_1H).$$

Also, F is equivariant, because

$$F(g'gH) = (g'g) \cdot p = g' \cdot F(gH).$$

It is smooth because it is obtained from the orbit map $\theta^{(p)} \colon G \to M$ by passing to the quotient (see Proposition 7.17).

Next we show that F is bijective. Given any point $q \in M$ there is a group element $g \in G$ such that $F(gH) = g \cdot p = q$ by transitivity. On the other hand, if $F(g_1H) = F(g_2H)$, then $g_1 \cdot p = g_2 \cdot p$ implies $g_1^{-1}g_2 \cdot p = p$, so $g_1^{-1}g_2 \in H$, which implies $g_1H = g_2H$.

Because F is equivariant, it has constant rank; and because it is bijective, it is a diffeomorphism by Theorem 7.14. □

This theorem shows that the study of homogeneous spaces can be reduced to the largely algebraic problem of understanding closed Lie subgroups of Lie groups. Because of this, some authors *define* a homogeneous space to be a quotient manifold of the form G/H, where G is a Lie group and H is a closed Lie subgroup of G. One disadvantage of this definition is that it suggests that there is something special about the coset of H (the image of $e \in G$ under the quotient map).

Applying the characterization theorem to the examples of transitive group actions we developed earlier, we see that some familiar spaces can be expressed as quotients of Lie groups by closed Lie subgroups.

Example 9.21 (Homogeneous Spaces Revisited).

(a) Consider again the natural action of $\mathrm{O}(n)$ on $\mathbb{S}^{n-1}$. If we choose our base point in $\mathbb{S}^{n-1}$ to be the "north pole" $N = (0, \dots, 0, 1)$, then it is easy to check that the isotropy group is $\mathrm{O}(n-1)$, thought of as orthogonal transformations of $\mathbb{R}^n$ that fix the last variable. Thus $\mathbb{S}^{n-1}$ is diffeomorphic to the quotient manifold $\mathrm{O}(n)/\mathrm{O}(n-1)$. For the action of $\mathrm{SO}(n)$ on $\mathbb{S}^{n-1}$, the isotropy group is $\mathrm{SO}(n-1)$, so $\mathbb{S}^{n-1}$ is also diffeomorphic to $\mathrm{SO}(n)/\mathrm{SO}(n-1)$.

(b) Because the Euclidean group $E(n)$ acts smoothly and transitively on $\mathbb{R}^n$, and the isotropy group of the origin is the subgroup $\mathrm{O}(n) \subset E(n)$ (identified with the $(n+1)\times(n+1)$ matrices of the form $\left(\begin{smallmatrix} A & 0 \\ 0 & 1 \end{smallmatrix}\right)$ with $A \in \mathrm{O}(n)$), $\mathbb{R}^n$ is diffeomorphic to $E(n)/\mathrm{O}(n)$.

(c) Next consider the transitive action of $\mathrm{SL}(2,\mathbb{R})$ on the upper half-plane by Möbius transformations. Direct computation shows that the isotropy group of the point $i \in \mathbb{H}$ consists of matrices of the form $\left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right)$ with $a^2 + b^2 = 1$. This subgroup is exactly $\mathrm{SO}(2) \subset \mathrm{SL}(2,\mathbb{R})$, so the characterization theorem gives rise to a diffeomorphism $\mathbb{H} \approx \mathrm{SL}(2,\mathbb{R})/\mathrm{SO}(2)$.

(d) Using the result of Exercise 9.5, we conclude that $\mathbb{S}^{2n-1} \approx \mathrm{U}(n)/\mathrm{U}(n-1) \approx \mathrm{SU}(n)/\mathrm{SU}(n-1)$.

If G is a group, a subgroup $K \subset G$ is said to be *normal* if $gkg^{-1} \in K$ whenever $k \in K$ and $g \in G$. Normal subgroups play an important role in group theory, as illustrated by the following lemma. Proofs can be found in most abstract algebra texts, such as [Hun90] or [Her75].

Lemma 9.22. *Suppose G is any group.*

(*a*) *If $K \subset G$ is a normal subgroup, then the set G/K of left cosets is a group with multiplication given by $(g_1K)(g_2K) = (g_1g_2)K$. The projection $\pi\colon G \to G/K$ sending each element of G to its coset is a surjective homomorphism whose kernel is K.*

(*b*) *If* $F: G \to H$ *is any surjective group homomorphism, then the kernel of* F *is a normal subgroup of* G*, and* F *descends to a group isomorphism* $\widetilde{F}: G/\operatorname{Ker} F \to H$*.*

The following proposition gives a smooth analogue of these group-theoretic results.

Proposition 9.23. *Suppose* G *is a Lie group.*

(*a*) *If* $K \subset G$ *is a closed normal Lie subgroup, then* G/K *is a Lie group and the quotient map* $\pi: G \to G/K$ *is a Lie group homomorphism.*

(*b*) *If* $F: G \to H$ *is a surjective Lie group homomorphism, then* F *descends to a Lie group isomorphism* $\widetilde{F}: G/\operatorname{Ker} F \to H$*.*

Proof. Suppose K is a closed normal Lie subgroup of G. By the quotient manifold theorem, G/K is a smooth manifold and π is a smooth submersion; and by Lemma 9.22, G/K is a group and $\pi: G \to G/K$ is a homomorphism. Thus to prove (a), the only thing that needs to be verified is that multiplication and inversion in G/K are smooth. Smoothness of both maps follows easily from Proposition 7.17.

For (b), observe that $\operatorname{Ker} F$ is a closed Lie subgroup by Proposition 9.8, and $G/\operatorname{Ker} F$ is a Lie group by part (a) above. Lemma 9.22 shows that F descends to a group isomorphism $\widetilde{F}: G/\operatorname{Ker} F \to H$. By Proposition 7.17, both F and its inverse are smooth. □

Application: Sets with Transitive Group Actions

A highly useful application of the characterization theorem is to put smooth structures on *sets* that admit transitive Lie group actions.

Proposition 9.24. *Suppose* X *is a set, and we are given a transitive action of a Lie group* G *on* X*, such that the isotropy group of a point* $p \in X$ *is a closed Lie subgroup of* G*. Then* X *has a unique smooth manifold structure such that the given action is smooth.*

Proof. Let H denote the isotropy group of p, so that G/H is a smooth manifold by Theorem 9.18. The map $F: G/H \to X$ defined by $F(gH) = g \cdot p$ is an equivariant bijection by exactly the same argument as we used in the proof of the characterization theorem, Theorem 9.20. (That part did not use the assumption that M was a manifold or that the action was smooth.) If we define a topology and smooth structure on X by declaring F to be a diffeomorphism, then the given action of G on X is smooth because it can be written $(g, x) \mapsto F(g \cdot F^{-1}(x))$.

If $\widetilde{X}$ denotes the set X with any smooth manifold structure such that the given action is smooth, then by the homogeneous space characterization theorem, $\widetilde{X}$ is equivariantly diffeomorphic to G/H and therefore to X, so the topology and smooth structure are unique. □

Example 9.25 (Grassmannians). Let $\mathrm{G}_k(\mathbb{R}^n)$ denote the Grassmannian of k-dimensional subspaces of $\mathbb{R}^n$ as in Example 1.24. The general linear group $\mathrm{GL}(n,\mathbb{R})$ acts transitively on $\mathrm{G}_k(\mathbb{R}^n)$: Given two subspaces A and A', choose bases for both subspaces and extend them to bases for $\mathbb{R}^n$, and then the linear transformation taking the first basis to the second also takes A to A'. The isotropy group of the subspace $\mathbb{R}^k \subset \mathbb{R}^n$ is

$$H = \left\{ \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} : A \in \mathrm{GL}(k,\mathbb{R}),\ D \in \mathrm{GL}(n-k,\mathbb{R}), \right.$$
$$\left. B \in \mathrm{M}(k \times (n-k),\mathbb{R}) \right\},$$

which is easily seen to be a closed Lie subgroup of $\mathrm{GL}(n,\mathbb{R})$. Therefore $\mathrm{G}_k(\mathbb{R}^n)$ has a unique smooth manifold structure making the natural $\mathrm{GL}(n,\mathbb{R})$ action smooth. Problem 9-13 shows that this is the same as the smooth structure we defined in Example 1.24.

Example 9.26 (Flag Manifolds). Let V be a real vector space of dimension $n > 1$, and let $K = (k_1, \dots, k_m)$ be a finite sequence of integers satisfying $0 < k_1 < \cdots < k_m < n$. A *flag* in V of type K is a nested sequence of linear subspaces $S_1 \subset S_2 \subset \cdots \subset S_m \subset V$, with $\dim S_i = k_i$ for each i. The set of all flags of type K in V is denoted $\mathrm{F}_K(V)$. (For example, if $K = (k)$, then $\mathrm{F}_K(V)$ is the Grassmannian $\mathrm{G}_k(V)$.) It is not hard to show that $\mathrm{GL}(V)$ acts transitively on $\mathrm{F}_K(V)$ with a closed Lie subgroup as isotropy group (see Problem 9-17), so $\mathrm{F}_K(V)$ has a unique smooth manifold structure making it into a homogeneous $\mathrm{GL}(V)$-space. With this structure, $\mathrm{F}_K(V)$ is called a *flag manifold*.

Application: Connectivity of Lie Groups

Another application of homogeneous space theory is to identify the connected components of many familiar Lie groups. The key result is the following proposition.

Proposition 9.27. *Suppose a Lie group G acts smoothly, freely, and properly on a manifold M. If G and M/G are connected, then M is connected.*

Proof. Suppose G and M/G are connected, but M is not. This means that there are nonempty, disjoint open sets $U, V \subset M$ whose union is M. Each G-orbit in M is the image of G under a smooth map of the form $g \mapsto g \cdot p$; since G is connected, each orbit must lie entirely in one of the sets U or V.

Because the quotient map $\pi \colon M \to M/G$ is an open map by Lemma 9.15, $\pi(U)$ and $\pi(V)$ are nonempty open subsets of M/G. If $\pi(U) \cap \pi(V)$ were not empty, some G-orbit in M would contain points of both U and V, which we have just shown is impossible. Thus $\{\pi(U), \pi(V)\}$ is a separation of M/G, which contradicts the assumption that M/G is connected. $\square$

Proposition 9.28. *For any $n \geq 1$, the Lie groups $\mathrm{SO}(n)$, $\mathrm{U}(n)$, and $\mathrm{SU}(n)$ are connected. The group $\mathrm{O}(n)$ has exactly two components, one of which is $\mathrm{SO}(n)$.*

Proof. First we prove by induction on n that $\mathrm{SO}(n)$ is connected. For $n = 1$ this is obvious, because $\mathrm{SO}(1)$ is the trivial group. Now suppose we have shown that $\mathrm{SO}(n-1)$ is connected for some $n \geq 2$. Because the homogeneous space $\mathrm{SO}(n)/\mathrm{SO}(n-1)$ is diffeomorphic to $\mathbb{S}^{n-1}$ and therefore is connected, Proposition 9.27 and the induction hypothesis imply that $\mathrm{SO}(n)$ is connected. A similar argument applies to $\mathrm{U}(n)$ and $\mathrm{SU}(n)$, using the facts that $\mathrm{U}(n)/\mathrm{U}(n-1) \approx \mathrm{SU}(n)/\mathrm{SU}(n-1) \approx \mathbb{S}^{2n-1}$.

Note that $\mathrm{O}(n)$ is equal to the union of the two open sets $\mathrm{O}^+(n)$ and $\mathrm{O}^-(n)$ consisting of orthogonal matrices whose determinant is $+1$ or -1, respectively. As we noted earlier, $\mathrm{O}^+(n) = \mathrm{SO}(n)$, which is connected. On the other hand, if A is any orthogonal matrix whose determinant is -1, then left translation L_A is a diffeomorphism from $\mathrm{O}^+(n)$ to $\mathrm{O}^-(n)$, so $\mathrm{O}^-(n)$ is connected as well. Therefore $\{\mathrm{O}^+(n), \mathrm{O}^-(n)\}$ are exactly the components of $\mathrm{O}(n)$. □

Determining the components of the general linear groups is a bit more involved. Let $\mathrm{GL}^+(n,\mathbb{R})$ and $\mathrm{GL}^-(n,\mathbb{R})$ denote the subsets of $\mathrm{GL}(n,\mathbb{R})$ consisting of matrices with positive determinant and negative determinant, respectively.

Proposition 9.29. *The components of $\mathrm{GL}(n,\mathbb{R})$ are $\mathrm{GL}^+(n,\mathbb{R})$ and $\mathrm{GL}^-(n,\mathbb{R})$.*

Proof. By continuity of the determinant function, $\mathrm{GL}^+(n,\mathbb{R})$ and $\mathrm{GL}^-(n,\mathbb{R})$ are nonempty, disjoint, open subsets of $\mathrm{GL}(n,\mathbb{R})$ whose union is $\mathrm{GL}(n,\mathbb{R})$, so all we need to prove is that both subsets are connected. We begin by showing that $\mathrm{GL}^+(n,\mathbb{R})$ is connected. It suffices to show that it is path connected, which will follow once we show that there is a continuous path in $\mathrm{GL}^+(n,\mathbb{R})$ from any $A \in \mathrm{GL}^+(n,\mathbb{R})$ to the identity matrix I_n.

Let $A \in \mathrm{GL}^+(n,\mathbb{R})$ be arbitrary, and let $(A_1,\dots,A_n)$ denote the columns of A, considered as vectors in $\mathbb{R}^n$. The Gram-Schmidt algorithm (Proposition A.47 in the Appendix) shows that there is an orthonormal basis $(Q_1,\dots,Q_n)$ for $\mathbb{R}^n$ with the property that $\operatorname{span}(Q_1,\dots,Q_k) = \operatorname{span}(A_1,\dots,A_k)$ for each $k = 1,\dots,n$. Thus we can write

$$\begin{aligned}
A_1 &= R_1^1 Q_1,\\
A_2 &= R_2^1 Q_1 + R_2^2 Q_2,\\
&\vdots\\
A_n &= R_n^1 Q_1 + R_n^2 Q_2 + \cdots + R_n^n Q_n,
\end{aligned}$$

for some constants R_i^j. Replacing each Q_i by $-Q_i$ if necessary, we may assume that $R_i^i > 0$ for each i. In matrix notation, this is equivalent to

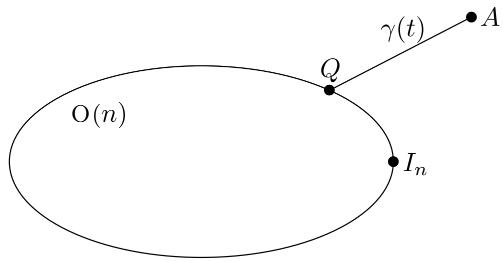


Figure 9.6. Proof that $\mathrm{GL}^+(n, \mathbb{R})$ is connected.

$A = QR$, where R is upper triangular with positive entries on the diagonal. Since the determinant of R is the product of its diagonal entries and $(\det Q)(\det R) = \det A > 0$, it follows that $Q \in \mathrm{SO}(n)$. (This *QR decomposition* plays an important role in numerical linear algebra.)

Let $R_t = tI_n + (1-t)R$. It is immediate that R_t is upper triangular with positive diagonal entries for all $t \in [0,1]$, so $R_t \in \mathrm{GL}^+(n, \mathbb{R})$. Therefore, the path $\gamma\colon [0,1] \to \mathrm{GL}^+(n, \mathbb{R})$ given by $\gamma(t) = QR_t$ satisfies $\gamma(0) = A$ and $\gamma(1) = Q \in \mathrm{SO}(n)$ (Figure 9.6). Because $\mathrm{SO}(n)$ is connected, there is a path in $\mathrm{SO}(n)$ from Q to the identity matrix. This shows that $\mathrm{GL}^+(n, \mathbb{R})$ is path connected.

Now, as in the case of $\mathrm{O}(n)$, any matrix B with $\det B < 0$ yields a diffeomorphism $L_B\colon \mathrm{GL}^+(n, \mathbb{R}) \to \mathrm{GL}^-(n, \mathbb{R})$, so $\mathrm{GL}^-(n, \mathbb{R})$ is connected as well. This completes the proof. □

Covering Manifolds

Proposition 2.12 showed that any covering space of a smooth manifold is a smooth manifold. It is often important to know when a space *covered by* a smooth manifold is itself a smooth manifold. To understand the answer to this question, we need to study the covering group of a covering space. In this section, we assume knowledge of the basic properties of topological covering maps as summarized in the Appendix (pages 556–557).

Let $\widetilde{M}$ and M be topological spaces, and let $\pi\colon \widetilde{M} \to M$ be a (topological) covering map. A *covering transformation* (or *deck transformation*) of π is a homeomorphism $\varphi\colon \widetilde{M} \to \widetilde{M}$ such that $\pi \circ \varphi = \pi$:

$$\begin{tikzcd} \widetilde{M} \arrow[rr, "\varphi"] \arrow[dr, "\pi"'] & & \widetilde{M} \arrow[dl, "\pi"] \\ & M. & \end{tikzcd} \tag{9.9}$$

The set $\mathcal{C}_\pi(\widetilde{M})$ of all covering transformations, called the *covering group* of π, is a group under composition, acting on $\widetilde{M}$ on the left. The covering group is the key to constructing smooth manifolds covered by $\widetilde{M}$.

We will see below that for any smooth covering $\pi\colon \widetilde{M} \to M$, the covering group is a zero-dimensional Lie group acting smoothly, freely, and properly on the covering space $\widetilde{M}$. Before proceeding, it is useful to have an alternative characterization of properness for actions of discrete groups.

Lemma 9.30. *Suppose a discrete group Γ acts continuously on a manifold $\widetilde{M}$. The action is proper if and only if the following condition holds:*

$$\begin{array}{l}\textit{Any two points } p, p' \in \widetilde{M} \textit{ have neighborhoods } U, U' \textit{ such}\\ \textit{that the set } \{\varphi \in \Gamma : (\varphi \cdot U) \cap U' \neq \varnothing\} \textit{ is finite.}\end{array} \tag{9.10}$$

Proof. First suppose that the action is proper. Let $p, p' \in \widetilde{M}$ be arbitrary, and let U, U' be precompact neighborhoods of p and p', respectively. If (9.10) does not hold, then there exist infinitely many distinct elements $\varphi_i \in \Gamma$ and points $p_i \in U$ such that $\varphi_i \cdot p_i \in U'$. Since $\overline{U}$ and $\overline{U}'$ are compact, by passing to a subsequence, we may assume that the sequences $\{p_i\}$ and $\{\varphi_i \cdot p_i\}$ converge. By Proposition 9.13, $\{\varphi_i\}$ has a convergent subsequence. But this is impossible, because $\{\varphi_i\}$ is an infinite sequence of distinct points in a discrete space.

Conversely, assume that (9.10) holds. Suppose $\{(\varphi_i, p_i)\}$ is a sequence in $\Gamma \times \widetilde{M}$ such that $p_i \to p$ and $\varphi_i \cdot p_i \to p'$. Let U, U' be neighborhoods of p and p', respectively, satisfying property (9.10). For all sufficiently large i, $p_i \in U$ and $\varphi_i \cdot p_i \in U'$. Since there are only finitely many $\varphi \in \Gamma$ for which $(\varphi \cdot U) \cap U' \neq \varnothing$, this means that there is some $\varphi \in \Gamma$ such that $\varphi_i = \varphi$ for infinitely many i; in particular, some subsequence of $\{\varphi_i\}$ converges. By Proposition 9.13, the action is proper. □

◇ **Exercise 9.6.** Suppose Γ is a discrete group acting continuously on a manifold $\widetilde{M}$. Show that the action is proper if and only if both of the following conditions are satisfied:

(i) Each $p \in \widetilde{M}$ has a neighborhood U such that $(\varphi \cdot U) \cap U = \varnothing$ for all but finitely many $\varphi \in \Gamma$.
(ii) If $p, p' \in \widetilde{M}$ are not in the same Γ-orbit, there exist neighborhoods U of p and U' of p' such that $(\varphi \cdot U) \cap U' = \varnothing$ for all $\varphi \in \Gamma$.

A continuous discrete group action satisfying conditions (i) and (ii) of the preceding exercise (or condition (9.10) of Lemma 9.30, or something closely related to these) has traditionally been called *properly discontinuous*. Because the term "properly discontinuous" is self-contradictory (properly discontinuous group actions are, after all, continuous!), and because there is no general agreement about exactly what the term should mean, we will avoid using this terminology and stick with the more general term "proper action" in this book.

Proposition 9.31. *Let $\pi\colon \widetilde{M} \to M$ be a smooth covering map. With the discrete topology, the covering group $\mathcal{C}_\pi(\widetilde{M})$ is a zero-dimensional Lie group acting smoothly, freely, and properly on $\widetilde{M}$.*

Proof. Suppose $\varphi \in \mathcal{C}_\pi(\widetilde{M})$ is a covering transformation that fixes a point $p \in \widetilde{M}$. Simply by rotating diagram (9.9), we can consider φ as a lift of π:

$$\begin{tikzcd} & \widetilde{M} \arrow[d, "\pi"] \\ \widetilde{M} \arrow[ur, "\varphi"] \arrow[r, "\pi"'] & M. \end{tikzcd}$$

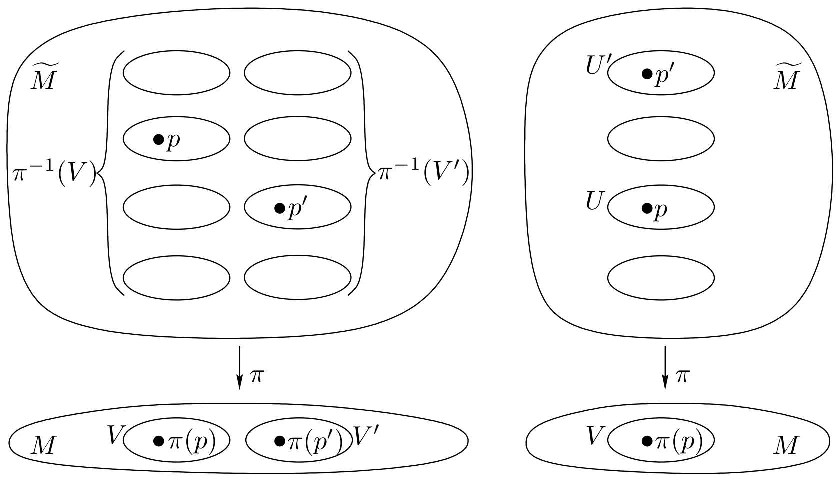
Since the identity map of $\widetilde{M}$ is another such lift that agrees with φ at p, the unique lifting property of covering maps (Proposition A.26(a) in the Appendix) guarantees that $\varphi = \mathrm{Id}_{\widetilde{M}}$. Thus the action of $\mathcal{C}_\pi(\widetilde{M})$ is free.

To show that $\mathcal{C}_\pi(\widetilde{M})$ is a Lie group, we need only verify that it is countable. Let $p \in \widetilde{M}$ be arbitrary, let $q = \pi(p) \in M$, and let $U \subset M$ be an evenly covered neighborhood of q. Because $\widetilde{M}$ is a manifold, $\pi^{-1}(U)$ has countably many components, and because each component contains exactly one point of $\pi^{-1}(q)$, it follows that $\pi^{-1}(q)$ is countable. Define a map $\theta^{(p)}\colon \mathcal{C}_\pi(\widetilde{M}) \to \pi^{-1}(q)$ by $\theta^{(p)}(\varphi) = \varphi(p)$. The fact that the action is free implies that $\theta^{(p)}$ is injective, and therefore $\mathcal{C}_\pi(\widetilde{M})$ is countable.

Smoothness of the action follows from the fact that any covering transformation φ can be written locally as $\varphi = \sigma \circ \pi$ for a suitable smooth local section σ.

To show that the action is proper, we will show that it satisfies (9.10). Let $p, p' \in \widetilde{M}$ be arbitrary. If $\pi(p) \neq \pi(p')$, then there are disjoint open sets $V \subset M$ containing $\pi(p)$ and $V' \subset M$ containing $\pi(p')$, and $U = \pi^{-1}(V)$ and $U' = \pi^{-1}(V')$ are disjoint open sets satisfying (9.10) (Figure 9.7(a)). On the other hand, if $\pi(p) = \pi(p')$, let V be an evenly covered neighborhood of $\pi(p)$, and let U, U' be the components of $\pi^{-1}(V)$ containing p and p', respectively. Suppose φ is a covering transformation such that $\varphi(U) \cap U' \neq \varnothing$. Since U is connected, $\varphi(U)$ must be contained in a single component of $\pi^{-1}(V)$, so $\varphi(U) \subset U'$. Since U' contains at most one point in each fiber, and both $\varphi(p)$ and p' lie in $\pi^{-1}(q) \cap U'$, it follows that $\varphi(p) = p'$. Since the action is free, there is at most one covering transformation taking p to p', and thus (9.10) is satisfied. □

The quotient manifold theorem yields an important converse to this proposition. A covering map $\pi\colon \widetilde{M} \to M$ is said to be *normal* if the covering group $\mathcal{C}_\pi(\widetilde{M})$ acts transitively on the fibers of π. (It can be shown that this is equivalent to $\pi_*(\pi_1(\widetilde{M}, p))$ being a normal subgroup of $\pi_1(M, \pi(p))$ for any $p \in \widetilde{M}$; see, for example, [Lee00, Proposition 11.29]).



(a) $\pi(p) \neq \pi(p')$. (b) $\pi(p) = \pi(p')$.

Figure 9.7. The covering group acts properly.

Theorem 9.32. *Suppose $\widetilde{M}$ is a connected smooth manifold, and Γ is a discrete group acting smoothly, freely, and properly on $\widetilde{M}$. Then the quotient space $\widetilde{M}/\Gamma$ is a topological manifold and has a unique smooth structure such that $\pi\colon \widetilde{M} \to \widetilde{M}/\Gamma$ is a smooth normal covering map.*

Proof. It follows from the quotient manifold theorem that $\widetilde{M}/\Gamma$ has a unique smooth manifold structure such that π is a smooth submersion. Because a smooth covering map is in particular a submersion, any other smooth manifold structure on $\widetilde{M}$ making π into a smooth covering map must be equal to this one. Because $\dim \widetilde{M}/\Gamma = \dim \widetilde{M} - \dim \Gamma = \dim \widetilde{M}$, π is a local diffeomorphism. Thus to prove the theorem, it suffices to show that π is a normal covering map.

Let $p \in \widetilde{M}$. By Exercise 9.6, there exists a neighborhood U_0 of p in $\widetilde{M}$ such that $(\varphi \cdot U_0) \cap U_0 = \varnothing$ for all $\varphi \in \Gamma$ except possibly finitely many elements $\varphi_1, \dots, \varphi_k$. Reordering if necessary, we may assume that $\varphi_1 = e$ and $\varphi_i \neq e$ for $2 \leq i \leq k$. Shrinking U_0 if necessary, we may assume that U_0 is connected and $\varphi_i^{-1} \cdot p \notin \overline{U}_0$ (which implies $p \notin \varphi_i \cdot \overline{U}_0$) for $i = 2, \dots, k$. Replacing U_0 by

$$U = U_0 \smallsetminus (\varphi_2 \cdot \overline{U_0} \cup \cdots \cup \varphi_k \cdot \overline{U_0}),$$

we obtain a neighborhood U of p satisfying

$$(\varphi \cdot U) \cap U = \varnothing \quad \text{for all } \varphi \in \Gamma \text{ except } \varphi = e. \tag{9.11}$$

Let $V = \pi(U)$, which is open in $\widetilde{M}/\Gamma$ by Lemma 9.15. Because $\pi^{-1}(V)$ is the union of the disjoint connected open sets $\varphi \cdot U$ for $\varphi \in \Gamma$, to show that

π is a covering map we need only show that π is a homeomorphism from each such set onto V. For each $\varphi \in \Gamma$, the following diagram commutes:

$$\begin{tikzcd} U \arrow[rr, "\varphi"] \arrow[dr, "\pi"'] & & \varphi \cdot U \arrow[dl, "\pi"] \\ & V. & \end{tikzcd}$$

Since $\varphi\colon U \to \varphi \cdot U$ is a homeomorphism (in fact a diffeomorphism), it suffices to show that $\pi\colon U \to V$ is a homeomorphism. We already know that it is surjective, continuous, and open. To see that it is injective, suppose $\pi(q) = \pi(q')$ for $q, q' \in U$, which means that $q' = \varphi \cdot q$ for some $\varphi \in \Gamma$. By (9.11), this can happen only if $\varphi = e$, which is to say that $q = q'$. This completes the proof that π is a smooth covering map. Because elements of Γ act as covering transformations, and Γ acts transitively on fibers by definition, the covering is normal. $\square$

Example 9.33 (Proper Discrete Group Actions).

(a) The discrete Lie group $\mathbb{Z}^n$ acts smoothly and freely on $\mathbb{R}^n$ by translation (Example 9.1(h)). To check that the action is proper, one can verify that condition (9.10) is satisfied by sufficiently small balls around p and p'. The quotient manifold $\mathbb{R}^n/\mathbb{Z}^n$ is homeomorphic to the n-torus $\mathbb{T}^n$, and Theorem 9.32 says that there is a unique smooth structure on $\mathbb{T}^n$ making the quotient map into a smooth covering map. To verify that this smooth structure on $\mathbb{T}^n$ is the same as the one we defined previously (thinking of $\mathbb{T}^n$ as the product manifold $\mathbb{S}^1 \times \cdots \times \mathbb{S}^1$), we just check that the covering map $\mathbb{R}^n \to \mathbb{T}^n$ given by $(x^1, \dots, x^n) \mapsto (e^{2\pi i x^1}, \dots, e^{2\pi i x^n})$ is a local diffeomorphism with respect to the product smooth structure on $\mathbb{T}^n$, and makes the same identifications as the quotient map $\mathbb{R}^n \to \mathbb{R}^n/\mathbb{Z}^n$; thus Proposition 7.18 implies that $\mathbb{R}^n/\mathbb{Z}^n$ is diffeomorphic to $\mathbb{T}^n$.

(b) The two-element group $\{\pm 1\}$ acts on $\mathbb{S}^n$ by multiplication. This action is obviously smooth and free, and it is proper because the group is compact. This defines a smooth structure on $\mathbb{S}^n/\{\pm 1\}$. In fact, this quotient manifold is diffeomorphic to $\mathbb{RP}^n$ with the smooth structure we defined in Chapter 1, which can be seen as follows. Let $p\colon \mathbb{S}^n \to \mathbb{RP}^n$ be the smooth covering map obtained by restricting the canonical projection $\mathbb{R}^{n+1} \smallsetminus \{0\} \to \mathbb{RP}^n$ to the sphere (see Problem 2-9). This map makes the same identifications as the quotient map $\pi\colon \mathbb{S}^n \to \mathbb{S}^n/\{\pm 1\}$. By Proposition 7.18, therefore, $\mathbb{S}^n/\{\pm 1\}$ is diffeomorphic to $\mathbb{RP}^n$.

Application: Discrete Subgroups

Let G be a Lie group. A *discrete subgroup* of G is any subgroup that is discrete in the subspace topology (and is thus a closed zero-dimensional Lie subgroup).

Proposition 9.34. *If G is a connected Lie group and $\Gamma \subset G$ is a discrete subgroup, then the quotient map $\pi\colon G \to G/\Gamma$ is a smooth covering map.*

Proof. The proof of Theorem 9.18 showed that Γ acts smoothly, freely, and properly on G on the right, and its quotient is the coset space G/Γ. The proposition is then an immediate consequence of Theorem 9.32. $\square$

Example 9.35. Let C be the unit cube centered at the origin in $\mathbb{R}^3$. The set Γ of positive-determinant orthogonal transformations of $\mathbb{R}^3$ that take C to itself is a finite subgroup of $\mathrm{SO}(3)$, and the quotient $\mathrm{SO}(3)/\Gamma$ is a connected smooth 3-manifold whose universal covering space is $\mathbb{S}^3$ (see Problem 9-9). Similar examples are obtained from the symmetry groups of other regular polyhedra, such as regular tetrahedra, dodecahedra, or icosahedra.

As a consequence of the preceding proposition, we have the following characterization of Lie group homomorphisms with discrete kernels.

Proposition 9.36. *Let G and H be connected Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, respectively. For any Lie group homomorphism $F\colon G \to H$, the following are equivalent:*

(*a*) *F is surjective and has discrete kernel.*

(*b*) *F is a smooth covering map.*

(*c*) *The induced homomorphism $F_*\colon \mathfrak{g} \to \mathfrak{h}$ is an isomorphism.*

(*d*) *F is a local diffeomorphism.*

Proof. We will show that (a) $\implies$ (b) $\implies$ (c) $\implies$ (d) $\implies$ (a). First assume that F is surjective with discrete kernel $\Gamma \subset G$. Then Proposition 9.34 implies that $\pi\colon G \to G/\Gamma$ is a smooth covering map, and Proposition 9.23 shows that F descends to a Lie group isomorphism $\widetilde{F}\colon G/\Gamma \to H$. This means that $F = \widetilde{F} \circ \pi$, which is a composition of a smooth covering map followed by a diffeomorphism and therefore is itself a smooth covering map. This proves that (a) $\implies$ (b). The next implication, (b) $\implies$ (c), is the content of Problem 4-18.

Under the assumption that F_* is an isomorphism, the inverse function theorem implies that F is a local diffeomorphism in a neighborhood of $e \in G$. Because Lie group homomorphisms have constant rank, this means that $\operatorname{rank} F = \dim G = \dim H$, which implies that F is a local diffeomorphism everywhere, and thus (c) $\implies$ (d).

Finally, if F is a local diffeomorphism, then each level set is an embedded 0-dimensional manifold by the inverse function theorem, so $\operatorname{Ker} F$ is discrete. Since a local diffeomorphism is an open map, $F(G)$ is an open subgroup of H, and thus by Problem 2-12, it is all of H. This shows that (d) $\implies$ (a) and completes the proof. $\square$

Problems

9-1. Prove that the set of maps from $\mathbb{R}^n$ to itself given by the action of $E(n)$ on $\mathbb{R}^n$ described in Example 9.17(b) is exactly the set of all diffeomorphisms of $\mathbb{R}^n$ that preserve the Euclidean distance function.

9-2. Suppose a Lie group acts smoothly on a manifold M.

(a) Show that each orbit is an immersed submanifold of M.
(b) Give an example of a Lie group acting smoothly on a manifold M in which two different orbits have different dimensions even though neither orbit has dimension equal to zero or to the dimension of M.

9-3. Prove the following partial converse to the quotient manifold theorem: If a Lie group G acts smoothly and freely on a smooth manifold M and the orbit space M/G has a smooth manifold structure such that the quotient map $\pi\colon M \to M/G$ is a smooth submersion, then G acts properly.

9-4. Give an example of a smooth, proper action of a Lie group on a smooth manifold such that the orbit space is not a topological manifold.

9-5. Suppose a connected Lie group G acts smoothly on a discrete space K. Show that the action is trivial.

9-6. Show that $\mathrm{SO}(2)$, $\mathrm{U}(1)$, and $\mathbb{S}^1$ are all isomorphic as Lie groups.

9-7. (a) Show that there exists a Lie group homomorphism $\rho : \mathrm{U}(1) \to \mathrm{U}(n)$ such that $\det \circ \rho = \mathrm{Id}_{\mathrm{U}(1)}$.
(b) Show that $\mathrm{U}(n)$ is diffeomorphic to $\mathrm{U}(1) \times \mathrm{SU}(n)$.
(c) Show that $\mathrm{U}(n)$ and $\mathrm{U}(1) \times \mathrm{SU}(n)$ are not isomorphic Lie groups when $n > 1$.

9-8. Show that $\mathrm{SU}(2)$ is isomorphic to the group $\mathcal{S}$ of unit quaternions (Problem 8-19) and diffeomorphic to $\mathbb{S}^3$.

9-9. Show that $\mathrm{SO}(3)$ is isomorphic to $\mathrm{SU}(2)/\{\pm e\}$ and diffeomorphic to $\mathbb{RP}^3$, as follows.

(a) Let $\mathcal{S} \subset \mathbb{H}$ denote the group of unit quaternions, and let $E \subset \mathbb{H}$ be the subspace of imaginary quaternions (see Problems 8-19 and 8-20). If $q \in \mathcal{S}$, show that the linear map $\mathbb{H} \to \mathbb{H}$ given by $v \mapsto qvq^*$ takes E to itself and preserves the inner product $\langle v, w\rangle = \frac{1}{2}(v^*w + w^*v)$ on E.

(b) For each $q \in \mathcal{S}$, let $\rho(q)$ be the matrix representation of the map $v \mapsto qvq^*$ with respect to the basis $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ for E. Show that $\rho(q) \in \mathrm{SO}(3)$, and the map $\rho\colon \mathcal{S} \to \mathrm{SO}(3)$ is a surjective Lie group homomorphism whose kernel is $\{\pm 1\}$.

(c) Prove the result.

9-10. Determine which of the following Lie groups are compact: $\mathrm{GL}(n,\mathbb{R})$, $\mathrm{SL}(n,\mathbb{R})$, $\mathrm{GL}(n,\mathbb{C})$, $\mathrm{SL}(n,\mathbb{C})$, $\mathrm{U}(n)$, $\mathrm{SU}(n)$.

9-11. Under the canonical isomorphism of $\mathrm{Lie}(\mathrm{GL}(n,\mathbb{C}))$ with the matrix algebra $\mathfrak{gl}(n,\mathbb{C})$ (Proposition 8.38), show that $\mathrm{Lie}(\mathrm{SL}(n,\mathbb{C})) \cong \mathfrak{sl}(n,\mathbb{C})$, $\mathrm{Lie}(\mathrm{U}(n)) \cong \mathfrak{u}(n)$, and $\mathrm{Lie}(\mathrm{SU}(n)) \cong \mathfrak{su}(n)$, where

$$\begin{aligned}\mathfrak{sl}(n,\mathbb{C}) &= \{A \in \mathfrak{gl}(n,\mathbb{C}) : \operatorname{tr} A = 0\},\\ \mathfrak{u}(n) &= \{A \in \mathfrak{gl}(n,\mathbb{C}) : A^* + A = 0\},\\ \mathfrak{su}(n) &= \mathfrak{u}(n) \cap \mathfrak{sl}(n,\mathbb{C}).\end{aligned}$$

9-12. Show by giving an explicit isomorphism that $\mathfrak{su}(2)$ and $\mathfrak{o}(3)$ are isomorphic Lie algebras, and that both are isomorphic to $\mathbb{R}^3$ with the cross product.

9-13. Show that the smooth structure on the Grassmannian $\mathrm{G}_k(\mathbb{R}^n)$ defined in Example 9.25 is the same as the one defined in Example 1.24.

9-14. Let V be a finite-dimensional vector space. Prove that the Grassmannian $\mathrm{G}_k(V)$ is compact for each k. [Hint: Show that it is a quotient space of a compact Lie group.]

9-15. Show that the image of a Lie group homomorphism is a Lie subgroup.

9-16. Define an action of $\mathbb{Z}$ on $\mathbb{R}^2$ by

$$n \cdot (x, y) = (x + n, (-1)^n y).$$

(a) Show that the action is smooth, free and proper. Let $E = \mathbb{R}^2/\mathbb{Z}$ denote the quotient manifold (Figure 9.8).

(b) Show that the projection on the first coordinate $\pi_1\colon \mathbb{R}^2 \to \mathbb{R}$ descends to a smooth map $\pi\colon E \to \mathbb{S}^1$.

(c) Show that E is a nontrivial rank-1 vector bundle over $\mathbb{S}^1$ with projection π.

(d) Show that E is isomorphic to the Möbius bundle constructed in Example 5.2.

9-17. Let $\mathrm{F}_K(V)$ be the set of flags of type K in a finite-dimensional vector space V as in Example 9.26. Show that $\mathrm{GL}(V)$ acts transitively on $\mathrm{F}_K(V)$, and that the isotropy group of a particular flag is a closed Lie subgroup of $\mathrm{GL}(V)$. For which K is $\mathrm{F}_K(V)$ compact?

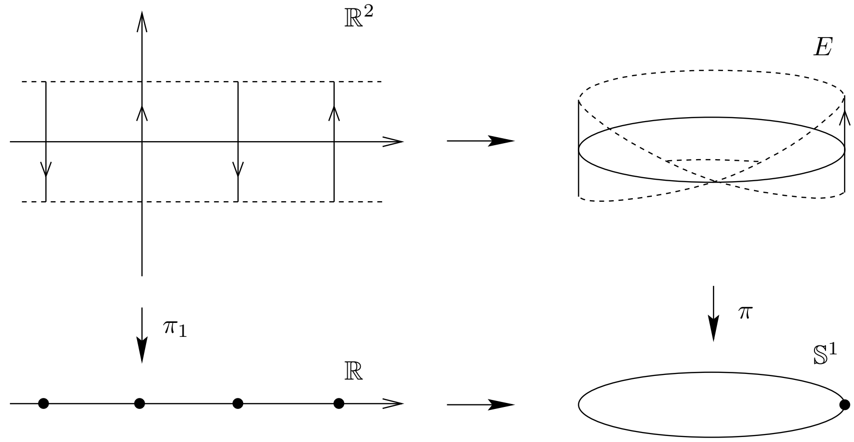


Figure 9.8. The Möbius bundle as a quotient of $\mathbb{R}^2$.

9-18. Let $\mathbb{CP}^n$ denote n-dimensional complex projective space (Problem 1-7). Show that the natural action of $\mathrm{U}(n+1)$ on $\mathbb{C}^{n+1}$ descends to a smooth, transitive action on $\mathbb{CP}^n$, so $\mathbb{CP}^n$ is a homogeneous $\mathrm{U}(n+1)$-space.

9-19. The set of k-dimensional complex linear subspaces of $\mathbb{C}^n$ is denoted by $\mathrm{G}_k(\mathbb{C}^n)$. Show that $\mathrm{G}_k(\mathbb{C}^n)$ has a unique smooth manifold structure making it into a compact homogeneous $\mathrm{U}(n)$-space (where the action of $\mathrm{U}(n)$ is induced from its usual action on $\mathbb{C}^n$). What is the dimension of $\mathrm{G}_k(\mathbb{C}^n)$?

9-20. Considering $\mathbb{S}^{2n+1}$ as the unit sphere in $\mathbb{C}^{n+1}$, define an action of $\mathbb{S}^1$ on $\mathbb{S}^{2n+1}$ by

$$z \cdot (w^1, \dots, w^{n+1}) = (zw^1, \dots, zw^{n+1}).$$

Show that this action is smooth, free, and proper, and that the orbit space $\mathbb{S}^{2n+1}/\mathbb{S}^1$ is diffeomorphic to $\mathbb{CP}^n$. [Hint: Consider the restriction of the natural quotient map $\mathbb{C}^{n+1} \smallsetminus \{0\} \to \mathbb{CP}^n$ to $\mathbb{S}^{2n+1}$. The quotient map $\pi\colon \mathbb{S}^{2n+1} \to \mathbb{CP}^n$ is known as the *Hopf map*.]

9-21. Let c be an irrational number, and let $\mathbb{R}$ act on $\mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1$ by

$$t \cdot (w, z) = (e^{2\pi i t} w, e^{2\pi i c t} z).$$

Show that this is a smooth free action, but $\mathbb{T}^2/\mathbb{R}$ is not Hausdorff.

9-22. Show that $\mathrm{GL}^+(n,\mathbb{R})$ is diffeomorphic to $\mathrm{SO}(n) \times \mathbb{R}^{n(n+1)/2}$. [Hint: Use the QR decomposition introduced in Proposition 9.29 to construct a diffeomorphism from $\mathrm{SO}(n) \times \mathrm{T}^+(n,\mathbb{R})$ to $\mathrm{GL}^+(n,\mathbb{R})$, where

$\mathrm{T}^+(n,\mathbb{R})$ is the Lie group of $n \times n$ upper triangular real matrices with positive diagonal entries.]

9-23. Show that $\mathrm{GL}(n,\mathbb{C})$ is diffeomorphic to $\mathrm{U}(n) \times \mathbb{R}^{n^2}$. [Hint: Argue as in Problem 9-22, but use the Hermitian dot product $z \cdot w = \sum_i z^i \overline{w^i}$ in place of the Euclidean dot product.]

9-24. Show that $\mathrm{SL}(n,\mathbb{R})$ and $\mathrm{SL}(n,\mathbb{C})$ are diffeomorphic to $\mathrm{SO}(n) \times \mathbb{R}^{n(n+1)/2-1}$ and $\mathrm{SU}(n) \times \mathbb{R}^{n^2}$, respectively.

9-25. The *center* of a group G is the set of all elements that commute with every element of G; a subgroup of G is said to be *central* if it is contained in the center of G. If G is a connected Lie group, show that every discrete normal subgroup of G is contained in the center of G. [Hint: Use the result of Problem 9-5.]

9-26. Use the results of Theorem 2.13 and Problem 9-25 to show that the fundamental group of a connected Lie group is abelian. You may use without proof the fact that if $\pi\colon \widetilde{G} \to G$ is a universal covering map, then the covering group $\mathcal{C}_\pi(\widetilde{G})$ is isomorphic to $\pi_1(G,e)$.

9-27. Show that the adjoint representation of $\mathrm{GL}(n,\mathbb{R})$ is given by $\mathrm{Ad}(A)Y = AYA^{-1}$ for $A \in \mathrm{GL}(n,\mathbb{R})$ and $Y \in \mathfrak{gl}(n,\mathbb{R})$. Show that it is not faithful.

9-28. Let $\mathfrak{g}$ be a finite-dimensional Lie algebra. For each $X \in \mathfrak{g}$, define a map $\mathrm{ad}(X)\colon \mathfrak{g} \to \mathfrak{g}$ by $\mathrm{ad}(X)Y = [X,Y]$. Show that $\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is a representation, called the *adjoint representation of* $\mathfrak{g}$. [Remark: Its relationship with the adjoint representation of a Lie group will be explained in Chapter 20.]

10
Embedding and Approximation Theorems

The purpose of this chapter is to address two fundamental questions about smooth manifolds. The questions may seem unrelated at first, but their solutions are closely related.

The first question is "Which smooth manifolds can be smoothly embedded in Euclidean spaces?" The answer, as we will see, is that they all can. This justifies our habit of visualizing manifolds as subsets of $\mathbb{R}^n$.

The second question is "To what extent can continuous maps between manifolds be approximated by smooth ones?" We will give two different answers, each of which is useful in certain contexts. Stated simply, we will show that any continuous function from a smooth manifold into $\mathbb{R}^k$ can be uniformly approximated by a smooth function, and that any continuous map from one smooth manifold to another is homotopic to a smooth map.

The essential strategy for answering both questions is the same: First use analysis in $\mathbb{R}^n$ to construct a "local" solution in a fixed coordinate chart; then use partitions of unity to piece together the local solutions into a global one.

Before we begin, we need to extend the notion of sets of measure zero to manifolds. These are sets that are "small" in a sense that is closely related to having zero volume (even though we do not yet have a way to measure volume quantitatively on manifolds), and include things like countable sets and submanifolds of positive codimension.

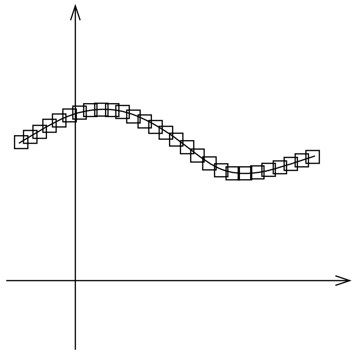


Figure 10.1. A set of measure zero.

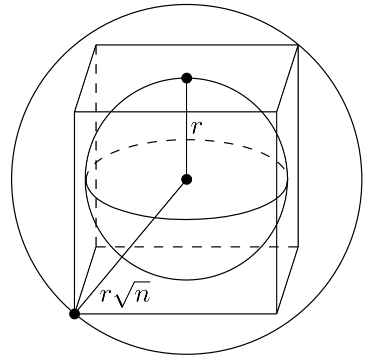


Figure 10.2. Balls and cubes.

Sets of Measure Zero in Manifolds

Recall what it means for a set $A \subset \mathbb{R}^n$ to have measure zero (see the Appendix, page 589): For any $\delta > 0$, A can be covered by a countable collection of open cubes, the sum of whose volumes is less than δ (Figure 10.1). The next lemma shows that cubes can be replaced by balls in the definition.

Lemma 10.1. *A subset $A \subset \mathbb{R}^n$ has measure zero if and only if for every $\delta > 0$, A can be covered by a countable collection of open balls, the sum of whose volumes is less than δ.*

Proof. This is based on the easily verified geometric fact that an open cube of side $2r$ contains an open ball of radius r and is contained in an open ball of radius $r\sqrt{n}$ (Figure 10.2). Since volumes of balls and cubes in $\mathbb{R}^n$ are proportional to the nth power of their diameters, it follows that every open cube of volume v is contained in an open ball of volume $c_n v$, and every open ball of volume v is contained in an open cube of volume $c'_n v$, where c_n and c'_n are constants depending only on n. Thus if A has measure zero, there is a countable cover of A by open cubes with total volume less than δ. Enclosing each cube in a ball whose volume is c_n times that of the cube, we obtain an open cover of A by open balls of total volume less than $c_n\delta$, which can be made as small as desired by taking δ sufficiently small. The converse is similar. □

We wish to extend the notion of measure zero in a diffeomorphism-invariant fashion to subsets of manifolds. Because a manifold does not come with a metric, volumes of cubes or balls do not make sense, so we cannot simply use the same definition. However, the key is provided by the next lemma, which implies that the condition of having measure zero is diffeomorphism-invariant for subsets of $\mathbb{R}^n$.

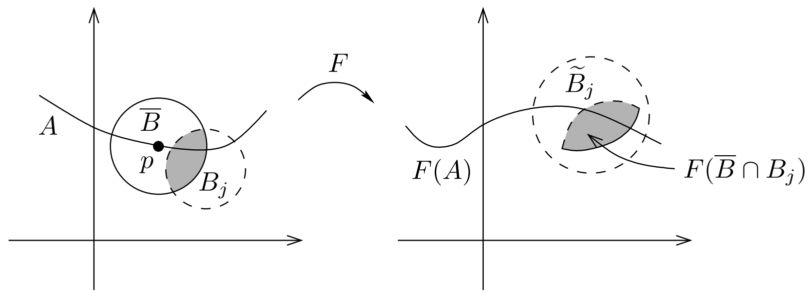


Figure 10.3. The image of a set of measure zero.

Lemma 10.2. *Suppose $A \subset \mathbb{R}^n$ has measure zero and $F\colon A \to \mathbb{R}^n$ is a smooth map. Then $F(A)$ has measure zero.*

Proof. By definition, for each $p \in A$, F has an extension to a smooth map, which we still denote by F, on a neighborhood of p in $\mathbb{R}^n$. Shrinking this neighborhood if necessary, we may assume F extends smoothly to a closed ball $\overline{B}$ centered at p. Since $\overline{B}$ is compact, there is a constant C such that $|DF(x)| \le C$ for all $x \in \overline{B}$. Using the Lipschitz estimate for smooth functions (Proposition A.69), we have

$$|F(x) - F(x')| \le C|x - x'| \tag{10.1}$$

for all $x, x' \in \overline{B}$.

Given $\delta > 0$, we can choose a countable cover $\{B_j\}$ of $A \cap \overline{B}$ by open balls satisfying

$$\sum_j \operatorname{Vol}(B_j) < \delta.$$

Then by (10.1), $F(\overline{B} \cap B_j)$ is contained in a ball $\widetilde{B}_j$ whose radius is no more than C times that of B_j (Figure 10.3). We conclude that $F(A \cap \overline{B})$ is contained in the collection of balls $\{\widetilde{B}_j\}$, whose total volume is no greater than

$$\sum_j \operatorname{Vol}(\widetilde{B}_j) < C^n \delta.$$

Since this can be made as small as desired, it follows that $F(A \cap \overline{B})$ has measure zero. Since $F(A)$ is the union of countably many such sets, it too has measure zero. □

It is important to be aware that the preceding lemma is false if F is merely assumed to be continuous. For example, the subset $A = [0,1] \times \{0\} \subset \mathbb{R}^2$ has measure zero in $\mathbb{R}^2$, but there is a continuous map $F\colon A \to \mathbb{R}^2$ whose image is the entire unit square $[0,1] \times [0,1]$. (Such a map is called a *space filling curve*. See [Rud76, p. 168] for an example.)

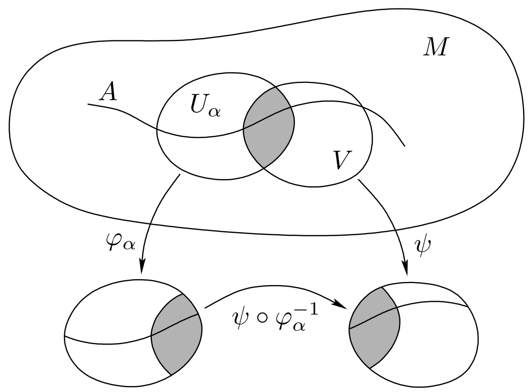


Figure 10.4. Proof of Lemma 10.4.

Lemma 10.3. *Suppose $F\colon U \to \mathbb{R}^n$ is a smooth map, where U is an open subset of $\mathbb{R}^m$ and $m < n$. Then $F(U)$ has measure zero in $\mathbb{R}^n$.*

Proof. Let $\pi\colon \mathbb{R}^n \to \mathbb{R}^m$ denote the projection onto the first m coordinates, and let $\widetilde{U} = \pi^{-1}(U)$. The result follows by applying the preceding lemma to $\widetilde{F} = F \circ \pi\colon \widetilde{U} \to \mathbb{R}^n$, because $F(U) = \widetilde{F}(\widetilde{U} \cap \mathbb{R}^m)$, which is the image of a set of measure zero. □

We say a subset A of a smooth n-manifold M has measure zero if for every smooth chart (U, φ) for M, the set $\varphi(A \cap U)$ has measure zero in $\mathbb{R}^n$. It follows immediately from Lemma A.60(c) that any set of measure zero has dense complement, because if $M \smallsetminus A$ is not dense then A contains a nonempty open set, which would imply that $\psi(A \cap V)$ contains a nonempty open set for some smooth chart (V, ψ).

The following lemma shows that we need only check this condition for a single collection of smooth charts whose domains cover A.

Lemma 10.4. *Suppose A is a subset of a smooth n-manifold M, and for some collection $\{(U_\alpha, \varphi_\alpha)\}$ of smooth charts whose domains cover A, $\varphi_\alpha(A \cap U_\alpha)$ has measure zero in $\mathbb{R}^n$ for each α. Then A has measure zero in M.*

Proof. Let (V, ψ) be an arbitrary smooth chart. We need to show that $\psi(A \cap V)$ has measure zero. Some countable collection of the U_αs covers $A \cap V$. For each such U_α, we have

$$\psi(A \cap V \cap U_\alpha) = (\psi \circ \varphi_\alpha^{-1}) \circ \varphi_\alpha(A \cap V \cap U_\alpha).$$

(See Figure 10.4.) Now $\varphi_\alpha(A \cap V \cap U_\alpha)$ is a subset of $\varphi_\alpha(A \cap U_\alpha)$, which has measure zero by hypothesis. By Lemma 10.2 applied to $\psi \circ \varphi_\alpha^{-1}$, therefore, $\psi(A \cap V \cap U_\alpha)$ has measure zero. Since $\psi(A \cap V)$ is the union of countably many such sets, it too has measure zero. □

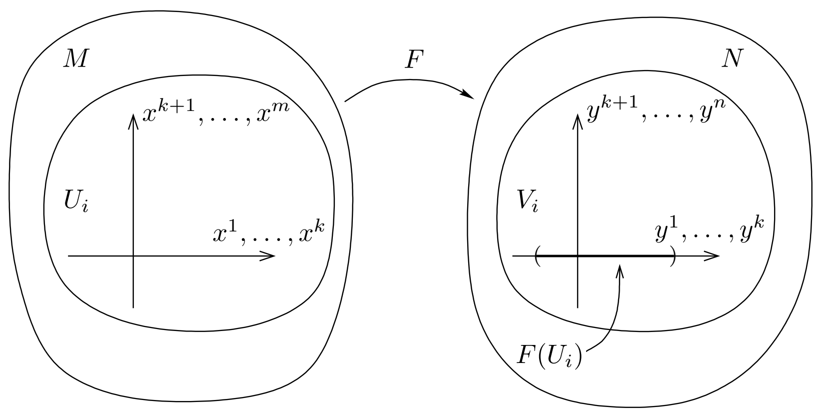


Figure 10.5. Proof of Theorem 7.14.

◇ **Exercise 10.1.** Let M be a smooth manifold. Show that a countable union of sets of measure zero in M has measure zero.

As our first application of the concept of measure zero in manifolds, we will finish the proof of Theorem 7.14.

Completion of the proof of Theorem 7.14. It remains only to prove that a smooth surjective map $F\colon M \to N$ of constant rank is a submersion. Let $m = \dim M$, $n = \dim N$, and $k = \operatorname{rank} F$. Suppose that F is not a submersion, so that $k < n$. By the rank theorem, each point has a smooth coordinate neighborhood in which F has the coordinate representation

$$F(x^1, \dots, x^k, x^{k+1}, \dots, x^m) = (x^1, \dots, x^k, 0, \dots, 0). \tag{10.2}$$

Since any open cover of a manifold has a countable subcover, we can choose countably many smooth charts $\{(U_i, \varphi_i)\}$ for M and corresponding smooth charts $\{(V_i, \psi_i)\}$ for N such that the sets $\{U_i\}$ cover M, F maps U_i into V_i, and the coordinate representation of $F\colon U_i \to V_i$ is as in (10.2) (Figure 10.5). Since $F(U_i)$ is contained in a k-dimensional slice of V_i, it has measure zero in N. Because $F(M)$ is equal to the countable union of sets $F(U_i)$ of measure zero, $F(M)$ itself has measure zero in N, which implies that F cannot be surjective. □

The next theorem is the main result of this section.

Theorem 10.5. *Suppose M and N are smooth manifolds with $\dim M < \dim N$, and $F\colon M \to N$ is a smooth map. Then $F(M)$ has measure zero in N. In particular, $N \smallsetminus F(M)$ is dense in N.*

Proof. Write $m = \dim M$ and $n = \dim N$, and let $\{(U_i, \varphi_i)\}$ be a countable cover of M by smooth charts. Given any smooth chart (V, ψ) for N, we need to show that $\psi(F(M) \cap V)$ has measure zero in $\mathbb{R}^n$. Observe that this set is the countable union of sets of the form $\psi \circ F \circ \varphi_i^{-1}(\varphi_i(F^{-1}(V) \cap U_i))$, each of which has measure zero by Lemma 10.3. □

Corollary 10.6. *If M is a smooth manifold and $N \subset M$ is an immersed submanifold of positive codimension, then N has measure zero in M.*

Theorem 10.5 can be considered as a special case of the following deeper (and somewhat harder to prove) theorem due to Arthur Sard.

Theorem 10.7 (Sard's Theorem). *If $F\colon M \to N$ is any smooth map, the set of critical values of F has measure zero in N.*

We will neither use nor prove this theorem in this book. For a proof, see [Mil65], [Ste64], or [Bre93].

The Whitney Embedding Theorem

Our first major task in this chapter is to show that every smooth n-manifold can be embedded in $\mathbb{R}^{2n+1}$. We will begin by proving that if $m \geq 2n$, any smooth map into $\mathbb{R}^m$ can be perturbed slightly to be an immersion.

Theorem 10.8. *Let $F\colon M \to \mathbb{R}^m$ be any smooth map, where M is a smooth n-manifold and $m \geq 2n$. For any $\varepsilon > 0$, there is a smooth immersion $\widetilde{F}\colon M \to \mathbb{R}^m$ such that $\sup_M |\widetilde{F} - F| \leq \varepsilon$.*

Proof. Let $\{W_i\}$ be any regular open cover of M as defined in Chapter 2 (for example, a regular refinement of the trivial cover consisting of M alone). Then each W_i is the domain of a smooth chart $\psi_i\colon W_i \to B_3(0)$, and the precompact sets $U_i = \psi_i^{-1}(B_1(0))$ still cover M. For each integer $k \geq 1$, let $M_k = \bigcup_{i=1}^k U_i$. We interpret M_0 to be the empty set. We will modify F inductively on one set W_i at a time.

For each i, let $\varphi_i \in C^\infty(M)$ be a smooth bump function supported in W_i that is equal to 1 on $\overline{U}_i$. Let $F_0 = F$, and suppose by induction we have defined smooth maps $F_j\colon M \to \mathbb{R}^m$ for $j = 0, \ldots, k-1$ satisfying

(i) $\sup_M |F_j - F| < \varepsilon$;

(ii) If $j \geq 1$, $F_j(x) = F_{j-1}(x)$ unless $x \in W_j$;

(iii) $(F_j)_*$ is injective at each point of $\overline{M}_j$.

For any $m \times n$ matrix A, define a new map $F_A\colon M \to \mathbb{R}^m$ as follows: On $M \smallsetminus \operatorname{supp}\varphi_k$, $F_A = F_{k-1}$; and on W_k, F_A is the map given in coordinates by

$$F_A(x) = F_{k-1}(x) + \varphi_k(x)Ax,$$

where $A\colon \mathbb{R}^n \to \mathbb{R}^m$ is thought of as a linear map. (When computing in W_k, we simplify the notation by identifying maps with their coordinate

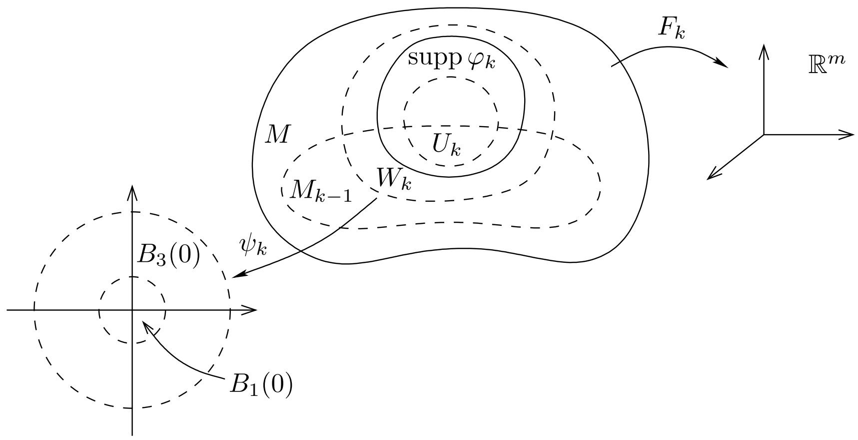


Figure 10.6. Perturbing a map to be an immersion.

representations as usual.) Since both definitions agree on the set $W_k \smallsetminus \operatorname{supp}\varphi_k$ where they overlap, this defines a smooth map. We will eventually set $F_k = F_A$ for a suitable choice of A (Figure 10.6).

Because (i) holds for $j = k-1$, there is a constant $\varepsilon_0 < \varepsilon$ such that $|F_{k-1}(x) - F(x)| \leq \varepsilon_0$ for x in the compact set $\operatorname{supp}\varphi_k$. By continuity, therefore, there is some $\delta > 0$ such that $|A| < \delta$ implies

$$\sup_M |F_A - F_{k-1}| = \sup_{x \in \operatorname{supp}\varphi_k} |\varphi_k(x)Ax| < \varepsilon - \varepsilon_0,$$

and therefore

$$\sup_M |F_A - F| \leq \sup_M |F_A - F_{k-1}| + \sup_M |F_{k-1} - F| < (\varepsilon - \varepsilon_0) + \varepsilon_0 = \varepsilon.$$

Let $P\colon W_k \times \mathrm{M}(m \times n, \mathbb{R}) \to \mathrm{M}(m \times n, \mathbb{R})$ be the matrix-valued function

$$P(x, A) = DF_A(x).$$

By the inductive hypothesis, $P(x, A)$ has rank n when (x, A) is in the compact set $(\operatorname{supp}\varphi_k \cap \overline{M}_{k-1}) \times \{0\}$. By choosing δ even smaller if necessary, we may also ensure that $\operatorname{rank} P(x, A) = n$ whenever $x \in \operatorname{supp}\varphi_k \cap \overline{M}_{k-1}$ and $|A| < \delta$.

The last condition we need to ensure is that $\operatorname{rank}(F_A)_* = n$ on $\overline{U}_k$ and therefore on $\overline{M}_k = \overline{M}_{k-1} \cup \overline{U}_k$. Notice that $DF_A(x) = DF_{k-1}(x) + A$ for $x \in \overline{U}_k$ because $\varphi_k \equiv 1$ there, and therefore $DF_A(x)$ has rank n in $\overline{U}_k$ if and only if A is not of the form $B - DF_{k-1}(x)$ for any $x \in \overline{U}_k$ and any matrix B of rank less than n. To ensure this, let $Q\colon W_k \times \mathrm{M}(m \times n, \mathbb{R}) \to \mathrm{M}(m \times n, \mathbb{R})$ be the smooth map

$$Q(x, B) = B - DF_{k-1}(x).$$

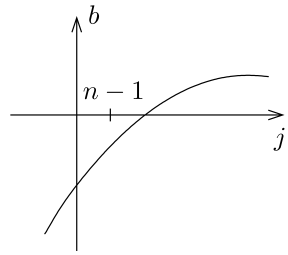


Figure 10.7. The graph of $b = n - (m - j)(n - j)$.

We need to show that there is some matrix A with $|A| < \delta$ that is not of the form $Q(x, B)$ for any $x \in \overline{U}_k$ and any matrix B of rank less than n. For each $j = 0, \dots, n-1$, the set $\mathrm{M}_j(m \times n, \mathbb{R})$ of $m \times n$ matrices of rank j is an embedded submanifold of $\mathrm{M}(m \times n, \mathbb{R})$ of codimension $(m-j)(n-j)$ by Example 8.14. By Theorem 10.5, therefore, $Q(W_k \times \mathrm{M}_j(m \times n, \mathbb{R}))$ has measure zero in $\mathrm{M}(m \times n, \mathbb{R})$ provided the dimension of $W_k \times \mathrm{M}_j(m \times n, \mathbb{R})$ is strictly less than the dimension of $\mathrm{M}(m \times n, \mathbb{R})$, which is to say

$$n + mn - (m - j)(n - j) < mn$$

or equivalently

$$n - (m - j)(n - j) < 0. \tag{10.3}$$

When $j = n-1$, $n-(m-j)(n-j) = 2n-m-1$, which is negative because we are assuming $m \geq 2n$. For $j \leq n - 1$, $n - (m - j)(n - j)$ is increasing in j because its derivative with respect to j is positive there (Figure 10.7). Thus (10.3) holds whenever $0 \leq j \leq n - 1$. This implies that for each $j = 0, \dots, n - 1$, the image under Q of $W_k \times \mathrm{M}_j(m \times n, \mathbb{R})$ has measure zero in $\mathrm{M}(m \times n, \mathbb{R})$. Choosing A such that $|A| < \delta$ and A is not in the union of these image sets, and setting $F_k = F_A$, we obtain a map satisfying the three conditions of the inductive hypothesis for $j = k$.

Now let $\widetilde{F}(x) = \lim_{k\to\infty} F_k(x)$. By local finiteness of the cover $\{W_j\}$, for each k there is some $N(k) > k$ such that $W_k \cap W_j = \varnothing$ for all $j \geq N(k)$, and then condition (ii) implies that $F_{N(k)} = F_{N(k)+1} = \cdots = F_i$ on W_k for all $i \geq N(k)$. Thus the sequence $\{F_k(x)\}$ is eventually constant for x in a neighborhood of any point, and so $\widetilde{F}\colon M \to \mathbb{R}^m$ is a smooth map. It is an immersion because $\widetilde{F} = F_{N(k)}$ on M_k, which has rank n by (iii). □

Corollary 10.9 (Whitney Immersion Theorem). *Every smooth n-manifold admits an immersion into $\mathbb{R}^{2n}$.*

Proof. Just apply the preceding theorem to any smooth map $F\colon M \to \mathbb{R}^{2n}$, for example a constant map. □

Next we show how to perturb our immersion to be injective. The intuition behind this theorem is that, due to the rank theorem, the image of an immersion looks locally like an n-dimensional affine subspace (after a suitable change of coordinates), so if $F(M) \subset \mathbb{R}^m$ has self-intersections, they will look locally like the intersection between two n-dimensional affine subspaces. If m is at least $2n+1$, such affine subspaces of $\mathbb{R}^m$ can be translated slightly so as to be disjoint, so we might hope to remove the self-intersections by perturbing F a little. The details of the proof are a bit more involved, but the idea is the same.

Theorem 10.10. *Let M be a smooth n-manifold, and suppose $m \geq 2n+1$ and $F\colon M \to \mathbb{R}^m$ is an immersion. Then for any $\varepsilon > 0$ there is an injective immersion $\widetilde{F}\colon M \to \mathbb{R}^m$ such that $\sup_M |\widetilde{F} - F| \leq \varepsilon$.*

Proof. Because an immersion is locally an embedding, there is an open cover $\{W_i\}$ of M such that the restriction of F to each W_i is injective. Passing to a refinement, we may assume that it is a regular cover. As in the proof of the previous theorem, for each i, let $\psi_i\colon W_i \to B_3(0)$ be the associated chart, let $U_i = \psi_i^{-1}(B_1(0))$, and let $\varphi_i \in C^\infty(M)$ be a smooth bump function supported in W_i that is equal to 1 on $\overline{U}_i$. Let $M_k = \bigcup_{i=1}^k U_k$ (see Figure 10.6 again).

As before, we will modify F inductively to make it injective on successively larger sets. Let $F_0 = F$, and suppose by induction we have defined smooth maps $F_j\colon M \to \mathbb{R}^m$ for $j = 0, \dots, k-1$ satisfying

(i) F_j is an immersion;

(ii) $\sup_M |F_j - F| < \varepsilon$;

(iii) If $j \geq 1$, then $F_j(x) = F_{j-1}(x)$ unless $x \in W_j$;

(iv) F_j is injective on $\overline{M}_j$;

(v) F_j is injective on W_i for every i.

Define the next map $F_k\colon M \to \mathbb{R}^m$ by

$$F_k(x) = F_{k-1}(x) + \varphi_k(x)b,$$

where $b \in \mathbb{R}^m$ is to be determined.

We wish to choose b such that $F_k(x) \neq F_k(y)$ when x and y are distinct points of $\overline{M}_k$. To begin, by an argument analogous to that of Theorem 10.8, there exists δ such that $|b| < \delta$ implies

$$\sup_M |F_k - F| \leq \sup_{\operatorname{supp}\varphi_k} |F_k - F_{k-1}| + \sup_M |F_{k-1} - F| < \varepsilon.$$

Choosing δ smaller if necessary, we may also ensure that $(F_k)_*$ is injective at each point of the compact set $\operatorname{supp}\varphi_k$; since $(F_k)_* = (F_{k-1})_*$ is already injective on the rest of M, this implies that F_k is an immersion.

Next, observe that if $F_k(x) = F_k(y)$, then exactly one of the following two cases must hold:

Case I: $\varphi_k(x) \neq \varphi_k(y)$ and

$$b = -\frac{F_{k-1}(x) - F_{k-1}(y)}{\varphi_k(x) - \varphi_k(y)}. \tag{10.4}$$

Case II: $\varphi_k(x) = \varphi_k(y)$ and therefore also $F_{k-1}(x) = F_{k-1}(y)$.

Define an open subset $U \subset M \times M$ by

$$U = \{(x,y) : \varphi_k(x) \neq \varphi_k(y)\},$$

and let $R\colon U \to \mathbb{R}^m$ be the smooth map

$$R(x,y) = -\frac{F_{k-1}(x) - F_{k-1}(y)}{\varphi_k(x) - \varphi_k(y)}.$$

Because $\dim U = \dim(M \times M) = 2n < m$, Theorem 10.5 implies that $R(U)$ has measure zero in $\mathbb{R}^m$. Therefore there exists $b \in \mathbb{R}^m$ with $|b| < \delta$ such that (10.4) does not hold for any $(x,y) \in U$. With this b, (i)–(iii) hold with $j = k$. We need to show that (iv) and (v) hold as well.

If $F_k(x) = F_k(y)$ for a pair of distinct points $x, y \in \overline{M}_k$, case I above cannot hold by our choice of b. Therefore we are in case II: $\varphi_k(x) = \varphi_k(y)$ and $F_{k-1}(x) = F_{k-1}(y)$. If $\varphi_k(x) = \varphi_k(y) = 0$, then $x, y \in \overline{M}_k \smallsetminus \overline{U}_k \subset \overline{M}_{k-1}$, contradicting the fact that F_{k-1} is injective on $\overline{M}_{k-1}$ by the inductive hypothesis. On the other hand, if $\varphi_k(x)$ and $\varphi_k(y)$ are nonzero, then $x, y \in \operatorname{supp} \varphi_k \subset W_k$, which contradicts the fact that F_{k-1} is injective on each W_i by the inductive hypothesis. This shows that F_k satisfies (iv). Similarly, if $F_k(x) = F_k(y)$ for a pair of distinct points x, y in *any one* of the sets W_i, the same argument shows that $F_{k-1}(x) = F_{k-1}(y)$, contradicting the fact that F_{k-1} is injective on W_i. This shows that F_{k-1} satisfies (v), and completes the induction.

Now we let $\widetilde{F}(x) = \lim_{j\to\infty} F_j(x)$. As before, for any k, this sequence is constant on W_k for j sufficiently large, so defines a smooth function. If $\widetilde{F}(x) = \widetilde{F}(y)$, choose k such that $x, y \in \overline{M}_k$. For sufficiently large j, $\widetilde{F} = F_j$ on $\overline{M}_k$, so the injectivity of F_j on $\overline{M}_k$ implies that $x = y$. □

We can now prove the main result of this section.

Theorem 10.11 (Whitney Embedding Theorem). *Every smooth n-manifold admits a proper smooth embedding into $\mathbb{R}^{2n+1}$.*

Proof. Let M be a smooth n-manifold. By Proposition 7.4, a proper injective immersion is a smooth embedding. We will start with a smooth proper map $F_0\colon M \to \mathbb{R}^{2n+1}$, and then use the two preceding theorems to perturb it to a proper injective immersion.

Let $f\colon M \to \mathbb{R}$ be a smooth exhaustion function (see Proposition 2.28). It is easy to check that the map $F_0\colon M \to \mathbb{R}^{2n+1}$ defined by $F_0(x) =$

$(f(x), 0, \dots, 0)$ is smooth and proper. Now by Theorem 10.8, there is an immersion $F_1\colon M \to \mathbb{R}^{2n+1}$ satisfying $\sup_M |F_1 - F_0| \le 1$. And by Theorem 10.10, there is an injective immersion $F_2\colon M \to \mathbb{R}^{2n+1}$ satisfying $\sup_M |F_2 - F_1| \le 1$. If $K \subset \mathbb{R}^{2n+1}$ is any compact set, it is contained in some closed ball $\overline{B}_R(0)$, and thus if $F_2(p) \in K$ we have

$$|F_0(p)| \le |F_0(p) - F_1(p)| + |F_1(p) - F_2(p)| + |F_2(p)| \le 1 + 1 + R,$$

which implies $F_2^{-1}(K)$ is a closed subset of $F_0^{-1}(\overline{B}_{2+R}(0))$, which is compact because F_0 is proper. Thus F_2 is a proper injective immersion and hence a smooth embedding. □

Corollary 10.12. *Every smooth n-manifold is diffeomorphic to a closed embedded submanifold of $\mathbb{R}^{2n+1}$.*

Proof. By the preceding theorem, every smooth n-manifold admits a proper smooth embedding into $\mathbb{R}^{2n+1}$, the image of which is an embedded submanifold by Theorem 8.3 and closed in $\mathbb{R}^{2n+1}$ because proper continuous maps are closed (Proposition 2.18). □

A topological space is said to be *metrizable* if it admits a distance function whose metric topology is the same as the given topology. Since any subset of a metric space inherits a metric that determines its subspace topology, the following corollary is an immediate consequence of the previous one.

Corollary 10.13. *Every smooth manifold is metrizable.*

Theorem 10.11 and Corollary 10.12, first proved by Hassler Whitney in 1936 [Whi36], answered a question that had been nagging mathematicians since the notion of an abstract manifold was first introduced: Are there abstract smooth manifolds that are not diffeomorphic to embedded submanifolds of Euclidean space? Although this version of the theorem will be quite sufficient for our purposes, it is interesting to note that eight years later [Whi44a, Whi44b], using much more sophisticated techniques of algebraic topology, Whitney was able to obtain the following improvements.

Theorem 10.14 (Strong Whitney Immersion Theorem). *If $n > 1$, every smooth n-manifold admits an immersion into $\mathbb{R}^{2n-1}$.*

Theorem 10.15 (Strong Whitney Embedding Theorem). *If $n > 0$, every smooth n-manifold admits a smooth embedding into $\mathbb{R}^{2n}$.*

The Whitney Approximation Theorems

In this section we prove the two theorems mentioned at the beginning of the chapter on approximation of continuous maps by smooth ones. Both of these theorems, like the embedding theorems we just proved, are due to Hassler Whitney [Whi36].

We begin with a theorem about smoothly approximating functions into Euclidean spaces. Our first theorem shows, in particular, that any continuous function from a smooth manifold M into $\mathbb{R}^k$ can be uniformly approximated by a smooth function. In fact, we will prove something stronger. If $\delta\colon M \to \mathbb{R}$ is a positive continuous function, we say that two functions $F, \widetilde{F}\colon M \to \mathbb{R}^k$ are *δ-close* if $|F(x) - \widetilde{F}(x)| < \delta(x)$ for all $x \in M$.

Theorem 10.16 (Whitney Approximation Theorem). *Let M be a smooth manifold and let $F\colon M \to \mathbb{R}^k$ be a continuous function. Given any positive continuous function $\delta\colon M \to \mathbb{R}$, there exists a smooth function $\widetilde{F}\colon M \to \mathbb{R}^k$ that is δ-close to F. If F is smooth on a closed subset $A \subset M$, then $\widetilde{F}$ can be chosen to be equal to F on A.*

Proof. If F is smooth on the closed set A, then by the extension lemma there is a smooth function $F_0\colon M \to \mathbb{R}^k$ that agrees with F on A. Let

$$U_0 = \{y \in M : |F_0(y) - F(y)| < \delta(y)\}.$$

It is easy to verify that U_0 is an open set containing A. (If there is no such set A, we just take $U_0 = A = \varnothing$.)

We will show that there are countably many points $\{x_i\}_{i=1}^\infty$ in $M \smallsetminus A$ and neighborhoods U_i of x_i in $M \smallsetminus A$ such that

$$|F(y) - F(x_i)| < \delta(y) \text{ for all } y \in U_i. \tag{10.5}$$

To see this, for any $x \in M \smallsetminus A$, let U_x be a neighborhood of x contained in $M \smallsetminus A$ and small enough that

$$\delta(y) > \tfrac{1}{2}\delta(x) \qquad \text{and} \qquad |F(y) - F(x)| < \tfrac{1}{2}\delta(x)$$

for all $y \in U_x$. (Such a neighborhood exists by continuity of δ and F.) Then if $y \in U_x$, we have

$$|F(y) - F(x)| < \tfrac{1}{2}\delta(x) < \delta(y).$$

The collection of all such sets U_x as x ranges over points of $M \smallsetminus A$ is an open cover of $M \smallsetminus A$. Choosing a countable subcover $\{U_{x_i}\}_{i=1}^\infty$ and setting $U_i = U_{x_i}$, we have (10.5).

Let $\{\varphi_0, \varphi_i\}$ be a smooth partition of unity subordinate to the cover $\{U_0, U_i\}$ of M, and define $\widetilde{F}\colon M \to \mathbb{R}^k$ by

$$\widetilde{F}(y) = \varphi_0(y)F_0(y) + \sum_{i \ge 1} \varphi_i(y)F(x_i).$$

Then clearly $\widetilde{F}$ is smooth, and is equal to F on A. For any $y \in M$, the fact that $\sum_{i\geq 0} \varphi_i \equiv 1$ implies that

$$\begin{aligned}
&|\widetilde{F}(y) - F(y)| \\
&\qquad = \left| \varphi_0(y)F_0(y) + \sum_{i\geq 1} \varphi_i(y)F(x_i) - \Big(\varphi_0(y) + \sum_{i\geq 1}\varphi_i(y)\Big)F(y) \right| \\
&\qquad \leq \varphi_0(y)|F_0(y) - F(y)| + \sum_{i\geq 1} \varphi_i(y)|F(x_i) - F(y)| \\
&\qquad < \varphi_0(y)\delta(y) + \sum_{i\geq 1}\varphi_i(y)\delta(y) = \delta(y)
\end{aligned}$$

□

Tubular Neighborhoods

We would like to find a way to apply the Whitney approximation theorem to produce smooth approximations to continuous maps between smooth manifolds. If $F\colon N \to M$ is such a map, then by the Whitney embedding theorem we can consider M as an embedded submanifold of some Euclidean space $\mathbb{R}^n$, and approximate F by a smooth map into $\mathbb{R}^n$. However, in general, the image of this smooth map will not lie in M. To correct for this, we need to know that there is a smooth retraction from some neighborhood of M onto M. For this purpose, we introduce a few more definitions.

Let $M \subset \mathbb{R}^n$ be an embedded m-dimensional submanifold. At any $x \in M$, our usual identifications allow us to view the tangent space T_xM as a subspace of $T_x\mathbb{R}^n$, which inherits a Euclidean dot product courtesy of its canonical identification with $\mathbb{R}^n$. We define the *normal space* to M at x to be the subspace $N_xM \subset T_x\mathbb{R}^n$ consisting of all vectors that are orthogonal to T_xM with respect to the Euclidean dot product. The *normal bundle* of M is the subset $NM \subset T\mathbb{R}^n$ defined by

$$NM = \coprod_{x\in M} N_xM = \{(x,v) \in T\mathbb{R}^n : x \in M \text{ and } v \in N_xM\}.$$

There is a natural projection $\pi_{NM}\colon NM \to M$ defined as the restriction to NM of $\pi\colon T\mathbb{R}^n \to \mathbb{R}^n$, and each fiber N_xM is a vector space of dimension $n - m$.

A local frame $(E_1, \dots, E_n)$ for $\mathbb{R}^n$ on an open set $U \subset \mathbb{R}^n$ is said to be *orthonormal* if the vectors $(E_1|_x, \dots, E_n|_x)$ are orthonormal at each point $x \in U$. It is said to be *adapted to M* if the first m vectors $(E_1|_x, \dots, E_m|_x)$ span T_xM at each $x \in U \cap M$.

Lemma 10.17 (Existence of Adapted Orthonormal Frames). *Let $M \subset \mathbb{R}^n$ be an embedded submanifold. For each $p \in M$, there is a smooth adapted orthonormal frame on a neighborhood U of p in $\mathbb{R}^n$.*

Proof. Let $(y^1, \dots, y^n)$ be slice coordinates for M on a neighborhood U of x, so that $M \cap U$ is the set where $y^{m+1} = \dots = y^n = 0$. Applying the Gram-Schmidt algorithm (Proposition A.47) to the smooth frame $(\partial/\partial y^i)$, we obtain an orthonormal frame $(E_1, \dots, E_n)$ given inductively by the formula

$$E_j = \frac{\partial/\partial y^j - \sum_{i=1}^{j-1}(\partial/\partial y^j \cdot E_i)E_i}{\left|\partial/\partial x^j - \sum_{i=1}^{j-1}(\partial/\partial y^j \cdot E_i)E_i\right|}.$$

Because $\operatorname{span}(E_1, \dots, E_{j-1}) = \operatorname{span}(\partial/\partial y^1, \dots, \partial/\partial y^{j-1})$, the vector whose norm appears in the denominator above is nowhere zero on U. Thus this formula defines (E_j) as a smooth orthonormal frame on U. In particular, for each $x \in M$, $\operatorname{span}(E_1|_x, \dots, E_m|_x) = \operatorname{span}(\partial/\partial y^1|_x, \dots, \partial/\partial y^m|_x) = T_xM$, so this frame is adapted to M. □

Proposition 10.18. *For any embedded m-dimensional submanifold $M \subset \mathbb{R}^n$, the normal bundle $NM \subset T\mathbb{R}^n$ is a smooth vector bundle of rank $n-m$ over M, and an embedded submanifold of $T\mathbb{R}^n$.*

Proof. We define local trivializations for NM as follows. For any $p \in M$, let $(E_1, \dots, E_n)$ be an adapted smooth orthonormal frame on a neighborhood U of p, and define $\Phi\colon \pi_{NM}^{-1}(M \cap U) \to (M \cap U) \times \mathbb{R}^{n-m}$ by

$$\Phi(a^i E_i|_x) = (x, (a^{m+1}, \dots, a^n)).$$

This is a bijection because N_xM is spanned by $(E_{m+1}|_x, \dots, E_n|_x)$. If $(\widetilde{E}_j)$ is another such adapted frame on an open set $\widetilde{U} \subset \mathbb{R}^n$, there is a smooth matrix-valued function $(A_i^j)\colon U \cap \widetilde{U} \to \mathrm{GL}(n, \mathbb{R})$ such that $E_i = A_i^j \widetilde{E}_j$. The corresponding local trivializations Φ and $\widetilde{\Phi}$ are related by

$$\widetilde{\Phi} \circ \Phi^{-1}(x, (a^{m+1}, \dots, a^n)) = \left(x, \left(\sum_{i=m+1}^{n} A_i^{m+1} a^i, \dots, \sum_{i=m+1}^{n} A_i^n a^i\right)\right).$$

Because $\operatorname{span}(E_{m+1}|_x, \dots, E_n|_x) = \operatorname{span}(\widetilde{E}_{m+1}|_x, \dots, \widetilde{E}_n|_x) = N_xM$ at each point $x \in M \cap U \cap \widetilde{U}$, the lower right $(n-m) \times (n-m)$ minor of (A_j^i) is nonsingular on $M \cap U \cap \widetilde{U}$. It follows from the vector bundle construction lemma that NM has a unique smooth structure making $\pi_{NM}\colon NM \to M$ into a smooth vector bundle.

It remains only to show that NM is embedded in $T\mathbb{R}^n$. Given $p \in M$, let (y^i) be slice coordinates for M on a neighborhood U of p in $\mathbb{R}^n$, and (shrinking U if necessary), let (E_i) be an adapted smooth orthonormal frame on U. By Corollary 5.12, the map $\widetilde{\varphi}\colon \pi^{-1}(U) \to \varphi(U) \times \mathbb{R}^n$ given by

$$\widetilde{\varphi}(a^i E_i|_x) = (y^1(x), \dots, y^n(x), a^1, \dots, a^n)$$

is a smooth coordinate chart for $T\mathbb{R}^n$. These are in fact slice coordinates for NM, because $\varphi(NM \cap U)$ is equal to the slice

$$\{(y, a) : y^{m+1} = \dots = y^n = a^1 = \dots = a^m = 0\}. \qquad \square$$

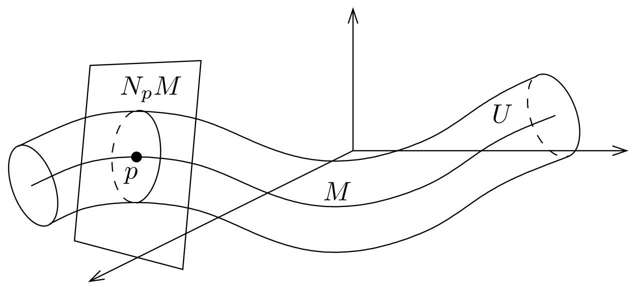


Figure 10.8. A tubular neighborhood.

Define a map $E\colon NM \to \mathbb{R}^n$ by

$$E(x, v) = x + v,$$

where we regard a vector $v \in N_xM \subset T_x\mathbb{R}^n$ as an element of $\mathbb{R}^n$ by means of the usual identification $T_x\mathbb{R}^n \cong \mathbb{R}^n$. This just maps each normal space N_xM affinely onto the affine subspace through x and orthogonal to T_xM. Clearly E is smooth because it is the restriction of the addition map $T\mathbb{R}^n \approx \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ to NM. A *tubular neighborhood* of M is a neighborhood U of M in $\mathbb{R}^n$ that is the diffeomorphic image under E of an open subset $V \subset NM$ of the form

$$V = \{(x, v) \in NM : |v| < \delta(x)\}, \tag{10.6}$$

for some positive continuous function $\delta\colon M \to \mathbb{R}$ (Figure 10.8).

Theorem 10.19 (Tubular Neighborhood Theorem). *Every embedded submanifold of $\mathbb{R}^n$ has a tubular neighborhood.*

Proof. Let $M_0 \subset NM$ be the subset $\{(x, 0) : x \in M\}$ (the image of the zero section of NM). We begin by showing that E is a local diffeomorphism in a neighborhood of M_0. Because NM and $\mathbb{R}^n$ have the same dimension, it suffices to show that E_* is surjective at each point of M_0. If $v \in T_xM$, there is a smooth curve $\gamma\colon (-\varepsilon, \varepsilon) \to M$ such that $\gamma(0) = x$ and $\gamma'(0) = v$. Let $\widetilde{\gamma}\colon (-\varepsilon, \varepsilon) \to NM$ be the curve $\widetilde{\gamma}(t) = (\gamma(t), 0)$. Then

$$E_*\widetilde{\gamma}'(0) = (E \circ \widetilde{\gamma})'(0) = \left.\frac{d}{dt}\right|_{t=0} (\gamma(t) + 0) = v.$$

On the other hand, if $w \in N_xM$, then defining $\sigma\colon (-\varepsilon, \varepsilon) \to NM$ by $\sigma(t) = (x, tw)$, we obtain

$$E_*\sigma'(t) = (E \circ \sigma)'(0) = \left.\frac{d}{dt}\right|_{t=0} (x + tw) = w.$$

Since T_xM and N_xM span $T_x\mathbb{R}^n$, this shows E_* is surjective. By the inverse function theorem, E is a diffeomorphism on a neighborhood of $(x, 0)$ in

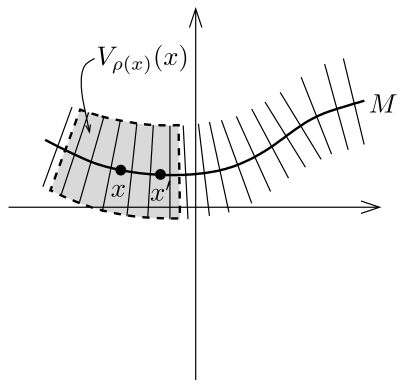


Figure 10.9. Continuity of ρ.

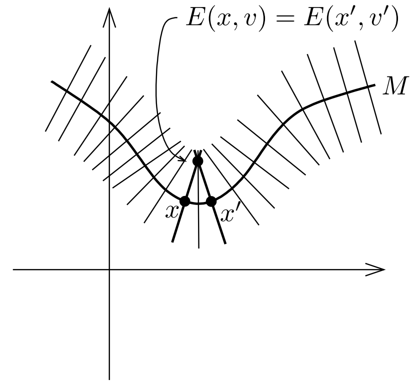


Figure 10.10. Injectivity of E.

NM, which we can take to be of the form $V_\delta(x) = \{(x',v') : |x - x'| < \delta, |v'| < \delta\}$ for some $\delta > 0$. (This uses the fact that NM is embedded and therefore its topology is induced by the Euclidean metric.)

To complete the proof, we need to show that there is an open set V of the form (10.6) on which E is a global diffeomorphism. For each point $x \in M$, let $\rho(x)$ be the supremum of all $\delta \leq 1$ such that E is a diffeomorphism on $V_\delta(x)$. The argument in the preceding paragraph implies that $\rho\colon M \to \mathbb{R}$ is positive. It is also continuous: Given $x, x' \in M$, if $|x - x'| < \rho(x)$, then by the triangle inequality $V_\delta(x')$ is contained in $V_{\rho(x)}(x)$ for $\delta = \rho(x) - |x - x'|$ (Figure 10.9), which implies that $\rho(x') \geq \rho(x) - |x - x'|$, or $\rho(x) - \rho(x') \leq |x - x'|$. The same is true trivially if $|x - x'| \geq \rho(x)$. Reversing the roles of x and x' yields the opposite inequality, which shows that $|\rho(x) - \rho(x')| \leq |x - x'|$, so ρ is continuous.

Now let $V = \{(x,v) \in NM : |v| < \frac{1}{2}\rho(x)\}$. We will show that E is injective on V. Suppose that (x,v) and (x',v') are points in V such that $E(x,v) = E(x',v')$ (Figure 10.10). Assume without loss of generality that $\rho(x') \leq \rho(x)$. It follows from $x + v = x' + v'$ that

$$|x - x'| = |v - v'| \leq |v| + |v'| < \tfrac{1}{2}\rho(x) + \tfrac{1}{2}\rho(x') \leq \rho(x).$$

Choosing $\delta \in \mathbb{R}$ such that $\max(|v|, |v'|, |x - x'|) < \delta < \rho(x)$, we see that both (x,v) and (x',v') are in $V_\delta(x)$. Since E is injective on this set, this implies $(x,v) = (x',v')$.

The set $U = E(V)$ is open in $\mathbb{R}^n$ because $E|_V$ is a local diffeomorphism and thus an open map. It follows that $E\colon V \to U$ is a smooth bijection and a local diffeomorphism, hence a diffeomorphism by Exercise 2.9. Therefore U is a tubular neighborhood of M. □

The primary reason we are interested in tubular neighborhoods is because of the next proposition. A *retraction* of a topological space X onto a subspace $M \subset X$ is a continuous map $r\colon X \to M$ such that $r|_M$ is the identity map of M.

Proposition 10.20. *Let $M \subset \mathbb{R}^n$ be an embedded submanifold. If U is any tubular neighborhood of M, there exists a smooth retraction of U onto M.*

Proof. By definition, there is an open subset $V \subset NM$ containing M such that $E\colon V \to U$ is a diffeomorphism. Just define $r\colon U \to M$ by $r = \pi \circ E^{-1}$, where $\pi\colon NM \to M$ is the natural projection. Clearly r is smooth. For $x \in M$, note that $E(x, 0) = x$, so $r(x) = \pi \circ E^{-1}(x) = \pi(x, 0) = x$, which shows that r is a retraction. □

Smooth Approximation of Maps Between Manifolds

The next theorem gives a form of smooth approximation for continuous maps between manifolds. It will have important applications later when we study de Rham cohomology.

Theorem 10.21 (Whitney Approximation on Manifolds). *Let N and M be smooth manifolds, and let $F\colon N \to M$ be a continuous map. Then F is homotopic to a smooth map $\widetilde{F}\colon N \to M$. If F is smooth on a closed subset $A \subset N$, then the homotopy can be taken to be relative to A.*

Proof. By the Whitney embedding theorem, we may as well assume that M is an embedded submanifold of $\mathbb{R}^n$. Let U be a tubular neighborhood of M in $\mathbb{R}^n$, and let $r\colon U \to M$ be the smooth retraction given by Lemma 10.20. For any $x \in M$, let

$$\delta(x) = \sup\{\varepsilon \le 1 : B_\varepsilon(x) \subset U\}.$$

By a triangle-inequality argument entirely analogous to the one in the proof of the tubular neighborhood theorem, $\delta\colon M \to \mathbb{R}$ is continuous.

Let $\widetilde{\delta} = \delta \circ F\colon N \to \mathbb{R}$. By the Whitney approximation theorem, there exists a smooth map $\widetilde{F}\colon N \to \mathbb{R}^n$ that is $\widetilde{\delta}$-close to F, and is equal to F on A (which might be the empty set). Define a homotopy $H\colon N \times I \to M$ by

$$H(p,t) = r((1-t)F(p) + t\widetilde{F}(p)).$$

This is well-defined, because our condition on $\widetilde{F}$ guarantees that for each p, $|\widetilde{F}(p) - F(p)| < \widetilde{\delta}(p) = \delta(F(p))$, which means that $\widetilde{F}(p)$ is contained in the ball of radius $\delta(F(p))$ around $F(p)$; since this ball is contained in U, so is the entire line segment from $F(p)$ to $\widetilde{F}(p)$.

Thus H is a homotopy between $H(p,0) = F(p)$ and $H(p,1) = r(\widetilde{F}(p))$, which is smooth. It satisfies $H(p,t) = F(p)$ for all $p \in A$, since $F = \widetilde{F}$ there. □

We end the chapter with an application to homotopy theory. If M and N are smooth manifolds, two smooth maps $F, G\colon M \to N$ are said to be *smoothly homotopic* if there is a smooth map $H\colon M \times I \to N$ that is a homotopy between F and G.

Proposition 10.22. *If $F, G\colon M \to N$ are homotopic smooth maps, then they are smoothly homotopic. If F is homotopic to G relative to some closed subset $A \subset M$, then they are smoothly homotopic relative to A.*

Proof. Let $H\colon M \times I \to N$ be a homotopy from F to G (relative to A, which may be empty). We wish to show that H can be replaced by a smooth homotopy.

Because Theorem 10.21 does not apply directly to manifolds with boundary, we first need to extend H to a manifold without boundary containing $M \times I$. Let $J = (-\varepsilon, 1+\varepsilon)$ for some $\varepsilon > 0$, and define $\overline{H}\colon M \times J \to M$ by

$$\overline{H}(x,t) = \begin{cases} H(x,t) & t \in [0,1] \\ H(x,0) & t \le 0 \\ H(x,1) & t \ge 1. \end{cases}$$

This is continuous by the gluing lemma. Moreover, the restriction of $\overline{H}$ to $M \times \{0\} \cup M \times \{1\}$ is smooth, because it is equal to $F \circ \pi_1$ on $M \times \{0\}$ and $G \circ \pi_1$ on $M \times \{1\}$ (where $\pi_1\colon M \times I \to M$ is the projection on the first factor). If $F \simeq G$ relative to A, $\overline{H}$ is also smooth on $A \times I$. Therefore, Theorem 10.21 implies that there is a smooth map $\widetilde{H}\colon M \times J \to N$ (homotopic to $\overline{H}$, but we do not need that here) whose restriction to $M \times \{0\} \cup M \times \{1\} \cup A \times I$ equals $\overline{H}$ (and therefore H). Restricting back to $M \times I$ again, we see that $\widetilde{H}|_{M \times I}$ is a smooth homotopy (relative to A) between F and G. □

Problems

10-1. Show that any two points in a connected smooth manifold can be joined by a smooth curve segment.

10-2. Let $M \subset \mathbb{R}^m$ be an embedded submanifold, let U be a tubular neighborhood of M, and let $r\colon U \to M$ be the retraction defined in Proposition 10.20. Show that U can be chosen small enough that for each $x \in U$, $r(x)$ is the point in M closest to x. [Hint: First show that U can be chosen so that each point $x \in U$ has a closest point $y \in M$, and this point satisfies $(x - y) \perp T_yM$.]

10-3. If $M \subset \mathbb{R}^m$ is an embedded submanifold and $\varepsilon > 0$, let M_ε be the set of points in $\mathbb{R}^m$ whose distance from M is less than ε. If M is compact, show that for sufficiently small ε, ∂M_ε is a compact embedded submanifold of $\mathbb{R}^m$, and $\overline{M}_\varepsilon$ is a smooth manifold with boundary.

10-4. Suppose $M \subset \mathbb{R}^m$ is a closed embedded submanifold. If M admits a global defining function, show that its normal bundle is trivial.

Conversely, if M has trivial normal bundle, show that there is a neighborhood U of M in $\mathbb{R}^n$ and a submersion $\Phi\colon U \to \mathbb{R}^k$ such that $M = \Phi^{-1}(0)$.

10-5. Suppose M is a smooth, compact manifold that admits a nowhere vanishing vector field. Show that there exists a smooth map $F\colon M \to M$ that is homotopic to the identity and has no fixed points. [Hint: Use the tubular neighborhood theorem.]

10-6. Let M be a smooth manifold, let B be a closed subset of M, and let $\delta\colon M \to \mathbb{R}$ be a positive continuous function.

(a) If $f\colon M \to \mathbb{R}^k$ is any continuous function, show that there is a continuous function $\widetilde{f}\colon M \to \mathbb{R}^k$ that is smooth on $M \smallsetminus B$, agrees with f on B, and is δ-close to f. [Hint: Use Problem 2-18.]

(b) If $F\colon M \to N$ is a continuous map to a smooth manifold N, show that F is homotopic relative to B to a map that is smooth on $M \smallsetminus B$.

11
Tensors

Much of the machinery of smooth manifold theory is designed to allow the concepts of linear algebra to be applied to smooth manifolds. Calculus tells us how to approximate smooth objects by linear ones, and the abstract definitions of manifold theory give a way to interpret these linear approximations in a coordinate-independent way. In this chapter, we carry this idea much further, by generalizing from linear objects to multilinear ones. This leads to the concepts of tensors and tensor fields on manifolds.

We begin with tensors on a vector space, which are multilinear generalizations of covectors; a covector is the special case of a tensor of rank one. We give two alternative definitions of tensors on a vector space: On the one hand, they are real-valued multilinear functions of several vectors; on the other hand, they are elements of the abstract "tensor product" of the dual vector space with itself. Each definition is useful in certain contexts. We then discuss the difference between covariant and contravariant tensors, and give a brief introduction to tensors of mixed variance.

We then move to smooth manifolds, and define tensors, tensor fields, and tensor bundles. After describing the coordinate representations of tensor fields, we describe how they can be pulled back by smooth maps. We introduce a special class of tensors, the symmetric ones, whose values are unchanged by permutations of their arguments.

The last section of the chapter is an introduction to one of the most important kinds of tensor fields, Riemannian metrics. A thorough treatment of Riemannian geometry is beyond the scope of this book, but we can at least lay the groundwork by giving the basic definitions and proving that every manifold admits Riemannian metrics.

The Algebra of Tensors

Suppose $V_1, \dots, V_k$ and W are vector spaces. A map $F\colon V_1 \times \cdots \times V_k \to W$ is said to be *multilinear* if it is linear as a function of each variable separately:

$$F(v_1, \dots, av_i + a'v_i', \dots, v_k) \\ = aF(v_1, \dots, v_i, \dots, v_k) + a'F(v_1, \dots, v_i', \dots, v_k).$$

(A multilinear function of two variables is generally called *bilinear.*)

Although linear maps are paramount in differential geometry, there are many situations in which multilinear maps play an important geometric role. Here are a few examples to keep in mind:

- *The dot product* in $\mathbb{R}^n$ is a scalar-valued bilinear function of two vectors, used to compute lengths of vectors and angles between them.
- *The cross product* in $\mathbb{R}^3$ is a vector-valued bilinear function of two vectors, used to compute areas of parallelograms and to find a third vector orthogonal to two given ones.
- *The determinant* is a real-valued multilinear function of n vectors in $\mathbb{R}^n$, used to detect linear independence and to compute the volume of the parallelepiped spanned by the vectors.

In this section, we will develop a unified language for talking about multilinear functions—the language of tensors. In a little while, we will give a very general and abstract definition of tensors. But it will help to clarify matters if we start with a more concrete definition.

Let V be a finite-dimensional real vector space, and let k be a natural number. (Many of the concepts we will introduce in this section—at least the parts that do not refer explicitly to finite bases—work equally well in the infinite-dimensional case; but we will restrict our attention to the finite-dimensional case in order to keep things simple.)

A *covariant k-tensor* on V is a real-valued multilinear function of k elements of V:

$$T\colon \underbrace{V \times \cdots \times V}_{k \text{ copies}} \to \mathbb{R}.$$

The number k is called the *rank* of T. A 0-tensor is, by convention, just a real number (a real-valued function depending multilinearly on no vectors!). The set of all covariant k-tensors on V, denoted by $T^k(V)$, is a vector space under the usual operations of pointwise addition and scalar multiplication:

$$(aT)(X_1, \dots, X_k) = a(T(X_1, \dots, X_k)),$$
$$(T + T')(X_1, \dots, X_k) = T(X_1, \dots, X_k) + T'(X_1, \dots, X_k).$$

Let us look at some examples.

Example 11.1 (Covariant Tensors).

(a) Every linear map $\omega\colon V \to \mathbb{R}$ is multilinear, so a covariant 1-tensor is just a covector. Thus $T^1(V)$ is naturally identified with V^*.

(b) A covariant 2-tensor on V is a real-valued bilinear function of two vectors, also called a *bilinear form*. One example is the dot product on $\mathbb{R}^n$. More generally, any inner product on V is a covariant 2-tensor.

(c) The determinant, thought of as a function of n vectors, is a covariant n-tensor on $\mathbb{R}^n$.

(d) Suppose $\omega, \eta \in V^*$. Define a map $\omega \otimes \eta\colon V \times V \to \mathbb{R}$ by

$$\omega \otimes \eta(X, Y) = \omega(X)\eta(Y),$$

where the product on the right is just ordinary multiplication of real numbers. The linearity of ω and η guarantees that $\omega \otimes \eta$ is a bilinear function of X and Y, i.e., a covariant 2-tensor.

The last example can be generalized to tensors of any rank as follows. Let V be a finite-dimensional real vector space and let $S \in T^k(V)$, $T \in T^l(V)$. Define a map

$$S \otimes T\colon \underbrace{V \times \cdots \times V}_{k+l \text{ copies}} \to \mathbb{R}$$

by

$$S \otimes T(X_1, \dots, X_{k+l}) = S(X_1, \dots, X_k)T(X_{k+1}, \dots, X_{k+l}).$$

It is immediate from the multilinearity of S and T that $S \otimes T$ depends linearly on each argument X_i separately, so it is a covariant $(k+l)$-tensor, called the *tensor product* of S and T.

◇ **Exercise 11.1.** Show that the tensor product operation is bilinear and associative. More precisely, show that $S \otimes T$ depends linearly on each of the tensors S and T, and that $(R \otimes S) \otimes T = R \otimes (S \otimes T)$.

Because of the result of the preceding exercise, we can write the tensor product of three or more tensors unambiguously without parentheses. If $T_1, \dots, T_l$ are tensors of ranks $k_1, \dots, k_l$ respectively, their tensor product $T_1 \otimes \cdots \otimes T_l$ is a tensor of rank $k = k_1 + \cdots + k_l$, whose action on k vectors is given by inserting the first k_1 vectors into T_1, the next k_2 vectors into T_2, and so forth, and multiplying the results together. For example, if R and S are 2-tensors and T is a 3-tensor, then

$$R \otimes S \otimes T(X_1, \dots, X_7) = R(X_1, X_2)S(X_3, X_4)T(X_5, X_6, X_7).$$

Proposition 11.2. *Let V be a real vector space of dimension n, let (E_i) be any basis for V, and let (ε^i) be the dual basis. The set of all k-tensors of the form $\varepsilon^{i_1} \otimes \cdots \otimes \varepsilon^{i_k}$ for $1 \le i_1, \dots, i_k \le n$ is a basis for $T^k(V)$, which therefore has dimension n^k.*

Proof. Let $\mathcal{B}$ denote the set $\{\varepsilon^{i_1} \otimes \cdots \otimes \varepsilon^{i_k} : 1 \leq i_1, \ldots, i_k \leq n\}$. We need to show that $\mathcal{B}$ is independent and spans $T^k(V)$. Suppose $T \in T^k(V)$ is arbitrary. For any k-tuple $(i_1, \ldots, i_k)$ of integers such that $1 \leq i_j \leq n$, define a number $T_{i_1 \ldots i_k}$ by

$$T_{i_1 \ldots i_k} = T(E_{i_1}, \ldots, E_{i_k}). \tag{11.1}$$

We will show that

$$T = T_{i_1 \ldots i_k} \varepsilon^{i_1} \otimes \cdots \otimes \varepsilon^{i_k}$$

(with the summation convention in effect as usual), from which it follows that $\mathcal{B}$ spans $T^k(V)$. We compute

$$\begin{aligned} T_{i_1 \ldots i_k} \varepsilon^{i_1} \otimes \cdots \otimes \varepsilon^{i_k}(E_{j_1}, \ldots, E_{j_k}) &= T_{i_1 \ldots i_k} \varepsilon^{i_1}(E_{j_1}) \cdots \varepsilon^{i_k}(E_{j_k}) \\ &= T_{i_1 \ldots i_k} \delta^{i_1}_{j_1} \cdots \delta^{i_k}_{j_k} \\ &= T_{j_1 \ldots j_k} \\ &= T(E_{j_1}, \ldots, E_{j_k}). \end{aligned}$$

By multilinearity, a tensor is determined by its action on sequences of basis vectors, so this proves the claim.

To show that $\mathcal{B}$ is independent, suppose some linear combination equals zero:

$$T_{i_1 \ldots i_k} \varepsilon^{i_1} \otimes \cdots \otimes \varepsilon^{i_k} = 0.$$

Apply this to any sequence $(E_{j_1}, \ldots, E_{j_k})$ of basis vectors. By the same computation as above, this implies that each coefficient $T_{j_1 \ldots j_k}$ is zero. Thus the only linear combination of elements of $\mathcal{B}$ that sums to zero is the trivial one. □

This proof shows, by the way, that the components $T_{i_1 \ldots i_k}$ of a tensor T in terms of the basis tensors in $\mathcal{B}$ are given by (11.1).

It is useful to see explicitly what this proposition means for tensors of low rank.

- $k = 0$: $T^0(V)$ is just $\mathbb{R}$, so $\dim T^0(V) = 1 = n^0$.
- $k = 1$: $T^1(V) = V^*$ has dimension $n = n^1$.
- $k = 2$: $T^2(V)$ is the space of bilinear forms on V. Any bilinear form can be written uniquely as $T = T_{ij} \varepsilon^i \otimes \varepsilon^j$, where (T_{ij}) is an arbitrary $n \times n$ matrix. Thus $\dim T^2(V) = n^2$.

Abstract Tensor Products of Vector Spaces

Because every covariant k-tensor can be written as a linear combination of tensor products of covectors, it is suggestive to write

$$T^k(V) = V^* \otimes \cdots \otimes V^*,$$

where we think of the expression on the right-hand side as a shorthand for the set of all linear combinations of tensor products of elements of V^*.

We will now give a construction that makes sense of this notation in a much more general setting. The construction is a bit involved, but the idea is simple: Given vector spaces V and W, we will construct a vector space $V \otimes W$ that consists of linear combinations of objects of the form $v \otimes w$ for $v \in V$, $w \in W$, defined in such a way that $v \otimes w$ depends bilinearly on v and w.

Let S be a set. The *free vector space* on S, denoted $\mathbb{R}\langle S\rangle$, is the set of all finite formal linear combinations of elements of S with real coefficients. More precisely, a finite formal linear combination is a function $\mathcal{F}\colon S \to \mathbb{R}$ such that $\mathcal{F}(s) = 0$ for all but finitely many $s \in S$. Under pointwise addition and scalar multiplication, $\mathbb{R}\langle S\rangle$ becomes a real vector space. Identifying each element $x \in S$ with the function that takes the value 1 on x and zero on all other elements of S, any element $\mathcal{F} \in \mathbb{R}\langle S\rangle$ can be written uniquely in the form $\mathcal{F} = \sum_{i=1}^m a_i x_i$, where $x_1, \dots, x_m$ are the elements of S for which $\mathcal{F}(x_i) \neq 0$, and $a_i = \mathcal{F}(x_i)$. Thus S is a basis for $\mathbb{R}\langle S\rangle$, which is therefore finite-dimensional if and only if S is a finite set.

◇ **Exercise 11.2 (Characteristic Property of Free Vector Spaces).** Let S be a set and W a vector space. Show that any map $F\colon S \to W$ has a unique extension to a linear map $\overline{F}\colon \mathbb{R}\langle S\rangle \to W$.

Now let V and W be finite-dimensional real vector spaces, and let $\mathcal{R}$ be the subspace of the free vector space $\mathbb{R}\langle V \times W\rangle$ spanned by all elements of the following forms:

$$\begin{gathered} a(v,w) - (av,w), \\ a(v,w) - (v,aw), \\ (v,w) + (v',w) - (v+v',w), \\ (v,w) + (v,w') - (v,w+w'), \end{gathered} \tag{11.2}$$

for $a \in \mathbb{R}$, $v, v' \in V$, and $w, w' \in W$. Define the *tensor product* of V and W, denoted by $V \otimes W$, to be the quotient space $\mathbb{R}\langle V \times W\rangle/\mathcal{R}$. The equivalence class of an element (v,w) in $V \otimes W$ is denoted by $v \otimes w$, and is called the *tensor product* of v and w. From the definition, tensor products satisfy

$$\begin{gathered} a(v \otimes w) = av \otimes w = v \otimes aw, \\ v \otimes w + v' \otimes w = (v+v') \otimes w, \\ v \otimes w + v \otimes w' = v \otimes (w+w'). \end{gathered}$$

Note that the definition implies that every element of $V \otimes W$ can be written as a linear combination of elements of the form $v \otimes w$ for $v \in V$, $w \in W$; but it is not true in general that *every* element of $V \otimes W$ is of the form $v \otimes w$.

Proposition 11.3 (Characteristic Property of Tensor Products). *Let V and W be finite-dimensional real vector spaces. If $A\colon V \times W \to X$ is a bilinear map into any vector space X, there is a unique linear map $\widetilde{A}\colon V \otimes W \to X$ such that the following diagram commutes:*

$$\begin{tikzcd} V \times W \arrow[r, "A"] \arrow[d, "\pi"'] & X \\ V \otimes W, \arrow[ur, "\widetilde{A}"'] & \end{tikzcd} \tag{11.3}$$

where $\pi(v,w) = v \otimes w$.

Proof. First note that any map $A\colon V \times W \to X$ extends uniquely to a linear map $\overline{A}\colon \mathbb{R}\langle V \times W\rangle \to X$ by the characteristic property of the free vector space. This map is characterized by the fact that $\overline{A}(v,w) = A(v,w)$ whenever $(v,w) \in V \times W \subset \mathbb{R}\langle V \times W\rangle$. The fact that A is bilinear means precisely that the subspace $\mathcal{R}$ is contained in the kernel of $\overline{A}$, because

$$\begin{aligned} \overline{A}(av,w) &= A(av,w) \\ &= aA(v,w) \\ &= a\overline{A}(v,w) \\ &= \overline{A}(a(v,w)), \end{aligned}$$

with similar considerations for the other expressions in (11.2). Therefore, $\overline{A}$ descends to a linear map $\widetilde{A}\colon V \otimes W = \mathbb{R}\langle V \times W\rangle/\mathcal{R} \to X$ satisfying $\widetilde{A} \circ \Pi = \overline{A}$, where $\Pi\colon \mathbb{R}\langle V \times W\rangle \to V \otimes W$ is the natural projection. This is easily seen to be equivalent to $\widetilde{A} \circ \pi = A$, which is (11.3). Uniqueness follows from the fact that every element of $V \otimes W$ can be written as a linear combination of elements of the form $v \otimes w$, and $\widetilde{A}$ is uniquely determined on such elements by $\widetilde{A}(v \otimes w) = \overline{A}(v,w) = A(v,w)$. □

The reason this is called the characteristic property is that it uniquely characterizes the tensor product up to isomorphism; see Problem 11-1.

Proposition 11.4 (Other Properties of Tensor Products). *Let V, W, and X be finite-dimensional real vector spaces.*

(*a*) *The tensor product $V^* \otimes W^*$ is canonically isomorphic to the space $B(V,W)$ of bilinear maps from $V \times W$ into $\mathbb{R}$.*

(*b*) *If (E_i) is a basis for V and (F_j) is a basis for W, then the set of all elements of the form $E_i \otimes F_j$ is a basis for $V \otimes W$, which therefore has dimension equal to $(\dim V)(\dim W)$.*

(*c*) *There is a unique isomorphism $V \otimes (W \otimes X) \to (V \otimes W) \otimes X$ sending $v \otimes (w \otimes x)$ to $(v \otimes w) \otimes x$.*

Proof. The canonical isomorphism between $V^* \otimes W^*$ and $B(V,W)$ is constructed as follows. First, define a map $\Phi\colon V^* \times W^* \to B(V,W)$ by

$$\Phi(\omega,\eta)(v,w) = \omega(v)\eta(w).$$

It is easy to check that Φ is bilinear, so by the characteristic property it descends uniquely to a linear map $\widetilde{\Phi}\colon V^* \otimes W^* \to B(V,W)$.

To see that $\widetilde{\Phi}$ is an isomorphism, we will construct an inverse for it. Let (E_i) and (F_j) be any bases for V and W, respectively, with dual bases (ε^i) and (φ^j). Since $V^* \otimes W^*$ is spanned by elements of the form $\omega \otimes \eta$ for $\omega \in V^*$ and $\eta \in W^*$, every $\tau \in V^* \otimes W^*$ can be written in the form $\tau = \tau_{ij}\varepsilon^i \otimes \varphi^j$. (We are not claiming yet that this expression is unique.)

Define a map $\Psi\colon B(V,W) \to V^* \otimes W^*$ by setting

$$\Psi(b) = b(E_k,F_l)\varepsilon^k \otimes \varphi^l.$$

We will show that Ψ and $\widetilde{\Phi}$ are inverses. First, for $\tau = \tau_{ij}\varepsilon^i \otimes \varphi^j \in V^* \otimes W^*$,

$$\begin{aligned}
\Psi \circ \widetilde{\Phi}(\tau) &= \widetilde{\Phi}(\tau)(E_k,F_l)\varepsilon^k \otimes \varphi^l \\
&= \tau_{ij}\widetilde{\Phi}(\varepsilon^i \otimes \varphi^j)(E_k,F_l)\varepsilon^k \otimes \varphi^l \\
&= \tau_{ij}\Phi(\varepsilon^i, \varphi^j)(E_k,F_l)\varepsilon^k \otimes \varphi^l \\
&= \tau_{ij}\varepsilon^i(E_k)\varphi^j(F_l)\varepsilon^k \otimes \varphi^l \\
&= \tau_{ij}\varepsilon^i \otimes \varphi^j \\
&= \tau.
\end{aligned}$$

On the other hand, for $b \in B(V,W)$, $v \in V$, and $w \in W$,

$$\begin{aligned}
\widetilde{\Phi} \circ \Psi(b)(v,w) &= \widetilde{\Phi}\big(b(E_k,F_l)\varepsilon^k \otimes \varphi^l\big)(v,w) \\
&= b(E_k,F_l)\widetilde{\Phi}(\varepsilon^k \otimes \varphi^l)(v,w) \\
&= b(E_k,F_l)\varepsilon^k(v)\varphi^l(w) \\
&= b(E_k,F_l)v^k w^l \\
&= b(v,w).
\end{aligned}$$

Thus $\Psi = \widetilde{\Phi}^{-1}$. (Note that although we used bases to prove that $\widetilde{\Phi}$ is invertible, $\widetilde{\Phi}$ itself is canonically defined without reference to any basis.)

We have already observed above that the elements of the form $\varepsilon^i \otimes \varphi^j$ span $V^* \otimes W^*$. On the other hand, it is easy to check that $\dim B(V,W) = (\dim V)(\dim W)$ (because any bilinear form is uniquely determined by its action on pairs of basis elements), so for dimensional reasons the set $\{\varepsilon^i \otimes \varphi^j\}$ is a basis for $V^* \otimes W^*$. Applying this observation to $V = (V^*)^*$ and $W = (W^*)^*$ proves (b).

Finally, the isomorphism between $V \otimes (W \otimes X)$ and $(V \otimes W) \otimes X$ is constructed as follows. For each $x \in X$, the map $\alpha_x\colon V \times W \to V \otimes (W \otimes X)$ defined by

$$\alpha_x(v,w) = v \otimes (w \otimes x)$$

is obviously bilinear, and thus by the characteristic property of the tensor product it descends uniquely to a linear map $\widetilde{\alpha}_x\colon V\otimes W\to V\otimes(W\otimes X)$ satisfying $\widetilde{\alpha}_x(v\otimes w)=v\otimes(w\otimes x)$. Similarly, the map $\beta\colon (V\otimes W)\times X\to V\otimes(W\otimes X)$ given by

$$\beta(\tau,x)=\widetilde{\alpha}_x(\tau)$$

determines a linear map $\widetilde{\beta}\colon (V\otimes W)\otimes X\to V\otimes(W\otimes X)$ satisfying

$$\widetilde{\beta}((v\otimes w)\otimes x)=v\otimes(w\otimes x).$$

Because $V\otimes(W\otimes X)$ is spanned by elements of the form $v\otimes(w\otimes x)$, $\widetilde{\beta}$ is clearly surjective, and therefore it is an isomorphism for dimensional reasons. It is clearly the unique such isomorphism, because any other would have to agree with $\widetilde{\beta}$ on the set of elements of the form $(v\otimes w)\otimes x$, which spans $(V\otimes W)\otimes X$. □

The next corollary explains the relationship between this abstract tensor product of vector spaces and the more concrete covariant k-tensors that we defined earlier.

Corollary 11.5. *If V is a finite-dimensional real vector space, the space $T^k(V)$ of covariant k-tensors on V is canonically isomorphic to the k-fold tensor product $V^*\otimes\cdots\otimes V^*$.*

◇ **Exercise 11.3.** Prove Corollary 11.5.

Using these results, we can generalize the notion of covariant tensors on a vector space as follows. For any finite-dimensional real vector space V, define the space of *contravariant tensors* of rank k to be

$$T_k(V)=\underbrace{V\otimes\cdots\otimes V}_{k\text{ copies}}.$$

Because of the canonical identification $V=V^{**}$ and Corollary 11.5, an element of $T_k(V)$ can be thought of as a multilinear function from $V^*\times\cdots\times V^*$ into $\mathbb{R}$. In particular, $T_1(V)\cong V^{**}\cong V$, so elements of V are sometimes called "contravariant vectors."

More generally, for any nonnegative integers k,l, the space of *mixed tensors* on V of type $\binom{k}{l}$ is defined as

$$T^k_l(V)=\underbrace{V^*\otimes\cdots\otimes V^*}_{k\text{ copies}}\otimes\underbrace{V\otimes\cdots\otimes V}_{l\text{ copies}}.$$

From the discussion above, $T^k_l(V)$ can be identified with the set of real-valued multilinear functions of k vectors and l covectors.

In this book, we will be concerned primarily with covariant tensors, which we will think of primarily as multilinear functions of vectors, in keeping with our original definition. Thus tensors will always be understood to be covariant unless we explicitly specify otherwise. However, it is important to be aware that contravariant and mixed tensors play an important role in more advanced parts of differential geometry, especially Riemannian geometry.

Tensors and Tensor Fields on Manifolds

Now let M be a smooth manifold. We define the *bundle of covariant k-tensors* on M by

$$T^k M = \coprod_{p \in M} T^k(T_p M).$$

Similarly, we define the *bundle of contravariant l-tensors* by

$$T_l M = \coprod_{p \in M} T_l(T_p M),$$

and the *bundle of mixed tensors of type $\binom{k}{l}$* by

$$T^k_l M = \coprod_{p \in M} T^k_l(T_p M).$$

Clearly there are natural identifications

$$\begin{aligned}
T^0 M &= T_0 M = M \times \mathbb{R}, \\
T^1 M &= T^* M, \\
T_1 M &= TM, \\
T^k_0 M &= T^k M, \\
T^0_l M &= T_l M.
\end{aligned}$$

◇ **Exercise 11.4.** Show that $T^k M$, $T_l M$, and $T^k_l M$ have natural structures as smooth vector bundles over M, and determine their ranks.

Any one of these bundles is called a *tensor bundle* over M. (Thus the tangent and cotangent bundles are special cases of tensor bundles.) A section of a tensor bundle is called a (*covariant, contravariant, or mixed*) *tensor field* on M. A *smooth tensor field* is a section that is smooth in the usual sense of smooth sections of vector bundles. We denote the vector spaces of smooth sections of these bundles by

$$\begin{aligned}
\mathcal{T}^k(M) &= \{\text{smooth sections of } T^k M\}; \\
\mathcal{T}_l(M) &= \{\text{smooth sections of } T_l M\}; \\
\mathcal{T}^k_l(M) &= \{\text{smooth sections of } T^k_l M\}.
\end{aligned}$$

In any smooth local coordinates (x^i), sections of these bundles can be written (using the summation convention) as

$$\sigma = \begin{cases} \sigma_{i_1\dots i_k} dx^{i_1} \otimes \cdots \otimes dx^{i_k}, & \sigma \in \mathcal{T}^k(M); \\ \sigma^{j_1\dots j_l} \dfrac{\partial}{\partial x^{j_1}} \otimes \cdots \otimes \dfrac{\partial}{\partial x^{j_l}}, & \sigma \in \mathcal{T}_l(M); \\ \sigma^{j_1\dots j_l}_{i_1\dots i_k} dx^{i_1} \otimes \cdots \otimes dx^{i_k} \otimes \dfrac{\partial}{\partial x^{j_1}} \otimes \cdots \otimes \dfrac{\partial}{\partial x^{j_l}}, & \sigma \in \mathcal{T}^k_l(M). \end{cases}$$

The functions $\sigma_{i_1\dots i_k}$, $\sigma^{j_1\dots j_l}$, or $\sigma^{j_1\dots j_l}_{i_1\dots i_k}$ are called the *component functions* of σ in these coordinates.

Lemma 11.6. *Let M be a smooth manifold, and let $\sigma\colon M \to T^kM$ be a rough section. The following are equivalent.*

(*a*) *σ is smooth.*

(*b*) *In any smooth coordinate chart, the component functions of σ are smooth.*

(*c*) *If $X_1, \dots, X_k$ are smooth vector fields defined on any open subset $U \subset M$, then the function $\sigma(X_1, \dots, X_k)\colon U \to \mathbb{R}$, defined by*

$$\sigma(X_1, \dots, X_k)(p) = \sigma_p(X_1|_p, \dots, X_k|_p),$$

is smooth.

◇ **Exercise 11.5.** Prove Lemma 11.6.

◇ **Exercise 11.6.** Formulate and prove smoothness criteria analogous to those of Lemma 11.6 for contravariant and mixed tensor fields.

Covariant 1-tensor fields are just covector fields. Recalling that a 0-tensor is just a real number, a 0-tensor field is the same as a continuous real-valued function.

Lemma 11.7. *Let M be a smooth manifold, and suppose $\sigma \in \mathcal{T}^k(M)$, $\tau \in \mathcal{T}^l(M)$, and $f \in C^\infty(M)$. Then $f\sigma$ and $\sigma \otimes \tau$ are also smooth tensor fields, whose components in any smooth local coordinate chart are*

$$(f\sigma)_{i_1\dots i_k} = f\sigma_{i_1\dots i_k},$$
$$(\sigma \otimes \tau)_{i_1\dots i_{k+l}} = \sigma_{i_1\dots i_k}\tau_{i_{k+1}\dots i_{k+l}}.$$

◇ **Exercise 11.7.** Prove Lemma 11.7.

Pullbacks

Just like smooth covector fields, smooth covariant tensor fields can be pulled back by smooth maps to yield smooth tensor fields.

If $F\colon M \to N$ is a smooth map, for each integer $k \geq 0$ and each $p \in M$ we obtain a map $F^*\colon T^k(T_{F(p)}N) \to T^k(T_pM)$ called the *pullback* by

$$(F^*S)(X_1, \ldots, X_k) = S(F_*X_1, \ldots, F_*X_k).$$

Proposition 11.8 (Properties of Tensor Pullbacks). *Suppose $F\colon M \to N$ and $G\colon N \to P$ are smooth maps, $p \in M$, $S \in T^k(T_{F(p)}N)$, and $T \in T^l(T_{F(p)}N)$.*

(*a*) *$F^*\colon T^k(T_{F(p)}N) \to T^k(T_pM)$ is linear over $\mathbb{R}$.*

(*b*) *$F^*(S \otimes T) = F^*S \otimes F^*T$.*

(*c*) *$(G \circ F)^* = F^* \circ G^*\colon T^k(T_{G\circ F(p)}P) \to T^k(T_pM)$.*

(*d*) *$(\mathrm{Id}_N)^*S = S$.*

(*e*) *$F^*\colon T^kN \to T^kM$ is a smooth bundle map.*

◇ **Exercise 11.8.** Prove Proposition 11.9.

Observe that properties (c), (d), and (e) imply that the assignments $M \mapsto T^kM$ and $F \mapsto F^*$ yield a contravariant functor from the category of smooth manifolds and smooth maps to the category of smooth vector bundles and smooth bundle maps. Because of this, the convention of calling elements of T^kM covariant tensors is particularly unfortunate; but this terminology is so deeply entrenched that one has no choice but to go along with it.

Just as in the case of covector fields, the pullback operation extends to smooth tensor fields. Suppose as above that $F\colon M \to N$ is a smooth map. For any smooth covariant k-tensor field σ on N, we define a k-tensor field $F^*\sigma$ on M, called the pullback of σ by F, by

$$(F^*\sigma)_p = F^*(\sigma_{F(p)}).$$

(The F^* on the right-hand side is the pullback operator on *tensors* defined above.) This can be written more explicitly in terms of its action on tangent vectors: If $X_1, \ldots, X_k \in T_pM$, then

$$(F^*\sigma)_p(X_1, \ldots, X_k) = \sigma_{F(p)}(F_*X_1, \ldots, F_*X_k).$$

Proposition 11.9 (Properties of Tensor Field Pullbacks). *Suppose $F\colon M \to N$ and $G\colon N \to P$ are smooth maps, $\sigma \in \mathcal{T}^k(N)$, $\tau \in \mathcal{T}^l(N)$, and $f \in C^\infty(N)$.*

(*a*) *$F^*\sigma$ is a smooth tensor field.*

(*b*) *$F^*\colon \mathcal{T}^k(N) \to \mathcal{T}^k(M)$ is linear over $\mathbb{R}$.*

(*c*) *$F^*(f\sigma) = (f \circ F)F^*\sigma$.*

(*d*) *$F^*(\sigma \otimes \tau) = F^*\sigma \otimes F^*\tau$.*

(e) $(G \circ F)^* = F^* \circ G^*$.

(f) $(\mathrm{Id}_N)^*\sigma = \sigma$.

◇ **Exercise 11.9.** Prove Proposition 11.9.

If f is a smooth real-valued function (i.e., a smooth 0-tensor field) and σ is a smooth k-tensor field, then it is consistent with our definitions to interpret $f \otimes \sigma$ as $f\sigma$, and F^*f as $f \circ F$. With these interpretations, property (c) of this proposition is really just a special case of (d).

The following corollary is an immediate consequence of Proposition 11.9.

Corollary 11.10. *Let $F\colon M \to N$ be smooth, and let $\sigma \in \mathcal{T}^k(N)$. If $p \in M$ and (y^j) are smooth coordinates for N on a neighborhood of $F(p)$, then $F^*\sigma$ has the following expression near p:*

$$F^*(\sigma_{j_1 \dots j_k} dy^{j_1} \otimes \cdots \otimes dy^{j_k}) = (\sigma_{j_1 \dots j_k} \circ F) d(y^{j_1} \circ F) \otimes \cdots \otimes d(y^{j_k} \circ F).$$

Therefore $F^*\sigma$ is smooth.

In words, this corollary just says that $F^*\sigma$ is computed by the same technique we described in Chapter 6 for computing the pullback of a covector field: Wherever you see y^j in the expression for σ, just substitute the jth component function of F and expand. We will see examples of this in the next section.

In general, there is neither a push-forward nor a pullback operation for mixed tensor fields. However, in the special case of a diffeomorphism, tensor fields of any variance can be pushed forward and pulled back at will (see Problem 11-6).

Symmetric Tensors

Symmetric tensors—those whose values are unchanged by rearranging their arguments—play an extremely important role in differential geometry. We will describe only covariant symmetric tensors, but similar considerations apply to contravariant ones. (However, there is no useful notion of symmetry for mixed tensors.)

It is useful to start, as usual, in the linear algebraic setting. Let V be a finite-dimensional vector space. A covariant k-tensor T on V is said to be *symmetric* if its value is unchanged by interchanging any pair of arguments:

$$T(X_1, \dots, X_i, \dots, X_j, \dots, X_k) = T(X_1, \dots, X_j, \dots, X_i, \dots, X_k)$$

whenever $1 \le i < j \le k$.

◇ **Exercise 11.10.** Show that the following are equivalent for a covariant k-tensor T:

(a) T is symmetric.

(b) For any vectors $X_1, \ldots, X_k \in V$, the value of $T(X_1, \ldots, X_k)$ is unchanged when $X_1, \ldots, X_k$ are rearranged in any order.
(c) The components $T_{i_1 \ldots i_k}$ of T with respect to any basis are unchanged by any permutation of the indices.

We denote the set of symmetric covariant k-tensors on V by $\Sigma^k(V)$. It is obviously a vector subspace of $T^k(V)$. There is a natural projection $\operatorname{Sym}\colon T^k(V) \to \Sigma^k(V)$ called *symmetrization*, defined as follows. First, let S_k denote the *symmetric group on k elements*, that is, the group of permutations of the set $\{1, \ldots, k\}$. Given a k-tensor T and a permutation $\sigma \in S_k$, we define a new k-tensor ${}^\sigma T$ by

$${}^\sigma T(X_1, \ldots, X_k) = T(X_{\sigma(1)}, \ldots, X_{\sigma(k)}).$$

Note that ${}^\tau({}^\sigma T) = {}^{\tau\sigma}T$. (This is the reason for putting σ before T in the notation ${}^\sigma T$, instead of after it.) We define $\operatorname{Sym} T$ by

$$\operatorname{Sym} T = \frac{1}{k!} \sum_{\sigma \in S_k} {}^\sigma T.$$

More explicitly, this means

$$\operatorname{Sym} T(X_1, \ldots, X_k) = \frac{1}{k!} \sum_{\sigma \in S_k} T(X_{\sigma(1)}, \ldots, X_{\sigma(k)}).$$

Lemma 11.11 (Properties of Symmetrization).

(*a*) *For any covariant tensor T,* $\operatorname{Sym} T$ *is symmetric.*

(*b*) *T is symmetric if and only if* $\operatorname{Sym} T = T$*.*

Proof. Suppose $T \in T^k(V)$. If $\tau \in S_k$ is any permutation, then

$$\begin{aligned}
(\operatorname{Sym} T)(X_{\tau(1)}, \ldots, X_{\tau(k)}) &= \frac{1}{k!} \sum_{\sigma \in S_k} {}^\sigma T(X_{\tau(1)}, \ldots, X_{\tau(k)}) \\
&= \frac{1}{k!} \sum_{\sigma \in S_k} {}^{\tau\sigma} T(X_1, \ldots, X_k) \\
&= \frac{1}{k!} \sum_{\eta \in S_k} {}^{\eta} T(X_1, \ldots, X_k) \\
&= (\operatorname{Sym} T)(X_1, \ldots, X_k),
\end{aligned}$$

where we have substituted $\eta = \tau\sigma$ in the second-to-last line and used the fact that η runs over all of S_k as σ does. This shows that $\operatorname{Sym} T$ is symmetric.

If T is symmetric, then it follows from Exercise 11.10(b) that ${}^\sigma T = T$ for every $\sigma \in S_k$, so it follows immediately that $\operatorname{Sym} T = T$. On the other hand, if $\operatorname{Sym} T = T$, then T is symmetric because part (a) shows that $\operatorname{Sym} T$ is. $\square$

we will see below. Because of this ambiguity, we will usually use the term "distance function" when considering a metric in the metric space sense, and reserve "metric" for a Riemannian metric. In any event, which type of metric is being considered should always be clear from the context.

If g is a Riemannian metric on M, then for each $p \in M$, g_p is an inner product on T_pM. Because of this, we will often use the notation $\langle X, Y\rangle_g$ to denote the real number $g_p(X, Y)$ for $X, Y \in T_pM$.

In any smooth local coordinates (x^i), a Riemannian metric can be written

$$g = g_{ij}dx^i \otimes dx^j,$$

where g_{ij} is a symmetric positive definite matrix of smooth functions. Observe that the symmetry of g allows us to write g also in terms of symmetric products as follows:

$$\begin{aligned}
g &= g_{ij}dx^i \otimes dx^j \\
&= \tfrac{1}{2}(g_{ij}dx^i \otimes dx^j + g_{ji}dx^i \otimes dx^j) && \text{(since } g_{ij} = g_{ji}\text{)} \\
&= \tfrac{1}{2}(g_{ij}dx^i \otimes dx^j + g_{ij}dx^j \otimes dx^i) && \text{(switch } i \leftrightarrow j \text{ in the second term)} \\
&= g_{ij}dx^i dx^j && \text{(by Proposition 11.12(b)).}
\end{aligned}$$

Example 11.13. The simplest example of a Riemannian metric is the *Euclidean metric* $\overline{g}$ on $\mathbb{R}^n$, defined in standard coordinates by

$$\overline{g} = \delta_{ij}dx^i dx^j.$$

It is common to use the abbreviation ω^2 for the symmetric product of a tensor ω with itself, so the Euclidean metric can also be written

$$\overline{g} = (dx^1)^2 + \cdots + (dx^n)^2.$$

Applied to vectors $v, w \in T_p\mathbb{R}^n$, this yields

$$\overline{g}_p(v, w) = \delta_{ij}v^i w^j = \sum_{i=1}^n v^i w^i = v \cdot w.$$

In other words, $\overline{g}$ is the 2-tensor field whose value at each point is the Euclidean dot product. (As you may recall, we warned in Chapter 1 that expressions involving the Euclidean dot product are likely to violate our index conventions and therefore to require explicit summation signs. This can usually be avoided by writing the metric coefficients δ_{ij} explicitly, as in $\delta_{ij}v^i w^j$.)

To transform a Riemannian metric under a change of coordinates, we use the same technique as we used for covector fields: Think of the change of coordinates as the identity map expressed in terms of different coordinates for the domain and range, and use the formula of Corollary 11.10. As before, in practice this just amounts to substituting the formulas for one set of coordinates in terms of the other.

If S and T are symmetric tensors on V, then $S \otimes T$ is not symmetric in general. However, using the symmetrization operator, it is possible to define a new product that takes symmetric tensors to symmetric tensors. If $S \in \Sigma^k(V)$ and $T \in \Sigma^l(V)$, we define their *symmetric product* to be the $(k+l)$-tensor ST (denoted by juxtaposition with no intervening product symbol) given by

$$ST = \operatorname{Sym}(S \otimes T).$$

More explicitly, the action of ST on vectors $X_1, \dots, X_{k+l}$ is given by

$$\begin{aligned}&ST(X_1, \dots, X_{k+l}) \\ &\qquad = \frac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} S(X_{\sigma(1)}, \dots, X_{\sigma(k)}) T(X_{\sigma(k+1)}, \dots, X_{\sigma(k+l)}).\end{aligned}$$

Proposition 11.12 (Properties of the Symmetric Product).

(a) *The symmetric product is symmetric and bilinear: For all symmetric tensors R, S, T and all $a, b \in \mathbb{R}$,*

$$ST = TS,$$
$$(aR + bS)T = aRT + bST = T(aR + bS)$$

(b) *If ω and η are covectors, then*

$$\omega\eta = \tfrac{1}{2}(\omega \otimes \eta + \eta \otimes \omega).$$

◇ **Exercise 11.11.** Prove Proposition 11.12.

A *symmetric tensor field* on a manifold is simply a covariant tensor field whose value at any point is a symmetric tensor. The symmetric product of two or more tensor fields is defined pointwise, just like the tensor product.

Riemannian Metrics

The most important examples of symmetric tensors on a vector space are inner products. Any inner product allows us to define lengths of vectors and angles between them, and thus to do Euclidean geometry.

Transferring these ideas to manifolds, we obtain one of the most important applications of tensors to differential geometry. Let M be a smooth manifold. A *Riemannian metric* on M is a smooth symmetric 2-tensor field that is positive definite at each point. A *Riemannian manifold* is a pair (M, g), where M is a smooth manifold and g is a Riemannian metric on M. One sometimes simply says "M is a Riemannian manifold" if M is understood to be endowed with a specific Riemannian metric.

Note that a Riemannian metric is not the same thing as a metric in the sense of metric spaces, although the two concepts are closely related, as

Example 11.14. To illustrate, let us compute the coordinate expression for the Euclidean metric on $\mathbb{R}^2$ in polar coordinates. The Euclidean metric is $\overline{g} = dx^2 + dy^2$. (By convention, the notation dx^2 means the symmetric product $dx\,dx$, not $d(x^2)$). Substituting $x = r\cos\theta$ and $y = r\sin\theta$ and expanding, we obtain

$$\begin{aligned}\overline{g} &= dx^2 + dy^2\\ &= d(r\cos\theta)^2 + d(r\sin\theta)^2\\ &= (\cos\theta\,dr - r\sin\theta\,d\theta)^2 + (\sin\theta\,dr + r\cos\theta\,d\theta)^2\\ &= (\cos^2\theta + \sin^2\theta)dr^2 + (r^2\sin^2\theta + r^2\cos^2\theta)d\theta^2\\ &\quad + (-2r\cos\theta\sin\theta + 2r\sin\theta\cos\theta)dr\,d\theta\\ &= dr^2 + r^2\,d\theta^2.\end{aligned} \tag{11.4}$$

Below are just a few of the geometric constructions that can be defined on a Riemannian manifold (M, g).

- The *length* or *norm* of a tangent vector $X \in T_pM$ is defined to be

$$|X|_g = \langle X, X\rangle_g^{1/2} = g_p(X, X)^{1/2}.$$

- The *angle* between two nonzero tangent vectors $X, Y \in T_pM$ is the unique $\theta \in [0, \pi]$ satisfying

$$\cos\theta = \frac{\langle X, Y\rangle_g}{|X|_g\,|Y|_g}.$$

- Two tangent vectors $X, Y \in T_pM$ are said to be *orthogonal* if $\langle X, Y\rangle_g = 0$.

- If $\gamma\colon [a, b] \to M$ is a piecewise smooth curve segment, the *length* of γ is

$$L_g(\gamma) = \int_a^b |\gamma'(t)|_g\,dt.$$

  Because $|\gamma'(t)|_g$ is continuous at all but finitely many values of t, and has well-defined left- and right-handed limits at those points, the integral is well-defined.

◇ **Exercise 11.12.** If $\gamma\colon [a, b] \to M$ is a piecewise smooth curve segment and $a < c < b$, show that

$$L_g(\gamma) = L_g\left(\gamma|_{[a,c]}\right) + L_g\left(\gamma|_{[c,b]}\right).$$

It is an extremely important fact that length is independent of parametrization in the following sense. In Chapter 6, we defined a reparametrization of a piecewise smooth curve segment $\gamma\colon [a, b] \to M$ to be a curve segment of the form $\widetilde{\gamma} = \gamma \circ \varphi$, where $\varphi\colon [c, d] \to [a, b]$ is a diffeomorphism.

Proposition 11.15 (Parameter Independence of Length). *Let (M, g) be a Riemannian manifold, and let $\gamma\colon [a, b] \to M$ be a piecewise smooth curve segment. If $\widetilde{\gamma}$ is any reparametrization of γ, then $L_g(\widetilde{\gamma}) = L_g(\gamma)$.*

Proof. First suppose that γ is smooth, and $\varphi\colon [c, d] \to [a, b]$ is a diffeomorphism such that $\widetilde{\gamma} = \gamma \circ \varphi$. The fact that φ is a diffeomorphism implies that either $\varphi' > 0$ or $\varphi' < 0$ everywhere. Let us assume first that $\varphi' > 0$. We have

$$\begin{aligned}
L_g(\widetilde{\gamma}) &= \int_c^d |\widetilde{\gamma}'(t)|_g \, dt \\
&= \int_c^d \left| \frac{d}{dt}(\gamma \circ \varphi)(t) \right|_g \, dt \\
&= \int_c^d \left|\varphi'(t)\gamma'(\varphi(t))\right|_g \, dt \\
&= \int_c^d |\gamma'(\varphi(t))|_g \, \varphi'(t) \, dt \\
&= \int_a^b |\gamma'(s)|_g \, ds \\
&= L_g(\gamma),
\end{aligned}$$

where the second-to-last equality follows from the change of variables formula for ordinary integrals.

In case $\varphi' < 0$, we just need to introduce two sign changes into the above calculation. The sign changes once when $\varphi'(t)$ is moved outside the absolute value signs, because $|\varphi'(t)| = -\varphi'(t)$. Then it changes again in the last step, because φ reverses the direction of the integral. Since the two sign changes cancel each other, the result is the same.

If γ is only piecewise smooth, we just apply the same argument on each subinterval on which it is smooth. □

Suppose (M, g) and $(\widetilde{M}, \widetilde{g})$ are Riemannian manifolds. A smooth map $F\colon M \to \widetilde{M}$ is called an *isometry* if it is a diffeomorphism that satisfies $F^*\widetilde{g} = g$. If there exists an isometry between M and $\widetilde{M}$, we say that M and $\widetilde{M}$ are *isometric* as Riemannian manifolds. More generally, F is called a *local isometry* if every point $p \in M$ has a neighborhood U such that $F|_U$ is an isometry of U onto an open subset of $\widetilde{M}$. A metric g on M is said to be *flat* if every point $p \in M$ has a neighborhood $U \subset M$ such that $(U, g|_U)$ is isometric to an open subset of $\mathbb{R}^n$ with the Euclidean metric.

Riemannian geometry is the study of properties of Riemannian manifolds that are invariant under isometries. See, for example, [Lee97] for an introduction to some of its main ideas and techniques.

◇ **Exercise 11.13.** Show that lengths of curves are isometry invariants of Riemannian manifolds. More precisely, suppose (M, g) and $(\widetilde{M}, \widetilde{g})$ are Riemannian manifolds, and $F\colon M \to \widetilde{M}$ is an isometry. Show that $L_{\widetilde{g}}(F \circ \gamma) = L_g(\gamma)$ for any piecewise smooth curve segment γ in M.

Another extremely useful tool on Riemannian manifolds is orthonormal frames. Let (M, g) be an n-dimensional Riemannian manifold. Just as we did in Chapter 10 for $\mathbb{R}^n$ (see page 252), we define an *orthonormal frame* for M to be a local frame $(E_1, \dots, E_n)$ defined on some open subset $U \subset M$ such that $(E_1|_p, \dots, E_n|_p)$ is an orthonormal basis for T_pM at each point $p \in U$, or equivalently such that $\langle E_i, E_j \rangle_g = \delta_{ij}$.

Example 11.16. The coordinate frame $(\partial/\partial x^i)$ is a global orthonormal frame on $\mathbb{R}^n$.

Proposition 11.17 (Existence of Orthonormal Frames). *Let (M, g) be a Riemannian manifold. For any $p \in M$, there is a smooth orthonormal frame on a neighborhood of p.*

Proof. Let (x^i) be any smooth coordinates on a neighborhood U of p, and apply the Gram-Schmidt algorithm to the coordinate frame $(\partial/\partial x^i)$. The same argument as in the proof of Proposition 10.17 shows that this yields a smooth orthonormal frame on U. □

Observe that Proposition 11.17 does *not* show that there are smooth coordinates near p for which the *coordinate frame* is orthonormal. Problem 11-14 shows that there are such coordinates in a neighborhood of each point only if the metric is flat.

The Riemannian Distance Function

Using curve segments as "measuring tapes," we can define a notion of distance between points on a Riemannian manifold. If (M, g) is a connected Riemannian manifold and $p, q \in M$, the (*Riemannian*) *distance* between p and q, denoted by $d_g(p, q)$, is defined to be the infimum of $L_g(\gamma)$ over all piecewise smooth curve segments γ from p to q. Because any pair of points in a connected smooth manifold can be joined by a piecewise smooth curve segment (Lemma 6.17), this is well-defined.

Example 11.18. On $\mathbb{R}^n$ with the Euclidean metric $\overline{g}$, one can show that any straight line segment is the shortest piecewise smooth curve segment between its endpoints (Problem 11-16). Therefore, the distance function $d_{\overline{g}}$ is equal to the usual Euclidean distance:

$$d_{\overline{g}}(x, y) = |x - y|.$$

◇ **Exercise 11.14.** If (M, g) and $(\widetilde{M}, \widetilde{g})$ are connected Riemannian manifolds and $F\colon M \to \widetilde{M}$ is an isometry, show that $d_{\widetilde{g}}(F(p), F(q)) = d_g(p, q)$ for all $p, q \in M$.

We will see below that the Riemannian distance function turns M into a metric space whose topology is the same as the given manifold topology. The key is the following technical lemma, which shows that any Riemannian metric is locally comparable to the Euclidean metric in coordinates.

Lemma 11.19. *Let g be any Riemannian metric on an open set $U \subset \mathbb{R}^n$. For any compact subset $K \subset U$, there exist positive constants c, C such that for all $x \in K$ and all $v \in T_x\mathbb{R}^n$,*

$$c|v|_{\overline{g}} \leq |v|_g \leq C|v|_{\overline{g}}. \tag{11.5}$$

Proof. For any compact subset $K \subset U$, let $L \subset T\mathbb{R}^n$ be the set

$$L = \{(x, v) \in T\mathbb{R}^n : x \in K,\ |v|_{\overline{g}} = 1\}.$$

Under the canonical identification of $T\mathbb{R}^n$ with $\mathbb{R}^n \times \mathbb{R}^n$, L is just the product set $K \times \mathbb{S}^{n-1}$ and therefore is compact. Because the norm $|v|_g$ is continuous and strictly positive on L, there are positive constants c, C such that $c \leq |v|_g \leq C$ whenever $(x, v) \in L$. If $x \in K$ and v is any nonzero vector in $T_x\mathbb{R}^n$, let $\lambda = |v|_{\overline{g}}$. Then $(x, \lambda^{-1}v) \in L$, so by homogeneity of the norm,

$$|v|_g = \lambda|\lambda^{-1}v|_g \leq \lambda C = C|v|_{\overline{g}}.$$

A similar computation shows that $|v|_g \geq c|v|_{\overline{g}}$. The same inequalities are trivially true when $v = 0$. □

Proposition 11.20 (Riemannian Manifolds as Metric Spaces). *Let (M, g) be a connected Riemannian manifold. With the Riemannian distance function, M is a metric space whose metric topology is the same as the original manifold topology.*

Proof. It is immediate from the definition that $d_g(p, q) \geq 0$ for any $p, q \in M$. Because any constant curve segment has length zero, it follows that $d_g(p, p) = 0$, and $d_g(p, q) = d_g(q, p)$ follows from the fact that any curve segment from p to q can be reparametrized to go from q to p. Suppose γ_1 and γ_2 are piecewise smooth curve segments from p to q and q to r, respectively (Figure 11.1), and let γ be a piecewise smooth curve segment that first follows γ_1 and then follows γ_2 (reparametrized if necessary). Then

$$d_g(p, r) \leq L_g(\gamma) = L_g(\gamma_1) + L_g(\gamma_2).$$

Taking the infimum over all such γ_1 and γ_2, we find that $d_g(p, r) \leq d_g(p, q) + d_g(q, r)$. (This is one reason why it is important to define the distance function using piecewise smooth curves instead of just smooth ones.)

To complete the proof that (M, d_g) is a metric space, we need only show that $d_g(p, q) > 0$ if $p \neq q$. For this purpose, let $p, q \in M$ be distinct points,

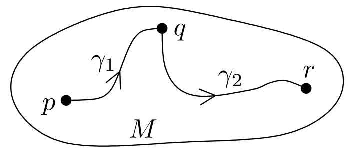


Figure 11.1. The triangle inequality.

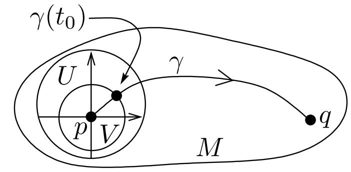


Figure 11.2. Positivity of d_g.

and let U be any smooth coordinate domain containing p but not q. Use the coordinate map as usual to identify U with an open subset in $\mathbb{R}^n$, and let $\overline{g}$ denote the Euclidean metric in these coordinates. If V is a smooth coordinate ball of radius ε centered at p such that $\overline{V} \subset U$, Lemma 11.19 shows that there are positive constants c, C such that

$$c|X|_{\overline{g}} \le |X|_g \le C|X|_{\overline{g}} \tag{11.6}$$

whenever $x \in \overline{V}$ and $X \in T_xM$. Then for any piecewise smooth curve segment γ lying entirely in $\overline{V}$, it follows that

$$cL_{\overline{g}}(\gamma) \le L_g(\gamma) \le CL_{\overline{g}}(\gamma).$$

Suppose $\gamma\colon [a,b] \to M$ is a piecewise smooth curve segment from p to q. Let t_0 be the infimum of all $t \in [a,b]$ such that $\gamma(t) \notin \overline{V}$. It follows that $\gamma(t_0) \in \partial V$ by continuity, and $\gamma(t) \in \overline{V}$ for $a \le t \le t_0$. Thus

$$L_g(\gamma) \ge L_g\left(\gamma|_{[a,t_0]}\right) \ge cL_{\overline{g}}\left(\gamma|_{[a,t_0]}\right) \ge cd_{\overline{g}}(p, \gamma(t_0)) = c\varepsilon.$$

Taking the infimum over all such γ, we conclude that $d_g(p,q) \ge c\varepsilon > 0$.

Finally, to show that the metric topology generated by d_g is the same as the given manifold topology on M, we will show that the open sets in the manifold topology are open in the metric topology and vice versa. Suppose first that $U \subset M$ is open in the manifold topology. Let p be any point of U, and let V be a smooth coordinate ball of radius ε around p such that $\overline{V} \subset U$ as above. The argument in the previous paragraph shows that $d_g(p,q) \ge c\varepsilon$ whenever $q \notin \overline{V}$. The contrapositive of this statement is that $d_g(p,q) < c\varepsilon$ implies $q \in \overline{V} \subset U$, or in other words the metric ball of radius $c\varepsilon$ around p is contained in U. This shows that U is open in the metric topology.

Conversely, suppose that W is open in the metric topology, let $p \in W$, and choose ε small enough that the closed metric ball of radius $C\varepsilon$ around p is contained in W. Let $\overline{V}$ be any closed smooth coordinate ball around p, let $\overline{g}$ be the Euclidean metric on $\overline{V}$ determined by the given coordinates, and let c, C be positive constants such that (11.6) is satisfied for $X \in T_qM$, $q \in \overline{V}$. For any $\varepsilon > 0$, let V_ε be the set of points whose Euclidean distance from p is less than ε. If $q \in V_\varepsilon$, let γ be the straight-line segment in coordinates from p to q. Arguing as above, (11.5) implies

$$d_g(p,q) \le L_g(\gamma) \le C L_{\overline{g}}(\gamma) = C\varepsilon.$$

This shows that V_ε is contained in the metric ball of radius $C\varepsilon$ around p, and therefore in W. Since V_ε is a neighborhood of p in the manifold topology, this shows that W is open in the manifold topology as well. □

Riemannian Submanifolds

If (M,g) is a Riemannian manifold and $S \subset M$ is an immersed submanifold, we can define a smooth symmetric 2-tensor $g|_S$ on S by $g|_S = \iota^* g$, where $\iota\colon S \hookrightarrow M$ is the inclusion map. By definition, this means for $X, Y \in T_p S$

$$(g|_S)(X,Y) = \iota^* g(X,Y) = g(\iota_* X, \iota_* Y) = g(X,Y),$$

so $g|_S$ is just the restriction of g to vectors tangent to S. Since the restriction of an inner product to a subspace is still positive definite, $g|_S$ is a Riemannian metric on S, called the *induced metric*. With this metric, S is called a *Riemannian submanifold* of M.

Example 11.21. The metric $\mathring{g} = \overline{g}|_{\mathbb{S}^n}$ induced on $\mathbb{S}^n$ from the Euclidean metric by the usual inclusion $\mathbb{S}^n \hookrightarrow \mathbb{R}^{n+1}$ is called the *round metric* (or the *standard metric*) on the sphere.

If S is a k-dimensional Riemannian submanifold of (M,g), it is usually easiest to compute the induced metric $g|_S$ in terms of a smooth local parametrization $X\colon U \to S$ (where U is an open subset of $\mathbb{R}^k$). The coordinate representation of $g|_S$ with respect to the coordinate chart X^{-1} is then the pullback metric $X^* g$. The next two examples will illustrate the procedure.

Example 11.22 (Riemannian Metrics in Graph Coordinates). Let $U \subset \mathbb{R}^n$ be an open set, and let $M \subset \mathbb{R}^{n+1}$ be the graph of the smooth function $f\colon U \to \mathbb{R}$. Then the map $X\colon U \to \mathbb{R}^{n+1}$ given by $X(u^1,\dots,u^n) = (u^1,\dots,u^n,f(u))$ is a smooth (global) parametrization of M (Figure 11.3), and the induced metric on M is given in graph coordinates by

$$\begin{aligned} X^*\overline{g} &= X^*((dx^1)^2 + \cdots + (dx^{n+1})^2) \\ &= (du^1)^2 + \cdots + (du^n)^2 + df^2. \end{aligned}$$

For example, the upper hemisphere of $\mathbb{S}^2$ is parametrized by the map $X\colon \mathbb{B}^2 \to \mathbb{R}^3$ given by

$$X(u,v) = (u, v, \sqrt{1-u^2-v^2}).$$

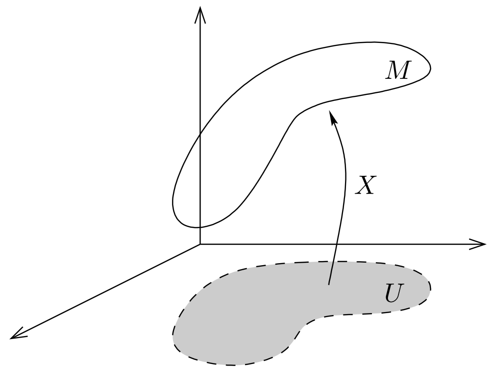


Figure 11.3. Graph coordinates.

In these coordinates, the round metric can be written

$$\begin{aligned}\mathring{g} = X^*\overline{g} &= du^2 + dv^2 + \left(\frac{u\,du + v\,dv}{\sqrt{1-u^2-v^2}}\right)^2 \\ &= \frac{(1-v^2)du^2 + (1-u^2)dv^2 + 2uv\,du\,dv}{1-u^2-v^2}.\end{aligned}$$

Example 11.23. Let $D \subset \mathbb{R}^3$ be the embedded torus obtained by revolving the circle $(y-2)^2 + z^2 = 1$ around the z-axis. If $X\colon \mathbb{R}^2 \to \mathbb{R}^3$ is the map

$$X(\varphi,\theta) = ((2+\cos\varphi)\cos\theta, (2+\cos\varphi)\sin\theta, \sin\varphi),$$

then the restriction of X to any sufficiently small open set $U \subset \mathbb{R}^2$ is a smooth local parametrization of D. The metric induced on D by the Euclidean metric is computed as follows:

$$\begin{aligned}X^*\overline{g} &= X^*(dx^2 + dy^2 + dz^2) \\ &= d((2+\cos\varphi)\cos\theta)^2 + d((2+\cos\varphi)\sin\theta)^2 + d(\sin\varphi)^2 \\ &= (-\sin\varphi\cos\theta\,d\varphi - (2+\cos\varphi)\sin\theta\,d\theta)^2 \\ &\quad + (-\sin\varphi\sin\theta\,d\varphi + (2+\cos\varphi)\cos\theta\,d\theta)^2 \\ &\quad + (\cos\varphi\,d\varphi)^2 \\ &= (\sin^2\varphi\cos^2\theta + \sin^2\varphi\sin^2\theta + \cos^2\varphi)d\varphi^2 \\ &\quad + ((2+\cos\varphi)\sin\varphi\cos\theta\sin\theta - (2+\cos\varphi)\sin\varphi\cos\theta\sin\theta)d\varphi\,d\theta \\ &\quad + ((2+\cos\varphi)^2\sin^2\theta + (2+\cos\varphi)^2\cos^2\theta)d\theta^2 \\ &= d\varphi^2 + (2+\cos\varphi)^2 d\theta^2.\end{aligned}$$

Suppose (M,g) is a Riemannian manifold and $S \subset M$ is a Riemannian submanifold. Just as for submanifolds of $\mathbb{R}^n$, for any $p \in S$, a vector $N \in$

T_pM is said to be *normal to* S if N is orthogonal to T_pS with respect to g. The set $N_pS \subset T_pM$ consisting of all vectors normal to S at p is a subspace of T_pM, called the *normal space* to S at p.

As in the Euclidean case, the most important tool for constructing normal vectors is adapted orthonormal frames. A local orthonormal frame $(E_1, \dots, E_n)$ for M on an open set $U \subset M$ is said to be *adapted to* S if the first k vectors $(E_1|_p, \dots, E_k|_p)$ span T_pS at each $p \in U \cap S$. It follows that $(E_{k+1}|_p, \dots, E_n|_p)$ span N_pS. The next proposition is proved in exactly the same way as its counterpart for submanifolds of $\mathbb{R}^n$ (Proposition 10.17).

Proposition 11.24 (Existence of Adapted Orthonormal Frames). *Let $S \subset M$ be an embedded Riemannian submanifold of a Riemannian manifold (M, g). For each $p \in S$, there is a smooth adapted orthonormal frame on a neighborhood U of p in M.*

◇ **Exercise 11.15.** Prove the preceding proposition.

If $S \subset M$ is a Riemannian submanifold, we define the *normal bundle* to S as

$$NS = \coprod_{p \in S} N_pS.$$

◇ **Exercise 11.16.** If $S \subset M$ is an embedded Riemannian submanifold, show that NS is a smooth vector bundle over S whose rank is equal to the codimension of S in M.

The Tangent-Cotangent Isomorphism

Another very important feature of any Riemannian metric is that it provides a natural isomorphism between the tangent and cotangent bundles. Given a Riemannian metric g on a manifold M, we define a bundle map $\widetilde{g} \colon TM \to T^*M$ as follows. For each $p \in M$ and each $X_p \in T_pM$, we let $\widetilde{g}(X_p) \in T_p^*M$ be the covector defined by

$$\widetilde{g}(X_p)(Y_p) = g(X_p, Y_p) \qquad \text{for all } Y_p \in T_pM,$$

To see that this is a smooth bundle map, it is easiest to consider its action on smooth vector fields:

$$\widetilde{g}(X)(Y) = g(X, Y) \qquad \text{for } X, Y \in \mathfrak{T}(M).$$

Because $\widetilde{g}(X)(Y)$ is linear over $C^\infty(M)$ as a function of Y, it follows from Problem 6-8 that $\widetilde{g}(X)$ is a smooth covector field; and because $\widetilde{g}(X)$ is linear over $C^\infty(M)$ as a function of X, this defines $\widetilde{g}$ as a smooth bundle map by Proposition 5.16. As usual, we use the same symbol for both the pointwise bundle map $\widetilde{g} \colon TM \to T^*M$ and the linear map on sections $\widetilde{g} \colon \mathfrak{T}(M) \to \mathfrak{T}^*(M)$.

Note that $\widetilde{g}$ is injective at each point, because $\widetilde{g}(X_p) = 0$ implies $0 = \widetilde{g}(X_p)(X_p) = \langle X_p, X_p\rangle_g$, which in turn implies $X_p = 0$. For dimensional reasons, therefore, $\widetilde{g}$ is bijective, and so it is a bundle isomorphism (see Problem 5-9).

If X and Y are smooth vector fields, in smooth coordinates we can write

$$\widetilde{g}(X)(Y) = g_{ij}X^iY^j,$$

which implies that the covector field $\widetilde{g}(X)$ has the coordinate expression

$$\widetilde{g}(X) = g_{ij}X^i dy^j.$$

In other words, $\widetilde{g}$ is the bundle map whose matrix with respect to coordinate frames for TM and T^*M is the same as the matrix of g itself.

It is customary to denote the components of the covector field $\widetilde{g}(X)$ by

$$X_j = g_{ij}X^i,$$

so that

$$\widetilde{g}(X) = X_j dy^j.$$

Because of this, one says that $\widetilde{g}(X)$ is obtained from X by *lowering an index*. The notation $X^\flat$ is frequently used for $\widetilde{g}(X)$, because the symbol $\flat$ ("flat") is used in musical notation to indicate that a tone is to be lowered.

The matrix of inverse map $\widetilde{g}^{-1}\colon T_p^*M \to T_pM$ is thus the inverse of (g_{ij}). (Because (g_{ij}) is the matrix of the isomorphism $\widetilde{g}$, it is invertible at each point.) We let (g^{ij}) denote the matrix-valued function whose value at $p \in M$ is the inverse of the matrix $(g_{ij}(p))$, so that

$$g^{ij}g_{jk} = g_{kj}g^{ji} = \delta^i_k.$$

Thus for a covector field $\omega \in \mathcal{T}^*M$, $\widetilde{g}^{-1}(\omega)$ has the coordinate representation

$$\widetilde{g}^{-1}(\omega) = \omega^i \frac{\partial}{\partial x^i}, \qquad \text{where } \omega^i = g^{ij}\omega_j.$$

We use the notation $\omega^\#$ ("ω-sharp") for $\widetilde{g}^{-1}(\omega)$, and say that $\omega^\#$ is obtained from ω by *raising an index*. A handy mnemonic device for keeping the flat and sharp operations straight is to remember that $\omega^\#$ is a vector, which we visualize as a (sharp) arrow; while $X^\flat$ is a covector, which we visualize by means of its (flat) level sets.

The most important use of the sharp operation is to reinstate the gradient as a vector field on Riemannian manifolds. For any smooth real-valued function f on a Riemannian manifold (M, g), we define a vector field called the *gradient* of f, by

$$\operatorname{grad} f = (df)^\# = \widetilde{g}^{-1}(df).$$

Unraveling the definitions, for any $X \in \mathcal{T}M$, it satisfies

$$\langle \operatorname{grad} f, X\rangle_g = \widetilde{g}(\operatorname{grad} f)(X) = df(X) = Xf.$$

Thus $\operatorname{grad} f$ is the unique vector field that satisfies

$$\langle \operatorname{grad} f, X\rangle_g = Xf \qquad \text{for every vector field } X,$$

or equivalently,

$$\langle \operatorname{grad} f, \cdot\rangle_g = df.$$

In smooth coordinates, $\operatorname{grad} f$ has the expression

$$\operatorname{grad} f = g^{ij}\frac{\partial f}{\partial x^i}\frac{\partial}{\partial x^j}.$$

In particular, this shows that $\operatorname{grad} f$ is smooth. On $\mathbb{R}^n$ with the Euclidean metric, this reduces to

$$\operatorname{grad} f = \delta^{ij}\frac{\partial f}{\partial x^i}\frac{\partial}{\partial x^j} = \sum_{i=1}^n \frac{\partial f}{\partial x^i}\frac{\partial}{\partial x^i}.$$

Thus our new definition of the gradient in this case coincides with the gradient from elementary calculus. In other coordinates, however, the gradient will not generally have the same form.

Example 11.25. Let us compute the gradient of a function $f \in C^\infty(\mathbb{R}^2)$ in polar coordinates. From (11.4), we see that the matrix of $\overline{g}$ in polar coordinates is $\left(\begin{smallmatrix}1 & 0\\ 0 & r^2\end{smallmatrix}\right)$, so its inverse matrix is $\left(\begin{smallmatrix}1 & 0\\ 0 & 1/r^2\end{smallmatrix}\right)$. Inserting this into the formula for the gradient, we obtain

$$\operatorname{grad} f = \frac{\partial f}{\partial r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial f}{\partial \theta}\frac{\partial}{\partial \theta}.$$

Existence of Riemannian Metrics

We conclude our discussion of Riemannian metrics by proving the following important result.

Proposition 11.26 (Existence of Riemannian Metrics). *Every smooth manifold admits a Riemannian metric.*

Proof. We give two proofs. For the first, we begin by covering M by smooth coordinate charts $(U_\alpha, \varphi_\alpha)$. In each coordinate domain, there is a Riemannian metric g_α given by the Euclidean metric $\delta_{ij}dx^i dx^j$ in coordinates. Now let $\{\psi_\alpha\}$ be a smooth partition of unity subordinate to the cover $\{U_\alpha\}$, and define

$$g = \sum_\alpha \psi_\alpha g_\alpha.$$

Because of the local finiteness condition for partitions of unity, there are only finitely many nonzero terms in a neighborhood of any point, so this expression defines a smooth tensor field. It is obviously symmetric, so only positivity needs to be checked. If $X \in T_pM$ is any nonzero vector, then

$$g_p(X,X) = \sum_{\alpha} \psi_\alpha(p) g_\alpha|_p(X,X).$$

This sum is nonnegative, because each term is nonnegative. At least one of the functions ψ_α is strictly positive at p (because they sum to 1). Because $g_\alpha|_p(X,X) > 0$, it follows that $g_p(X,X) > 0$.

The second proof is shorter, but relies on the Whitney embedding theorem, which is far less elementary. We simply embed M in $\mathbb{R}^N$ for some N, and then the Euclidean metric induces a Riemannian metric $\overline{g}|_M$ on M. □

It is worth remarking that because every Riemannian manifold is, in particular, a metric space, the first proof of this proposition yields another proof that smooth manifolds are metrizable, which does not depend on the Whitney embedding theorem.

Pseudo-Riemannian Metrics

An important generalization of Riemannian metrics is obtained by relaxing the requirement that the metric be positive definite. A 2-tensor g on a vector space V is said to be *nondegenerate* if $g(X,Y) = 0$ for all $Y \in V$ if and only if $X = 0$. Just as any inner product can be transformed to the Euclidean one by switching to an orthonormal basis, every nondegenerate symmetric 2-tensor can be transformed by a change of basis to one whose matrix is diagonal with all entries equal to ± 1. The numbers of positive and negative diagonal entries are independent of the choice of basis; thus the *signature* of g, defined as the sequence $(-1,\dots,-1,+1,\dots,+1)$ of diagonal entries in nondecreasing order, is an invariant of g.

A *pseudo-Riemannian metric* on a manifold M is a smooth symmetric 2-tensor field whose value is nondegenerate at each point. Pseudo-Riemannian metrics with signature $(-1,+1,\dots,+1)$ are called *Lorentz metrics*; they play a central role in physics, where they are used to model gravitation in Einstein's general theory of relativity.

We will not pursue the subject of pseudo-Riemannian metrics any further, except to note that neither of the proofs we gave of the existence of Riemannian metrics carries over to the pseudo-Riemannian case: In particular, it is not always true that the restriction of a nondegenerate 2-tensor to a subspace is nondegenerate, nor is it true that a linear combination of nondegenerate 2-tensors with positive coefficients is necessarily nondegenerate. Indeed, it is not true that every manifold admits a Lorentz metric.

Problems

11-1. Let V and W be finite-dimensional real vector spaces. Show that the tensor product $V \otimes W$ is uniquely determined up to canonical isomorphism by its characteristic property (Proposition 11.3). More precisely, suppose $\widetilde{\pi}\colon V \times W \to Z$ is a bilinear map into a vector space Z with following property: For any bilinear map $A\colon V \times W \to Y$, there is a unique linear map $\widetilde{A}\colon Z \to Y$ such that the following diagram commutes:

$$\begin{tikzcd} V \times W \arrow[r, "A"] \arrow[d, "\widetilde{\pi}"'] & Y \\ Z. \arrow[ur, "\widetilde{A}"'] & \end{tikzcd}$$

Then there is a unique isomorphism $\Phi\colon V \otimes W \to Z$ such that $\widetilde{\pi} = \Phi \circ \pi$, where $\pi\colon V \times W \to V \otimes W$ is the canonical projection. [Remark: This shows that the details of the construction used to define the tensor product are irrelevant, as long as the resulting space satisfies the characteristic property.]

11-2. If V is any finite-dimensional real vector space, prove that there are canonical isomorphisms $\mathbb{R} \otimes V \cong V \cong V \otimes \mathbb{R}$.

11-3. Let V and W be finite-dimensional real vector spaces. Prove that there is a canonical (basis-independent) isomorphism between $V^* \otimes W$ and the space $\operatorname{Hom}(V, W)$ of linear maps from V to W.

11-4. Let M be a smooth n-manifold, and let σ be a smooth covariant k-tensor field on M. If $(U, (x^i))$ and $(\widetilde{U}, (\widetilde{x}^j))$ are overlapping smooth charts on M, we can write

$$\sigma = \sigma_{i_1 \dots i_k} dx^{i_1} \otimes \cdots \otimes dx^{i_k} = \widetilde{\sigma}_{j_1 \dots j_k} d\widetilde{x}^{j_1} \otimes \cdots \otimes d\widetilde{x}^{j_k}.$$

Compute a transformation law analogous to (6.7) expressing the component functions $\sigma_{i_1 \dots i_k}$ in terms of $\widetilde{\sigma}_{j_1 \dots j_k}$.

11-5. Generalize the coordinate transformation law of Problem 11-4 to mixed tensors of any rank.

11-6. Suppose $F\colon M \to N$ is a diffeomorphism. For any pair of nonnegative integers k, l, show that there are smooth bundle isomorphisms $F_*\colon T^k_l M \to T^k_l N$ and $F^*\colon T^k_l N \to T^k_l M$ satisfying

$$\begin{aligned} &F_* S(X_1, \dots, X_k, \omega^1, \dots, \omega^l) \\ &\qquad = S(F_*^{-1} X_1, \dots, F_*^{-1} X_k, F^* \omega^1, \dots, F^* \omega^l), \\ &F^* S(X_1, \dots, X_k, \omega^1, \dots, \omega^l) \\ &\qquad = S(F_* X_1, \dots, F_* X_k, F^{-1*} \omega^1, \dots, F^{-1*} \omega^l). \end{aligned}$$

11-7. Let M be a smooth manifold.

(a) Given a smooth covariant k-tensor field $\tau \in \mathcal{T}^k(M)$, show that the map $\mathcal{T}(M) \times \cdots \times \mathcal{T}(M) \to C^\infty(M)$ defined by

$$(X_1, \dots, X_k) \mapsto \tau(X_1, \dots, X_k)$$

is multilinear over $C^\infty(M)$, in the sense that for any smooth functions $f, f' \in C^\infty(M)$ and smooth vector fields X_i, X_i',

$$\begin{aligned}\tau(X_1, &\dots, fX_i + f'X_i', \dots, X_k) \\ &= f\tau(X_1, \dots, X_i, \dots, X_k) + f'\tau(X_1, \dots, X_i', \dots, X_k).\end{aligned}$$

(b) Show that a map

$$\widetilde{\tau} \colon \mathcal{T}(M) \times \cdots \times \mathcal{T}(M) \to C^\infty(M)$$

is induced by a smooth tensor field as above if and only if it is multilinear over $C^\infty(M)$.

11-8. Let V be an n-dimensional real vector space. Show that

$$\dim \Sigma^k(V) = \binom{n+k-1}{k} = \frac{(n+k-1)!}{k!(n-1)!}.$$

11-9. (a) Let T be a covariant k-tensor on a finite-dimensional real vector space V. Show that $\operatorname{Sym} T$ is the unique symmetric k-tensor satisfying

$$(\operatorname{Sym} T)(X, \dots, X) = T(X, \dots, X)$$

for all $X \in V$.

(b) Show that the symmetric product is associative: For all symmetric tensors R, S, T,

$$(RS)T = R(ST).$$

[Hint: Use part (a).]

(c) If $\omega^1, \dots, \omega^k$ are covectors, show that

$$\omega^1 \cdots \omega^k = \frac{1}{k!} \sum_{\sigma \in S_k} \omega^{\sigma(1)} \otimes \cdots \otimes \omega^{\sigma(k)}.$$

11-10. Let $\mathring{g} = \overline{g}|_{\mathbb{S}^n}$ denote the round metric on the n-sphere, i.e., the metric induced from the Euclidean metric by the usual inclusion $\mathbb{S}^n \hookrightarrow \mathbb{R}^{n+1}$.

(a) Derive an expression for $\mathring{g}$ in stereographic coordinates by computing the pullback $(\sigma^{-1})^*\overline{g}$.

(b) In the case $n = 2$, do the analogous computation in spherical coordinates $(x, y, z) = (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi)$.

11-11. Let M be any smooth manifold.

(a) Show that TM and T^*M are isomorphic vector bundles.

(b) Show that the isomorphism of part (a) is *not* canonical, in the following sense: There does not exist a rule that assigns to every smooth manifold M a bundle isomorphism $\lambda_M : TM \to T^*M$ in such a way that for every smooth map $F : M \to N$, the following diagram commutes:

$$\begin{tikzcd} TM \arrow[r, "F_*"] \arrow[d, "\lambda_M"'] & TN \arrow[d, "\lambda_N"] \\ T^*M & T^*N. \arrow[l, "F^*"] \end{tikzcd}$$

11-12. This problem shows how to give a rigorous meaning to words like "natural" and "canonical" that are so often used informally in mathematics. Suppose C and D are categories, and $\mathcal{F}, \mathcal{G}$ are (covariant or contravariant) functors from C to D. A *natural transformation* λ from $\mathcal{F}$ to $\mathcal{G}$ is a rule that assigns to each object X of C a morphism $\lambda_X \in \operatorname{Hom}_{\mathsf{D}}(\mathcal{F}(X), \mathcal{G}(X))$, in such a way that for every pair of objects X, Y of C and every morphism $f \in \operatorname{Hom}_{\mathsf{C}}(X, Y)$, the following diagram commutes:

$$\begin{tikzcd} \mathcal{F}(X) \arrow[r, "\mathcal{F}(f)"] \arrow[d, "\lambda_X"'] & \mathcal{F}(Y) \arrow[d, "\lambda_Y"] \\ \mathcal{G}(X) \arrow[r, "\mathcal{G}(f)"'] & \mathcal{G}(Y). \end{tikzcd}$$

(If either $\mathcal{F}$ or $\mathcal{G}$ is contravariant, the corresponding horizontal arrow should be reversed.)

(a) Let $\mathsf{VECT}_{\mathbb{R}}$ denote the category of real vector spaces and linear maps, and let $\mathcal{D}$ be the contravariant functor from $\mathsf{VECT}_{\mathbb{R}}$ to itself that sends each vector space to its dual space and each linear map to its dual map. Show that the assignment $V \mapsto \xi_V$, where $\xi_V : V \to V^{**}$ is the map defined as in Chapter 6 by $\xi_V(X)\omega = \omega(X)$, is a natural transformation from the identity functor of $\mathsf{VECT}_{\mathbb{R}}$ to $\mathcal{D} \circ \mathcal{D}$.

(b) Show that there does not exist a natural transformation from the identity functor of $\mathsf{VECT}_{\mathbb{R}}$ to $\mathcal{D}$.

(c) Let SM and VB denote the categories of smooth manifolds and smooth vector bundles, respectively, and let $\mathcal{T}, \mathcal{T}^* : \mathsf{SM} \to \mathsf{VB}$ be the functors defined by

$$\begin{aligned} \mathcal{T}(M) &= TM, & \mathcal{T}(f) &= f_*; \\ \mathcal{T}^*(M) &= T^*M, & \mathcal{T}^*(f) &= f^*. \end{aligned}$$

Show that there does not exist a natural transformation from $\mathcal{T}$ to $\mathcal{T}^*$.

11-13. Let (M, g) and $(\widetilde{M}, \widetilde{g})$ be Riemannian manifolds. Suppose $F\colon M \to \widetilde{M}$ is a smooth map such that $F^*\widetilde{g} = g$. Show that F is an immersion.

11-14. Let (M, g) be a Riemannian manifold. Show that the following are equivalent:

(a) Each point of M has a smooth coordinate neighborhood in which the coordinate frame is orthonormal.
(b) g is flat.

11-15. Let $\mathbb{T}^n = \mathbb{S}^1 \times \cdots \times \mathbb{S}^1 \subset \mathbb{C}^n$, and let g be the metric on $\mathbb{T}^n$ induced from the Euclidean metric on $\mathbb{C}^n$ (identified with $\mathbb{R}^{2n}$). Show that g is flat.

11-16. Show that the shortest path between two points in Euclidean space is a straight line. More precisely, for $x, y \in \mathbb{R}^n$, let $\gamma\colon [0,1] \to \mathbb{R}^n$ be the curve segment

$$\gamma(t) = (1-t)x + ty,$$

and show that any other piecewise smooth curve segment $\widetilde{\gamma}$ from x to y satisfies $L_{\overline{g}}(\widetilde{\gamma}) \geq L_{\overline{g}}(\gamma)$. [Hint: First consider the case in which both x and y lie on the x^1-axis.]

11-17. Let $M = \mathbb{R}^2 \smallsetminus \{0\}$ with the Euclidean metric $\overline{g}$, and let $p = (1,0)$, $q = (-1,0)$. Show that there is no piecewise smooth curve segment γ from p to q in M such that $L_{\overline{g}}(\gamma) = d_{\overline{g}}(p,q)$.

11-18. Let (M, g) be a Riemannian manifold, and let $f \in C^\infty(M)$.

(a) For any $p \in M$, show that among all unit vectors $X \in T_pM$, the directional derivative Xf is greatest when X points in the same direction as $\operatorname{grad} f|_p$, and the length of $\operatorname{grad} f|_p$ is equal to the value of the directional derivative in that direction.
(b) If p is a regular point of f, show that $\operatorname{grad} f|_p$ is normal to the level set of f through p.

11-19. If $S \subset M$ is a regular level set of a smooth function $\Phi\colon M \to \mathbb{R}^k$, show that its normal bundle NS is trivial.

11-20. Let (M, g) be a Riemannian manifold, let Γ be a Lie group, and let $\theta\colon \Gamma \times M \to M$ be a group action. We say that Γ *acts by isometries* if for each $g \in \Gamma$, the map $\theta_g\colon M \to M$ is a Riemannian isometry, and Γ acts *discontinuously* if no Γ-orbit has a limit point in M. If Γ acts freely, smoothly, and discontinuously on M by isometries, show that the quotient map $M \to M/\Gamma$ is a smooth covering map.

11-21. Let Γ be a discrete group acting smoothly, freely, and properly on a smooth manifold $\widetilde{M}$, and let $M = \widetilde{M}/\Gamma$. Show that a Riemannian metric $\widetilde{g}$ on $\widetilde{M}$ is the pullback of a metric on M by the quotient map $\pi\colon \widetilde{M} \to M$ if and only Γ acts by isometries.

11-22. In our first proof of Proposition 11.26, we used a partition of unity to paste together locally defined Riemannian metrics to obtain a global one. A crucial part of the proof was verifying that the global tensor field so obtained was positive definite. The key to the success of this argument is the fact that the set of inner products on each tangent space is a convex subset of the vector space of all symmetric 2-tensors. This problem outlines a generalization of this construction to arbitrary vector bundles. Suppose that E is a smooth vector bundle over a smooth manifold M, and $V \subset E$ is an open set with the property that for each $p \in M$, the intersection of V with the fiber E_p is convex and nonempty. By a "section of V," we will mean a (local or global) section of E whose image lies in V.

(a) Show that there exists a smooth global section of V.

(b) Suppose $\sigma\colon A \to V$ is a smooth section of V defined on a closed subset $A \subset M$. (This means that σ extends to a smooth section of V in a neighborhood of each point of A). For any open set U containing A, show that there exists a smooth global section of V whose restriction to A is equal to σ and whose support is contained in U.

12
Differential Forms

In the previous chapter, we introduced symmetric tensors—those whose values are unchanged by interchanging any pair of arguments. In this chapter, we explore the complementary notion of alternating tensors, whose values change sign whenever two arguments are interchanged. The main focus of the chapter is differential forms, which are just alternating tensor fields. These innocent-sounding objects play an unexpectedly important role in smooth manifold theory, through two applications. First, as we will see in Chapter 14, they are the objects that can be integrated in a coordinate-independent way over manifolds or submanifolds; second, as we will see in Chapter 15, they provide a link between analysis and topology by way of de Rham cohomology.

We begin the chapter with a heuristic discussion of the measurement of volume, to motivate the central role played by alternating tensors. We then proceed to study the algebra of alternating tensors. The most important algebraic construction is a product operation called the wedge product, which takes alternating tensors to alternating tensors. Then we transfer this to manifolds, and introduce the exterior derivative, which is a natural differential operator on forms.

At the end of the chapter, we introduce symplectic forms, which are a particular type of differential form that play an important role in geometry, analysis, and mathematical physics.

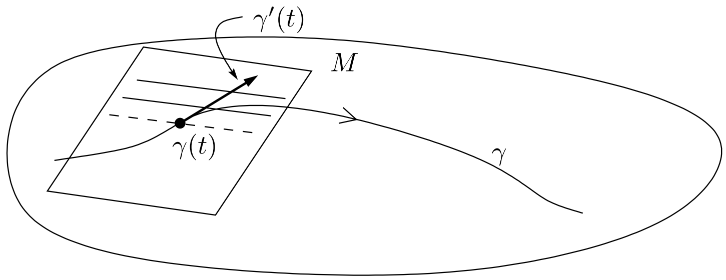


Figure 12.1. A covector field as a "signed length meter."

The Geometry of Volume Measurement

In Chapter 6, we introduced line integrals of covector fields, which generalize ordinary integrals to curves in manifolds. As we will see in subsequent chapters, it is also useful to generalize the theory of *multiple* integrals to manifolds.

How might we make coordinate-independent sense of multiple integrals? First, observe that there is no way to define integrals of real-valued *functions* in a coordinate-independent way on a manifold. It is easy to see why, even in the simplest possible case: Suppose $C \subset \mathbb{R}^n$ is an n-dimensional cube, and $f\colon C \to \mathbb{R}$ is the constant function $f(x) \equiv 1$. Then

$$\int_C f\, dV = \mathrm{Vol}(C),$$

which is clearly not invariant under coordinate transformations, even if we just restrict attention to linear ones.

Let us think a bit more geometrically about why covector fields are the natural fields to integrate along curves. A covector field on a manifold M assigns a number to each tangent vector, in such a way that multiplying the tangent vector by a constant has the effect of multiplying the resulting number by the same constant. Thus a covector field can be thought of as assigning a "signed length meter" to each 1-dimensional subspace of each tangent space (Figure 12.1), and it does so in a coordinate-independent way. Computing the line integral of a covector field, in effect, assigns a "length" to a curve by using this varying measuring scale along the points of the curve.

Now we wish to seek a kind of "field" that can be integrated in a coordinate-independent way over submanifolds of dimension $k > 1$. Its value at each point should be something that we can interpret as a "signed volume meter" on k-dimensional subspaces of the tangent space—a machine Ω that accepts any k tangent vectors $(X_1, \dots, X_k)$ at a point and returns a number $\Omega(X_1, \dots, X_k)$ that we might think of as the "signed vol-

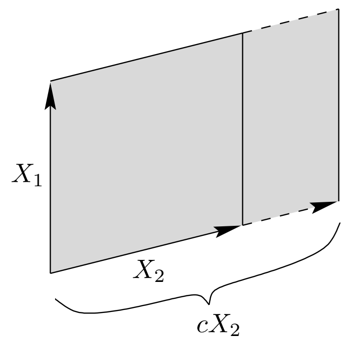


Figure 12.2. Scaling by a constant.

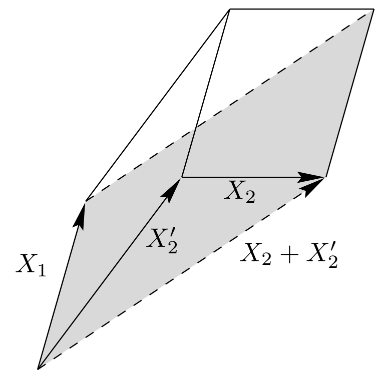


Figure 12.3. Sum of two vectors.

ume" of the parallelepiped spanned by those vectors, measured according to a scale determined by Ω.

The most obvious example of such a machine is the determinant in $\mathbb{R}^n$. For example, it is shown in most linear algebra texts that for any two vectors $X_1, X_2 \in \mathbb{R}^2$, $\det(X_1, X_2)$ is, up to a sign, the area of the parallelogram spanned by X_1, X_2. It is not hard to show (see Problem 12-1) that the analogous fact is true in all dimensions. The determinant, remember, is an example of a tensor. In fact, it is a tensor of a very specific type: It changes sign whenever two of its arguments are interchanged. A covariant k-tensor T on a finite-dimensional vector space V is said to be *alternating* if it has this property:

$$T(X_1, \dots, X_i, \dots, X_j, \dots, X_k) = -T(X_1, \dots, X_j, \dots, X_i, \dots, X_k).$$

An alternating k-tensor is sometimes called a *k-covector* or a *multicovector*.

Let us consider what properties we might expect a general "signed volume meter" Ω to have. To be consistent with our ordinary ideas of volume, we would expect that multiplying any one of the vectors by a constant c should cause the volume to be scaled by that same constant (Figure 12.2), and that the parallelepiped formed by adding together two vectors in the ith place results in a volume that is the sum of the volumes of the two parallelepipeds with the original vectors in the ith place (Figure 12.3):

$$\begin{aligned}\Omega(X_1, \dots, cX_i, \dots, X_n) &= c\Omega(X_1, \dots, X_i, \dots, X_n),\\ \Omega(X_1, \dots, X_i + X_i', \dots, X_n) &= \Omega(X_1, \dots, X_i, \dots, X_n)\\ &\quad + \Omega(X_1, \dots, X_i', \dots, X_n).\end{aligned}$$

These two requirements suggest that Ω should be multilinear, and thus should be a covariant k-tensor.

There is one more essential property that we should expect: Since n linearly dependent vectors span a parallelepiped of zero n-dimensional volume,

Ω should give the value zero whenever it is applied to n linearly dependent vectors. As the next lemma shows, this forces Ω to be an alternating tensor.

Lemma 12.1. *Suppose Ω is a k-tensor on a vector space V with the property that $\Omega(X_1, \ldots, X_k) = 0$ whenever $X_1, \ldots, X_k$ are linearly dependent. Then Ω is alternating.*

Proof. The hypothesis implies, in particular, that Ω gives the value zero whenever two of its arguments are the same. This in turn implies

$$\begin{aligned} 0 &= \Omega(X_1, \ldots, X_i + X_j, \ldots, X_i + X_j, \ldots, X_n) \\ &= \Omega(X_1, \ldots, X_i, \ldots, X_i, \ldots, X_n) + \Omega(X_1, \ldots, X_i, \ldots, X_j, \ldots, X_n) \\ &\quad + \Omega(X_1, \ldots, X_j, \ldots, X_i, \ldots, X_n) + \Omega(X_1, \ldots, X_j, \ldots, X_j, \ldots, X_n) \\ &= \Omega(X_1, \ldots, X_i, \ldots, X_j, \ldots, X_n) + \Omega(X_1, \ldots, X_j, \ldots, X_i, \ldots, X_n). \end{aligned}$$

Thus Ω is alternating. □

Because of these considerations, alternating tensor fields are promising candidates for objects that can be integrated in a coordinate-independent way. We will develop these ideas rigorously in the remainder of this chapter and the next; as we do, you should keep this geometric motivation in mind.

The Algebra of Alternating Tensors

In this section, we set aside heuristics and start developing the technical machinery for working with alternating tensors. For any finite-dimensional real vector space V, let $\Lambda^k(V)$ denote the subspace of $T^k(V)$ consisting of alternating tensors (k-covectors). (Warning: Some authors use the notation $\Lambda^k(V^*)$ in place of $\Lambda^k(V)$ for this space; see Problem 12-8 for a discussion of the reasons why.)

Recall that for any permutation $\sigma \in S_k$, the *sign* of σ, denoted by $\operatorname{sgn} \sigma$, is equal to $+1$ if σ is even (i.e., can be written as a composition of an even number of transpositions), and -1 if σ is odd.

The following exercise is an analogue of Exercise 11.10.

◇ **Exercise 12.1.** Show that the following are equivalent for a covariant k-tensor T:

(a) T is alternating.
(b) For any vectors $X_1, \ldots, X_k$ and any permutation $\sigma \in S_k$,
$$T(X_{\sigma(1)}, \ldots, X_{\sigma(k)}) = (\operatorname{sgn} \sigma) T(X_1, \ldots, X_k).$$
(c) T gives zero whenever two of its arguments are equal:
$$T(X_1, \ldots, Y, \ldots, Y, \ldots, X_k) = 0.$$
(d) $T(X_1, \ldots, X_k) = 0$ whenever the vectors $(X_1, \ldots, X_k)$ are linearly dependent.

(e) With respect to any basis, the components $T_{i_1 \dots i_k}$ of T change sign whenever two indices are interchanged.

Notice that part (d) implies that there are no nonzero alternating k-tensors on V if $k > \dim V$, for then every k-tuple of vectors is dependent.

Every 0-tensor (which is just a real number) is alternating, because there are no arguments to interchange. Similarly, every 1-tensor is alternating. An alternating 2-tensor is just a skew-symmetric bilinear form on V. It is interesting to note that any 2-tensor T can be expressed as the sum of an alternating tensor and a symmetric one, because

$$\begin{aligned} T(X,Y) &= \tfrac{1}{2}(T(X,Y) - T(Y,X)) + \tfrac{1}{2}(T(X,Y) + T(Y,X)) \\ &= A(X,Y) + S(X,Y), \end{aligned}$$

where $A(X,Y) = \frac{1}{2}(T(X,Y) - T(Y,X))$ is alternating, and $S(X,Y) = \frac{1}{2}(T(X,Y) + T(Y,X))$ is symmetric. This is not true for tensors of higher rank, as Problem 12-2 shows.

The tensor S defined above is just $\operatorname{Sym} T$, the symmetrization of T defined in the preceding chapter. We define a similar projection $\operatorname{Alt}\colon T^k(V) \to \Lambda^k(V)$, called the *alternating projection*, as follows:

$$\operatorname{Alt} T = \frac{1}{k!} \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma)({}^{\sigma}T).$$

More explicitly, this means

$$(\operatorname{Alt} T)(X_1, \dots, X_k) = \frac{1}{k!} \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma) T(X_{\sigma(1)}, \dots, X_{\sigma(k)}).$$

Example 12.2. If T is any 1-tensor, then $\operatorname{Alt} T = T$. If T is a 2-tensor, then

$$\operatorname{Alt} T(X,Y) = \tfrac{1}{2}(T(X,Y) - T(Y,X)).$$

For a 3-tensor T,

$$\begin{aligned} \operatorname{Alt} T(X,Y,Z) = \tfrac{1}{6}(&T(X,Y,Z) + T(Y,Z,X) + T(Z,X,Y) \\ &- T(Y,X,Z) - T(X,Z,Y) - T(Z,Y,X)). \end{aligned}$$

The next lemma is the analogue of Lemma 11.11.

Lemma 12.3 (Properties of the Alternating Projection).

(*a*) *For any tensor T,* $\operatorname{Alt} T$ *is alternating.*

(*b*) *T is alternating if and only if* $\operatorname{Alt} T = T$.

◇ **Exercise 12.2.** Prove Lemma 12.3.

Elementary Alternating Tensors

Let k be a positive integer. An ordered k-tuple $I = (i_1, \dots, i_k)$ of positive integers is called a *multi-index* of length k. If I is such a multi-index and $\sigma \in S_k$ is a permutation, we write I_σ for the multi-index

$$I_\sigma = (i_{\sigma(1)}, \dots, i_{\sigma(k)}).$$

Note that $I_{\sigma\tau} = (I_\sigma)_\tau$ for $\sigma, \tau \in S_k$. It is useful to extend the Kronecker delta notation in the following way. If I and J are multi-indices of length k, we define

$$\delta_I^J = \begin{cases} \operatorname{sgn}\sigma & \text{if neither } I \text{ nor } J \text{ has a repeated index} \\ & \qquad \text{and } J = I_\sigma \text{ for some } \sigma \in S_k, \\ 0 & \text{if } I \text{ or } J \text{ has a repeated index} \\ & \qquad \text{or } J \text{ is not a permutation of } I. \end{cases}$$

Let V be an n-dimensional vector space, and suppose $(\varepsilon^1, \dots, \varepsilon^n)$ is any basis for V^*. We will define a collection of alternating tensors on V that generalize the determinant function on $\mathbb{R}^n$. For each multi-index $I = (i_1, \dots, i_k)$ of length k such that $1 \le i_1, \dots, i_k \le n$, define a covariant k-tensor ε^I by

$$\begin{aligned} \varepsilon^I(X_1, \dots, X_k) &= \det \begin{pmatrix} \varepsilon^{i_1}(X_1) & \dots & \varepsilon^{i_1}(X_k) \\ \vdots & & \vdots \\ \varepsilon^{i_k}(X_1) & \dots & \varepsilon^{i_k}(X_k) \end{pmatrix} \\ &= \det \begin{pmatrix} X_1^{i_1} & \dots & X_k^{i_1} \\ \vdots & & \vdots \\ X_1^{i_k} & \dots & X_k^{i_k} \end{pmatrix}. \end{aligned} \tag{12.1}$$

In other words, if $\mathbb{X}$ denotes the matrix whose columns are the components of the vectors $X_1, \dots, X_k$ with respect to the basis (E_i) dual to (ε^i), then $\varepsilon^I(X_1, \dots, X_k)$ is the determinant of the $k \times k$ minor consisting of rows $i_1, \dots, i_k$ of $\mathbb{X}$. Because the determinant changes sign whenever two columns are interchanged, it is clear that ε^I is an alternating k-tensor. We will call ε^I an *elementary alternating tensor* or *elementary k-covector*.

For example, in terms of the standard dual basis (e^1, e^2, e^3) for $(\mathbb{R}^3)^*$, we have

$$\begin{aligned} e^{13}(X, Y) &= X^1 Y^3 - Y^1 X^3; \\ e^{123}(X, Y, Z) &= \det(X, Y, Z). \end{aligned}$$

Lemma 12.4. *Let (E_i) be a basis for V, let (ε^i) be the dual basis for V^*, and let ε^I be as defined above.*

(*a*) *If I has a repeated index, then $\varepsilon^I = 0$.*

(*b*) *If $J = I_\sigma$ for some $\sigma \in S_k$, then $\varepsilon^I = (\operatorname{sgn}\sigma)\varepsilon^J$.*

(*c*) *The result of evaluating* ε^I *on a sequence of basis vectors is*

$$\varepsilon^I(E_{j_1}, \dots, E_{j_k}) = \delta^I_J.$$

Proof. If I has a repeated index, then for any vectors $X_1, \dots, X_k$, the determinant in (12.1) has two identical rows and thus is equal to zero, which proves (a). On the other hand, if J is obtained from I by interchanging two indices, then the corresponding determinants have opposite signs; this implies (b).

To prove (c), we consider several cases. First, if I has a repeated index, then $\varepsilon^I = \delta^I_J = 0$ by part (a). If J has a repeated index, then $\varepsilon^I(E_{j_1}, \dots, E_{j_k}) = 0$ by Exercise 12.1(c). If neither multi-index has any repeated indices but J is not a permutation of I, then the determinant in the definition of $\varepsilon^I(E_{j_1}, \dots, E_{j_k})$ has at least one row of zeros, so it is zero. If $J = I$, then $\varepsilon^I(E_{j_1}, \dots, E_{j_k})$ is the determinant of the identity matrix, which is 1. Therefore, if $J = I_\sigma$, then

$$\varepsilon^I(E_{j_1}, \dots, E_{j_k}) = (\operatorname{sgn}\sigma)\varepsilon^J(E_{j_1}, \dots, E_{j_k}) = \operatorname{sgn}\sigma = \delta^I_J$$

by part (b). □

The significance of the elementary k-covectors is that they provide a convenient basis for $\Lambda^k(V)$. Of course, the ε^Is are not all independent, because some of them are zero and the ones corresponding to different permutations of the same multi-index are constant multiples of each other. But, as the next proposition shows, we can get a basis by restricting attention to an appropriate subset of multi-indices. A multi-index $I = (i_1, \dots, i_k)$ is said to be *increasing* if $i_1 < \dots < i_k$. It will be useful to use a primed summation sign to denote a sum over only increasing multi-indices, so that, for example,

$$\sum_I{}' T_I \varepsilon^I = \sum_{\{I : 1 \le i_1 < \dots < i_k \le n\}} T_I \varepsilon^I.$$

Proposition 12.5. *Let V be an n-dimensional vector space. If (ε^i) is any basis for V^*, then for each positive integer $k \le n$, the collection of k-covectors*

$$\mathcal{E} = \{\varepsilon^I : I \textit{ is an increasing multi-index of length } k\}$$

is a basis for $\Lambda^k(V)$. Therefore,

$$\dim \Lambda^k(V) = \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

If $k > n$, then $\dim \Lambda^k(V) = 0$.

Proof. The fact that $\Lambda^k(V)$ is the trivial vector space when $k > n$ follows immediately from Exercise 12.1(d), since any k vectors are dependent in that case. For the case $k \le n$, we need to show that the set $\mathcal{E}$ spans $\Lambda^k(V)$ and is independent. Let (E_i) be the basis for V dual to (ε^i).

To show that $\mathcal{E}$ spans, let $T \in \Lambda^k(V)$ be arbitrary. For each multi-index $I = (i_1, \dots, i_k)$, define a real number T_I by

$$T_I = T(E_{i_1}, \dots, E_{i_k}).$$

The fact that T is alternating implies that $T_I = 0$ if I contains a repeated multi-index, and $T_J = (\operatorname{sgn} \sigma) T_I$ if $J = I_\sigma$ for $\sigma \in S_k$. For any multi-index J, Lemma 12.4 gives

$$\sum_I{}' T_I \varepsilon^I(E_{j_1}, \dots, E_{j_k}) = \sum_I{}' T_I \delta^I_J = T_J = T(E_{j_1}, \dots, E_{j_k}).$$

Therefore, $\sum_I' T_I \varepsilon^I = T$, so $\mathcal{E}$ spans $\Lambda^k(V)$.

To show that $\mathcal{E}$ is an independent set, suppose

$$\sum_I{}' T_I \varepsilon^I = 0$$

for some coefficients T_I. Let J be any increasing multi-index. Applying both sides to $(E_{j_1}, \dots, E_{j_k})$ and using Lemma 12.4,

$$0 = \sum_I{}' T_I \varepsilon^I(E_{j_1}, \dots, E_{j_k}) = T_J.$$

Thus each coefficient T_J is zero. □

In particular, for an n-dimensional vector space V, this proposition implies that $\Lambda^n(V)$ is 1-dimensional, and is spanned by $\varepsilon^{1\dots n}$. By definition, this elementary n-covector acts on vectors $(X_1, \dots, X_n)$ by taking the determinant of their component matrix $\mathbb{X} = (X^i_j)$. For example, on $\mathbb{R}^n$ with the standard basis, $e^{1\dots n}$ is precisely the determinant function.

One consequence of this is the following useful description of the behavior of an n-covector under linear maps. Recall that if $T\colon V \to V$ is a linear map, the determinant of T is defined to be the determinant of the matrix representation of T with respect to any basis (see the Appendix, page 573).

Lemma 12.6. *Suppose V is an n-dimensional vector space and $\omega \in \Lambda^n(V)$. If $T\colon V \to V$ is any linear map and $X_1, \dots, X_n$ are arbitrary vectors in V, then*

$$\omega(TX_1, \dots, TX_n) = (\det T)\omega(X_1, \dots, X_n). \tag{12.2}$$

Proof. Let (E_i) be any basis for T, and let (ε^i) be the dual basis. Let (T^j_i) denote the matrix of T with respect to this basis, and let $T_i = TE_i = T^j_i E_j$. By Proposition 12.5, we can write $\omega = c\varepsilon^{1\dots n}$ for some real number c.

Since both sides of (12.2) are multilinear functions of $X_1, \dots, X_n$, it suffices to verify it in the special case $X_i = E_i$, $i = 1, \dots, n$. In this case, the right-hand side of (12.2) is

$$(\det T)c\varepsilon^{1\dots n}(E_1, \dots, E_n) = c \det T.$$

On the other hand, the left-hand side reduces to

$$\omega(TE_1, \ldots, TE_n) = c\varepsilon^{1\ldots n}(T_1, \ldots, T_n) = c\det(\varepsilon^j(T_i)) = c\det(T_i^j),$$

which is equal to the right-hand side. □

The Wedge Product

In Chapter 11, we defined the symmetric product, which takes a pair of symmetric tensors S, T and yields another symmetric tensor $ST = \mathrm{Sym}(S \otimes T)$ whose rank is the sum of the ranks of the original ones.

In this section, we will define a similar product operation for alternating tensors. One way to define it would be to mimic what we did in the symmetric case and define the product of alternating tensors ω and η to be $\mathrm{Alt}(\omega \otimes \eta)$. However, we will use a different definition that looks more complicated at first but turns out to be much better suited to computation.

If $\omega \in \Lambda^k(V)$ and $\eta \in \Lambda^l(V)$, we define the *wedge product* or *exterior product* of ω and η to be the alternating $(k+l)$-tensor

$$\omega \wedge \eta = \frac{(k+l)!}{k!l!} \mathrm{Alt}(\omega \otimes \eta). \tag{12.3}$$

The mysterious coefficient is motivated by the simplicity of the statement of the following lemma.

Lemma 12.7. *Let $(\varepsilon^1, \ldots, \varepsilon^n)$ be a basis for V^*. For any multi-indices $I = (i_1, \ldots, i_k)$ and $J = (j_1, \ldots, j_l)$,*

$$\varepsilon^I \wedge \varepsilon^J = \varepsilon^{IJ}, \tag{12.4}$$

where IJ is the multi-index $(i_1, \ldots, i_k, j_1, \ldots, j_l)$ obtained by concatenating I and J.

Proof. By multilinearity, it suffices to show that

$$\varepsilon^I \wedge \varepsilon^J(E_{p_1}, \ldots, E_{p_{k+l}}) = \varepsilon^{IJ}(E_{p_1}, \ldots, E_{p_{k+l}}) \tag{12.5}$$

for any sequence $(E_{p_1}, \ldots, E_{p_{k+l}})$ of basis vectors. We consider several cases.

Case I: *$P = (p_1, \ldots, p_{k+l})$ has a repeated index.* In this case, both sides of (12.5) are zero by Exercise 12.1(c).

Case II: *P contains an index that does not appear in either I or J.* In this case, the right-hand side is zero by Lemma 12.4(c). Similarly, each term in the expansion of the left-hand side involves either ε^I or ε^J evaluated on a sequence of basis vectors that is not a permutation of I or J, respectively, so the left-hand side is also zero.

Case III: *$P = IJ$ and P has no repeated indices.* In this case, the right-hand side of (12.5) is equal to 1 by Lemma 12.4(c), so we need to show that the left-hand side is also equal to 1. By definition,

$$\begin{aligned}
&\varepsilon^I \wedge \varepsilon^J(E_{p_1}, \dots, E_{p_{k+l}}) \\
&\quad = \frac{(k+l)!}{k!l!} \operatorname{Alt}(\varepsilon^I \otimes \varepsilon^J)(E_{p_1}, \dots, E_{p_{k+l}}) \\
&\quad = \frac{1}{k!l!} \sum_{\sigma \in S_{k+l}} (\operatorname{sgn} \sigma) \varepsilon^I(E_{p_{\sigma(1)}}, \dots, E_{p_{\sigma(k)}}) \varepsilon^J(E_{p_{\sigma(k+1)}}, \dots, E_{p_{\sigma(k+l)}}).
\end{aligned}$$

By Lemma 12.4 again, the only terms in the sum above that give nonzero values are those in which σ permutes the first k indices and the last l indices of P separately. In other words, σ must be of the form $\sigma = \tau\eta$, where $\tau \in S_k$ acts by permuting $\{1, \dots, k\}$ and $\eta \in S_l$ acts by permuting $\{k+1, \dots, k+l\}$. Since $\operatorname{sgn}(\tau\eta) = (\operatorname{sgn} \tau)(\operatorname{sgn} \eta)$, we have

$$\begin{aligned}
&\varepsilon^I \wedge \varepsilon^J(E_{p_1}, \dots, E_{p_{k+l}}) \\
&\quad = \frac{1}{k!l!} \sum_{\substack{\tau \in S_k \\ \eta \in S_l}} (\operatorname{sgn} \tau)(\operatorname{sgn} \eta) \varepsilon^I(E_{p_{\tau(1)}}, \dots, E_{p_{\tau(k)}}) \varepsilon^J(E_{p_{\eta(k+1)}}, \dots, E_{p_{\eta(k+l)}}) \\
&\quad = \left(\frac{1}{k!} \sum_{\tau \in S_k} (\operatorname{sgn} \tau) \varepsilon^I(E_{p_{\tau(1)}}, \dots, E_{p_{\tau(k)}}) \right) \times \\
&\qquad\qquad\qquad \left(\frac{1}{l!} \sum_{\eta \in S_l} (\operatorname{sgn} \eta) \varepsilon^J(E_{p_{\eta(k+1)}}, \dots, E_{p_{\eta(k+l)}}) \right) \\
&\quad = (\operatorname{Alt} \varepsilon^I)(E_{p_1}, \dots, E_{p_k}) \, (\operatorname{Alt} \varepsilon^J)(E_{p_{k+1}}, \dots, E_{p_{k+l}}) \\
&\quad = \varepsilon^I(E_{p_1}, \dots, E_{p_k}) \varepsilon^J(E_{p_{k+1}}, \dots, E_{p_{k+l}}) \\
&\quad = 1.
\end{aligned}$$

Case IV: *P is a permutation of IJ and P has no repeated indices.* In this case, applying a permutation to P brings us back to Case III. Since the effect of the permutation is to multiply both sides of (12.5) by the same sign, the result holds in this case as well. □

Proposition 12.8 (Properties of the Wedge Product).

(*a*) Bilinearity:

$$\begin{aligned}
(a\omega + a'\omega') \wedge \eta &= a(\omega \wedge \eta) + a'(\omega' \wedge \eta), \\
\eta \wedge (a\omega + a'\omega') &= a(\eta \wedge \omega) + a'(\eta \wedge \omega').
\end{aligned}$$

(*b*) Associativity:

$$\omega \wedge (\eta \wedge \xi) = (\omega \wedge \eta) \wedge \xi.$$

(*c*) Anticommutativity: *For $\omega \in \Lambda^k(V)$ and $\eta \in \Lambda^l(V)$,*

$$\omega \wedge \eta = (-1)^{kl} \eta \wedge \omega. \tag{12.6}$$

(*d*) *If* $(\varepsilon^1, \ldots, \varepsilon^n)$ *is any basis for* V^* *and* $I = (i_1, \ldots, i_k)$ *is any multi-index,*

$$\varepsilon^{i_1} \wedge \cdots \wedge \varepsilon^{i_k} = \varepsilon^I. \tag{12.7}$$

(*e*) *For any covectors* $\omega^1, \ldots, \omega^k$ *and vectors* $X_1, \ldots, X_k$,

$$\omega^1 \wedge \cdots \wedge \omega^k(X_1, \ldots, X_k) = \det(\omega^j(X_i)). \tag{12.8}$$

Proof. Bilinearity follows immediately from the definition, because the tensor product is bilinear and Alt is linear. To prove associativity, note that Lemma 12.7 gives

$$(\varepsilon^I \wedge \varepsilon^J) \wedge \varepsilon^K = \varepsilon^{IJ} \wedge \varepsilon^K = \varepsilon^{IJK} = \varepsilon^I \wedge \varepsilon^{JK} = \varepsilon^I \wedge (\varepsilon^J \wedge \varepsilon^K).$$

The general case follows from bilinearity. Similarly, using Lemma 12.7 again, we get

$$\varepsilon^I \wedge \varepsilon^J = \varepsilon^{IJ} = (\operatorname{sgn} \tau)\varepsilon^{JI} = (\operatorname{sgn} \tau)\varepsilon^J \wedge \varepsilon^I,$$

where τ is the permutation that sends IJ to JI. It is easy to check that $\operatorname{sgn} \tau = (-1)^{kl}$, because τ can be decomposed as a composition of kl transpositions (each index of I must be moved past each of the indices of J). Anticommutativity then follows from bilinearity.

Part (d) is an immediate consequence of Lemma 12.7 and induction. To prove part (e), we note that the special case in which each ω^j is one of the basis covectors ε^{i_j} just reduces to (12.7). Since both sides of (12.8) are multilinear in $(\omega^1, \ldots, \omega^k)$, this suffices. □

Because of part (d) of this lemma, we will generally use the notations ε^I and $\varepsilon^{i_1} \wedge \cdots \wedge \varepsilon^{i_k}$ interchangeably.

The definition and computational properties of the wedge product can seem daunting at first sight. However, the only properties that you need to remember for most practical purposes are that it is bilinear, associative, and anticommutative, and satisfies (12.8). In fact, these properties determine the wedge product uniquely, as the following exercise shows.

◇ **Exercise 12.3.** Show that the wedge product is the unique associative, bilinear, and anticommutative map $\Lambda^k(V) \times \Lambda^l(V) \to \Lambda^{k+l}(V)$ satisfying (12.8).

For any n-dimensional vector space V, define a vector space $\Lambda^*(V)$ by

$$\Lambda^*(V) = \bigoplus_{k=0}^{n} \Lambda^k V.$$

It follows from Proposition 12.5 that $\dim \Lambda^*(V) = 2^n$. Proposition 12.8 shows that the wedge product turns $\Lambda^*(V)$ into an associative algebra, called the *exterior algebra* of V. This algebra is not commutative, but it has a closely related property. An algebra A is said to be *graded* if it has a direct sum decomposition $A = \bigoplus A^k$ such that the product satisfies $(A^k)(A^l) \subset A^{k+l}$. A graded algebra is *anticommutative* if the product satisfies $ab = (-1)^{kl}ba$ for $a \in A^k$, $b \in A^l$. Proposition 12.8(c) shows that $\Lambda^*(V)$ is an anticommutative graded algebra.

As we observed at the beginning of this section, one could also define the wedge product without the unwieldy coefficient of (12.3). Many authors choose this alternative definition of the wedge product, which we denote by $\bar{\wedge}$:

$$\omega \mathbin{\bar{\wedge}} \eta = \operatorname{Alt}(\omega \otimes \eta). \tag{12.9}$$

Using this definition, (12.4) is replaced by

$$\varepsilon^I \mathbin{\bar{\wedge}} \varepsilon^J = \frac{k!l!}{(k+l)!}\varepsilon^{IJ}$$

and (12.8) is replaced by

$$\omega^1 \mathbin{\bar{\wedge}} \cdots \mathbin{\bar{\wedge}} \omega^k(X_1, \dots, X_k) = \frac{1}{k!}\det(\omega^i(X_j)) \tag{12.10}$$

whenever $\omega^1, \dots, \omega^k$ are covectors, as you can check.

Because of (12.8), we will call the wedge product defined by (12.3) the *determinant convention* for the wedge product, and the wedge product defined by (12.9) the *Alt convention*. Although the definition of the Alt convention is perhaps a bit more natural, the computational advantages of the determinant convention make it preferable for most applications, and we will use it exclusively in this book. (But see Problem 12-8 for another perspective.)

Differential Forms on Manifolds

Now we turn our attention to an n-dimensional smooth manifold M. The subset of T^kM consisting of alternating tensors is denoted by Λ^kM:

$$\Lambda^kM = \coprod_{p \in M} \Lambda^k(T_pM).$$

◇ **Exercise 12.4.** Show that Λ^kM is a smooth subbundle of T^kM, and therefore is a smooth vector bundle of rank $\binom{n}{k}$ over M.

A section of Λ^kM is called a *differential k-form*, or just a *k-form*; this is just a (continuous) tensor field whose value at each point is an alternating tensor. The integer k is sometimes called the *degree* of the form. We denote the vector space of smooth sections of Λ^kM by $\mathcal{A}^k(M)$. (We ordinarily denote the space of smooth sections of a vector bundle by the upper-case script letter corresponding to the name of the bundle; in this case, we use $\mathcal{A}$ because of the typographical similarity between A and Λ, and because

$\mathcal{A}$ suggests "alternating." Other notations in common use for the space of smooth k-forms are $\mathcal{E}^k(M)$ and $\Omega^k(M)$.)

In any smooth chart, a k-form ω can be written locally as

$$\omega = \sum_I{}' \omega_I dx^{i_1} \wedge \cdots \wedge dx^{i_k} = \sum_I{}' \omega_I dx^I,$$

where the coefficients ω_I are continuous functions defined on the coordinate domain, and we use dx^I as an abbreviation for $dx^{i_1} \wedge \cdots \wedge dx^{i_k}$ (not to be mistaken for the differential of a real-valued function x^I). In terms of differential forms, the result of Lemma 12.4(c) translates to

$$dx^{i_1} \wedge \cdots \wedge dx^{i_k} \left(\frac{\partial}{\partial x^{j_1}}, \ldots, \frac{\partial}{\partial x^{j_k}} \right) = \delta^I_J.$$

Thus the component functions ω^I of ω are determined by

$$\omega^I = \omega \left(\frac{\partial}{\partial x^{i_1}}, \ldots, \frac{\partial}{\partial x^{i_k}} \right).$$

Example 12.9. On $\mathbb{R}^3$, some examples of smooth 2-forms are given by

$$\begin{aligned} \omega &= (\sin xy)\, dy \wedge dz; \\ \eta &= dx \wedge dy + dx \wedge dz + dy \wedge dz. \end{aligned}$$

Every n-form on $\mathbb{R}^n$ is a continuous real-valued function times $dx^1 \wedge \cdots \wedge dx^n$, because there is only one increasing multi-index of length n.

A 0-form is just a continuous real-valued function, and a 1-form is a covector field. The wedge product of two differential forms is defined pointwise: $(\omega \wedge \eta)_p = \omega_p \wedge \eta_p$. Thus the wedge product of a k-form with an l-form is a $(k+l)$-form. If f is a 0-form and η is a k-form, we interpret the wedge product $f \wedge \eta$ to mean the ordinary product $f\eta$. If we define

$$\mathcal{A}^*(M) = \bigoplus_{k=0}^{n} \mathcal{A}^k(M), \tag{12.11}$$

the wedge product turns $\mathcal{A}^*(M)$ into an associative, anticommutative graded algebra.

If $F\colon M \to N$ is a smooth map and ω is a smooth differential form on N, the pullback $F^*\omega$ is a smooth differential form on M, defined as for any smooth tensor field:

$$F^*\omega(X_1, \ldots, X_k) = \omega(F_*X_1, \ldots, F_*X_k).$$

In particular, if $\iota\colon N \hookrightarrow M$ is the inclusion map of an immersed submanifold, then we usually use the notation $\omega|_N$ for $\iota^*\omega$.

Lemma 12.10. *Suppose $F\colon M \to N$ is smooth.*

(*a*) $F^*\colon \mathcal{A}^k(N) \to \mathcal{A}^k(M)$ *is linear.*

(*b*) $F^*(\omega \wedge \eta) = (F^*\omega) \wedge (F^*\eta)$.

(*c*) *In any smooth chart,*

$$F^*\left({\sum_I}' \omega_I dy^{i_1} \wedge \cdots \wedge dy^{i_k}\right) = {\sum_I}' (\omega_I \circ F)\, d(y^{i_1} \circ F) \wedge \cdots \wedge d(y^{i_k} \circ F).$$

◇ **Exercise 12.5.** Prove this lemma.

This lemma gives a computational rule for pullbacks of differential forms similar to the one we developed for arbitrary tensor fields in the preceding chapter. As before, it can also be used to compute the expression for a differential form in another smooth chart.

Example 12.11. Let ω be the 2-form $dx \wedge dy$ on $\mathbb{R}^2$. Thinking of the transformation to polar coordinates $x = r\cos\theta$, $y = r\sin\theta$ as an expression for the identity map with respect to different coordinates on the domain and range, we find

$$\begin{aligned}
\omega &= dx \wedge dy \\
&= d(r\cos\theta) \wedge d(r\sin\theta) \\
&= (\cos\theta\, dr - r\sin\theta\, d\theta) \wedge (\sin\theta\, dr + r\cos\theta\, d\theta) \\
&= r\cos^2\theta\, dr \wedge d\theta - r\sin^2\theta\, d\theta \wedge dr,
\end{aligned}$$

where we have used the fact that $dr \wedge dr = d\theta \wedge d\theta = 0$ by anticommutativity. Because $d\theta \wedge dr = -dr \wedge d\theta$, this simplifies to

$$dx \wedge dy = r\, dr \wedge d\theta.$$

The similarity between this formula and the formula for changing a double integral from Cartesian to polar coordinates is striking. The following lemma generalizes this.

Proposition 12.12. *Let $F\colon M \to N$ be a smooth map between n-manifolds. If (x^i) and (y^j) are smooth coordinates on open sets $U \subset M$ and $V \subset N$, respectively, and u is a smooth real-valued function on V, then the following holds on $U \cap F^{-1}(V)$:*

$$F^*\left(u\, dy^1 \wedge \cdots \wedge dy^n\right) = (u \circ F)(\det DF) dx^1 \wedge \cdots \wedge dx^n, \tag{12.12}$$

where DF represents the matrix of partial derivatives of F in coordinates.

Proof. Because the fiber of $\Lambda^n M$ is spanned by $dx^1 \wedge \cdots \wedge dx^n$ at each point, it suffices to show that both sides of (12.12) give the same result when evaluated on $(\partial/\partial x^1, \ldots, \partial/\partial x^n)$. From Lemma 12.10,

$$F^*\left(u\, dy^1 \wedge \cdots \wedge dy^n\right) = (u \circ F) dF^1 \wedge \cdots \wedge dF^n.$$

Proposition 12.8(e) shows that

$$dF^1 \wedge \cdots \wedge dF^n \left(\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n} \right) = \det \left(dF^j \left(\frac{\partial}{\partial x^i} \right) \right) = \det \left(\frac{\partial F^j}{\partial x^i} \right).$$

Therefore, the left-hand side of (12.12) gives $(u \circ F) \det DF$ when applied to $(\partial/\partial x^1, \ldots, \partial/\partial x^n)$. On the other hand, the right-hand side gives the same thing, because $dx^1 \wedge \cdots \wedge dx^n(\partial/\partial x^1, \ldots, \partial/\partial x^n) = 1$. □

Corollary 12.13. *If $(U,(x^i))$ and $(\widetilde{U},(\widetilde{x}^j))$ are overlapping smooth coordinate charts on M, then the following identity holds on $U \cap \widetilde{U}$:*

$$d\widetilde{x}^1 \wedge \cdots \wedge d\widetilde{x}^n = \det \left(\frac{\partial \widetilde{x}^j}{\partial x^i} \right) dx^1 \wedge \cdots \wedge dx^n. \tag{12.13}$$

Proof. Just apply the previous proposition with G equal to the identity map of $U \cap \widetilde{U}$, but using coordinates (x^i) in the domain and $(\widetilde{x}^j)$ in the range. □

Exterior Derivatives

In this section, we define a natural differential operator on smooth forms, called the exterior derivative. It is a generalization of the differential of a function.

To give some idea where the motivation for the exterior derivative comes from, let us look back at a question we addressed in Chapter 6. Recall that not all 1-forms are differentials of functions: Given a smooth 1-form ω, a necessary condition for the existence of a smooth function f such that $\omega = df$ is that ω be closed, which means that it satisfies

$$\frac{\partial \omega_j}{\partial x^i} - \frac{\partial \omega_i}{\partial x^j} = 0 \tag{12.14}$$

in every smooth coordinate system. Since this is a coordinate-independent property by Proposition 6.27, one might hope to find a more invariant way to express it. The key is that the expression in (12.14) is antisymmetric in the indices i and j, so it can be interpreted as the ij-component of an alternating tensor field, i.e., a 2-form. We will define a 2-form $d\omega$ by

$$d\omega = \sum_{i<j} \left(\frac{\partial \omega_j}{\partial x^i} - \frac{\partial \omega_i}{\partial x^j} \right) dx^i \wedge dx^j,$$

so it follows that ω is closed if and only if $d\omega = 0$.

This formula has a significant generalization to differential forms of all degrees. For any manifold, we will show that there is a differential operator $d\colon \mathcal{A}^k(M) \to \mathcal{A}^{k+1}(M)$ satisfying $d(d\omega) = 0$ for all ω. Thus it will follow that a necessary condition for a smooth k-form ω to be equal to $d\eta$ for some $(k-1)$-form η is that $d\omega = 0$.

The definition of d in coordinates is straightforward:

$$d\left(\sum_J{}' \omega_J dx^J\right) = \sum_J{}' d\omega_J \wedge dx^J, \tag{12.15}$$

where $d\omega_J$ is just the differential of the function ω_J. In somewhat more detail, this is

$$d\left(\sum_J{}' \omega_J dx^{j_1} \wedge \cdots \wedge dx^{j_k}\right) = \sum_J{}' \sum_i \frac{\partial \omega_J}{\partial x^i} dx^i \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_k}. \tag{12.16}$$

Observe that when ω is a 1-form, this becomes

$$\begin{aligned}
d(\omega_j\, dx^j) &= \frac{\partial \omega_j}{\partial x^i} dx^i \wedge dx^j \\
&= \sum_{i<j} \frac{\partial \omega_j}{\partial x^i} dx^i \wedge dx^j + \sum_{i>j} \frac{\partial \omega_j}{\partial x^i} dx^i \wedge dx^j \\
&= \sum_{i<j} \left(\frac{\partial \omega_j}{\partial x^i} - \frac{\partial \omega_i}{\partial x^j} \right) dx^i \wedge dx^j
\end{aligned}$$

after interchanging i and j in the second sum and using the fact that $dx^j \wedge dx^i = -dx^i \wedge dx^j$, so this is consistent with our earlier definition. For a smooth 0-form f (a real-valued function), (12.16) reduces to

$$df = \frac{\partial f}{\partial x^i} dx^i,$$

which is just the differential of f.

Proving that this definition is independent of the choice of coordinates and thus can be extended to smooth manifolds takes a little work. This is the content of the next theorem.

Theorem 12.14 (The Exterior Derivative). *For every smooth manifold M, there are unique linear maps $d\colon \mathcal{A}^k(M) \to \mathcal{A}^{k+1}(M)$ defined for each integer $k \geq 0$ and satisfying the following three conditions:*

(*i*) *If f is a smooth real-valued function* (*a 0-form*), *then df is the differential of f, defined as usual by*

$$df(X) = Xf.$$

(*ii*) *If $\omega \in \mathcal{A}^k(M)$ and $\eta \in \mathcal{A}^l(M)$, then*

$$d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta.$$

(*iii*) $d \circ d = 0$.

This operator also satisfies the following properties:

(*a*) *d is given in any smooth local coordinates by* (12.15).

(*b*) *d is local: If $\omega = \omega'$ on an open set $U \subset M$, then $d\omega = d\omega'$ on U.*

(*c*) *d commutes with restriction: If $U \subset M$ is any open set, then*

$$d(\omega|_U) = (d\omega)|_U. \tag{12.17}$$

Proof. We begin with a special case: Suppose M can be covered by a single smooth chart. Let $(x^1, \dots, x^n)$ be global smooth coordinates on M, and define $d\colon \mathcal{A}^k(M) \to \mathcal{A}^{k+1}(M)$ by (12.15). It is clearly linear and satisfies (i). We need to check that it satisfies (ii) and (iii). Before doing so, we need to know that d satisfies $d(f\,dx^I) = df \wedge dx^I$ for *any* multi-index I, not just increasing ones. If I has repeated indices, then clearly $d(f\,dx^I) = df \wedge dx^I = 0$. If not, let σ be the permutation sending I to an increasing multi-index J. Then

$$d(f\,dx^I) = (\operatorname{sgn}\sigma) d(f\,dx^J) = (\operatorname{sgn}\sigma) df \wedge dx^J = df \wedge dx^I.$$

To prove (ii), by linearity it suffices to consider terms of the form $\omega = f\,dx^I$ and $\eta = g\,dx^J$. We compute

$$\begin{aligned}
d(\omega \wedge \eta) &= d((f\,dx^I) \wedge (g\,dx^J)) \\
&= d(fg\,dx^I \wedge dx^J) \\
&= (g\,df + f\,dg) \wedge dx^I \wedge dx^J \\
&= (df \wedge dx^I) \wedge (g\,dx^J) + (-1)^k (f\,dx^I) \wedge (dg \wedge dx^J) \\
&= d\omega \wedge \eta + (-1)^k \omega \wedge d\eta,
\end{aligned}$$

where the $(-1)^k$ comes from the fact that $dg \wedge dx^I = (-1)^k dx^I \wedge dg$ because dg is a 1-form and dx^I is a k-form.

We will prove (iii) first for the special case of a 0-form, i.e., a real-valued function. In this case,

$$\begin{aligned}
d(df) &= d\left(\frac{\partial f}{\partial x^j} dx^j\right) \\
&= \frac{\partial^2 f}{\partial x^i \partial x^j} dx^i \wedge dx^j \\
&= \sum_{i<j} \left(\frac{\partial^2 f}{\partial x^i \partial x^j} - \frac{\partial^2 f}{\partial x^j \partial x^i}\right) dx^i \wedge dx^j \\
&= 0.
\end{aligned}$$

For the general case, we use the $k = 0$ case together with (ii) to compute

$$\begin{aligned} d(d\omega) &= d\Bigl({\sum_J}' d\omega_J \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_k}\Bigr) \\ &= {\sum_J}' d(d\omega_J) \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_k} \\ &\qquad + {\sum_J}' \sum_{i=1}^{k} (-1)^i d\omega_J \wedge dx^{j_1} \wedge \cdots \wedge d(dx^{j_i}) \wedge \cdots \wedge dx^{j_k} \\ &= 0. \end{aligned}$$

This proves that there exists an operator d satisfying (i)–(iii) in this special case. Properties (a)–(c) are immediate consequences of the definition (noting that if M is covered by a single smooth chart, then any open subset of M has the same property).

To show that d is unique, suppose $\widetilde{d}\colon \mathcal{A}^k(M) \to \mathcal{A}^{k+1}(M)$ is another linear operator defined for each $k \geq 0$ and satisfying (i), (ii), and (iii). Let $\omega = {\sum_J}' \omega_J \, dx^J \in \mathcal{A}^k(M)$ be arbitrary. Using linearity of $\widetilde{d}$ together with (ii), we compute

$$\begin{aligned} \widetilde{d}\omega &= \widetilde{d}\Bigl({\sum_J}' \omega_J dx^{j_1} \wedge \cdots \wedge dx^{j_k}\Bigr) \\ &= {\sum_J}' \widetilde{d}\omega_J \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_k} + (-1)^0 {\sum_J}' \omega_J \widetilde{d}(dx^{j_1} \wedge \cdots \wedge dx^{j_k}). \end{aligned}$$

Using (ii) again, the last term expands into a sum of terms, each of which contains a factor of the form $\widetilde{d}(dx^{j_i})$, which is equal to $\widetilde{d}(\widetilde{d}x^{j_i})$ by (i) and hence is zero by (iii). On the other hand, since each component function ω_J is a smooth function, (i) implies that $\widetilde{d}\omega_J = d\omega_J$, and thus $\widetilde{d}\omega$ is equal to $d\omega$ defined by (12.15). This implies, in particular, that we get the same operator no matter which (global) smooth coordinates we use to define it. This completes the proof of the existence and uniqueness of d in this special case.

Next, let M be an arbitrary smooth manifold. On any smooth coordinate domain $U \subset M$, the argument above yields a unique linear operator from smooth k-forms to smooth $(k+1)$-forms, which we denote by d_U, satisfying (i)–(iii). On any set $U \cap U'$ where two smooth charts overlap, the restrictions of $d_U\omega$ and $d_{U'}\omega$ to $U \cap U'$ satisfy

$$(d_U\omega)|_{U\cap U'} = d_{U\cap U'}\omega = (d_{U'}\omega)|_{U\cap U'}$$

by (12.17). Therefore, we can unambiguously define $d\colon \mathcal{A}^k(M) \to \mathcal{A}^{k+1}(M)$ by defining the value of $d\omega$ at $p \in M$ to be $(d\omega)_p = d_U(\omega|_U)_p$, where U is any smooth coordinate domain containing p. This operator satisfies (i), (ii), and (iii) because each d_U does. It also satisfies (a), (b), and (c) by definition.

Finally, we need to prove uniqueness in the general case. Suppose we have some other operator $\widetilde{d}\colon \mathcal{A}^k(M) \to \mathcal{A}^{k+1}(M)$ defined for each k and satisfying (i)–(iii). We begin by showing that $\widetilde{d}$ satisfies the locality property (b). Writing $\eta = \omega - \omega'$, it clearly suffices to show that $\widetilde{d}\eta = 0$ on U if η vanishes on U. Let $p \in U$ be arbitrary, and let $\varphi \in C^\infty(M)$ be a smooth bump function that is equal to 1 in a neighborhood of p and supported in U. Then $\varphi\eta$ is identically zero on M, so

$$0 = \widetilde{d}(\varphi\eta)_p = d\varphi_p \wedge \eta_p + \varphi(p)\widetilde{d}\eta_p = \widetilde{d}\eta_p,$$

because $\varphi \equiv 1$ in a neighborhood of p. Since p was an arbitrary point of U, this shows that $d\eta = 0$ on U.

Let $U \subset M$ be an arbitrary smooth coordinate domain. For each k, define an operator $\widetilde{d}_U\colon \mathcal{A}^k(U) \to \mathcal{A}^{k+1}(U)$ as follows. For each $p \in U$, choose an extension of ω to a smooth global k-form $\widetilde{\omega} \in \mathcal{A}^k(M)$ that agrees with ω on a neighborhood of p (such an extension exists by the extension lemma for vector bundles, Lemma 5.6), and set $(\widetilde{d}_U\omega)_p = (\widetilde{d}\widetilde{\omega})_p$. Because $\widetilde{d}$ is local, this definition is independent of the extension $\widetilde{\omega}$ chosen. The fact that $\widetilde{d}$ satisfies (i)–(iii) implies immediately that $\widetilde{d}_U$ satisfies the same properties. But by the uniqueness property we already proved for smooth coordinate domains, this implies that $\widetilde{d}_U = d_U$. In particular, if ω is the restriction to U of a smooth global form $\widetilde{\omega}$ on M, then we can use the same extension $\widetilde{\omega}$ near each point, so $d_U(\widetilde{\omega}|_U) = \widetilde{d}_U(\widetilde{\omega}|_U) = (\widetilde{d}\widetilde{\omega})|_U$. This shows that $\widetilde{d}$ is equal to the operator d we defined above, thus proving uniqueness. □

The operator d whose existence and uniqueness are asserted in this theorem is called *exterior differentiation*, and $d\omega$ is called the *exterior derivative* of ω. (Some authors use the term *exterior differential* for the same operator.) The exterior derivative of a real-valued function f is, of course, just its differential df.

If A is a graded algebra, a linear map $T\colon A \to A$ is said to be of *degree* m if $T(A^k) \subset A^{k+m}$. It is said to be an *antiderivation* if it satisfies $T(xy) = (Tx)y + (-1)^k x(Ty)$ whenever $x \in A^k$ and $y \in A^l$. The preceding theorem can be summarized by saying that the differential extends to a unique antiderivation of $\mathcal{A}^*(M)$ of degree 1 whose square is zero.

Example 12.15. Let us work out the exterior derivatives of arbitrary 1-forms and 2-forms on $\mathbb{R}^3$. Any smooth 1-form can be written

$$\omega = P\,dx + Q\,dy + R\,dz$$

for some smooth functions P, Q, R. Using (12.15) and the fact that the wedge product of any 1-form with itself is zero, we compute

$$
\begin{aligned}
d\omega &= dP \wedge dx + dQ \wedge dy + dR \wedge dz \\
&= \left(\frac{\partial P}{\partial x}dx + \frac{\partial P}{\partial y}dy + \frac{\partial P}{\partial z}dz\right) \wedge dx + \left(\frac{\partial Q}{\partial x}dx + \frac{\partial Q}{\partial y}dy + \frac{\partial Q}{\partial z}dz\right) \wedge dy \\
&\quad + \left(\frac{\partial R}{\partial x}dx + \frac{\partial R}{\partial y}dy + \frac{\partial R}{\partial z}dz\right) \wedge dz \\
&= \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx \wedge dy + \left(\frac{\partial R}{\partial x} - \frac{\partial P}{\partial z}\right) dx \wedge dz \\
&\quad + \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) dy \wedge dz.
\end{aligned}
$$

It is interesting to note that the components of this 2-form are exactly the components of the curl of the vector field with components (P, Q, R) (except perhaps in a different order and with different signs). We will explore this connection in more depth in Chapter 14.

An arbitrary 2-form on $\mathbb{R}^3$ can be written

$$\omega = \alpha\, dx \wedge dy + \beta\, dx \wedge dz + \gamma\, dy \wedge dz.$$

A similar computation shows

$$d\omega = \left(\frac{\partial \alpha}{\partial z} - \frac{\partial \beta}{\partial y} + \frac{\partial \gamma}{\partial x}\right) dx \wedge dy \wedge dz.$$

One important feature of the exterior derivative is that it behaves well with respect to pullbacks, as the next lemma shows.

Lemma 12.16. *If $G\colon M \to N$ is a smooth map, then the pullback map $G^*\colon \mathcal{A}^k(N) \to \mathcal{A}^k(M)$ commutes with d: For all $\omega \in \mathcal{A}^k(N)$,*

$$G^*(d\omega) = d(G^*\omega). \tag{12.18}$$

Proof. Let $\omega \in \mathcal{A}^k(N)$ be arbitrary. Because d is local, if (12.18) holds in a neighborhood of each point, then it holds on all of M. In a smooth coordinate neighborhood, ω can be written as a sum of terms like $f\, dx^{i_1} \wedge \cdots \wedge dx^{i_k}$, so by linearity it suffices to check (12.18) for a form of this type.

For such a form, the left-hand side of (12.18) is

$$
\begin{aligned}
G^* d(f\, dx^{i_1} \wedge \cdots \wedge dx^{i_k}) &= G^*(df \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}) \\
&= d(f \circ G) \wedge d(x^{i_1} \circ G) \wedge \cdots \wedge d(x^{i_k} \circ G),
\end{aligned}
$$

while the right-hand side is

$$
\begin{aligned}
dG^*(f\, dx^{i_1} \wedge \cdots \wedge dx^{i_k}) &= d((f \circ G)\, d(x^{i_1} \circ G) \wedge \cdots \wedge d(x^{i_k} \circ G)) \\
&= d(f \circ G) \wedge d(x^{i_1} \circ G) \wedge \cdots \wedge d(x^{i_k} \circ G). \qquad \square
\end{aligned}
$$

Extending the terminology that we introduced for covector fields in Chapter 6, we say that a smooth differential form $\omega \in \mathcal{A}^k(M)$ is *closed* if $d\omega = 0$, and *exact* if there exists a smooth $(k-1)$-form η on M such that $\omega = d\eta$. The fact that $d \circ d = 0$ implies that every exact form is closed. The converse may not be true, as we saw already in Chapter 6 for the case of 1-forms. We will return to these ideas in Chapter 15.

In addition to the coordinate formula (12.15) that we used in the definition of d, there is another formula for d that is often useful, not least because it is manifestly coordinate-independent. The formula for 1-forms is by far the most important, and is the easiest to state and prove.

Proposition 12.17 (Exterior Derivative of a 1-form). *For any smooth 1-form ω and smooth vector fields X and Y,*

$$d\omega(X,Y) = X(\omega(Y)) - Y(\omega(X)) - \omega([X,Y]). \tag{12.19}$$

Proof. Since any smooth 1-form can be expressed locally as a sum of terms of the form $u\,dv$ for smooth functions u and v, it suffices to consider that case. Suppose $\omega = u\,dv$, and X, Y are smooth vector fields. Then the left-hand side of (12.19) is

$$\begin{aligned} d(u\,dv)(X,Y) &= du \wedge dv(X,Y) \\ &= du(X)dv(Y) - dv(X)du(Y) \\ &= Xu\,Yv - Xv\,Yu. \end{aligned}$$

The right-hand side is

$$\begin{aligned} &X(u\,dv(Y)) - Y(u\,dv(X)) - u\,dv([X,Y]) \\ &\quad = X(u\,Yv) - Y(u\,Xv) - u\,[X,Y]v \\ &\quad = (Xu\,Yv + u\,XYv) - (Yu\,Xv + u\,YXv) - u(XYv - YXv). \end{aligned}$$

After the $u\,XYv$ and $u\,YXv$ terms are canceled, this is equal to the left-hand side. □

We will see a number of applications of (12.19) in later chapters. Here is our first one—it shows that the exterior derivative is in a certain sense dual to the Lie bracket. In particular, it shows that if we know all the Lie brackets of basis vector fields in a smooth local frame, we can compute the exterior derivatives of the dual covector fields, and vice versa.

Proposition 12.18. *Let M be a smooth n-manifold, let (E_i) be a smooth local frame for M, and let (ε^i) be the dual coframe. Let c^i_{jk}, $i = 1, \dots, n$, be the component functions of the Lie bracket $[E_j, E_k]$ in this frame:*

$$[E_j, E_k] = c^i_{jk} E_i.$$

Then the exterior derivative of each 1-form ε^i is given by

$$d\varepsilon^i = -c^i_{jk}\,\varepsilon^j \wedge \varepsilon^k.$$

◇ **Exercise 12.6.** Use (12.19) to prove the preceding proposition.

The generalization of (12.19) to higher-degree forms is more complicated.

Proposition 12.19 (Invariant Formula for Exterior Derivatives). *Let M be a smooth manifold and $\omega \in \mathcal{A}^k(M)$. For any smooth vector fields $X_1, \ldots, X_{k+1}$ on M,*

$$\begin{aligned} d\omega(X_1, \ldots, X_{k+1}) &= \sum_{1 \le i \le k+1} (-1)^{i-1} X_i\big(\omega(X_1, \ldots, \widehat{X}_i, \ldots, X_{k+1})\big) \\ &+ \sum_{1 \le i < j \le k+1} (-1)^{i+j} \omega([X_i, X_j], X_1, \ldots, \widehat{X}_i, \ldots, \widehat{X}_j, \ldots, X_{k+1}), \end{aligned} \tag{12.20}$$

where the hats indicate omitted arguments.

Proof. For this proof, let us denote the two sums on the right-hand side of (12.20) by $\mathrm{I}(X_1, \ldots, X_{k+1})$ and $\mathrm{II}(X_1, \ldots, X_{k+1})$, and the entire right-hand side by $D\omega(X_1, \ldots, X_{k+1})$. Note that $D\omega$ is obviously multilinear over $\mathbb{R}$. We will begin by showing that, like $d\omega$, it is actually multilinear over $C^\infty(M)$ (see Problem 11-7), which is to say that for $1 \le p \le k+1$ and $f \in C^\infty(M)$,

$$D\omega(X_1, \ldots, fX_p, \ldots, X_{k+1}) = f D\omega(X_1, \ldots, X_p, \ldots, X_{k+1}).$$

In the expansion of $\mathrm{I}(X_1, \ldots, fX_p, \ldots, X_{k+1})$, f obviously factors out of the $i = p$ term. The other terms expand as follows:

$$\begin{aligned} &\sum_{i \ne p} (-1)^{i-1} X_i\big(f\omega(X_1, \ldots, \widehat{X}_i, \ldots, X_{k+1})\big) \\ &\qquad = \sum_{i \ne p} (-1)^{i-1} \Big(f X_i\big(\omega(X_1, \ldots, \widehat{X}_i, \ldots, X_{k+1})\big) \\ &\qquad\qquad + (X_i f)\omega(X_1, \ldots, \widehat{X}_i, \ldots, X_{k+1}) \Big). \end{aligned}$$

Therefore

$$\begin{aligned} \mathrm{I}(X_1, \ldots, fX_p, \ldots, X_{k+1}) &= f\,\mathrm{I}(X_1, \ldots, X_p, \ldots, X_{k+1}) \\ &+ \sum_{i \ne p} (-1)^{i-1} (X_i f)\omega(X_1, \ldots, \widehat{X}_i, \ldots, X_{k+1}). \end{aligned} \tag{12.21}$$

Consider next the expansion of II. Again, f factors out of all the terms in which $i \ne p$ and $j \ne p$. To expand the other terms, we use (4.7), which implies

$$\begin{aligned} [fX_p, X_j] &= f[X_p, X_j] - (X_j f)X_p, \\ [X_i, fX_p] &= f[X_i, X_p] + (X_i f)X_p. \end{aligned}$$

Inserting these formulas into the $i = p$ and $j = p$ terms, we obtain

$$\begin{aligned}
&\mathrm{II}(X_1, \dots, fX_p, \dots, X_{k+1}) \\
&\quad = f\,\mathrm{II}(X_1, \dots, X_p, \dots, X_{k+1}) \\
&\qquad + \sum_{p<j} (-1)^{p+j+1}(X_j f)\omega(X_p, X_1, \dots, \widehat{X}_p, \dots, \widehat{X}_j, \dots, X_{k+1}) \\
&\qquad + \sum_{i<p} (-1)^{i+p}(X_i f)\omega(X_p, X_1, \dots, \widehat{X}_i, \dots, \widehat{X}_p, \dots, X_{k+1}).
\end{aligned}$$

Rearranging the arguments in these two sums so as to put X_p into its original position, we see that they exactly cancel the sum in (12.21). This completes the proof that $D\omega$ is multilinear over $C^\infty(M)$.

By multilinearity, to verify that $D\omega = d\omega$, it suffices to show that both sides give the same result when applied to any sequence of basis vectors in an arbitrary local frame. The computations are greatly simplified by working in a coordinate frame, for which all the Lie brackets vanish. Thus let $(U, (x^i))$ be an arbitrary smooth chart on M. Because both $d\omega$ and $D\omega$ depend linearly on ω, we may assume that $\omega = f\,dx^I$ for some smooth function f and some increasing multi-index $I = (i_1, \dots, i_k)$, so

$$d\omega = df \wedge dx^I = \sum_l \frac{\partial f}{\partial x^l} dx^l \wedge dx^I.$$

If $J = (j_1, \dots, j_{k+1})$ is any multi-index of length $k+1$, therefore,

$$d\omega\left(\frac{\partial}{\partial x^{j_1}}, \dots, \frac{\partial}{\partial x^{j_{k+1}}}\right) = \sum_l \frac{\partial f}{\partial x^l} \delta^{lI}_J.$$

The only terms in this sum that can possibly be nonzero are those for which l is equal to one of the indices in J, say $l = j_p$. In this case, it is easy to check that $\delta^{lI}_J = (-1)^{p-1}\delta^I_{\widehat{J}_p}$, where $\widehat{J}_p = (j_1, \dots, \widehat{j_p}, \dots, j_{k+1})$, so

$$d\omega\left(\frac{\partial}{\partial x^{j_1}}, \dots, \frac{\partial}{\partial x^{j_{k+1}}}\right) = \sum_{1 \le p \le k+1} (-1)^{p-1} \frac{\partial f}{\partial x^{j_p}} \delta^I_{\widehat{J}_p}. \tag{12.22}$$

On the other hand, because all the Lie brackets are zero, we have

$$\begin{aligned}
&D\omega\left(\frac{\partial}{\partial x^{j_1}}, \dots, \frac{\partial}{\partial x^{j_{k+1}}}\right) \\
&\quad = \sum_{1 \le p \le k+1} (-1)^{p-1} \frac{\partial}{\partial x^{j_p}} \left(f\,dx^I \left(\frac{\partial}{\partial x^{j_1}}, \dots, \widehat{\frac{\partial}{\partial x^{j_p}}}, \dots, \frac{\partial}{\partial x^{j_{k+1}}}\right)\right) \\
&\quad = \sum_{1 \le p \le k+1} (-1)^{p-1} \frac{\partial f}{\partial x^{j_p}} \delta^I_{\widehat{J}_p},
\end{aligned}$$

which agrees with (12.22). $\square$

It is worth remarking that formula (12.20) can be used to give an invariant definition of d, as well as an alternative proof of Theorem 12.14 on the existence, uniqueness, and properties of d. As the proof of Proposition 12.19 showed, the right-hand side of (12.20) is multilinear over $C^\infty(M)$ as a function of $(X_1, \dots, X_{k+1})$. By the result of Problem 11-7, therefore, it defines a smooth covariant $(k+1)$-tensor field, which we could have used as a definition of $d\omega$. The rest of the proof of Proposition 12.19 then shows that $d\omega$ is actually given locally by the coordinate formula (12.15), and so the properties asserted in Theorem 12.14 follow just as before. We have chosen to define d by means of its coordinate formula because that formula is generally much easier to remember and to work with. Except in the $k = 1$ case, the invariant formula (12.20) is too complicated to be of much use for computation; in addition, it has the serious flaw that in order to compute the action of $d\omega$ on vectors $(X_1, \dots, X_k)$ at a point $p \in M$, one must first extend them to vector *fields* in a neighborhood of p. Nonetheless, it does have some important theoretical consequences, so it is useful to know that it exists.

Symplectic Forms

In this section, we introduce symplectic forms. These are special 2-forms that play a leading role in many applications of smooth manifold theory to analysis and physics.

We begin with some linear algebra. A 2-tensor ω on a finite-dimensional real vector space V is said to be *nondegenerate* if $\omega(X, Y) = 0$ for all $Y \in V$ implies $X = 0$.

◇ **Exercise 12.7.** Show that the following are equivalent for a 2-tensor ω on a finite-dimensional vector space V:

(a) ω is nondegenerate.
(b) The matrix (ω_{ij}) representing ω in terms of any basis is nonsingular.
(c) The linear map $\widetilde{\omega}\colon V \to V^*$ defined by $\widetilde{\omega}(X)(Y) = \omega(X, Y)$ is invertible.

A nondegenerate alternating 2-tensor is called a *symplectic tensor*. A vector space V endowed with a specific symplectic tensor is called a *symplectic vector space*. (A symplectic tensor is also often called a "symplectic form," because it is in particular a bilinear form. But to avoid confusion, we will reserve the name "symplectic form" for something slightly different, to be defined below.)

Example 12.20. Let V be a vector space of dimension $2n$. Choose any basis for V, and denote the basis by $(A_1, B_1, \dots, A_n, B_n)$ and the corresponding dual basis for V^* by $(\alpha^1, \beta^1, \dots, \alpha^n, \beta^n)$. Let $\omega \in \Lambda^2(V)$ be the

2-covector defined by

$$\omega = \sum_{i=1}^{n} \alpha^i \wedge \beta^i. \tag{12.23}$$

Note that the action of ω on basis vectors is given by

$$\begin{aligned} \omega(A_i, B_j) &= -\omega(B_j, A_i) = \delta_{ij}, \\ \omega(A_i, A_j) &= \omega(B_i, B_j) = 0. \end{aligned} \tag{12.24}$$

Suppose $X = a^i A_i + b^i B_i \in V$ satisfies $\omega(X, Y) = 0$ for all $Y \in V$. Then $0 = \omega(X, B_i) = a^i$ and $0 = \omega(X, A_i) = -b^i$, which implies that $X = 0$. Thus ω is nondegenerate, and so is a symplectic tensor.

It is interesting to consider the special case in which $\dim V = 2$. In this case, every 2-covector is a multiple of $\alpha^1 \wedge \beta^1$, which is nondegenerate by the argument above. Thus every nonzero 2-covector on a 2-dimensional vector space is symplectic.

If (V, ω) is a symplectic vector space and $S \subset V$ is any subspace, we define the *symplectic complement* of S, denoted by $S^\perp$, to be the subspace

$$S^\perp = \{X \in V : \omega(X, Y) = 0 \text{ for all } Y \in S\}.$$

As the notation suggests, the symplectic complement is analogous to the orthogonal complement in an inner product space. For example, just as in the inner product case, the dimension of $S^\perp$ is the codimension of S, as the next lemma shows.

Lemma 12.21. *Let (V, ω) be a symplectic vector space. For any subspace $S \subset V$, $\dim S + \dim S^\perp = \dim V$.*

Proof. Define a linear map $\Phi\colon V \to S^*$ by $\Phi(X) = \widetilde{\omega}(X)|_S$, or equivalently

$$\Phi(X)(Y) = \omega(X, Y) \quad \text{for } X \in V,\ Y \in S.$$

If $\varphi \in S^*$ is arbitrary, let $\widetilde{\varphi} \in V^*$ be any extension of φ to a linear functional on all of V. Since $\widetilde{\omega}\colon V \to V^*$ is an isomorphism, there exists $X \in V$ such that $\widetilde{\omega}(X) = \widetilde{\varphi}$. It follows that $\Phi(X) = \varphi$, and therefore Φ is surjective. By the rank-nullity law, $S^\perp = \operatorname{Ker} \Phi$ has dimension equal to $\dim V - \dim S^* = \dim V - \dim S$. $\square$

Symplectic complements differ from orthogonal complements in one important respect: Although it is always true that $S \cap S^\perp = \{0\}$ in an inner product space, this need not be true in a symplectic vector space. Indeed, if S is 1-dimensional, the fact that ω is alternating forces $\omega(X, X) = 0$ for every $X \in S$, so $S \subset S^\perp$. Carrying this idea a little further, subspaces of V can be classified in the following way. A subspace $S \subset V$ is said to be

- *symplectic* if $S \cap S^\perp = \{0\}$;
- *isotropic* if $S \subset S^\perp$;

- *coisotropic* if $S \supset S^{\perp}$;
- *Lagrangian* if $S = S^{\perp}$.

◇ **Exercise 12.8.** Let (V, ω) be a symplectic vector space, and let $S \subset V$ be a subspace.

(a) Show that $(S^{\perp})^{\perp} = S$.
(b) Show that S is symplectic if and only if $\omega|_S$ is nondegenerate.
(c) Show that S is isotropic if and only if $\omega|_S = 0$.
(d) Show that S is Lagrangian if and only if $\omega|_S = 0$ and $\dim S = n$.

The symplectic tensor ω defined in Example 12.20 turns out to be the prototype of all symplectic tensors, as the next proposition shows. This can be viewed as a symplectic version of the Gram-Schmidt algorithm.

Proposition 12.22 (Canonical Form for a Symplectic Tensor). *Let ω be a symplectic tensor on an m-dimensional vector space V. Then V has even dimension $m = 2n$, and there exists a basis for V in which ω has the form* (12.23).

Proof. It is easy to check that ω has the form (12.23) with respect to a basis $(A_1, B_1, \ldots, A_n, B_n)$ if and only if the action of ω on basis vectors is given by (12.24). Thus we will prove the theorem by induction on $m = \dim V$, by showing that there exists a basis with this property.

For $m = 0$ there is nothing to prove. Suppose (V, ω) is a symplectic vector space of dimension $m \geq 1$, and assume the proposition is true for all symplectic vector spaces of dimension less than m. Let A_1 be any nonzero vector in V. Since ω is nondegenerate, there exists $B_1 \in V$ such that $\omega(A_1, B_1) \neq 0$. Multiplying B_1 by a constant if necessary, we may assume that $\omega(A_1, B_1) = 1$. Because ω is alternating, B_1 cannot be a multiple of A_1, so the set $\{A_1, B_1\}$ is independent.

Let $S \subset V$ be the subspace spanned by $\{A_1, B_1\}$. Then $\dim S^{\perp} = m - 2$ by Lemma 12.21. Since $\omega|_S$ is obviously nondegenerate, by Exercise 12.8 it follows that S is symplectic. This means $S \cap S^{\perp} = \{0\}$, so $S^{\perp}$ is also symplectic. By induction $S^{\perp}$ is even-dimensional and there is a basis $(A_2, B_2, \ldots, A_n, B_n)$ for $S^{\perp}$ such that (12.24) is satisfied for $2 \leq i, j \leq n$. It follows easily that $(A_1, B_1, A_2, B_2, \ldots, A_n, B_n)$ is the required basis for V. □

Because of this proposition, if (V, ω) is a symplectic vector space, a basis $(A_1, B_1, \ldots, A_n, B_n)$ for V is called a *symplectic basis* if (12.24) holds, which is equivalent to ω being given by (12.23) in terms of the dual basis. The proposition then says that every symplectic vector space has a symplectic basis.

Now let us turn to manifolds. A *symplectic form* on a smooth manifold M is a smooth, closed, nondegenerate 2-form. In other words, a smooth 2-form ω is symplectic if and only if it is closed and ω_p is a symplectic tensor for each $p \in M$. A smooth manifold endowed with a specific choice of symplectic form is called a *symplectic manifold*. A choice of symplectic form is also sometimes called a *symplectic structure* on M. Proposition 12.22 implies that a symplectic manifold must be even-dimensional. If (M, ω) and $(\widetilde{M}, \widetilde{\omega})$ are symplectic manifolds, a diffeomorphism $F\colon M \to \widetilde{M}$ satisfying $F^*\widetilde{\omega} = \omega$ is called a *symplectomorphism*. The study of properties of symplectic manifolds that are invariant under symplectomorphisms is known as *symplectic geometry*.

Example 12.23 (Symplectic Manifolds).

(a) If we denote the standard coordinates on $\mathbb{R}^{2n}$ by $(x^1, y^1, \ldots, x^n, y^n)$, the 2-form

$$\omega = \sum_{i=1}^{n} dx^i \wedge dy^i$$

is symplectic: It is obviously closed, and it is nondegenerate because its value at each point is the standard symplectic tensor of Example 12.20. This is called the *standard symplectic form* on $\mathbb{R}^{2n}$. (When working with the standard symplectic form, like the Euclidean inner product, it is usually necessary to insert explicit summation signs, because the summation index i appears twice in the upper position.)

(b) Suppose Σ is any smooth 2-manifold and Ω is any nonvanishing smooth 2-form on Σ. Then Ω is closed because $d\Omega$ is a 3-form, and every 3-form on a 2-manifold is zero. Moreover, as we observed above, every nonvanishing 2-form is nondegenerate, so (Σ, Ω) is a symplectic manifold.

Suppose (M, ω) is a symplectic manifold. An (immersed or embedded) submanifold $N \subset M$ is said to be symplectic, isotropic, coisotropic, or Lagrangian if T_pN (thought of as a subspace of T_pM) has this property at each point $p \in N$. More generally, an immersion (or embedding) $F\colon N \to M$ is said to have one of these properties if the subspace $F_*(T_pN) \subset T_{F(p)}M$ has the corresponding property for every $p \in N$. Thus a submanifold is symplectic (isotropic, etc.) if and only if its inclusion map has the same property.

◇ **Exercise 12.9.** Suppose (M, ω) is a symplectic manifold, and $F\colon N \to M$ is an immersion. Show that F is isotropic if and only if $F^*\omega = 0$, and F is symplectic if and only if $F^*\omega$ is a symplectic form.

The Canonical Symplectic Form on the Cotangent Bundle

The most important example of a symplectic manifold is the total space of the cotangent bundle of any smooth manifold Q, which carries a canonical symplectic structure that we now define. First, there is a natural 1-form τ

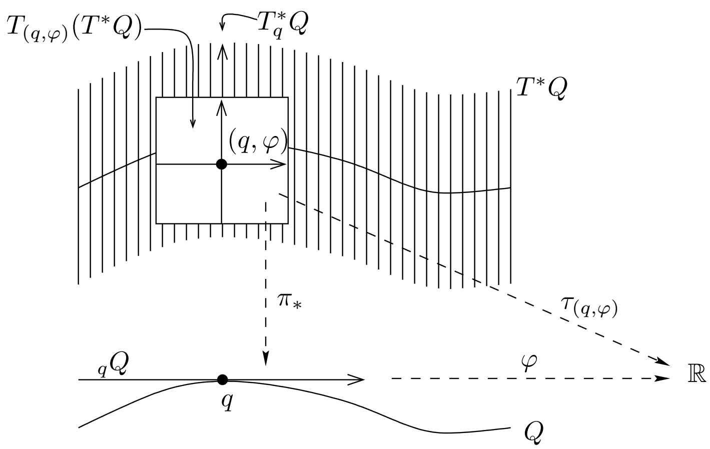


Figure 12.4. The tautological 1-form on T^*Q.

on $M = T^*Q$ (the total space of the cotangent bundle), called the *tautological* 1*-form*, defined as follows. A point in T^*Q is a covector $\varphi \in T_q^*Q$ for some $q \in Q$; we will denote such a point by the notation (q,φ). The natural projection $\pi\colon T^*Q \to Q$ is then just $\pi(q,\varphi) = q$, and its dual map is a linear map $\pi^*\colon T_q^*Q \to T_{(q,\varphi)}(T^*Q)$. We define $\tau \in \Lambda^1(T^*Q)$ by

$$\tau_{(q,\varphi)} = \pi^*\varphi.$$

(See Figure 12.4.) In other words, the value of τ at $(q,\varphi) \in T^*Q$ is the pullback with respect to π of the covector φ itself. If X is a tangent vector in $T_{(q,\varphi)}(T^*Q)$, then

$$\tau_{(q,\varphi)}(X) = \varphi(\pi_* X).$$

Proposition 12.24. *Let Q be a smooth manifold. The tautological* 1*-form τ is smooth, and $\omega = -d\tau$ is a symplectic form on the total space of T^*Q.*

Proof. Let (x^i) be any smooth coordinates on Q, and let (x^i,ξ_i) denote the corresponding standard coordinates on T^*Q as defined on page 130. Recall that the coordinates of $(q,\varphi) \in T^*Q$ are defined to be (x^i,ξ_i), where (x^i) is the coordinate representation of q and $\xi_i\,dx^i$ is the coordinate representation of φ. In terms of these coordinates, the projection $\pi\colon T^*Q \to Q$ has the coordinate expression $\pi(x,\xi) = x$, and therefore the coordinate representation of τ is

$$\tau_{(x,\xi)} = \pi^*(\xi_i\,dx^i) = \xi_i\,dx^i.$$

It follows immediately that τ is smooth, because its component functions are linear.

Clearly ω is closed, because it is exact. Moreover,

$$\omega = -d\tau = \sum_i dx^i \wedge d\xi_i.$$

Under the identification of an open subset of T^*Q with an open subset of $\mathbb{R}^{2n}$ by means of these coordinates, ω corresponds to the standard symplectic form on $\mathbb{R}^{2n}$ (with ξ_i substituted for y^i). It follows that ω is symplectic. $\square$

The symplectic form defined in this proposition is called the *canonical symplectic form* on T^*Q. One of its many uses is in giving a somewhat more "geometric" interpretation of what it means for a 1-form to be closed, as shown by the following proposition.

Proposition 12.25. *Let M be a smooth manifold, and let σ be a smooth 1-form on M. Thought of as a smooth map from M to T^*M, σ is a smooth embedding, and σ is closed if and only if its image $\sigma(M)$ is a Lagrangian submanifold of T^*M.*

Proof. Throughout this proof, we need to remember that σ is playing two roles: On the one hand, it is a 1-form on M, and on the other hand, it is a smooth map between manifolds. Since they are literally the same map, we will not use different notations to distinguish between them; but you should be careful to think about which role σ is playing at each step of the argument.

In terms of any smooth local coordinates (x^i) for M and the corresponding standard coordinates (x^i, ξ_i) for T^*M, the map $\sigma\colon M \to T^*M$ has the coordinate representation

$$\sigma(x^1, \dots, x^n) = (x^1, \dots, x^n, \sigma_1(x), \dots, \sigma_n(x)),$$

where $\sigma_i\, dx^i$ is the coordinate representation of σ as a 1-form. It follows immediately that σ is an immersion, and the fact that it is injective follows from $\pi \circ \sigma = \mathrm{Id}_M$.

To show that it is an embedding, it suffices by Proposition 7.4 to show that it is a proper map. This in turn follows from the fact that π is a left inverse for σ, by Lemma 2.16.

Because $\sigma(M)$ is n-dimensional, it is Lagrangian if and only if it is isotropic, which is the case if and only if $\sigma^*\omega = 0$. The pullback of the tautological form τ under σ is

$$\sigma^*\tau = \sigma^*(\xi_i\, dx^i) = \sigma_i\, dx^i = \sigma.$$

This can also be seen somewhat more invariantly from the computation

$$(\sigma^*\tau)_p(X) = \tau_{\sigma(p)}(\sigma_* X) = \sigma_p(\pi_*\sigma_* X) = \sigma_p(X),$$

which follows from the definition of τ and the fact that $\pi \circ \sigma = \mathrm{Id}_M$. Therefore,

$$\sigma^*\omega = -\sigma^* d\tau = -d(\sigma^*\tau) = -d\sigma.$$

It follows that σ is a Lagrangian embedding if and only if $d\sigma = 0$. □

Problems

12-1. Let $v_1, \dots, v_n$ be any n vectors in $\mathbb{R}^n$, and let P be the n-dimensional parallelepiped spanned by them:

$$P = \{t_1 v_1 + \cdots + t_n v_n : 0 \le t_i \le 1\}.$$

Show that $\mathrm{Vol}(P) = |\det(v_1, \dots, v_n)|$.

12-2. Let (e^1, e^2, e^3) be the standard dual basis for $(\mathbb{R}^3)^*$. Show that $e^1 \otimes e^2 \otimes e^3$ is not equal to a sum of an alternating tensor and a symmetric tensor.

12-3. Show that covectors $\omega^1, \dots, \omega^k$ on a finite-dimensional vector space are linearly dependent if and only if $\omega^1 \wedge \cdots \wedge \omega^k = 0$.

12-4. Show that two k-tuples $\{\omega^1, \dots, \omega^k\}$ and $\{\eta^1, \dots, \eta^k\}$ of independent covectors have the same span if and only if

$$\omega^1 \wedge \cdots \wedge \omega^k = c\, \eta^1 \wedge \cdots \wedge \eta^k$$

for some nonzero real number c.

12-5. A k-covector η on a finite-dimensional vector space V is said to be *decomposable* if it can be written

$$\eta = \omega^1 \wedge \cdots \wedge \omega^k,$$

where $\omega^1, \dots, \omega^k$ are covectors. For what values of n is it true that every 2-covector on $\mathbb{R}^n$ is decomposable?

12-6. Define a 2-form Ω on $\mathbb{R}^3$ by

$$\Omega = x\, dy \wedge dz + y\, dz \wedge dx + z\, dx \wedge dy.$$

(a) Compute Ω in spherical coordinates (ρ, φ, θ) defined by $(x, y, z) = (\rho \sin\varphi \cos\theta, \rho \sin\varphi \sin\theta, \rho \cos\varphi)$.
(b) Compute $d\Omega$ in both Cartesian and spherical coordinates and verify that both expressions represent the same 3-form.
(c) Compute the restriction $\Omega|_{\mathbb{S}^2} = \iota^*\Omega$, using coordinates (φ, θ), on the open subset where these coordinates are defined.
(d) Show that $\Omega|_{\mathbb{S}^2}$ is nowhere zero.

12-7. In each of the following problems, $g\colon M \to N$ is a smooth map between manifolds M and N, and ω is a smooth differential form on N. In each case, compute $g^*\omega$ and $d\omega$, and verify by direct computation that $g^*(d\omega) = d(g^*\omega)$.

(a) $g\colon \mathbb{R}^2 \to \mathbb{R}^2$ by

$$\begin{aligned}(x, y) &= g(s, t) = (st, e^t);\\ \omega &= x\,dy.\end{aligned}$$

(b) $g\colon \mathbb{R}^2 \to \mathbb{R}^3$ by

$$\begin{aligned}(x, y, z) &= g(\theta, \varphi) = ((\cos\varphi + 2)\cos\theta, (\cos\varphi + 2)\sin\theta, \sin\varphi);\\ \omega &= y\,dz \wedge dx.\end{aligned}$$

(c) $g\colon \{(u, v) \in \mathbb{R}^2 : u^2 + v^2 < 1\} \to \mathbb{R}^3 \smallsetminus \{0\}$ by

$$\begin{aligned}(x, y, z) &= (u, v, \sqrt{1 - u^2 - v^2});\\ \omega &= (x^2 + y^2 + z^2)^{-3/2}(x\,dy \wedge dz + y\,dz \wedge dx + z\,dx \wedge dy).\end{aligned}$$

12-8. Let V be a finite-dimensional real vector space. We have two ways to think about the tensor space $T^k(V)$: concretely, as the space of k-multilinear functionals on V; and abstractly, as the tensor product space $V^* \otimes \cdots \otimes V^*$. However, we have defined alternating and symmetric tensors only in terms of the concrete definition. This problem outlines an abstract approach to alternating tensors. (Symmetric tensors can be handled similarly.)

Let $\mathcal{A}$ denote the subspace of $V^* \otimes \cdots \otimes V^*$ spanned by all elements of the form $\alpha \otimes \varphi \otimes \varphi \otimes \beta$ for a covector φ and arbitrary tensors α, β, and let $A^k(V^*)$ denote the quotient vector space $V^* \otimes \cdots \otimes V^*/\mathcal{A}$. Define a wedge product on $A^k(V^*)$ by $\omega \wedge \eta = \pi(\widetilde{\omega} \otimes \widetilde{\eta})$, where $\pi\colon V^* \otimes \cdots \otimes V^* \to A^k(V^*)$ is the projection, and $\widetilde{\omega}$, $\widetilde{\eta}$ are arbitrary tensors such that $\pi(\widetilde{\omega}) = \omega$, $\pi(\widetilde{\eta}) = \eta$. Show that this wedge product is well-defined, and that there is a unique isomorphism $F\colon A^k(V^*) \to \Lambda^k(V)$ such that the following diagram commutes:

$$\begin{tikzcd}
V^* \otimes \cdots \otimes V^* \arrow[r, "\cong"] \arrow[d, "\pi"'] & T^k(V) \arrow[d, "\mathrm{Alt}"] \\
A^k(V^*) \arrow[r, "F"'] & \Lambda^k(V).
\end{tikzcd}$$

Show that F takes the wedge product we just defined on $A^k(V^*)$ to the Alt convention wedge product on $\Lambda^k(V)$. [Remark: This is one reason why some authors consider the Alt convention for the wedge product to be more natural than the determinant convention. It also explains why some authors prefer the notation $\Lambda^k(V^*)$ instead of $\Lambda^k(V)$ for the space of alternating covariant k-tensors, since it can be viewed as a quotient of the k-fold tensor product of V^* with itself.]

12-9. Let (V,ω) be a symplectic vector space of dimension $2n$. Show that for each symplectic, isotropic, coisotropic, or Lagrangian subspace $S \subset V$, there exists a symplectic basis (A_i, B_i) for V with the following property:

(a) If S is symplectic, $S = \operatorname{span}(A_1, B_1, \dots, A_k, B_k)$ for some k.
(b) If S is isotropic, $S = \operatorname{span}(A_1, \dots, A_k)$ for some k.
(c) If S is coisotropic, $S = \operatorname{span}(A_1, \dots, A_n, B_1, \dots, B_k)$ for some k.
(d) If S is Lagrangian, $S = \operatorname{span}(A_1, \dots, A_n)$.

12-10. Let (M,ω) be a symplectic manifold, and suppose $F\colon N \to M$ is a smooth map such that $F^*\omega$ is symplectic. Show that F is an immersion.

12-11. Let Q be a smooth manifold, and let S be an embedded submanifold of the total space of T^*Q. Show that S is the image of a smooth closed 1-form on Q if and only if S is Lagrangian, transverse to the fibers, and intersects each fiber in exactly one point. (Two submanifolds N_1, N_2 of a smooth manifold M are said to be *transverse* if $T_pN_1 + T_pN_2$ spans T_pM at each point $p \in N_1 \cap N_2$. See also Problem 8-17.)

12-12. Let M be a smooth manifold of dimension at least 1. Show that there is no 1-form σ on M such that the tautological form $\tau \in \Lambda^1(T^*M)$ is equal to the pullback $\pi^*\sigma$.

12-13. Let (M,ω) and $(\widetilde{M},\widetilde{\omega})$ be symplectic manifolds. Define a 2-form Ω on $M \times \widetilde{M}$ by

$$\Omega = \pi^*\omega - \widetilde{\pi}^*\widetilde{\omega}$$

where $\pi\colon M \times \widetilde{M} \to M$ and $\widetilde{\pi}\colon M \times \widetilde{M} \to \widetilde{M}$ are the projections.

(a) Show that Ω is symplectic.
(b) For a smooth map $F\colon M \to \widetilde{M}$, let $\Gamma(F) \subset M \times \widetilde{M}$ be the graph of F:

$$\Gamma(F) = \{(x,y) \in M \times \widetilde{M} : y = F(x)\}.$$

Show that F is a symplectomorphism if and only if $\Gamma(F)$ is a Lagrangian submanifold of $(M \times \widetilde{M}, \Omega)$.

12-14. The (*real*) *symplectic group* is the subgroup $\mathrm{Sp}(n,\mathbb{R}) \subset \mathrm{GL}(2n,\mathbb{R})$ consisting of $2n \times 2n$ matrices leaving the standard symplectic form $\omega = \sum_{i=1}^n dx^i \wedge dy^i$ invariant, that is, the set of invertible linear maps $A\colon \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ such that $A^*\omega = \omega$.

(a) Show that a matrix A is in $\mathrm{Sp}(n,\mathbb{R})$ if and only if it takes the standard basis to a symplectic basis.
(b) Show that $A \in \mathrm{Sp}(n,\mathbb{R})$ if and only if $A^T JA = J$, where J is given in block form as

$$J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}.$$

(c) Show that $\mathrm{Sp}(n,\mathbb{R})$ is an embedded Lie subgroup of $\mathrm{GL}(2n,\mathbb{R})$, and determine its dimension.
(d) Determine the Lie algebra of $\mathrm{Sp}(n,\mathbb{R})$ as a subalgebra of $\mathfrak{gl}(2n,\mathbb{R})$.
(e) Is $\mathrm{Sp}(n,\mathbb{R})$ compact?

12-15. Let $\Lambda_n \subset \mathrm{G}_n(\mathbb{R}^{2n})$ denote the set of Lagrangian subspaces of $\mathbb{R}^{2n}$.

(a) Show that $\mathrm{Sp}(n,\mathbb{R})$ acts transitively on Λ_n.
(b) Show that Λ_n has a unique smooth manifold structure such that the action of $\mathrm{Sp}(n,\mathbb{R})$ is smooth, and determine its dimension.
(c) Is Λ_n compact?

12-16. Let Q be a smooth manifold and let $S \subset Q$ be an embedded submanifold. Define the *conormal bundle* of S to be the subset $N^*S \subset T^*Q$ defined by

$$N^*S = \{(q,\eta) \in T^*Q : q \in S,\ \eta|_{T_qS} \equiv 0\}.$$

Show that N^*S is an embedded Lagrangian submanifold of T^*Q (with respect to the canonical symplectic structure on T^*Q).

12-17. Cartan's Lemma: Let M be a smooth n-manifold, and let $\omega^1,\dots,\omega^k$ be independent smooth 1-forms on an open subset $U \subset M$. If $\alpha^1,\dots,\alpha^k$ are 1-forms on U such that

$$\sum_{i=1}^{k} \alpha^i \wedge \omega^i = 0,$$

show that each α^i can be written as a linear combination of $\omega^1,\dots,\omega^k$ with smooth coefficients.

13
Orientations

When we introduced differential forms, we suggested that they should be objects that can be integrated over manifolds in an invariant way. Before we can do that, though, we need to address a serious issue that we have so far swept under the rug. This is the small matter of the positive and negative signs that arise when we try to interpret a k-covector as a machine for measuring k-dimensional volumes. In the previous chapter, we brushed this aside by saying that the value of a k-covector applied to a k-tuple of vectors has to be interpreted as a "signed volume" of the parallelepiped spanned by the vectors. These signs will cause problems, however, when we try to integrate differential forms on manifolds, for the simple reason that the transformation law for an n-form under a change of coordinates involves the determinant of the Jacobian, while the change of variables formula for multiple integrals involves the *absolute value* of the determinant.

In this chapter, we develop the theory of orientations, which is a systematic way to restrict to coordinate transformations with positive determinant, thus eliminating the sign problem. We begin with orientations of vector spaces, and then show how this theory can be carried over to manifolds. Along the way, we show that any nonorientable manifold has an orientable 2-sheeted covering space. Then we explore the ways in which orientations can be induced on hypersurfaces and on boundaries of manifolds with boundary.

In the last part of the chapter, we treat the special case of orientations on Riemannian manifolds and Riemannian hypersurfaces.

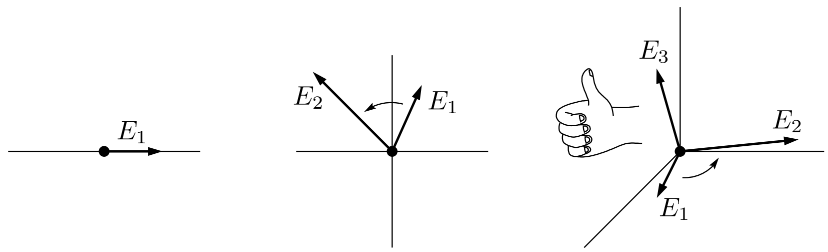


Figure 13.1. Oriented bases for $\mathbb{R}^1$, $\mathbb{R}^2$, and $\mathbb{R}^3$.

Orientations of Vector Spaces

The word "orientation" has some familiar meanings from our everyday experience, which can be interpreted as rules for singling out certain bases of $\mathbb{R}^1$, $\mathbb{R}^2$, and $\mathbb{R}^3$ (see Figure 13.1). For example, most people would understand an "orientation" of a line to mean a choice of preferred direction along the line, so we might declare an oriented basis for $\mathbb{R}^1$ to be one that points to the right (i.e., in the positive direction). A natural family of preferred bases for $\mathbb{R}^2$ is the ones for which the rotation from the first vector to the second is in the counterclockwise direction. And every student of vector calculus encounters "right-handed" bases in $\mathbb{R}^3$: These are the bases (E_1, E_2, E_3) with the property that when the fingers of your right hand curl from E_1 to E_2, your thumb points in the direction of E_3.

Although "to the right," "counterclockwise," and "right-handed" are not mathematical terms, it is easy to translate the rules for selecting oriented bases of $\mathbb{R}^1$, $\mathbb{R}^2$, and $\mathbb{R}^3$ into rigorous mathematical terms: You can check that in all three cases, the preferred bases are the ones whose transition matrix from the standard basis has positive determinant.

In an abstract vector space for which there is no canonical basis, we no longer have any way to determine which bases are "correctly oriented." For example, if V is the space of polynomials in one real variable of degree at most 2, who is to say which of the ordered bases $(1, x, x^2)$ or $(x^2, x, 1)$ is "right-handed"? All we can say in general is what it means for two bases to have the "same orientation."

Thus we are led to introduce the following definition. Let V be a vector space of dimension $n \geq 1$. We say two ordered bases $(E_1, \dots, E_n)$ and $(\widetilde{E}_1, \dots, \widetilde{E}_n)$ are *consistently oriented* if the transition matrix (B_i^j), defined by

$$E_i = B_i^j \widetilde{E}_j, \tag{13.1}$$

has positive determinant.

◇ **Exercise 13.1.** Show that being consistently oriented is an equivalence relation on the set of all ordered bases for V, and show that there are exactly two equivalence classes.

If $\dim V = n \geq 1$, we define an *orientation* for V as an equivalence class of ordered bases. If $(E_1, \ldots, E_n)$ is any ordered basis for V, we denote the orientation that it determines by $[E_1, \ldots, E_n]$. A vector space together with a choice of orientation is called an *oriented vector space*. If V is oriented, then any ordered basis $(E_1, \ldots, E_n)$ that is in the given orientation is said to be *oriented* or *positively oriented*. A basis that is not oriented is said to be *negatively oriented*.

For the special case of a zero-dimensional vector space V, we define an orientation of V to be simply a choice of one of the numbers ± 1.

Example 13.1. The orientation $[e_1, \ldots, e_n]$ of $\mathbb{R}^n$ determined by the standard basis is called the *standard orientation*. You should convince yourself that, in our usual way of representing the axes graphically, an oriented basis for $\mathbb{R}$ is one that points to the right; an oriented basis for $\mathbb{R}^2$ is one for which the rotation from the first vector to the second is counterclockwise; and an oriented basis for $\mathbb{R}^3$ is a right-handed one. (These can be taken as mathematical definitions for the words "right," "counterclockwise," and "right-handed.") The standard orientation for $\mathbb{R}^0$ is defined to be $+1$.

There is an important connection between orientations and alternating tensors, expressed in the following lemma.

Lemma 13.2. *Let V be a vector space of dimension $n \geq 1$, and suppose Ω is a nonzero element of $\Lambda^n(V)$. The set of ordered bases $(E_1, \ldots, E_n)$ such that $\Omega(E_1, \ldots, E_n) > 0$ is an orientation for V.*

Proof. Let $\mathcal{O}_\Omega$ denote the set of ordered bases on which Ω gives positive values. We need to show that $\mathcal{O}_\Omega$ is exactly one equivalence class.

Suppose (E_i) and $(\widetilde{E}_j)$ are any two bases for V, and let $B\colon V \to V$ be the linear map sending E_j to $\widetilde{E}_j$. This means that $\widetilde{E}_j = BE_j = B_j^i E_i$, so B is the transition matrix between the two bases. By Lemma 12.6,

$$\begin{aligned}\Omega(\widetilde{E}_1, \ldots, \widetilde{E}_n) &= \Omega(BE_1, \ldots, BE_n) \\ &= \det(B)\Omega(E_1, \ldots, E_n).\end{aligned}$$

It follows that $(\widetilde{E}_j)$ is consistently oriented with (E_i) if and only if $\Omega(E_1, \ldots, E_n)$ and $\Omega(\widetilde{E}_1, \ldots, \widetilde{E}_n)$ have the same sign, which is the same as saying that $\mathcal{O}_\Omega$ is one equivalence class. □

If V is an oriented vector space and Ω is an n-covector that determines the orientation of V as described in this lemma, we say that Ω is an *oriented* (or *positively oriented*) n-covector. For example, the n-covector $e^1 \wedge \cdots \wedge e^n$ is positively oriented for the standard orientation on $\mathbb{R}^n$.

Orientations of Manifolds

Let M be a smooth manifold. We define a *pointwise orientation* on M to be a choice of orientation of each tangent space. By itself, this is not a very useful concept, because the orientations of nearby points may have no relation to each other. For example, a pointwise orientation on $\mathbb{R}^n$ might switch randomly from point to point between the standard orientation and its opposite. In order for orientations to have some relationship with the smooth structure, we need an extra condition to ensure that the orientations of nearby tangent spaces are consistent with each other.

Suppose M is a smooth n-manifold with a given pointwise orientation. We say that a local frame (E_i) for M is (*positively*) *oriented* if $(E_1|_p, \dots, E_n|_p)$ is a positively oriented basis for T_pM at each point $p \in U$. A *negatively oriented* frame is defined analogously.

A pointwise orientation is said to be *continuous* if every point of M is in the domain of an oriented local frame. An *orientation* of M is a continuous pointwise orientation. An *oriented manifold* is a smooth manifold together with a choice of orientation. We say M is *orientable* if there exists an orientation for it, and *nonorientable* if not.

If M is zero-dimensional, this definition just means that an orientation of M is a choice of ± 1 attached to each of its points. The continuity condition is vacuous in this case, and the notion of oriented frames is not useful. Clearly every 0-manifold is orientable.

◇ **Exercise 13.2.** If M is an oriented manifold of dimension $n \geq 1$, show that every local frame with connected domain is either positively oriented or negatively oriented. Show that the connectedness assumption is necessary.

The next two propositions give ways of specifying orientations on manifolds that are somewhat more practical to use than the definition. A smooth coordinate chart on an oriented manifold is said to be (*positively*) *oriented* if the coordinate frame $(\partial/\partial x^i)$ is positively oriented, and *negatively oriented* if the coordinate frame is negatively oriented. A collection of smooth charts $\{(U_\alpha, \varphi_\alpha)\}$ is said to be *consistently oriented* if for each α, β, the transition map $\varphi_\beta \circ \varphi_\alpha^{-1}$ has positive Jacobian determinant everywhere on $\varphi_\alpha(U_\alpha \cap U_\beta)$.

Proposition 13.3. *Let M be a smooth positive-dimensional manifold. Given any open cover of M by consistently oriented smooth charts $\{(U_\alpha, \varphi_\alpha)\}$, there is a unique orientation for M with the property that each chart φ_α is oriented. Conversely, if M is oriented, then the collection of all oriented smooth charts is a consistently oriented cover of M.*

Proof. Suppose $\{(U_\alpha, \varphi_\alpha)\}$ is an open cover of M by consistently oriented smooth charts. For any $p \in M$, the consistency condition means that the transition matrix between the coordinate bases determined by any two of the charts has positive determinant. Thus the coordinate bases for all of the given charts determine the same orientation on T_pM. This defines a pointwise orientation on M. Each point of M is in the domain of at least one of the given charts, and the corresponding coordinate frame is oriented by definition, so this pointwise orientation is continuous. The converse is similar, and is left as an exercise. □

◇ **Exercise 13.3.** Complete the proof of Proposition 13.3.

Proposition 13.4. *Let M be a smooth manifold of dimension $n \geq 1$. Any nonvanishing n-form $\Omega \in \mathcal{A}^n(M)$ determines a unique orientation of M for which Ω is positively oriented at each point. Conversely, if M is given an orientation, then there is a smooth nonvanishing n-form on M that is positively oriented at each point.*

Remark. Because of this proposition, any nonvanishing n-form on an n-manifold is called an *orientation form*. If M is an oriented manifold and Ω is an orientation form determining the given orientation, we also say that Ω is (*positively*) *oriented*. It is easy to check that if Ω and $\widetilde{\Omega}$ are two positively oriented smooth forms on the same oriented manifold M, then $\widetilde{\Omega} = f\Omega$ for some strictly positive smooth real-valued function f.

If M is a 0-manifold, the proposition remains true if we interpret an orientation form as a nonvanishing real-valued function Ω, which assigns the orientation $+1$ to points where $\Omega > 0$ and -1 to points where $\Omega < 0$.

Proof. Let Ω be a nonvanishing n-form on M. Then Ω defines a pointwise orientation by Lemma 13.2, so all we need to check is that it is continuous. Let (x^i) be any smooth local coordinates on a connected domain $U \subset M$. On U, Ω has the coordinate expression $\Omega = f\, dx^1 \wedge \cdots \wedge dx^n$ for some continuous function f. The fact that Ω is nonvanishing means that f is nonvanishing, and therefore

$$\Omega\left(\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}\right) = f \neq 0$$

at all points of U. Since U is connected, it follows that this expression is either always positive or always negative on U, and therefore the coordinate chart is either positively oriented or negatively oriented. If negatively, we can replace x^1 by $-x^1$ to obtain a new coordinate chart for which the coordinate frame is positively oriented. Thus the pointwise orientation determined by Ω is continuous.

Conversely, suppose M is oriented, and let $\Lambda^n_+ M \subset \Lambda^n M$ be the open subset consisting of positively oriented n-covectors at all points of M. At any point $p \in M$, the intersection of $\Lambda^n_+ M$ with the fiber $\Lambda^n(T_pM)$ is an open half-line, and therefore convex. By the result of Problem 11-22, therefore, there exists a smooth global section of $\Lambda^n_+ M$ (i.e., a positively oriented smooth global n-form). □

◇ **Exercise 13.4.** Show that any open subset of an orientable manifold is orientable, and any product of orientable manifolds is orientable.

Suppose M and N are oriented positive-dimensional manifolds, and $F\colon M \to N$ is a local diffeomorphism. We say F is *orientation-preserving* if for each $p \in M$, F_* takes oriented bases of T_pM to oriented bases of $T_{F(p)}N$, and *orientation-reversing* if it takes oriented bases of T_pM to negatively oriented bases of $T_{F(p)}N$.

◇ **Exercise 13.5.** Show that a smooth map $F\colon M \to N$ is orientation-preserving if and only if its Jacobian matrix with respect to any oriented smooth charts for M and N has positive determinant, and orientation-reversing if and only if it has negative determinant.

Recall that a smooth manifold is said to be parallelizable if it admits a smooth global frame.

Proposition 13.5. *Every parallelizable manifold is orientable.*

Proof. Suppose M is parallelizable, and let $(E_1, \dots, E_n)$ be a global smooth frame for M. Define a pointwise orientation by declaring $(E_1|_p, \dots, E_n|_p)$ to be positively oriented at each $p \in M$. This pointwise orientation is continuous, because every point of M is in the domain of the (global) oriented frame (E_i). □

Example 13.6. The preceding proposition implies that Euclidean space $\mathbb{R}^n$, the n-torus $\mathbb{T}^n$, the spheres $\mathbb{S}^1$ and $\mathbb{S}^3$, and products of them are all orientable, because they are all parallelizable. Therefore any open subset of one of these manifolds is also orientable. Likewise, any Lie group is orientable because it is parallelizable.

In the case of Lie groups, we can say more. If G is a Lie group, an orientation of G is said to be *left-invariant* if L_g is orientation-preserving for every $g \in G$.

Lemma 13.7. *Every Lie group has precisely two left-invariant orientations, corresponding to the two choices of orientation of* $\operatorname{Lie}(G)$*.*

◇ **Exercise 13.6.** Prove the preceding lemma.

The Orientation Covering

There is a close relationship between orientability and covering maps. In this section we show that every nonorientable smooth manifold has an orientable two-sheeted covering manifold. The key idea is Lemma 13.2, which shows that each nonzero n-covector at a point $p \in M$ determines an orientation of T_pM, and two such n-covectors determine the same orientation if and only if they differ by a positive multiple. Thus we will construct a two-sheeted covering space of M in which the two points in the fiber over $p \in M$ are the two orientations of T_pM, considered as equivalence classes of nonzero n-covectors.

Let M be a smooth n-manifold, and let $\Lambda^n_* M$ denote the subset of $\Lambda^n M$ consisting of nonzero n-covectors. Observe that the Lie group $\mathbb{R}^+$ acts on $\Lambda^n_* M$ by multiplication in each fiber.

Lemma 13.8. *Let M be a smooth n-manifold. The natural action of $\mathbb{R}^+$ on $\Lambda^n_* M$ is smooth, free, and proper.*

Proof. The $\mathbb{R}^+$ action on $\Lambda^n_* M$ is clearly free, because if $c \in \mathbb{R}^+$ and Ω_p is a nonzero n-covector at a point $p \in M$, then $c\Omega_p = \Omega_p$ if and only if $c = 1$. To see that it is smooth, let (x^i) be smooth coordinates on an open set $U \subset M$. Because $\Lambda^n M$ is a rank 1 bundle, the nonvanishing smooth local section $dx^1 \wedge \cdots \wedge dx^n$ forms a smooth local frame for $\Lambda^n M$ over U. By Corollary 5.12, the map

$$u\, dx^1 \wedge \cdots \wedge dx^n|_x \mapsto (x^1, \ldots, x^n, u)$$

is a smooth coordinate chart on $\pi^{-1}(U)$ (where $\pi\colon \Lambda^n M \to M$ denotes the usual projection). In terms of these coordinates, the $\mathbb{R}^+$ action $\theta\colon \mathbb{R}^+ \times \Lambda^n_* M \to \Lambda^n_* M$ is given by

$$\theta(c, (x^1, \ldots, x^n, u)) = (x^1, \ldots, x^n, cu),$$

which is clearly smooth.

To see that the action is proper, we will use Proposition 9.13, which characterizes proper group actions in terms of convergent sequences. Suppose $\{(p_j, \Omega_j)\}$ is a convergent sequence of points in $\Lambda^n_* M$, and $\{c_j\}$ is a sequence in $\mathbb{R}^+$ such that $\{(p_j, c_j\Omega_j)\}$ also converges in $\Lambda^n_* M$. Then $\{p_j\}$ converges to some point $p \in M$. Choosing smooth coordinates (x^i) around p, we can write

$$\begin{aligned} \Omega_j &= u_j\, dx^1 \wedge \cdots \wedge dx^n|_{p_j}, \\ c_j\Omega_j &= c_j u_j\, dx^1 \wedge \cdots \wedge dx^n|_{p_j}. \end{aligned}$$

The hypothesis implies that $u_j \to u \neq 0$ and $c_j u_j \to u' \neq 0$, so c_j converges to u'/u. □

As a consequence of the preceding lemma, the quotient space $\widehat{M} = \Lambda^n_* M/\mathbb{R}^+$ is a smooth manifold of dimension n (because the total space of $\Lambda^n_* M$ has dimension $n+1$). We call $\widehat{M}$ the *orientation covering* of M. We will see below that it is a covering space of M when it is connected, which occurs precisely when M is not orientable.

The projection map $\pi\colon \Lambda^n_* M \to M$ is constant on $\mathbb{R}^+$-orbits, and so by Proposition 7.17 it descends to a smooth surjective map $\widehat{\pi}\colon \widehat{M} \to M$ such that the following diagram commutes:

$$\begin{tikzcd} \Lambda^n_* M \arrow[d, "q"'] \arrow[dr, "\pi"] & \\ \widehat{M} \arrow[r, "\widehat{\pi}"'] & M, \end{tikzcd}$$

where q is the quotient map defining $\widehat{M}$.

In order to handle the orientable and nonorientable cases in a uniform way, it is useful to expand our definition of covering maps slightly. If $\widehat{M}$ and M are smooth manifolds, let us say that a smooth map $\widehat{\pi}\colon \widehat{M} \to M$ is a *generalized covering map* if M is connected and every point of M has an evenly covered neighborhood U (in the sense that each component of $\widehat{\pi}^{-1}(U)$ is mapped diffeomorphically by $\widehat{\pi}$ onto U). In other words, $\widehat{\pi}$ satisfies all of the hypotheses of a smooth covering map except that $\widehat{M}$ might not be connected.

Theorem 13.9 (The Orientation Covering). *Let M be a smooth connected manifold, and let $\widehat{\pi}\colon \widehat{M} \to M$ be its orientation covering.*

(*a*) *$\widehat{\pi}$ is a generalized covering map.*

(*b*) *$\widehat{M}$ has a canonical orientation.*

(*c*) *M is orientable if and only if there exists a global section of $\widehat{\pi}$.*

(*d*) *$\widehat{M}$ is connected if and only if M is nonorientable, in which case $\widehat{\pi}$ is a smooth two-sheeted covering map.*

Proof. Because $\widehat{\pi}_* \circ q_* = \pi_*$ is surjective at each point, it follows that $\widehat{\pi}_*$ is also surjective and therefore $\widehat{\pi}$ is a submersion. For dimensional reasons, therefore, it is a local diffeomorphism.

To see that $\widehat{\pi}$ is a generalized covering map, let $(U, (x^i))$ be any connected smooth coordinate chart on M, so that $\Omega = dx^1 \wedge \cdots \wedge dx^n$ is a smooth local frame for $\Lambda^n M$ over U. Because of the correspondence between local frames and local trivializations (Proposition 5.10), it follows that $\pi^{-1}(U)$ is diffeomorphic to $U \times \mathbb{R}$, and thus $\pi^{-1}(U) \cap \Lambda^n_* M$ (the set of nonzero elements of $\pi^{-1}(U)$) has exactly two components:

$$U_+ = \{\omega_p : p \in U,\ \omega_p = c\Omega_p \text{ with } c > 0\},$$
$$U_- = \{\omega_p : p \in U,\ \omega_p = c\Omega_p \text{ with } c < 0\}.$$

Let $\widehat{U}_\pm = q(U_\pm) \subset \widehat{M}$ (Figure 13.2). Since the quotient map $q\colon \Lambda^n_* M \to \widehat{M}$ is an open map by Lemma 9.15, $\widehat{U}_+$ and $\widehat{U}_-$ are open in $\widehat{M}$. Because they are also connected (as continuous images of connected sets) and disjoint (as images of disjoint saturated sets), they are exactly the components of $\widehat{\pi}^{-1}(U)$.

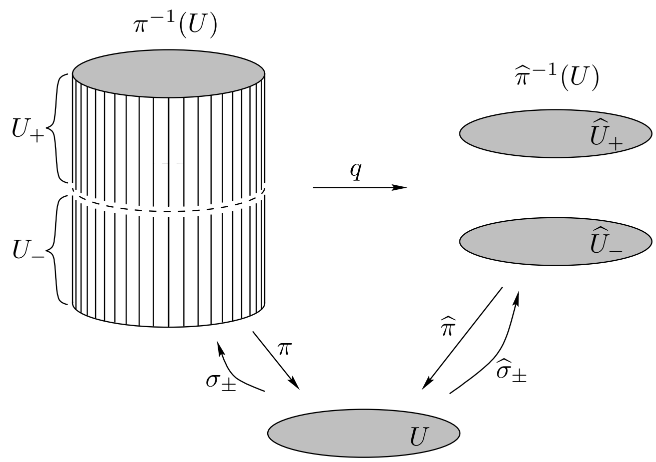


Figure 13.2. An evenly covered neighborhood in $\widehat{M}$.

We will show that U is evenly covered by showing that the restriction of $\widehat{\pi}$ to either $\widehat{U}_+$ or $\widehat{U}_-$ is a diffeomorphism onto U. We will do this by constructing local sections. First consider the maps $\sigma_\pm \colon U \to \Lambda^n_* M$ defined by $\sigma_\pm(p) = \pm\Omega_p = \pm dx^1 \wedge \cdots \wedge dx^n|_p$. These are both smooth sections of $\Lambda^n_* M$. Let $\widehat{\sigma}_\pm = q \circ \sigma_\pm \colon U \to \widehat{U}_\pm$. More explicitly,

$$\widehat{\sigma}_+(p) = [\Omega_p], \qquad \widehat{\sigma}_-(p) = [-\Omega_p],$$

where the brackets represent equivalence classes under the $\mathbb{R}^+$ action. The maps $\widehat{\sigma}_\pm$ are smooth because they are compositions of smooth maps, and they are local sections of $\widehat{\pi}$ because $\widehat{\pi} \circ \widehat{\sigma}_\pm(p) = p$. On the other hand, an arbitrary point in $\widehat{U}_+$ is of the form $[c\Omega_p]$ for some $c > 0$ and some $p \in U$, and thus

$$\widehat{\sigma}_+ \circ \widehat{\pi}[c\Omega_p] = \widehat{\sigma}_+(p) = [\Omega_p].$$

Because Ω_p is in the same orbit as $c\Omega_p$, this shows that $\widehat{\sigma}_+ \circ \widehat{\pi}$ is the identity on $\widehat{U}_+$. A similar argument shows that $\widehat{\sigma}_- \circ \widehat{\pi}$ is the identity on $\widehat{U}_-$. Therefore $\widehat{\pi}$ is a diffeomorphism on each component of $\widehat{\pi}^{-1}(U)$, which completes the proof that $\widehat{\pi}$ is a generalized covering map.

Next we will prove (b). Let $\widehat{q}$ be a point in $\widehat{M}$, and let $q = \widehat{\pi}(\widehat{q}) \in M$. Because $\widehat{q}$ is by definition an equivalence class of n-covectors on T_qM under multiplication by positive reals, it determines a unique orientation of T_qM. We give $T_{\widehat{q}}\widehat{M}$ the orientation such that $\widehat{\pi}_* \colon T_{\widehat{q}}\widehat{M} \to T_qM$ is orientation-preserving. To show that this pointwise orientation is continuous, let $(U, (x^i))$ be a smooth chart containing q, which we may assume without loss of generality determines the same orientation on T_qM as $\widehat{q}$; and let $\widehat{\sigma}_+ \colon U \to \widehat{M}$ be the local section determined by $\Omega = dx^1 \wedge \cdots \wedge dx^n$

as in the preceding paragraph. By definition, for each $p \in U$, $\widehat{\sigma}_+(p) = [\Omega_p]$ is the orientation determined by Ω_p, and thus the canonical orientation on $T_{\widehat{\sigma}_+(p)}\widehat{M}$ is the one determined by $\widehat{\pi}^*(\Omega_p)$. It follows that $\widehat{\pi}^*\Omega$ is a positively oriented smooth n-form on a neighborhood of $\widehat{q}$, so the orientation is continuous.

To prove (c), suppose we are given an orientation of M. Then there exists a smooth nonvanishing global n-form Ω on M, which we can think of as a global section of $\Lambda^n_* M$. The composite map $q \circ \Omega$ is therefore a global section of $\widehat{M}$. Conversely, given a section σ of $\widehat{M}$, let $\widehat{\Omega}$ be an orientation form on $\widehat{M}$. Since σ is a local diffeomorphism, it follows that $\sigma^*\widehat{\Omega}$ is an orientation form on M.

Finally, we will prove (d), by showing that M is orientable if and only if $\widehat{M}$ is not connected. First suppose that M is orientable. By (c), there exists a global section $\sigma\colon M \to \widehat{M}$. Because σ is a local homeomorphism, it is an open map; and because it is a section of $\widehat{\pi}$, it is a proper map by Lemma 2.16 and hence closed. Therefore, $\sigma(M)$ is an open and closed subset of $\widehat{M}$. Since it contains exactly one point in each fiber, it is neither empty nor equal to $\widehat{M}$ itself, so $\widehat{M}$ is not connected.

Conversely, suppose that $\widehat{M}$ is not connected. Let W be a component of $\widehat{M}$, and let $N\colon W \to \mathbb{R}$ be the function that assigns to each $p \in M$ the cardinality of the set $\widehat{\pi}^{-1}(p) \cap W$ (which must be equal to 0, 1, or 2). We will show that N is constant on M. If $p \in M$ and U is an evenly covered neighborhood of p, let $\widehat{U}_1$, $\widehat{U}_2$ denote the two components of $\widehat{\pi}^{-1}(U)$. By connectivity, each set $\widehat{U}_i$ must be contained in a single component of $\widehat{M}$; thus N is identically equal to 0, 1, or 2 on U. It follows that N is constant in a neighborhood of each point, and because M is connected, it is constant on M.

If N is zero everywhere, then W is empty, which is clearly impossible because W is a component of $\widehat{M}$; while if N is equal to 2 everywhere, then $W = \widehat{M}$, which contradicts the assumption that $\widehat{M}$ is disconnected. The only remaining possibility is that $N \equiv 1$. Thus W contains exactly one point in each fiber, and we can define a global section $\sigma\colon M \to \widehat{M}$ by letting $\sigma(p)$ be the unique point in $\pi^{-1}(p) \cap W$. By the result of part (c), this shows that M is orientable. $\square$

By invoking some covering space theory, we obtain the following sufficient topological condition for orientability. If G is a group and $H \subset G$ is a subgroup, the *index* of H in G is the cardinality of the set of cosets of H in G. (If H is normal, it is just the cardinality of the quotient group G/H.)

Corollary 13.10. *Let M be a connected smooth manifold, and suppose that the fundamental group of M has no subgroup of index 2. Then M is orientable. In particular, this is the case if M is simply connected.*

Proof. Suppose M is not orientable, and let $\widehat{\pi}\colon \widehat{M} \to M$ be its orientation covering, which is an honest covering map in this case. Choose any point $\widehat{q} \in$

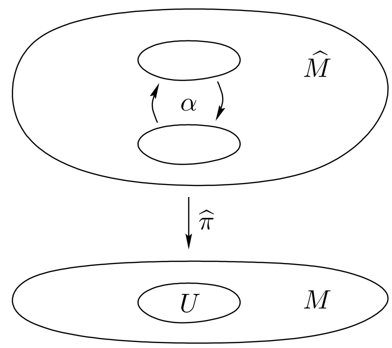


Figure 13.3. The nontrivial covering transformation of $\widehat{M}$.

$\widehat{M}$, and let $q = \widehat{\pi}(\widehat{q}) \in M$. Let $\alpha\colon \widehat{M} \to \widehat{M}$ be the map that interchanges the two points in each fiber of $\widehat{\pi}$ (Figure 13.3). If $U \subset M$ is any evenly covered open set, then α just maps each component of $\widehat{\pi}^{-1}(U)$ diffeomorphically onto the other, so α is a smooth covering transformation. In fact, since a covering transformation is determined by what it does to one point, α is the unique nontrivial element of the covering group $\mathcal{C}_{\widehat{\pi}}(\widehat{M})$, which is therefore equal to the two-element group $\{\mathrm{Id}_{\widehat{M}}, \alpha\}$. Because the covering group acts transitively on fibers, $\widehat{\pi}$ is a normal covering map. Let H denote the subgroup $\widehat{\pi}_*(\pi_1(\widehat{M}, \widehat{q}))$ of $\pi_1(M, q)$. A fundamental result in the theory of covering spaces (see, e.g., [Lee00, Corollary 11.31]) is that the quotient group $\pi_1(M, q)/H$ is isomorphic to $\mathcal{C}_{\widehat{\pi}}(\widehat{M})$. Therefore H has index 2 in $\pi_1(M, q)$. □

Orientations of Hypersurfaces

If M is an oriented manifold and N is a submanifold of M, N may not inherit an orientation from M, even if N is embedded. Clearly it is not sufficient to restrict an orientation form from M to N, since the restriction of an n-form to a manifold of lower dimension must necessarily be zero. A useful example to consider is the Möbius band, which is not orientable (see Problem 13-13), even though it can be embedded in $\mathbb{R}^3$.

In this section, we will restrict our attention to immersed or embedded hypersurfaces. With one extra piece of information (a certain kind of vector field along the hypersurface), we can use an orientation on M to induce an orientation on any hypersurface $S \subset M$.

We start with some definitions. Let V be a finite-dimensional vector space, and let $X \in V$. We define a linear map $i_X\colon \Lambda^k(V) \to \Lambda^{k-1}(V)$, called *interior multiplication* or *contraction* with X, by

$$i_X\omega(Y_1,\dots,Y_{k-1}) = \omega(X,Y_1,\dots,Y_{k-1}).$$

In other words, $i_X\omega$ is obtained from ω by inserting X into the first slot. By convention, we interpret $i_X\omega$ to be zero when ω is a 0-covector (i.e., a number). Another common notation is

$$X \lrcorner\, \omega = i_X\omega.$$

Interior multiplication shares two important properties with exterior differentiation: They are both antiderivations whose square is zero, as the following lemma shows.

Lemma 13.11. *Let V be a finite-dimensional vector space and $X \in V$.*

(*a*) $i_X \circ i_X = 0$.

(*b*) i_X *is an antiderivation: If ω is a k-covector and η is an l-covector,*

$$i_X(\omega\wedge\eta) = (i_X\omega)\wedge\eta + (-1)^k\omega\wedge(i_X\eta). \tag{13.2}$$

Proof. On k-covectors for $k \geq 2$, part (a) is immediate from the definition, because any alternating tensor gives zero when two of its arguments are identical. On 1-covectors and 0-covectors, it follows from the fact that $i_X \equiv 0$ on 0-covectors.

To prove (b), it suffices to consider the case in which both ω and η are wedge products of 1-covectors (such an alternating tensor is said to be *decomposable*), since every alternating tensor can be written as a linear combination of decomposable ones. It is easy to verify that (b) will follow in this special case from the following general formula for covectors $\omega^1,\dots,\omega^k$:

$$X \lrcorner\, (\omega^1\wedge\cdots\wedge\omega^k) = \sum_{i=1}^{k}(-1)^{i-1}\omega^i(X)\omega^1\wedge\cdots\wedge\widehat{\omega^i}\wedge\cdots\wedge\omega^k, \tag{13.3}$$

where the hat indicates that ω^i is omitted.

To prove (13.3), let us write $X_1 = X$ and apply both sides to vectors $(X_2,\dots,X_k)$; then what we have to prove is

$$\begin{aligned}(\omega^1\wedge\cdots\wedge\omega^k)&(X_1,\dots,X_k)\\ &= \sum_{i=1}^{k}(-1)^{i-1}\omega^i(X_1)(\omega^1\wedge\cdots\wedge\widehat{\omega^i}\wedge\cdots\wedge\omega^k)(X_2,\dots,X_k).\end{aligned} \tag{13.4}$$

The left-hand side of (13.4) is the determinant of the matrix $\mathbb{X}$ whose (i,j)-entry is $\omega^i(X_j)$. To simplify the right-hand side, let $\mathbb{X}^i_j$ denote the $(k-1)\times(k-1)$ minor of $\mathbb{X}$ obtained by deleting the ith row and jth column. Then the right-hand side of (13.4) is

$$\sum_{i=1}^{k}(-1)^{i-1}\omega^i(X_1)\det\mathbb{X}^i_1.$$

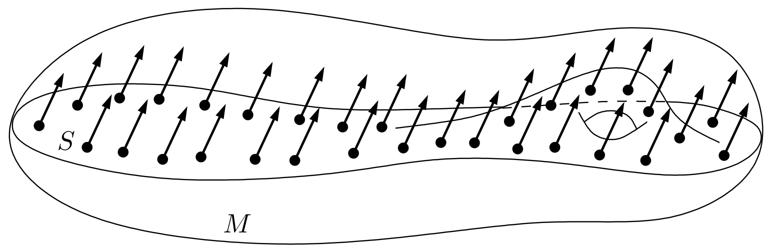


Figure 13.4. A vector field along a submanifold.

This is just the expansion of $\det \mathbb{X}$ by minors along the first column, and therefore is equal to $\det \mathbb{X}$. $\square$

It should be noted that when the wedge product is defined using the Alt convention, interior multiplication has to be defined with an extra factor of k:

$$\bar{i}_X \omega(Y_1, \ldots, Y_{k-1}) = k\omega(X, Y_1, \ldots, Y_{k-1}).$$

This definition ensures that interior multiplication i_X is still an antiderivation; the factor of k is needed to compensate for the difference between the factors of $1/k!$ and $1/(k-1)!$ that occur when the left-hand and right-hand sides of (13.4) are evaluated using the Alt convention.

On a smooth manifold M, interior multiplication extends naturally to vector fields and differential forms, simply by letting it act pointwise: if $X \in \mathfrak{T}(M)$ and $\omega \in \mathcal{A}^k(M)$, define a $(k-1)$-form $X \lrcorner \omega = i_X\omega$ by

$$(X \lrcorner \omega)_p = X_p \lrcorner \omega_p.$$

◇ **Exercise 13.7.** Let X be a smooth vector field on M.

(a) Verify that $i_X : \mathcal{A}^k(M) \to \mathcal{A}^{k-1}(M)$ is linear over $C^\infty(M)$ and therefore corresponds to a smooth bundle map $i_X : \Lambda^k M \to \Lambda^{k-1} M$.

(b) Show that i_X is an antiderivation of $\mathcal{A}^*(M)$ of degree -1 whose square is zero.

Now suppose M is a smooth manifold and $S \subset M$ is a submanifold (immersed or embedded). A *vector field along* S is a continuous map $N : S \to TM$ with the property that $N_p \in T_pM$ for each $p \in S$ (Figure 13.4). (Note the difference between this and a vector field *on* S, which would have to satisfy $N_p \in T_pS$ at each point.) A vector $N_p \in T_pM$ at some point $p \in S$ is said to be *transverse to* S if T_pM is spanned by N_p and T_pS. Similarly, a vector field N along S is transverse to S if N_p is transverse to T_pS at each $p \in S$.

For example, any smooth vector field on M restricts to a smooth vector field along S. If S is a hypersurface, such a vector field is transverse if and only if it is nowhere tangent to S.

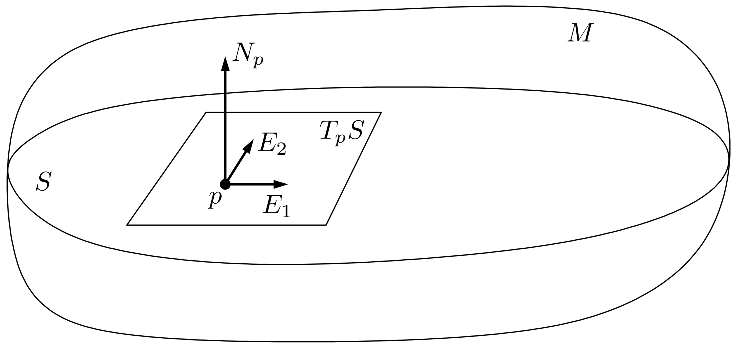


Figure 13.5. The orientation induced by a transverse vector field.

Proposition 13.12. *Suppose M is an oriented smooth n-manifold, S is an immersed hypersurface in M, and N is a transverse vector field along S. Then S has a unique orientation with the property that $(E_1, \dots, E_{n-1})$ is an oriented basis for T_pS if and only if $(N_p, E_1, \dots, E_{n-1})$ is an oriented basis for T_pM. If Ω is an orientation form for M, then $(N \lrcorner \Omega)|_S$ is an orientation form for S with respect to this orientation.*

Remark. See Figure 13.5 for an illustration of the $n = 3$ case. When $n = 1$, since S is a 0-manifold, this proposition should be interpreted as follows: At each point $p \in S$, we assign the orientation $+1$ to p if N_p is an oriented basis for T_pM, and -1 if N_p is negatively oriented. With this understanding, the proof below goes through in this case without modification.

Proof. Let Ω be an orientation form for M. Then $\omega = (N \lrcorner \Omega)|_S$ is an $(n-1)$-form on S. It will be an orientation form for S if we can show that it never vanishes. Given any basis $(E_1, \dots, E_{n-1})$ for T_pS, the fact that N is transverse to S implies that $(N_p, E_1, \dots, E_{n-1})$ is a basis for T_pM. The fact that Ω is nonvanishing implies that

$$\omega_p(E_1, \dots, E_{n-1}) = \Omega_p(N_p, E_1, \dots, E_{n-1}) \neq 0.$$

Since $\omega_p(E_1, \dots, E_{n-1}) > 0$ if and only if $\Omega_p(N_p, E_1, \dots, E_n) > 0$, the orientation determined by ω is the one defined in the statement of the proposition. □

Example 13.13. Considering $\mathbb{S}^n$ as a hypersurface in $\mathbb{R}^{n+1}$, the vector field $N = x^i \partial/\partial x^i$ along $\mathbb{S}^n$ is easily seen to be transverse, so it induces an orientation on $\mathbb{S}^n$. This shows that all spheres are orientable. We define the *standard orientation* of $\mathbb{S}^n$ to be the orientation determined by N. Unless otherwise specified, we will always use the standard orientation on $\mathbb{S}^n$. (The standard orientation on $\mathbb{S}^0$ is the one that assigns the orientation $+1$ to the point $+1 \in \mathbb{S}^0$ and -1 to $-1 \in \mathbb{S}^0$.)

Not every hypersurface admits a transverse vector field. (See Problem 13-15.) However, the following lemma gives a sufficient condition that holds in many cases.

Lemma 13.14. *Let M be an oriented smooth manifold, and suppose $S \subset M$ is a regular level set of a smooth function $f\colon M \to \mathbb{R}$. Then S is orientable.*

Proof. Let g be any Riemannian metric on M, and let $N = \operatorname{grad} f|_S$. The hypotheses imply that N is a transverse vector field along S, so the result follows from Proposition 13.12. □

Boundary Orientations

An important application of the construction of the preceding section is to define a canonical orientation on the boundary of any oriented manifold with boundary. First, we note that an orientation of a smooth manifold with boundary can be defined exactly as in the case of a smooth manifold.

One situation that arises frequently is the following. If M is a smooth n-manifold, a smooth, compact, embedded n-dimensional submanifold with boundary $D \subset M$ is called a *regular domain* in M. An orientation on M immediately yields an orientation on D, for example by restricting an orientation n-form to D. Examples are the closed unit ball in $\mathbb{R}^n$ and the closed upper hemisphere in $\mathbb{S}^n$, each of which inherits an orientation from its containing manifold.

If M is a smooth manifold with boundary, ∂M is easily seen to be an embedded hypersurface in M. Any point $p \in \partial M$ is in the domain of a smooth boundary chart (U, φ) such that $\varphi(p) \in \partial\mathbb{H}^n$, and it follows from the result of Problem 7-7 that φ takes *every* point in $\partial M \cap U$ to $\partial\mathbb{H}^n$. Thus $\varphi(U \cap \partial M)$ is the "slice" $\varphi(U) \cap \partial\mathbb{H}^n$, and boundary charts play a role for ∂M analogous to that played by slice charts for ordinary embedded submanifolds.

Let $p \in \partial M$. A vector $N \in T_pM$ is said to be *inward-pointing* if $N \notin T_p\partial M$ and for some $\varepsilon > 0$ there exists a smooth curve segment $\gamma\colon [0, \varepsilon] \to M$ such that $\gamma(0) = p$ and $\gamma'(0) = N$. It is said to be *outward-pointing* if $-N$ is inward-pointing. The following lemma gives another characterization of inward-pointing vectors, which is usually much easier to check. (See Figure 13.6.)

Lemma 13.15. *Suppose M is a smooth manifold with boundary, $p \in \partial M$, and (x^i) are any smooth boundary coordinates in a neighborhood of p. The inward-pointing vectors in T_pM are precisely those with positive x^n-component, and the outward-pointing ones are those with negative x^n-component.*

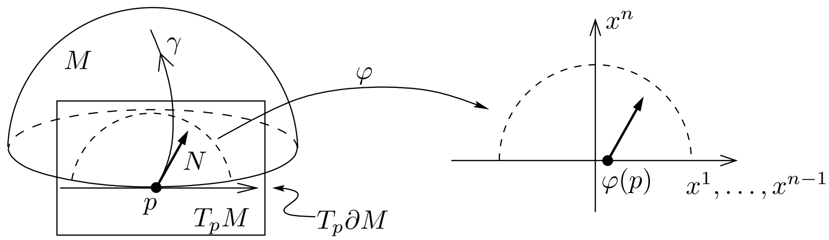


Figure 13.6. An inward-pointing vector field.

◇ **Exercise 13.8.** Prove Lemma 13.15.

A vector field along ∂M (defined just as for ordinary hypersurfaces) is said to be inward-pointing or outward-pointing if its value at each point has that property.

Lemma 13.16. *If M is any smooth manifold with boundary, there is a smooth outward-pointing vector field along ∂M.*

Proof. Cover a neighborhood of ∂M by smooth boundary charts $\{(U_\alpha, \varphi_\alpha)\}$. In each such chart, $N_\alpha = -\partial/\partial x^n|_{\partial M \cap U_\alpha}$ is a smooth vector field along $\partial M \cap U_\alpha$, which is outward-pointing by Lemma 13.15. Let $\{\psi_\alpha\}$ be a smooth partition of unity subordinate to the cover $\{U_\alpha \cap \partial M\}$ of ∂M, and define a global vector field N along ∂M by

$$N = \sum_\alpha \psi_\alpha N_\alpha.$$

Clearly N is a smooth vector field along ∂M. To show that it is outward-pointing, let $(y^1, \dots, y^n)$ be any smooth boundary coordinates in a neighborhood of $p \in \partial M$. Because each N_α is outward-pointing, it satisfies $dy^n(N_\alpha) < 0$. Therefore, the y^n-component of N at p satisfies

$$dy^n(N_p) = \sum_\alpha \psi_\alpha(p) dy^n(N_\alpha|_p).$$

This sum is strictly negative, because each term is nonpositive and at least one term is negative. □

Proposition 13.17 (The Induced Orientation on a Boundary). *Let M be an oriented smooth manifold with boundary. Then ∂M is orientable, and the orientation determined by any outward-pointing vector field along ∂M is independent of the choice of vector field.*

Remark. The orientation on ∂M determined by any outward-pointing vector field is called the *induced orientation* or the *Stokes orientation* on ∂M. (The second term is chosen because of the role this orientation will play in Stokes's theorem, to be described in Chapter 14.)

Proof. Let $n = \dim M$, and let Ω be an orientation form for M. By Lemma 13.16, there exists a smooth outward-pointing vector field N along ∂M. The $(n-1)$-form $(N \lrcorner \Omega)|_{\partial M}$ is an orientation form for ∂M by Proposition 13.12, so ∂M is orientable. It remains only to show that this orientation is independent of the choice of N.

Let $(x^1, \ldots, x^n)$ be smooth boundary coordinates for M in a neighborhood of $p \in \partial M$. Replacing x^1 by $-x^1$ if necessary, we may assume they are oriented coordinates, which implies that $\Omega = f\, dx^1 \wedge \cdots \wedge dx^n$ (locally) for some strictly positive function f. Because $x^n = 0$ along ∂M, the restriction $dx^n|_{\partial M}$ is equal to zero (Exercise 8.6). Therefore, using the antiderivation property of i_N,

$$\begin{aligned}(N \lrcorner \Omega)|_{\partial M} &= f \sum_{i=1}^{n} (-1)^{i-1} dx^i(N) dx^1|_{\partial M} \wedge \cdots \wedge \widehat{dx^i}|_{\partial M} \wedge \cdots \wedge dx^n|_{\partial M} \\ &= (-1)^{n-1} f dx^n(N) dx^1|_{\partial M} \wedge \cdots \wedge dx^{n-1}|_{\partial M}.\end{aligned}$$

Because $dx^n(N) = N^n < 0$, this is a positive multiple of $(-1)^n dx^1|_{\partial M} \wedge \cdots \wedge dx^{n-1}|_{\partial M}$. If $\widetilde{N}$ is any other outward-pointing vector field, the same computation shows that $(\widetilde{N} \lrcorner \Omega)|_{\partial M}$ is a positive multiple of the same $(n-1)$-form, and thus a positive multiple of $(N \lrcorner \Omega)|_{\partial M}$. This proves that N and $\widetilde{N}$ determine the same orientation on ∂M. □

Example 13.18. This proposition gives a simpler proof that $\mathbb{S}^n$ is orientable, because it is the boundary of the closed unit ball. The orientation thus induced on $\mathbb{S}^n$ is the standard one, as you can check.

Example 13.19. Let us determine the induced orientation on $\partial\mathbb{H}^n$ when $\mathbb{H}^n$ itself has the standard orientation inherited from $\mathbb{R}^n$. We can identify $\partial\mathbb{H}^n$ with $\mathbb{R}^{n-1}$ under the correspondence $(x^1, \ldots, x^{n-1}, 0) \leftrightarrow (x^1, \ldots, x^{n-1})$. Since the vector field $-\partial/\partial x^n$ is outward-pointing along $\partial\mathbb{H}^n$, the standard coordinate frame for $\mathbb{R}^{n-1}$ is positively oriented for $\partial\mathbb{H}^n$ if and only if $[-\partial/\partial x^n, \partial/\partial x^1, \ldots, \partial/\partial x^{n-1}]$ is the standard orientation for $\mathbb{R}^n$. This orientation satisfies

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

Thus the induced orientation on $\partial\mathbb{H}^n$ is equal to the standard orientation on $\mathbb{R}^{n-1}$ when n is even, but it is *opposite* to the standard orientation when n is odd. In particular, the standard coordinates on $\partial\mathbb{H}^n \approx \mathbb{R}^{n-1}$ are positively oriented if and only if n is even. (This fact will play an important role in the proof of Stokes's theorem below.)

For many purposes, the most useful way of describing submanifolds is by means of local parametrizations. The next lemma gives a useful criterion for checking whether a local parametrization of a boundary is orientation-preserving.

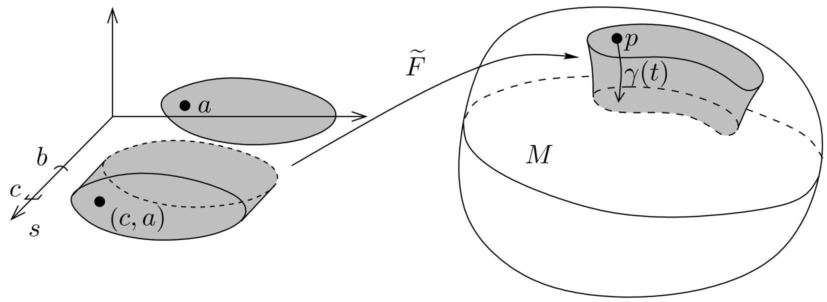


Figure 13.7. Orientation criterion for a boundary parametrization.

Lemma 13.20. *Let M be an oriented n-manifold with boundary, and let $F\colon U \to M$ be a smooth local parametrization of ∂M, where U is a connected open subset of $\mathbb{R}^{n-1}$. Suppose that for some $b < c \in \mathbb{R}$, F admits an extension to a smooth immersion $\widetilde{F}\colon (b, c] \times U \to M$ such that $\widetilde{F}(c, x) = F(x)$. Then F is orientation-preserving for ∂M (with the induced orientation) if and only if $\widetilde{F}$ is orientation-preserving for M.*

Proof. Let a be an arbitrary point of U, and let $p = F(a) = \widetilde{F}(c, a) \in \partial M$ (Figure 13.7). The hypothesis that $\widetilde{F}$ is an immersion means that $\widetilde{F}_*\colon (T_c\mathbb{R} \oplus T_a\mathbb{R}^{n-1}) \to T_pM$ is injective (actually bijective for dimensional reasons). Since the restriction of $\widetilde{F}_*$ to $T_a\mathbb{R}^{n-1}$ is equal to $F_*\colon T_a\mathbb{R}^{n-1} \to T_p\partial M$, which is already injective, it follows that $\widetilde{F}_*(\partial/\partial s|_{(c,a)}) \notin T_p\partial M$ (where s denotes the coordinate on $(b, c]$).

Define a smooth curve $\gamma\colon [0, \varepsilon) \to M$ by

$$\gamma(t) = \widetilde{F}(c - t, a).$$

This curve satisfies $\gamma(0) = p$ and $\gamma'(0) = -\widetilde{F}_*(\partial/\partial s|_{(c,a)}) \notin T_p\partial M$. It follows that $-\widetilde{F}_*(\partial/\partial s|_{(c,a)})$ is inward-pointing, and therefore $\widetilde{F}_*(\partial/\partial s|_{(c,a)})$ is outward-pointing at p, and by continuity on all of $F(U)$.

By definition of the induced orientation on ∂M, we have the following equivalences:

$$\begin{aligned}
&\widetilde{F} \text{ is orientation-preserving for } M \\
&\iff (\widetilde{F}_*\partial/\partial s, \widetilde{F}_*\partial/\partial x^1, \dots, \widetilde{F}_*\partial/\partial x^{n-1}) \text{ is oriented for } TM \\
&\iff (F_*\partial/\partial x^1, \dots, F_*\partial/\partial x^{n-1}) \text{ is oriented for } T\partial M \\
&\iff F \text{ is orientation-preserving for } \partial M,
\end{aligned}$$

which was to be proved. □

Here is an illustration of how the lemma can be used.

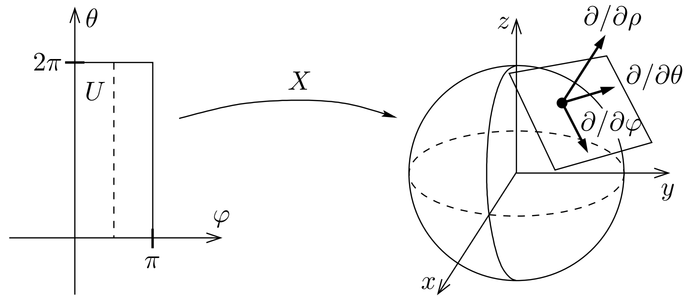


Figure 13.8. Parametrizing the sphere via spherical coordinates.

Example 13.21. The spherical coordinate parametrization of $\mathbb{R}^3$ (Example 7.11) restricts to a smooth local parametrization of $\mathbb{S}^2$ as follows. Let U be the open rectangle $(0,\pi)\times(0,2\pi)\subset\mathbb{R}^2$, and let $X\colon U\to\mathbb{S}^2$ be the following map:

$$X(\varphi,\theta)=(\sin\varphi\cos\theta,\sin\varphi\sin\theta,\cos\varphi)$$

(Figure 13.8). We can check whether X preserves or reverses orientation by using the fact that it is the restriction of the 3-dimensional spherical coordinate parametrization. If we define $\widetilde{X}\colon(0,1]\times U\to\overline{\mathbb{B}^3}$ by

$$\widetilde{X}(\rho,\varphi,\theta)=(\rho\sin\varphi\cos\theta,\rho\sin\varphi\sin\theta,\rho\cos\varphi),$$

then $\widetilde{X}(1,\theta,\rho)=X(\theta,\rho)$, so the hypotheses of Lemma 13.20 are satisfied. By direct computation, the Jacobian determinant of $\widetilde{X}$ is $\rho^2\sin\varphi$, which is positive on D. By virtue of Lemma 13.20, X is orientation-preserving.

The Riemannian Volume Form

In this section, we explore how the theory of orientations can be specialized to Riemannian manifolds. Let (M,g) be an oriented Riemannian manifold. We know from Proposition 11.17 that there is a smooth orthonormal frame $(E_1,\dots,E_n)$ in a neighborhood of each point of M. By replacing E_1 by $-E_1$ if necessary, we can find an *oriented* orthonormal frame in a neighborhood of each point.

Proposition 13.22. *Suppose (M,g) is an oriented Riemannian n-manifold. There is a unique smooth orientation form $\Omega\in\mathcal{A}^n(M)$ such that*

$$\Omega(E_1,\dots,E_n)=1 \tag{13.5}$$

for every oriented local orthonormal frame (E_i) for M.

Remark. The n-form whose existence and uniqueness are guaranteed by this proposition is called the *Riemannian volume form*, or sometimes the *Riemannian volume element*. Because of the role it plays in integration on Riemannian manifolds, as we will see in the next chapter, it is often denoted by dV_g (or dA_g or ds_g in the 2-dimensional or 1-dimensional case, respectively). Be warned, however, that this notation is *not* meant to imply that the volume form is the exterior derivative of an $(n-1)$-form; in fact, as we will see in Chapter 15, this is never the case on a compact manifold. You should just interpret dV_g as a notational convenience.

Proof. Suppose first that such a form Ω exists. If $(E_1, \dots, E_n)$ is any local oriented orthonormal frame and $(\varepsilon^1, \dots, \varepsilon^n)$ is the dual coframe, we can write $\Omega = f\,\varepsilon^1 \wedge \dots \wedge \varepsilon^n$ locally. The condition (13.5) then reduces to $f = 1$, so

$$\Omega = \varepsilon^1 \wedge \dots \wedge \varepsilon^n. \tag{13.6}$$

This proves that such a form is uniquely determined.

To prove existence, we would like to *define* Ω in a neighborhood of each point by (13.6). If $(\widetilde{E}_1, \dots, \widetilde{E}_n)$ is another oriented orthonormal frame, with dual coframe $(\widetilde{\varepsilon}^1, \dots, \widetilde{\varepsilon}^n)$, let

$$\widetilde{\Omega} = \widetilde{\varepsilon}^1 \wedge \dots \wedge \widetilde{\varepsilon}^n.$$

We can write

$$\widetilde{E}_i = A_i^j E_j$$

for some matrix (A_i^j) of smooth functions. The fact that both frames are orthonormal means that $(A_i^j(p)) \in \mathrm{O}(n)$ for each p, so $\det(A_i^j) = \pm 1$, and the fact that the two frames are consistently oriented forces the positive sign. We compute

$$\begin{aligned} \Omega(\widetilde{E}_1, \dots, \widetilde{E}_n) &= \det(\varepsilon^j(\widetilde{E}_i)) \\ &= \det(A_i^j) \\ &= 1 \\ &= \widetilde{\Omega}(\widetilde{E}_1, \dots, \widetilde{E}_n). \end{aligned}$$

Thus $\Omega = \widetilde{\Omega}$, so defining Ω in a neighborhood of each point by (13.6) with respect to some smooth oriented orthonormal frame yields a global n-form. The resulting form is clearly smooth and satisfies (13.5) for every oriented orthonormal basis. □

Although the expression for the Riemannian volume form with respect to an oriented orthonormal frame is particularly simple, it is also useful to have an expression for it in coordinates.

Lemma 13.23. *Let (M, g) be an oriented Riemannian manifold. In any oriented smooth coordinates (x^i), the Riemannian volume form has the local coordinate expression*

$$dV_g = \sqrt{\det(g_{ij})}\, dx^1 \wedge \cdots \wedge dx^n,$$

where g_{ij} are the components of g in these coordinates.

Proof. Let $(U, (x^i))$ be an oriented smooth chart, and let $p \in M$. In these coordinates, $dV_g = f\, dx^1 \wedge \cdots \wedge dx^n$ for some positive coefficient function f. To compute f, let (E_i) be any smooth oriented orthonormal frame defined on a neighborhood of p, and let (ε^i) be the dual coframe. If we write the coordinate frame in terms of the orthonormal frame as

$$\frac{\partial}{\partial x^i} = A_i^j E_j,$$

then we can compute

$$\begin{aligned}
f &= dV_g\left(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}\right)\\
&= \varepsilon^1 \wedge \cdots \wedge \varepsilon^n \left(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}\right)\\
&= \det\left(\varepsilon^j\left(\frac{\partial}{\partial x^i}\right)\right)\\
&= \det(A_i^j).
\end{aligned}$$

On the other hand, observe that

$$\begin{aligned}
g_{ij} &= \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle_g\\
&= \langle A_i^k E_k, A_j^l E_l \rangle_g\\
&= A_i^k A_j^l \langle E_k, E_l \rangle_g\\
&= \sum_k A_i^k A_j^k.
\end{aligned}$$

This last expression is the (i, j)-entry of the matrix product $A^T A$, where $A = (A_i^j)$. Thus

$$\det(g_{ij}) = \det(A^T A) = \det A^T \det A = (\det A)^2,$$

from which it follows that $f = \det A = \pm\sqrt{\det(g_{ij})}$. Since both frames $(\partial/\partial x^i)$ and (E_j) are oriented, the sign must be positive. □

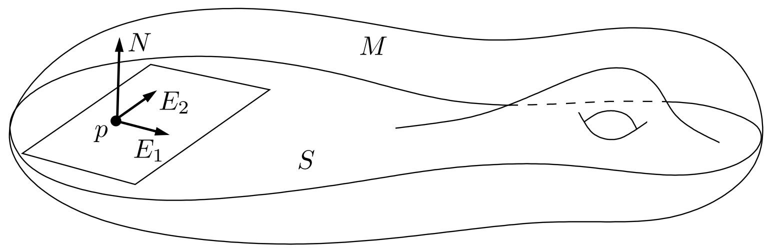


Figure 13.9. A hypersurface in a Riemannian manifold.

Hypersurfaces in Riemannian Manifolds

Let (M, g) be an oriented Riemannian manifold, and suppose $S \subset M$ is an immersed hypersurface. Any unit normal vector field along S is clearly transverse to S, so it determines an orientation of S by Proposition 13.12. The next proposition gives a very simple formula for the volume form of the induced metric on S with respect to this orientation.

Proposition 13.24. *Let (M, g) be an oriented Riemannian manifold, let $S \subset M$ be an immersed hypersurface, and let $\widetilde{g}$ denote the induced metric on S. Suppose N is a smooth unit normal vector field along S. With respect to the orientation of S determined by N, the volume form of $(S, \widetilde{g})$ is given by*

$$dV_{\widetilde{g}} = (N \lrcorner\, dV_g)|_S.$$

Proof. By Proposition 13.12, the $(n-1)$-form $N \lrcorner dV_g$ is an orientation form for S. To prove that it is the volume form for the induced Riemannian metric, we need only show that it gives the value 1 whenever it is applied to an oriented orthonormal frame for S. Thus let $(E_1, \dots, E_{n-1})$ be such a frame. At each point $p \in S$, the basis $(N_p, E_1|_p, \dots, E_{n-1}|_p)$ is orthonormal (Figure 13.9), and is oriented for T_pM (this is the definition of the orientation determined by N). Thus

$$(N \lrcorner\, dV_g)|_S(E_1, \dots, E_{n-1}) = dV_g(N, E_1, \dots, E_{n-1}) = 1,$$

which proves the result. □

The following lemma will be useful in our proofs of the classical theorems of vector analysis below.

Lemma 13.25. *With notation as in Proposition 13.24, if X is any vector field along S, we have*

$$X \lrcorner\, dV_g|_S = \langle X, N \rangle_g \, dV_{\widetilde{g}}. \tag{13.7}$$

Proof. Define two vector fields $X^\top$ and $X^\perp$ along S by

$$X^\perp = \langle X, N\rangle_g N,$$
$$X^\top = X - X^\perp.$$

Then $X = X^\perp + X^\top$, where $X^\perp$ is normal to S and $X^\top$ is tangent to it. Using this decomposition,

$$X \lrcorner\, dV_g = X^\perp \lrcorner\, dV_g + X^\top \lrcorner\, dV_g.$$

Using Proposition 13.24, the first term simplifies to

$$(X^\perp \lrcorner\, dV_g)|_S = \langle X, N\rangle_g (N \lrcorner\, dV_g)|_S = \langle X, N\rangle_g dV_{\widetilde{g}}.$$

Thus (13.7) will be proved if we can show that $(X^\top \lrcorner\, dV_g)|_S = 0$. If $X_1, \dots, X_{n-1}$ are any vectors tangent to S, then

$$(X^\top \lrcorner\, dV_g)(X_1, \dots, X_{n-1}) = dV_g(X^\top, X_1, \dots, X_{n-1}) = 0,$$

because any n vectors in an $(n-1)$-dimensional vector space are linearly dependent. □

The result of Proposition 13.24 takes on particular importance in the case of a Riemannian manifold with boundary, because of the following proposition.

Proposition 13.26. *Suppose M is any Riemannian manifold with boundary. There is a unique smooth outward-pointing unit normal vector field N along ∂M.*

Proof. First we prove uniqueness. At any point $p \in \partial M$, the vector space $(T_p\partial M)^\perp \subset T_pM$ is 1-dimensional, so there are exactly two unit vectors at p that are normal to ∂M. Since any unit normal vector N is obviously transverse to ∂M, it must have nonzero x^n-component in any smooth boundary chart. Thus exactly one of the two choices of unit normal has negative x^n-component, which is equivalent to being outward-pointing.

To prove existence, we will show that there exists a smooth outward unit normal field in a neighborhood of each point. By the uniqueness result above, these vector fields all agree where they overlap, so the resulting vector field is globally defined.

Let $p \in \partial M$. By Proposition 11.24, there exists a smooth adapted orthonormal frame $(E_1, \dots, E_n)$ in a neighborhood U of p. In this frame, E_n is a smooth unit normal vector field along ∂M. If we assume (by shrinking U if necessary) that U is connected, then E_n must be either inward-pointing or outward-pointing on all of $\partial M \cap U$. Replacing E_n by $-E_n$ if necessary, we obtain a smooth outward-pointing unit normal vector field defined near p. This completes the proof. □

The next corollary is immediate.

Corollary 13.27. *If (M, g) is an oriented Riemannian manifold with boundary and $\widetilde{g}$ is the induced Riemannian metric on ∂M, then the volume form of $\widetilde{g}$ is*

$$dV_{\widetilde{g}} = (N \lrcorner\, dV_g)|_{\partial M},$$

where N is the outward unit normal vector field along ∂M.

Problems

13-1. Prove that every smooth 1-manifold is orientable.

13-2. Suppose M is a smooth manifold that is the union of two orientable open submanifolds with connected intersection. Show that M is orientable. Use this to give another proof that $\mathbb{S}^n$ is orientable.

13-3. Suppose $\pi\colon \widetilde{M} \to M$ is a smooth covering map and M is orientable. Show that $\widetilde{M}$ is also orientable.

13-4. Suppose M and N are oriented smooth manifolds and $F\colon M \to N$ is a local diffeomorphism. If M is connected, show that F is either orientation-preserving or orientation-reversing.

13-5. Suppose M is a connected, oriented, smooth manifold and Γ is a discrete group acting smoothly, freely, and properly on M. We say the action is orientation-preserving if for each $\gamma \in \Gamma$, the diffeomorphism $x \mapsto \gamma \cdot x$ is orientation-preserving. Show that M/Γ is orientable if and only if the action of Γ is orientation-preserving.

13-6. Let $\alpha\colon \mathbb{S}^n \to \mathbb{S}^n$ be the antipodal map: $\alpha(x) = -x$. Show that α is orientation-preserving if and only if n is odd.

13-7. Prove that $\mathbb{RP}^n$ is orientable if and only if n is odd.

13-8. If ω is a symplectic form on a $2n$-manifold, show that $\omega \wedge \cdots \wedge \omega$ (the n-fold wedge product of ω with itself) is a nonvanishing $2n$-form on M, and thus every symplectic manifold is orientable.

13-9. Suppose M is a smooth orientable Riemannian manifold and $S \subset M$ is an immersed or embedded submanifold.

(a) If S has trivial normal bundle (see page 281), show that S is orientable.

(b) If S is an orientable hypersurface, show that S has trivial normal bundle.

13-10. Suppose M is an oriented Riemannian manifold, and $S \subset M$ is an oriented hypersurface (with or without boundary). Show that there is a unique smooth unit normal vector field along S that determines the given orientation of S.

13-11. Let M be a connected, nonorientable smooth manifold, and let $\widehat{\pi}\colon \widehat{M} \to M$ be its orientation covering.

(a) If $\widetilde{M}$ is an orientable smooth manifold and $\pi\colon \widetilde{M} \to M$ is a smooth covering map, show that there exists a smooth map $\varphi\colon \widetilde{M} \to \widehat{M}$ such that $\widehat{\pi} \circ \varphi = \pi$.

(b) Uniqueness of the Orientation Covering: If $\pi\colon \widetilde{M} \to M$ is as above and in addition π is a two-sheeted covering, show that φ is a diffeomorphism.

13-12. Suppose S is an oriented embedded 2-manifold with boundary in $\mathbb{R}^3$, and let $C = \partial S$ with the induced orientation. By Problem 13-10, there is a unique smooth unit normal vector field N on S that determines the orientation. Let T be the oriented unit tangent vector field on C, and let V be the unique unit vector field tangent to S along C that is orthogonal to T and inward-pointing. Show that (T_p, V_p, N_p) is an oriented orthonormal basis for $\mathbb{R}^3$ at each $p \in C$.

13-13. Let E be the total space of the Möbius bundle, which is the quotient of $\mathbb{R}^2$ by the $\mathbb{Z}$-action $n \cdot (x, y) = (x + n, (-1)^n y)$ (see Problem 9-16). The *Möbius band* is the subset $M \subset E$ that is the image under the quotient map of the set $\{(x, y) \in \mathbb{R}^2 : |y| \leq 1\}$. (It is a smooth 1-manifold with boundary.) Show that neither E nor M is orientable.

13-14. Let E be as in Problem 13-13. Show that the orientation covering of E is diffeomorphic to $\mathbb{S}^1 \times \mathbb{R}$.

13-15. Let $U \subset \mathbb{R}^3$ be the open subset $\{(x, y, z) : (\sqrt{x^2 + y^2} - 2)^2 + z^2 < 1\}$ (the solid torus of revolution bounded by the doughnut surface of Example 8.13). Define a map $F\colon \mathbb{R}^2 \to U$ by

$$F(u, v) = (\cos 2\pi u(2 + \tanh v \cos \pi u), \\ \sin 2\pi u(2 + \tanh v \cos \pi u), \tanh v \sin \pi u).$$

(a) Show that F descends to a smooth embedding of E into U, where E is the total space of the Möbius bundle of Problem 9-16.

(b) Let S be the image of F. Show that S is a closed embedded submanifold of U.

(c) Show that there is no normal vector field along S.

(d) Show that S has no global defining function in U.

14
Integration on Manifolds

We introduced differential forms with a promise that they would turn out to be objects that can be integrated on manifolds in a coordinate-independent way. In this chapter, we fulfill that promise by defining the integrals of n-forms over smooth n-manifolds.

First we define the integral of a differential form over a domain in Euclidean space, and then we show how to use diffeomorphism invariance and partitions of unity to extend this definition to the integral of a compactly supported n-form over a smooth oriented n-manifold. The key feature of the definition is that it is invariant under orientation-preserving diffeomorphisms.

Next we prove one of the most fundamental theorems in all of differential geometry. This is Stokes's theorem, which is a generalization of the fundamental theorem of calculus, as well as of the three great classical theorems of vector analysis: Green's theorem for vector fields in the plane; the divergence theorem for vector fields in space; and (the classical version of) Stokes's theorem for surface integrals in $\mathbb{R}^3$. We also describe an extension of Stokes's theorem to manifolds with corners, which will be useful in our treatment of de Rham cohomology in Chapters 15 and 16.

Next we show how these ideas play out on a Riemannian manifold. In particular, we prove Riemannian versions of the divergence theorem and of Stokes's theorem for surface integrals, of which the classical theorems are special cases. At the end of the chapter, we show how to extend the theory of integration to nonorientable manifolds by introducing densities, which are fields that can be integrated on *any* manifold, not just oriented ones.

Integration of Differential Forms on Euclidean Space

In this section, we will define integrals of differential forms over subsets of $\mathbb{R}^n$. For the time being, let us restrict attention to the case $n \geq 1$. You should be sure that you are familiar with the basic properties of multiple integrals in $\mathbb{R}^n$, as summarized in the Appendix.

Recall that a *domain of integration* is a bounded subset of $\mathbb{R}^n$ whose boundary has n-dimensional measure zero (see page 591). Let $D \subset \mathbb{R}^n$ be a compact domain of integration, and let ω be an n-form on D. Any such form can be written as

$$\omega = f\, dx^1 \wedge \cdots \wedge dx^n$$

for a continuous real-valued function f on D. We define the *integral of* ω *over* D to be

$$\int_D \omega = \int_D f\, dV.$$

This can be written more suggestively as

$$\int_D f\, dx^1 \wedge \cdots \wedge dx^n = \int_D f\, dx^1 \cdots dx^n.$$

In simple terms, to compute the integral of a form such as $f\, dx^1 \wedge \cdots \wedge dx^n$, we just "erase the wedges"!

Somewhat more generally, let U be an open set in $\mathbb{R}^n$. We would like to define the integral of any compactly supported n-form ω over U. However, since neither U nor $\operatorname{supp} \omega$ may be a domain of integration in general, we need the following lemma.

Lemma 14.1. *Suppose $K \subset U \subset \mathbb{R}^n$, where U is an open set and K is compact. Then there is a compact domain of integration D such that $K \subset D \subset U$.*

Proof. For each $p \in K$, there is an open ball containing p whose closure is contained in U. By compactness, finitely many such open balls $B_1, \dots, B_m$ cover K (Figure 14.1). Since the boundary of an open ball is a codimension-1 submanifold, it has measure zero by Theorem 10.5, and so each ball is a domain of integration. The set $D = \overline{B}_1 \cup \cdots \cup \overline{B}_m$ is the required domain of integration. □

Now if $U \subset \mathbb{R}^n$ is open and ω is a compactly supported n-form on U, we define

$$\int_U \omega = \int_D \omega,$$

where D is any domain of integration such that $\operatorname{supp} \omega \subset D \subset U$. It is an easy matter to verify that this definition does not depend on the choice of

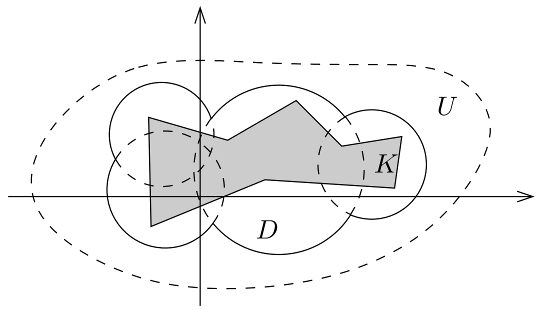


Figure 14.1. A domain of integration containing a compact set.

D. Similarly, if V is an open subset of the upper half-space $\mathbb{H}^n$ and ω is a compactly supported n-form on V, we define

$$\int_V \omega = \int_{D\cap\mathbb{H}^n} \omega,$$

where D is chosen in the same way.

It is worth remarking that it is possible to extend the definition to integrals of noncompactly supported forms, and integrals of such forms play an important role in many applications. However, in such cases the resulting multiple integrals are improper, so one must pay close attention to convergence issues. For the purposes we have in mind, the compactly supported case will be quite sufficient.

The next proposition explains the motivation for this definition.

Proposition 14.2. *Let D and E be compact domains of integration in $\mathbb{R}^n$, and let ω be an n-form on E. If $G\colon D \to E$ is a smooth map whose restriction to* $\operatorname{Int} D$ *is an orientation-preserving or orientation-reversing diffeomorphism onto* $\operatorname{Int} E$, *then*

$$\int_E \omega = \begin{cases} \displaystyle\int_D G^*\omega & \text{if } G \text{ is orientation-preserving,} \\ \displaystyle -\int_D G^*\omega & \text{if } G \text{ is orientation-reversing.} \end{cases}$$

Proof. Let us use $(y^1, \ldots, y^n)$ to denote standard coordinates on E, and $(x^1, \ldots, x^n)$ to denote those on D. Suppose first that G is orientation-preserving. Writing $\omega = f\, dy^1 \wedge \cdots \wedge dy^n$, the change of variables formula

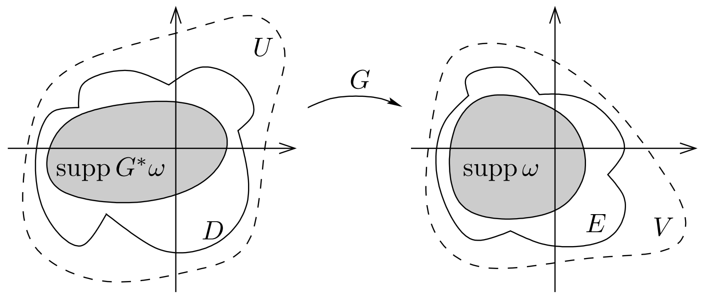


Figure 14.2. Diffeomorphism invariance of the integral of a form.

together with formula (12.12) for pullbacks of n-forms yields

$$\begin{aligned}\int_E \omega &= \int_E f\,dV \\ &= \int_D (f\circ G)\,|\det DG|\;dV \\ &= \int_D (f\circ G)(\det DG)\,dV \\ &= \int_D (f\circ G)(\det DG)\,dx^1\wedge\cdots\wedge dx^n \\ &= \int_D G^*\omega.\end{aligned}$$

If G is orientation-reversing, the same computation holds except that a negative sign is introduced when the absolute value signs are removed. □

Corollary 14.3. *Suppose U,V are open subsets of $\mathbb{R}^n$, $G\colon U\to V$ is an orientation-preserving diffeomorphism, and ω is a compactly supported n-form on V. Then*

$$\int_V \omega = \int_U G^*\omega.$$

Proof. Let $E\subset V$ be a compact domain of integration containing $\operatorname{supp}\omega$ (Figure 14.2). Since smooth maps take interiors to interiors, boundaries to boundaries, and sets of measure zero to sets of measure zero, $D = G^{-1}(E)\subset U$ is a domain of integration containing $\operatorname{supp} G^*\omega$. Therefore, the result follows from the preceding proposition. □

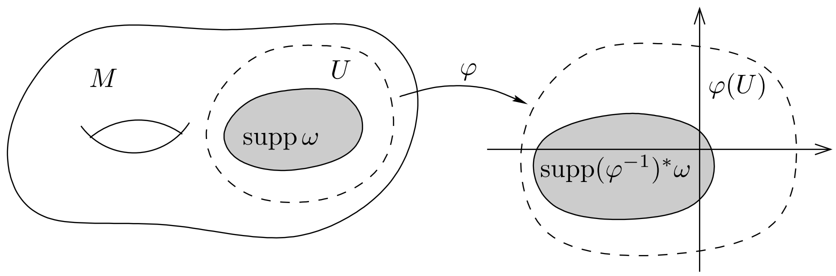


Figure 14.3. The integral of a form over a manifold.

Integration on Manifolds

Using the results of the previous section, it is easy to make invariant sense of the integral of a differential form over an oriented manifold. Let M be a smooth, oriented n-manifold, and let ω be an n-form on M. Suppose first that ω is compactly supported in the domain of a single oriented smooth coordinate chart (U,φ). We define the *integral of* ω *over* M to be

$$\int_M \omega = \int_{\varphi(U)} (\varphi^{-1})^*\omega.$$

(See Figure 14.3.) Since $(\varphi^{-1})^*\omega$ is a compactly supported n-form on the open subset $\varphi(U) \subset \mathbb{R}^n$, its integral is defined as discussed above.

Proposition 14.4. *With ω as above, $\int_M \omega$ does not depend on the choice of oriented smooth chart whose domain contains* $\operatorname{supp}\omega$.

Proof. Suppose $(\widetilde{U},\widetilde{\varphi})$ is another oriented smooth chart such that $\operatorname{supp}\omega \subset \widetilde{U}$ (Figure 14.4). Because $\widetilde{\varphi}\circ\varphi^{-1}$ is an orientation-preserving diffeomorphism from $\varphi(U\cap\widetilde{U})$ to $\widetilde{\varphi}(U\cap\widetilde{U})$, Corollary 14.3 implies that

$$\begin{aligned}\int_{\widetilde{\varphi}(\widetilde{U})} (\widetilde{\varphi}^{-1})^*\omega &= \int_{\widetilde{\varphi}(U\cap\widetilde{U})} (\widetilde{\varphi}^{-1})^*\omega\\ &= \int_{\varphi(U\cap\widetilde{U})} (\widetilde{\varphi}\circ\varphi^{-1})^*(\widetilde{\varphi}^{-1})^*\omega\\ &= \int_{\varphi(U\cap\widetilde{U})} (\varphi^{-1})^*(\widetilde{\varphi})^*(\widetilde{\varphi}^{-1})^*\omega\\ &= \int_{\varphi(U)} (\varphi^{-1})^*\omega.\end{aligned}$$

Thus the two definitions of $\int_M \omega$ agree. □

If M is an oriented smooth n-manifold with boundary, and ω is an n-form on M that is compactly supported in a smooth coordinate domain,

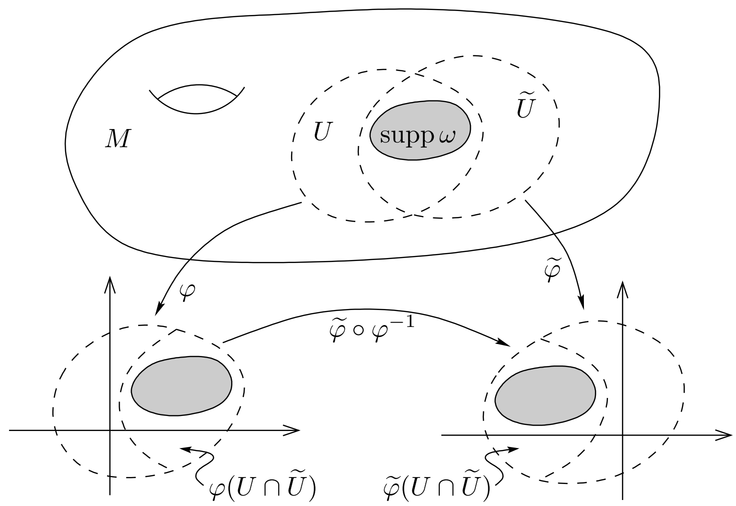


Figure 14.4. Coordinate independence of the integral.

the definition of $\int_M \omega$ and the statement and proof of Proposition 14.4 go through unchanged, provided we compute the integrals over open subsets of $\mathbb{H}^n$ in the way we described above.

To integrate over an entire manifold, we simply apply this same definition together with a partition of unity. Suppose M is an oriented smooth n-manifold (possibly with boundary) and ω is a compactly supported n-form on M. Let $\{(U_i, \varphi_i)\}$ be a finite cover of $\operatorname{supp}\omega$ by oriented smooth charts, and let $\{\psi_i\}$ be a smooth partition of unity subordinate to this cover. We define the *integral of* ω *over* M to be

$$\int_M \omega = \sum_i \int_M \psi_i \omega. \tag{14.1}$$

Since for each i, the n-form $\psi_i\omega$ is compactly supported in U_i, each of the terms in this sum is well-defined according to our discussion above. To show that the integral is well-defined, therefore, we need only examine the dependence on the charts and the partition of unity.

Lemma 14.5. *The definition of* $\int_M \omega$ *given above does not depend on the choice of oriented charts or partition of unity.*

Proof. Suppose $\{(\widetilde{U}_j, \widetilde{\varphi}_j)\}$ is another finite collection of oriented smooth charts whose domains cover $\operatorname{supp}\omega$, and $\{\widetilde{\psi}_j\}$ is a subordinate smooth partition of unity. For each i, we compute

$$\begin{aligned}\int_M \psi_i\omega &= \int_M \Bigl(\sum_j \widetilde{\psi}_j\Bigr)\psi_i\omega \\ &= \sum_j \int_M \widetilde{\psi}_j\psi_i\omega.\end{aligned}$$

Summing over i, we obtain

$$\sum_i \int_M \psi_i\omega = \sum_{i,j} \int_M \widetilde{\psi}_j\psi_i\omega.$$

Observe that each term in this last sum is the integral of a form compactly supported in a single smooth chart (U_i, for example), so by Proposition 14.4 each term is well-defined, regardless of which coordinate map we use to compute it. The same argument, starting with $\int_M \widetilde{\psi}_j\omega$, shows that

$$\sum_j \int_M \widetilde{\psi}_j\omega = \sum_{i,j} \int_M \widetilde{\psi}_j\psi_i\omega.$$

Thus both definitions yield the same value for $\int_M \omega$. □

As usual, we have a special definition in the zero-dimensional case. The integral of a compactly supported 0-form (i.e., a real-valued function) f over an oriented 0-manifold M is defined to be the sum

$$\int_M f = \sum_{p\in M} \pm f(p),$$

where we take the positive sign at points where the orientation is positive and the negative sign at points where it is negative. The assumption that f is compactly supported implies that there are only finitely many nonzero terms in this sum.

If $N \subset M$ is an oriented immersed k-dimensional submanifold (with or without boundary), and ω is a k-form on M whose restriction to N is compactly supported, we interpret $\int_N \omega$ to mean $\int_N (\omega|_N)$. In particular, if M is a compact, oriented, smooth n-manifold with boundary and ω is an $(n-1)$-form on M, we can interpret $\int_{\partial M} \omega$ unambiguously as the integral of $\omega|_{\partial M}$ over ∂M, where ∂M is always understood to have the induced orientation.

Proposition 14.6 (Properties of Integrals of Forms). *Suppose M and N are oriented smooth n-manifolds with or without boundaries, and ω, η are compactly supported n-forms on M.*

(*a*) Linearity: *If $a, b \in \mathbb{R}$, then*

$$\int_M a\omega + b\eta = a\int_M \omega + b\int_M \eta.$$

(*b*) Orientation Reversal: *If $\overline{M}$ denotes M with the opposite orientation, then*

$$\int_{\overline{M}} \omega = -\int_M \omega.$$

(*c*) Positivity: *If ω is an orientation form for M, then $\int_M \omega > 0$.*

(*d*) Diffeomorphism Invariance: *If $F\colon N \to M$ is an orientation-preserving diffeomorphism, then $\int_M \omega = \int_N F^*\omega$.*

Proof. Parts (a) and (b) are left as an exercise. Suppose that ω is an orientation form for M. This means that for any oriented smooth chart (U, φ), $(\varphi^*)^{-1}\omega$ is a positive function times $dx^1 \wedge \cdots \wedge dx^n$. Thus each term in the sum (14.1) defining $\int_M \omega$ is nonnegative, and at least one term is strictly positive, thus proving (c).

To prove (d), it suffices to assume that ω is compactly supported in a single smooth chart, because any n-form on M can be written as a finite sum of such forms by means of a partition of unity. Thus suppose (U, φ) is an oriented smooth chart on M whose domain contains the support of ω. It is easy to check that $(F^{-1}(U), \varphi \circ F)$ is an oriented smooth chart on N whose domain contains the support of $F^*\omega$, and the result then follows immediately from Corollary 14.3. □

◇ **Exercise 14.1.** Prove parts (a) and (b) of the preceding proposition.

Although the definition of the integral of a form based on partitions of unity is very convenient for theoretical purposes, it is useless for doing actual computations. It is generally quite difficult to write down a smooth partition of unity explicitly, and even when one can be written down, one would have to be exceptionally lucky to be able to compute the resulting integrals (think of trying to integrate $e^{-1/x}$).

For computational purposes, it is much more convenient to "chop up" the manifold into a finite number of pieces whose boundaries are sets of measure zero, and compute the integral on each one separately by means of local parametrizations. One way to do this is described below.

A subset $E \subset M$ is called a *domain of integration* if $\overline{E}$ is compact and ∂E has measure zero (in the sense described in Chapter 10). For example, any regular domain (i.e., compact embedded n-submanifold with boundary) in an n-manifold is a domain of integration.

Proposition 14.7 (Integration Over Parametrizations). *Let M be an oriented smooth n-manifold with or without boundary. Suppose $E_1, \dots, E_k$ are compact domains of integration in M; $D_1, \dots, D_k$ are compact domains of integration in $\mathbb{R}^n$; and for $i = 1, \dots, k$, $F_i\colon D_i \to M$ are smooth maps satisfying*

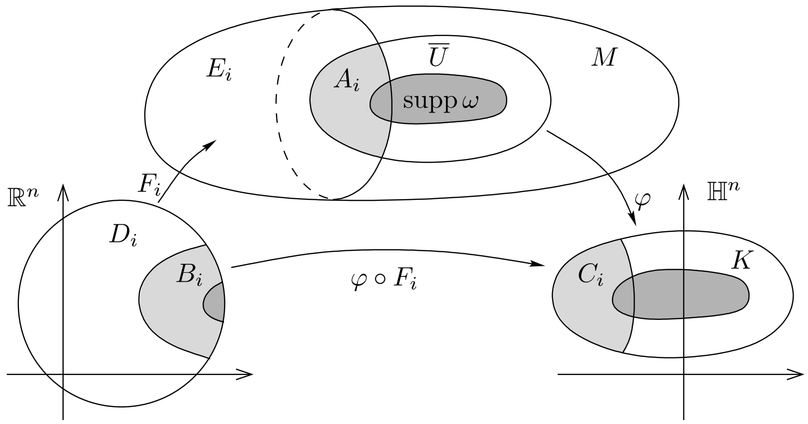


Figure 14.5. Integrating over parametrizations.

(*i*) $F_i(D_i) = E_i$, *and* $F_i|_{\operatorname{Int} D_i}$ *is an orientation-preserving diffeomorphism from* $\operatorname{Int} D_i$ *onto* $\operatorname{Int} E_i$.

(*ii*) $M = E_1 \cup \cdots \cup E_k$.

(*iii*) *For each* $i \neq j$, E_i *and* E_j *intersect only on their boundaries.*

Then for any n*-form* ω *on* M *whose support is contained in* $E_1 \cup \cdots \cup E_k$,

$$\int_M \omega = \sum_i \int_{D_i} F_i^* \omega.$$

Proof. As in the preceding proof, it suffices to assume that ω is compactly supported in the domain of a single oriented smooth chart (U, φ). In fact, by starting with a cover of M by sufficiently nice charts, we may assume that ∂U has measure zero, and that φ extends to a diffeomorphism from $\overline{U}$ to a compact domain of integration $K \subset \mathbb{H}^n$.

For each i, let

$$A_i = \overline{U} \cap E_i \subset M.$$

(See Figure 14.5.) Then A_i is a compact subset of M whose boundary has measure zero, since $\partial A_i \subset \partial U \cup \partial E_i$. Define compact subsets $B_i, C_i \subset \mathbb{R}^n$ by

$$B_i = F_i^{-1}(A_i),$$
$$C_i = \varphi(A_i).$$

Since smooth maps take sets of measure zero to sets of measure zero, both B_i and C_i are domains of integration, and $\varphi \circ F_i$ maps B_i to C_i smoothly and restricts to a diffeomorphism from $\operatorname{Int} B_i$ to $\operatorname{Int} C_i$. Therefore, Proposition

14.2 implies that

$$\int_{C_i} (\varphi^{-1})^*\omega = \int_{B_i} F_i^*\omega.$$

Summing over i, and noting that the interiors of the various sets A_i (and thus also C_i) are disjoint, we obtain

$$\begin{aligned}\int_M \omega &= \int_K (\varphi^{-1})^*\omega \\ &= \sum_i \int_{C_i} (\varphi^{-1})^*\omega \\ &= \sum_i \int_{B_i} F_i^*\omega \\ &= \sum_i \int_{D_i} F_i^*\omega. \end{aligned} \qquad \square$$

Example 14.8. Let us use this technique to compute the integral of a 2-form over $\mathbb{S}^2$, oriented as the boundary of $\overline{\mathbb{B}^3}$. Let ω be the following 2-form on $\mathbb{R}^3 \smallsetminus \{0\}$:

$$\omega = x\,dy \wedge dz + y\,dz \wedge dx + z\,dx \wedge dy.$$

Let D be the rectangle $[0,\pi] \times [0,2\pi]$, and let $F\colon D \to \mathbb{S}^2$ be the spherical coordinate parametrization

$$F(\varphi,\theta) = (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi).$$

Example 13.21 showed that F is orientation-preserving on $\operatorname{Int} D$. Let $D_1 = [0,\pi] \times [0,\pi]$ and $D_2 = [0,\pi] \times [\pi, 2\pi]$, and let $F_i = F|_{D_i}$ for $i = 1,2$. The two maps $F_1\colon D_1 \to \mathbb{S}^2$ and $F_2\colon D_2 \to \mathbb{S}^2$ satisfy the hypotheses of Proposition 14.7. (The only reason we cannot use F on the whole domain D is because $F(D)$ is all of $\mathbb{S}^2$, so $F(\operatorname{Int} D)$ is not equal to the interior of $F(D)$. Cutting the domain in half avoids this problem.) Note that

$$\begin{aligned} F^*dx &= \cos\varphi\cos\theta\,d\varphi - \sin\varphi\sin\theta\,d\theta, \\ F^*dy &= \cos\varphi\sin\theta\,d\varphi + \sin\varphi\cos\theta\,d\theta, \\ F^*dz &= -\sin\varphi\,d\varphi. \end{aligned}$$

Therefore,

$$\begin{aligned}\int_{\mathbb{S}^2} \omega &= \int_{D_1} F_1^*\omega + \int_{D_2} F_2^*\omega \\ &= \int_D F^*\omega \\ &= \int_D \big(-\sin^3\varphi\cos^2\theta\, d\theta\wedge d\varphi + \sin^3\varphi\sin^2\theta\, d\varphi\wedge d\theta \\ &\qquad + \cos^2\varphi\sin\varphi\cos^2\theta\, d\varphi\wedge d\theta - \cos^2\varphi\sin\varphi\sin^2\theta\, d\theta\wedge d\varphi\big) \\ &= \int_D \sin\varphi\, d\varphi\wedge d\theta \\ &= \int_0^{2\pi}\int_0^{\pi} \sin\varphi\, d\varphi\, d\theta \\ &= 4\pi.\end{aligned}$$

It is worth remarking that the hypotheses of Proposition 14.7 can be relaxed somewhat. For example, the maps F_i need not be smooth up to the boundaries of the domains D_i, provided they still map $\operatorname{Int} D_i$ diffeomorphically onto $\operatorname{Int} E_i$. For example, if the closed upper hemisphere of $\mathbb{S}^2$ is parametrized by the map $F\colon \overline{\mathbb{B}^2} \to \mathbb{S}^2$ given by $F(u,v) = (u, v, \sqrt{1-u^2-v^2})$, then F is continuous but not smooth up to the boundary. It turns out that, even though the resulting integrand is unbounded, the proposition still holds in this case provided the integral is interpreted in an appropriate limiting sense (see Problem 14-5). We leave it to the interested reader to work out reasonable conditions under which such a generalization of Proposition 14.7 holds.

Stokes's Theorem

In this section we will state and prove the central result in the theory of integration on manifolds: Stokes's theorem for manifolds. This is a far-reaching generalization of the fundamental theorem of calculus and of the classical theorems of vector calculus.

Theorem 14.9 (Stokes's Theorem). *Let M be a smooth, oriented n-dimensional manifold with boundary, and let ω be a compactly supported smooth $(n-1)$-form on M. Then*

$$\int_M d\omega = \int_{\partial M} \omega. \tag{14.2}$$

The statement of this theorem is concise and elegant, but it requires a bit of interpretation. First, as usual, ∂M is understood to have the induced (Stokes) orientation, and the ω on the right-hand side is to be interpreted

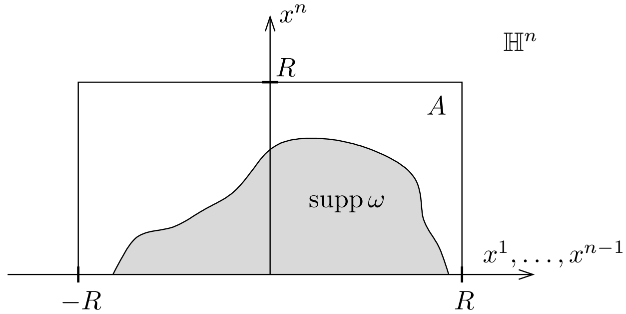


Figure 14.6. Proof of Stokes's theorem.

as $\omega|_{\partial M}$. If $\partial M = \varnothing$, then the right-hand side is to be interpreted as zero. When M is 1-dimensional, the right-hand integral is really just a finite sum.

With these understandings, we proceed with the proof of the theorem. You should check as you read through the proof that it works correctly when $n = 1$ and when $\partial M = \varnothing$.

Proof. We begin by considering a very special case: Suppose M is the upper half space $\mathbb{H}^n$ itself. Then the fact that ω has compact support means that there is a number $R > 0$ such that $\operatorname{supp}\omega$ is contained in the rectangle $A = [-R, R] \times \cdots \times [-R, R] \times [0, R]$ (Figure 14.6). We can write ω in standard coordinates as

$$\omega = \sum_{i=1}^{n} \omega_i \, dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n,$$

where the hat means that dx^i is omitted. Therefore,

$$\begin{aligned} d\omega &= \sum_{i=1}^{n} d\omega_i \wedge dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n \\ &= \sum_{i,j=1}^{n} \frac{\partial \omega_i}{\partial x^j} dx^j \wedge dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n \\ &= \sum_{i=1}^{n} (-1)^{i-1} \frac{\partial \omega_i}{\partial x^i} dx^1 \wedge \cdots \wedge dx^n. \end{aligned}$$

Thus we compute

$$\begin{aligned} \int_{\mathbb{H}^n} d\omega &= \sum_{i=1}^{n} (-1)^{i-1} \int_A \frac{\partial \omega_i}{\partial x^i} dx^1 \wedge \cdots \wedge dx^n \\ &= \sum_{i=1}^{n} (-1)^{i-1} \int_0^R \int_{-R}^{R} \cdots \int_{-R}^{R} \frac{\partial \omega_i}{\partial x^i}(x) dx^1 \cdots dx^n. \end{aligned}$$

We can rearrange the order of integration in each term so as to do the x^i integration first. By the fundamental theorem of calculus, the terms for which $i \neq n$ reduce to

$$\begin{aligned}
&\sum_{i=1}^{n-1}(-1)^{i-1}\int_0^R\int_{-R}^R\cdots\int_{-R}^R \frac{\partial\omega_i}{\partial x^i}(x)dx^1\cdots dx^n\\
&\qquad=\sum_{i=1}^{n-1}(-1)^{i-1}\int_0^R\int_{-R}^R\cdots\int_{-R}^R \frac{\partial\omega_i}{\partial x^i}(x)dx^i\,dx^1\cdots\widehat{dx^i}\cdots dx^n\\
&\qquad=\sum_{i=1}^{n-1}(-1)^{i-1}\int_0^R\int_{-R}^R\cdots\int_{-R}^R\Big[\omega_i(x)\Big]_{x^i=-R}^{x^i=R} dx^1\cdots\widehat{dx^i}\cdots dx^n\\
&\qquad=0,
\end{aligned}$$

because we have chosen R large enough that $\omega = 0$ when $x^i = \pm R$. The only term that might not be zero is the one for which $i = n$. For that term we have

$$\begin{aligned}
\int_{\mathbb{H}^n} d\omega &= (-1)^{n-1}\int_{-R}^R\cdots\int_{-R}^R\int_0^R \frac{\partial\omega_n}{\partial x^n}(x)dx^n\,dx^1\cdots dx^{n-1}\\
&= (-1)^{n-1}\int_{-R}^R\cdots\int_{-R}^R\Big[\omega_n(x)\Big]_{x^n=0}^{x^n=R} dx^1\cdots dx^{n-1} \qquad (14.3)\\
&= (-1)^n\int_{-R}^R\cdots\int_{-R}^R \omega_n(x^1,\dots,x^{n-1},0)\,dx^1\cdots dx^{n-1},
\end{aligned}$$

because $\omega_n = 0$ when $x^n = R$. (Note that this term too will vanish if $\operatorname{supp}\omega$ does not meet $\partial\mathbb{H}^n$.)

To compare this to the other side of (14.2), we compute as follows:

$$\int_{\partial\mathbb{H}^n}\omega = \sum_i \int_{A\cap\partial\mathbb{H}^n} \omega_i(x^1,\dots,x^{n-1},0)\,dx^1\wedge\cdots\wedge\widehat{dx^i}\wedge\cdots\wedge dx^n.$$

Because x^n vanishes on $\partial\mathbb{H}^n$, the restriction of dx^n to the boundary is identically zero (see Exercise 8.6). Thus the only term above that is nonzero is the one for which $i = n$, which becomes

$$\int_{\partial\mathbb{H}^n}\omega = \int_{A\cap\partial\mathbb{H}^n} \omega_n(x^1,\dots,x^{n-1},0)\,dx^1\wedge\cdots\wedge dx^{n-1}.$$

Taking into account the fact that the coordinates $(x^1,\dots,x^{n-1})$ are positively oriented for $\partial\mathbb{H}^n$ when n is even and negatively oriented when n is odd (Example 13.19), this becomes

$$\int_{\partial\mathbb{H}^n}\omega = (-1)^n\int_{-R}^R\cdots\int_{-R}^R \omega_n(x^1,\dots,x^{n-1},0)\,dx^1\cdots dx^{n-1},$$

which is equal to (14.3).

Next let M be an arbitrary smooth manifold with boundary, but consider an $(n-1)$-form ω that is compactly supported in the domain of a single smooth chart (U, φ). Assuming without loss of generality that φ is an oriented chart, the definition yields

$$\int_M d\omega = \int_{\mathbb{H}^n} (\varphi^{-1})^* d\omega = \int_{\mathbb{H}^n} d((\varphi^{-1})^* \omega),$$

since $(\varphi^{-1})^* d\omega$ is compactly supported on $\mathbb{H}^n$. By the computation above, this is equal to

$$\int_{\partial\mathbb{H}^n} (\varphi^{-1})^* \omega, \tag{14.4}$$

where $\partial\mathbb{H}^n$ is given the induced orientation. Since φ_* takes outward-pointing vectors on ∂M to outward-pointing vectors on $\mathbb{H}^n$ (by Lemma 13.15), it follows that $\varphi|_{U\cap\partial M}$ is an orientation-preserving diffeomorphism onto $\varphi(U)\cap\partial\mathbb{H}^n$, and thus (14.4) is equal to $\int_{\partial M}\omega$. This proves the theorem in this case.

Finally, let ω be an arbitrary compactly supported smooth $(n-1)$-form. Choosing a cover of $\operatorname{supp}\omega$ by finitely many oriented smooth charts $\{(U_i, \varphi_i)\}$, and choosing a subordinate smooth partition of unity $\{\psi_i\}$, we can apply the preceding argument to $\psi_i\omega$ for each i and obtain

$$\begin{aligned}
\int_{\partial M} \omega &= \sum_i \int_{\partial M} \psi_i \omega \\
&= \sum_i \int_M d(\psi_i \omega) \\
&= \sum_i \int_M d\psi_i \wedge \omega + \psi_i d\omega \\
&= \int_M d\Bigl(\sum_i \psi_i\Bigr) \wedge \omega + \int_M \Bigl(\sum_i \psi_i\Bigr) d\omega \\
&= 0 + \int_M d\omega,
\end{aligned}$$

because $\sum_i \psi_i \equiv 1$. □

Example 14.10. Let N be a smooth manifold and suppose $\gamma\colon [a,b] \to N$ is a smooth embedding, so that $M = \gamma[a,b]$ is an embedded 1-submanifold with boundary in N. If we give M the orientation such that γ is orientation-preserving, then for any smooth function $f \in C^\infty(N)$, Stokes's theorem says

$$\int_\gamma df = \int_{[a,b]} \gamma^* df = \int_M df = \int_{\partial M} f = f(\gamma(b)) - f(\gamma(b)).$$

Thus Stokes's theorem reduces to the fundamental theorem for line integrals (Theorem 6.22) in this case. In particular, when $\gamma\colon [a,b] \to \mathbb{R}$ is the inclusion map, then Stokes's theorem is just the ordinary fundamental theorem of calculus.

Two special cases of Stokes's theorem arise so frequently that they are worthy of special note. The proofs are immediate.

Corollary 14.11. *Suppose M is a compact manifold without boundary. Then the integral of every exact form over M is zero:*

$$\int_M d\omega = 0 \quad \textit{if } \partial M = \varnothing.$$

Corollary 14.12. *Suppose M is a compact smooth manifold with boundary. If ω is a closed form on M, then the integral of ω over ∂M is zero:*

$$\int_{\partial M} \omega = 0 \quad \textit{if } d\omega = 0 \textit{ on } M.$$

The next corollary is essentially a restatement of the previous two.

Corollary 14.13. *Suppose M is a smooth manifold, $S \subset M$ is a compact k-dimensional submanifold (without boundary), and ω is a closed k-form on M. If $\int_S \omega \neq 0$, then ω is not exact and S is not the boundary of a compact submanifold with boundary in M.*

Example 14.14. It follows from the computation of Example 6.20 that the closed 1-form $\omega = (x\,dy - y\,dx)/(x^2+y^2)$ has nonzero integral over $\mathbb{S}^1$. We already observed that ω is not exact on $\mathbb{R}^2 \smallsetminus \{0\}$. The preceding corollary tells us in addition that there is no smooth, compact, 2-dimensional submanifold with boundary in $\mathbb{R}^2 \smallsetminus \{0\}$ whose boundary is equal to $\mathbb{S}^1$.

One important application of Stokes's theorem is to prove the classical result known as Green's theorem.

Theorem 14.15 (Green's Theorem). *Suppose D is a regular domain in $\mathbb{R}^2$, and P, Q are smooth real-valued functions on D. Then*

$$\int_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx\,dy = \int_{\partial D} P\,dx + Q\,dy.$$

Proof. This is just Stokes's theorem applied to the 1-form $P\,dx + Q\,dy$. □

We will see other applications of Stokes's theorem later in this chapter.

Manifolds with Corners

In many applications of Stokes's theorem, it is necessary to deal with geometric objects such as triangles, squares, or cubes that are topological manifolds with boundary, but are not smooth manifolds with boundary because they have "corners." It is easy to generalize Stokes's theorem to

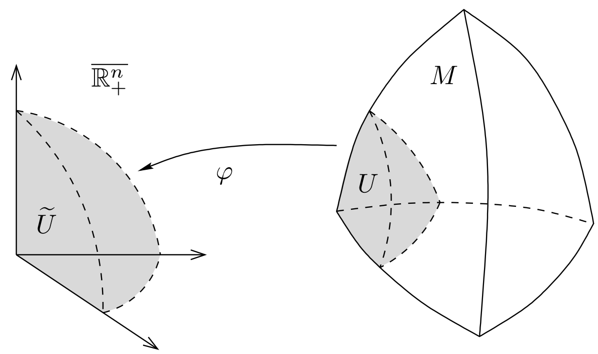


Figure 14.7. A chart with corners.

this situation, and we do so in this section. We will use this generalization only in our discussion of de Rham cohomology in Chapters 15 and 16.

Let $\overline{\mathbb{R}^n_+}$ denote the closed positive "quadrant" of $\mathbb{R}^n$:

$$\overline{\mathbb{R}^n_+} = \{(x^1, \dots, x^n) \in \mathbb{R}^n : x^1 \geq 0, \dots, x^n \geq 0\}.$$

This space is the model for the type of corners we will be concerned with.

◇ **Exercise 14.2.** Prove that $\overline{\mathbb{R}^n_+}$ is homeomorphic to the upper half-space $\mathbb{H}^n$.

Suppose M is a topological n-manifold with boundary. A *chart with corners* for M is a pair (U, φ), where U is an open subset of M and φ is a homeomorphism from U to a (relatively) open set $\widetilde{U} \subset \overline{\mathbb{R}^n_+}$ (Figure 14.7). Two charts with corners (U, φ), (V, ψ) are said to be smoothly compatible if the composite map $\varphi \circ \psi^{-1} \colon \psi(U \cap V) \to \varphi(U \cap V)$ is smooth. (As usual, this means that it admits a smooth extension in an open neighborhood of each point.)

A *smooth structure with corners* on a topological manifold with boundary is a maximal collection of smoothly compatible charts with corners whose domains cover M. A topological manifold with boundary together with a smooth structure with corners is called a *smooth manifold with corners*.

If M is a smooth manifold with corners, any chart with corners (U, φ) in the given smooth structure with corners is called a *smooth chart with corners* for M.

Example 14.16. Any closed rectangle in $\mathbb{R}^n$ is a smooth n-manifold with corners.

Because of the result of Exercise 14.2, charts with corners are topologically indistinguishable from boundary charts. Thus from the topological point of view there is no difference between manifolds with boundary and

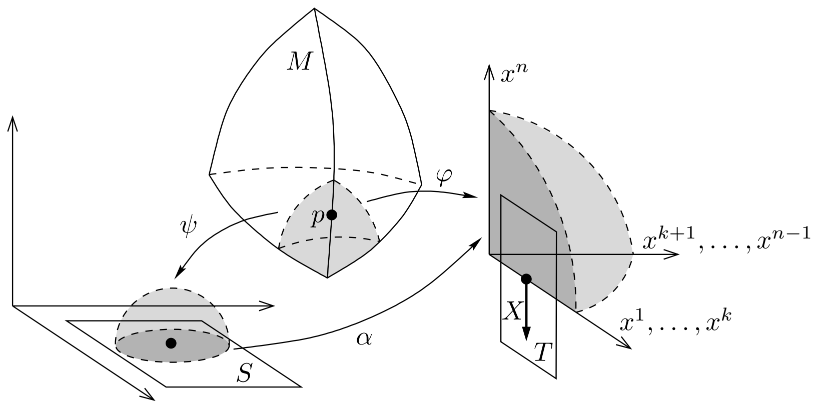


Figure 14.8. Invariance of corner points.

manifolds with corners. The difference is in the smooth structure, because the compatibility condition for charts with corners is different from that for boundary charts.

It is easy to check that the boundary of $\overline{\mathbb{R}^n_+}$ in $\mathbb{R}^n$ is the set of points at which at least one coordinate vanishes. The points in $\overline{\mathbb{R}^n_+}$ at which more than one coordinate vanishes are called its *corner points*. For example, the corner points of $\overline{\mathbb{R}^3_+}$ are the origin together with all the points on the positive x, y, and z-axes.

Lemma 14.17 (Invariance of Corner Points). *Let M be a smooth n-manifold with corners, and let $p \in M$. If $\varphi(p)$ is a corner point for some smooth chart with corners (U, φ), then the same is true for every such chart whose domain contains p.*

Proof. Suppose (U, φ) and (V, ψ) are two smooth charts with corners such that $\varphi(p)$ is a corner point but $\psi(p)$ is not (Figure 14.8). To simplify notation, let us assume without loss of generality that $\varphi(p)$ has coordinates $(x^1, \dots, x^k, 0, \dots, 0)$ with $k \le n-2$. Then $\psi(V)$ contains an open subset of some $(n-1)$-dimensional linear subspace $S \subset \mathbb{R}^n$, with $\psi(p) \in S$. (If $\psi(p)$ is a boundary point, S can be taken to be the unique subspace defined by an equation of the form $x^i = 0$ that contains $\psi(p)$. If $\psi(p)$ is an interior point, any $(n-1)$-dimensional subspace containing $\psi(p)$ will do.)

Let $\alpha\colon S \cap \psi(V) \to \mathbb{R}^n$ be the restriction of $\varphi \circ \psi^{-1}$ to $S \cap \psi(V)$. Because $\varphi \circ \psi^{-1}$ is a diffeomorphism, α is a smooth immersion. Let $T = \alpha_*(T_{\psi(p)}S) \subset \mathbb{R}^n$. Because T is $(n-1)$-dimensional, it must contain a vector X such that one of the last two components X^{n-1} or X^n is nonzero (otherwise T would be contained in a codimension-2 subspace). Renumbering the coordinates and replacing X by $-X$ if necessary, we may assume that $X^n < 0$.

Now let $\gamma\colon (-\varepsilon, \varepsilon) \to S$ be a smooth curve such that $\gamma(0) = p$ and $\alpha_*\gamma'(0) = X$. Then $\alpha \circ \gamma(t)$ has negative x^n coordinate for small $t > 0$, which contradicts the fact that α takes its values in $\overline{\mathbb{R}^n_+}$. $\square$

If M is a smooth manifold with corners, a point $p \in M$ is called a *corner point* if $\varphi(p)$ is a corner point in $\overline{\mathbb{R}^n_+}$ with respect to some (and hence every) smooth chart with corners (U, φ). Similarly p is called a *boundary point* if $\varphi(p) \in \partial\overline{\mathbb{R}^n_+}$ with respect to some (hence every) such chart. For example, the set of corner points of the unit cube $[0,1]^3 \subset \mathbb{R}^3$ is the union of its eight vertices and twelve edges.

It is clear that every smooth manifold with or without boundary is also a smooth manifold with corners (but with no corner points). Conversely, a smooth manifold with corners is a smooth manifold with boundary if and only if it has no corner points. The boundary of a smooth manifold with corners, however, is in general not a smooth manifold with corners (think of the boundary of a cube, for example). In fact, even the boundary of $\overline{\mathbb{R}^n_+}$ itself is not a smooth manifold with corners. It is, however, a union of finitely many such: $\partial\overline{\mathbb{R}^n_+} = H_1 \cup \cdots \cup H_n$, where

$$H_i = \{(x^1, \ldots, x^n) \in \overline{\mathbb{R}^n_+} : x^i = 0\}$$

is an $(n-1)$-dimensional smooth manifold with corners contained in the subspace defined by $x^i = 0$.

The usual flora and fauna of smooth manifolds—smooth maps, partitions of unity, tangent vectors, covectors, tensors, differential forms, orientations, and integrals of differential forms—can be defined on smooth manifolds with corners in exactly the same way as we have done for smooth manifolds and smooth manifolds with boundary, using smooth charts with corners in place of smooth boundary charts. The details are left to the reader.

In addition, for Stokes's theorem we will need to integrate a differential form over the boundary of a smooth manifold with corners. Since the boundary is not itself a smooth manifold with corners, this requires a special definition. Let M be an oriented smooth n-manifold with corners, and suppose ω is an $(n-1)$-form on ∂M that is compactly supported in the domain of a single oriented smooth chart with corners (U, φ). We define the integral of ω over ∂M by

$$\int_{\partial M} \omega = \sum_{i=1}^{n} \int_{H_i} (\varphi^{-1})^* \omega,$$

where each H_i is given the induced orientation as part of the boundary of the set where $x^i \geq 0$. In other words, we simply integrate ω in coordinates over the codimension-1 portion of the boundary. Finally, if ω is an arbitrary compactly supported $(n-1)$-form on M, we define the integral of ω over ∂M by piecing together with a partition of unity just as in the case of a manifold with boundary.

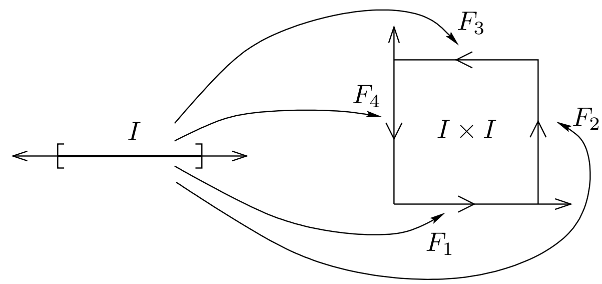


Figure 14.9. Parametrizing the boundary of the square.

In practice, of course, one does not evaluate such integrals by using partitions of unity. Instead, one "chops up" the boundary into pieces that can be parametrized by compact Euclidean domains of integration, just as for ordinary manifolds with or without boundary. If M is a smooth manifold with corners, we say a subset $A \subset \partial M$ has measure zero in ∂M if for every smooth chart with corners (U, φ), each set $\varphi(A) \cap H_i$ has measure zero in H_i for $i = 1, \ldots, n$. A domain of integration in ∂M is a subset $E \subset \partial M$ whose boundary has measure zero in ∂M. The following proposition is an analogue of Proposition 14.7.

Proposition 14.18. *The statement of Proposition 14.7 is true if M is replaced by the boundary of a compact, oriented, smooth n-manifold with corners.*

◇ **Exercise 14.3.** Show how the proof of Proposition 14.7 needs to be adapted to prove Proposition 14.18.

Example 14.19. Let $I \times I = [0,1] \times [0,1]$ be the unit square in $\mathbb{R}^2$, and suppose ω is a 1-form on $\partial(I \times I)$. Then it is not hard to check that the maps $F_i \colon I \to I \times I$ given by

$$\begin{aligned} F_1(t) &= (t, 0), \\ F_2(t) &= (1, t), \\ F_3(t) &= (1 - t, 1), \\ F_4(t) &= (0, 1 - t) \end{aligned} \tag{14.5}$$

satisfy the hypotheses of Proposition 14.18. (These four curve segments in sequence traverse the boundary of $I \times I$ in the counterclockwise direction—see Figure 14.9.) Therefore,

$$\int_{\partial(I \times I)} \omega = \int_{F_1} \omega + \int_{F_2} \omega + \int_{F_3} \omega + \int_{F_4} \omega. \tag{14.6}$$

◇ **Exercise 14.4.** Verify the claims of the preceding example.

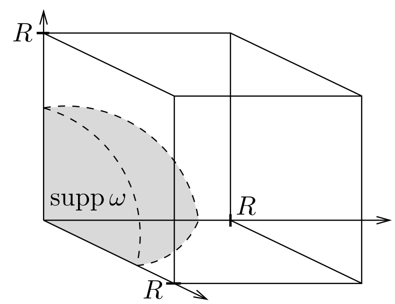


Figure 14.10. Stokes's theorem for manifolds with corners.

The next theorem is the main result of this section.

Theorem 14.20 (Stokes's Theorem on Manifolds with Corners). *Let M be a smooth, oriented n-manifold with corners, and let ω be a compactly supported $(n-1)$-form on M. Then*

$$\int_M d\omega = \int_{\partial M} \omega.$$

Proof. The proof is nearly identical to the proof of Stokes's theorem proper, so we will just indicate where changes need to be made. By means of smooth charts with corners and a partition of unity just as in that proof, we may reduce the theorem to the case in which $M = \overline{\mathbb{R}^n_+}$ and ω is supported in the cube $[0,R]^n$ (Figure 14.10). In that case, calculating exactly as in the proof of Theorem 14.9, we obtain

$$\begin{aligned}
\int_{\overline{\mathbb{R}^n_+}} d\omega &= \sum_{i=1}^n (-1)^{i-1} \int_0^R \cdots \int_0^R \frac{\partial \omega_i}{\partial x^i}(x) dx^1 \cdots dx^n \\
&= \sum_{i=1}^n (-1)^{i-1} \int_0^R \cdots \int_0^R \frac{\partial \omega_i}{\partial x^i}(x) dx^i \, dx^1 \cdots \widehat{dx^i} \cdots dx^n \\
&= \sum_{i=1}^n (-1)^{i-1} \int_0^R \cdots \int_0^R \Big[\omega_i(x)\Big]_{x^i=0}^{x^i=R} dx^1 \cdots \widehat{dx^i} \cdots dx^n \\
&= \sum_{i=1}^n (-1)^{i} \int_0^R \cdots \int_0^R \omega_i(x^1, \dots, 0, \dots, x^n)\, dx^1 \cdots \widehat{dx^i} \cdots dx^n \\
&= \sum_{i=1}^n \int_{H_i} \omega \\
&= \int_{\partial \overline{\mathbb{R}^n_+}} \omega.
\end{aligned}$$

(The factor $(-1)^i$ disappeared because the induced orientation on H_i is $(-1)^i$ times that of the standard coordinates $(x^1, \dots, \widehat{x^i}, \dots, x^n)$.) This completes the proof. □

Here is an immediate application of this result, which we will use when we study de Rham cohomology in the next chapter.

Theorem 14.21. *Suppose M is a smooth manifold, and $\gamma_0, \gamma_1 \colon [a,b] \to M$ are path homotopic piecewise smooth curve segments. For every closed 1-form ω on M,*

$$\int_{\gamma_0} \omega = \int_{\gamma_1} \omega.$$

Proof. By means of an affine reparametrization, we may as well assume for simplicity that $[a,b] = [0,1]$. Assume first that γ_0 and γ_1 are smooth. By Proposition 10.22, γ_0 and γ_1 are smoothly homotopic relative to $\{0,1\}$. Let $H \colon I \times I \to M$ be such a smooth homotopy. Since ω is closed, we have

$$\int_{I \times I} d(H^*\omega) = \int_{I \times I} H^* d\omega = 0.$$

On the other hand, $I \times I$ is a manifold with corners, so Stokes's theorem implies

$$0 = \int_{I \times I} d(H^*\omega) = \int_{\partial(I \times I)} H^*\omega.$$

Using the parametrization of $\partial(I \times I)$ given in Example 14.19 together with the diffeomorphism invariance of line integrals (Exercise 6-9), we obtain

$$\begin{aligned}
0 &= \int_{\partial(I\times I)} H^*\omega \\
&= \int_{F_1} H^*\omega + \int_{F_2} H^*\omega + \int_{F_3} H^*\omega + \int_{F_4} H^*\omega \\
&= \int_{H\circ F_1} \omega + \int_{H\circ F_2} \omega + \int_{H\circ F_3} \omega + \int_{H\circ F_4} \omega,
\end{aligned}$$

where F_1, F_2, F_3, F_4 are defined by (14.5). The fact that H is a homotopy relative to $\{0,1\}$ means that $H \circ F_2$ and $H \circ F_4$ are constant maps, and therefore the second and fourth terms above are zero. The theorem then follows from the facts that $H \circ F_1 = \gamma_0$ and $H \circ F_3$ is a backward reparametrization of γ_1.

Next we consider the general case of piecewise smooth curves. We cannot simply apply the preceding result on each subinterval where γ_0 and γ_1 are smooth, because the restricted curves may not start and end at the same points. Instead, we will prove the following more general claim: *Let $\gamma_0, \gamma_1 \colon I \to M$ be piecewise smooth curve segments (not necessarily with*

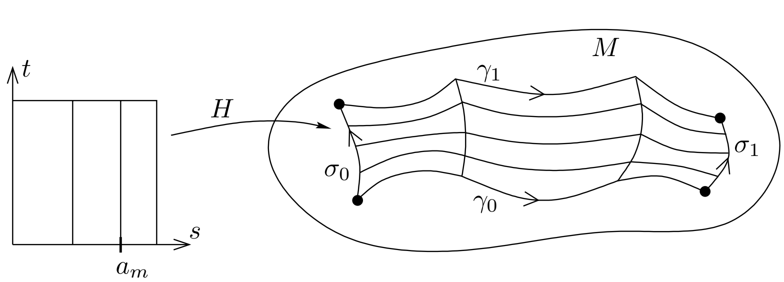


Figure 14.11. Homotopic piecewise smooth curve segments.

the same endpoints), and suppose $H\colon I \times I \to M$ *is any homotopy between them (Figure 14.11). Define curve segments* $\sigma_0, \sigma_1 \colon I \to M$ *by*

$$\sigma_0(t) = H(0,t),$$
$$\sigma_1(t) = H(1,t),$$

and let $\widetilde{\sigma}_0, \widetilde{\sigma}_1$ *be any smooth curve segments that are path homotopic to* σ_0, σ_1 *respectively. Then*

$$\int_{\gamma_1} \omega - \int_{\gamma_0} \omega = \int_{\widetilde{\sigma}_1} \omega - \int_{\widetilde{\sigma}_0} \omega. \tag{14.7}$$

When specialized to the case in which γ_0 and γ_1 are path homotopic, this implies the theorem, because σ_0 and σ_1 are constant maps in that case.

Since γ_0 and γ_1 are piecewise smooth, there are only finitely many points $(a_1, \dots, a_m)$ in $(0,1)$ at which either γ_0 or γ_1 is not smooth. We will prove the claim by induction on the number m of such points. When $m = 0$, both curves are smooth, and by Proposition 10.22 we may replace the given homotopy H by a smooth homotopy $\widetilde{H}$. Recall from the proof of Proposition 10.22 that the smooth homotopy $\widetilde{H}$ can actually be taken to be homotopic to H relative to $I \times \{0\} \cup I \times \{1\}$. Thus for $i = 0, 1$, the curve $\widetilde{\sigma}_i(t) = \widetilde{H}(i,t)$ is a smooth curve segment that is path homotopic to σ_i. In this setting, (14.7) just reduces to (14.6). Note that the integrals over $\widetilde{\sigma}_0$ and $\widetilde{\sigma}_1$ do not depend on which smooth curves path homotopic to σ_0 and σ_1 are chosen, by the smooth case of the theorem proved above.

Now let γ_0, γ_1 be homotopic piecewise smooth curves with m nonsmooth points $(a_1, \dots, a_m)$, and suppose the claim is true for curves with fewer than m such points. For $i = 0, 1$, let γ_i' be the restriction of γ_i to $[0, a_m]$, and let γ_i'' be its restriction to $[a_m, 1]$. Let $\sigma\colon I \to M$ be the curve segment

$$\sigma(t) = H(a_m, t),$$

and let $\widetilde{\sigma}$ by any smooth curve segment that is path homotopic to σ. Then, since γ_i' and γ_i'' have fewer than m nonsmooth points, the inductive hypothesis implies

$$\begin{aligned}\int_{\gamma_1}\omega-\int_{\gamma_0}\omega&=\left(\int_{\gamma_1'}\omega-\int_{\gamma_0'}\omega\right)+\left(\int_{\gamma_1''}\omega-\int_{\gamma_0''}\omega\right)\\&=\left(\int_{\widetilde{\sigma}}\omega-\int_{\widetilde{\sigma}_0}\omega\right)+\left(\int_{\widetilde{\sigma}_1}\omega-\int_{\widetilde{\sigma}}\omega\right)\\&=\int_{\widetilde{\sigma}_1}\omega-\int_{\widetilde{\sigma}_0}\omega.\end{aligned}$$ □

Integration on Riemannian Manifolds

We noted earlier that real-valued functions cannot be integrated in a coordinate-independent way on an arbitrary manifold. However, with the additional structures of a Riemannian metric and an orientation, we can recover the notion of the integral of a real-valued function.

Suppose (M, g) is an oriented Riemannian manifold (with or without boundary), and let dV_g denote its Riemannian volume form. If f is a compactly supported continuous real-valued function on M, then $f\,dV_g$ is a compactly supported n-form, so we can define the *integral of f over M* to be $\int_M f\,dV_g$. (This, of course, is the reason we chose the notation dV_g for the Riemannian volume form.) If M itself is compact, we define the *volume* of M by

$$\mathrm{Vol}(M)=\int_M dV_g.$$

Lemma 14.22. *Let (M, g) be an oriented Riemannian manifold, and suppose f is a compactly supported continuous real-valued function on M satisfying $f \geq 0$. Then $\int_M f\,dV_g \geq 0$, with equality if and only if $f \equiv 0$.*

Proof. If f is supported in the domain of a single oriented smooth chart (U, φ), then Lemma 13.23 shows that

$$\int_M f\,dV_g=\int_{\varphi(U)} f(x)\sqrt{\det(g_{ij})}\,dx^1\cdots dx^n \geq 0.$$

The general case follows from this one, because $\int_M f\,dV_g$ is equal to a sum of terms like $\int_M \psi_i f\,dV_g$, where each integrand $\psi_i f$ is nonnegative and supported in a single smooth chart. If in addition f is positive somewhere, then it is positive on an open set by continuity, so at least one of the integrals in this sum will be positive. On the other hand, if f is identically zero, then clearly $\int_M f\,dV_g = 0$. □

The Divergence Theorem

Let (M, g) be an oriented Riemannian manifold. Multiplication by the Riemannian volume form defines a linear map $*\colon C^\infty(M) \to \mathcal{A}^n(M)$:

$$* f = f\, dV_g.$$

It is easy to check that $*$ is an isomorphism.

Define the *divergence operator* $\operatorname{div}\colon \mathfrak{T}(M) \to C^\infty(M)$ by

$$\operatorname{div} X = *^{-1} d(X \lrcorner\, dV_g),$$

or equivalently,

$$d(X \lrcorner\, dV_g) = (\operatorname{div} X) dV_g.$$

Its geometric meaning will be discussed in Chapter 18.

In the special case of a regular domain in $\mathbb{R}^3$, the following theorem is due to Gauss and is often referred to as *Gauss's theorem.*

Theorem 14.23 (The Divergence Theorem). *Let M be an oriented Riemannian manifold with boundary. For any compactly supported smooth vector field X on M,*

$$\int_M (\operatorname{div} X)\, dV_g = \int_{\partial M} \langle X, N \rangle_g\, dV_{\widetilde{g}},$$

where N is the outward-pointing unit normal vector field along ∂M and $\widetilde{g}$ is the induced Riemannian metric on ∂M.

Proof. By Stokes's theorem,

$$\begin{aligned} \int_M (\operatorname{div} X)\, dV_g &= \int_M d(X \lrcorner\, dV_g) \\ &= \int_{\partial M} X \lrcorner\, dV_g. \end{aligned}$$

The theorem then follows from Lemma 13.25. □

Surface Integrals

The original theorem that bears the name of Stokes concerned "surface integrals" of vector fields over surfaces in $\mathbb{R}^3$. Using the version of Stokes's theorem that we have proved, this can be generalized to surfaces in Riemannian 3-manifolds. (For reasons that will be explained later, the restriction to dimension 3 cannot be removed.)

Let (M, g) be an oriented Riemannian 3-manifold. We define a bundle map $\beta\colon TM \to \Lambda^2 M$ by letting it act on smooth vector fields as follows:

$$\beta(X) = X \lrcorner\, dV_g. \tag{14.8}$$

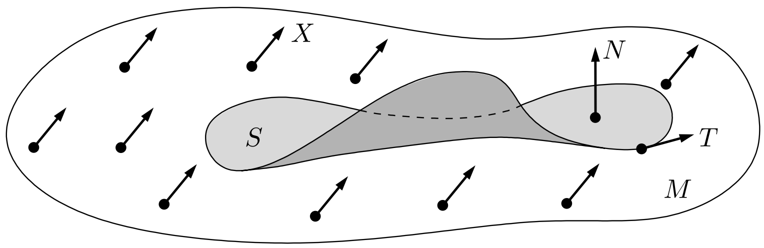


Figure 14.12. The setup for a surface integral.

It is easy to check that β is linear over $C^\infty(M)$, so it is a smooth bundle map, and it is an isomorphism because both bundles TM and $\Lambda^2 M$ have the same fiber dimension and β is injective on each fiber.

Define an operator $\operatorname{curl}\colon \mathcal{T}(M) \to \mathcal{T}(M)$ by

$$\operatorname{curl} X = \beta^{-1} d(X^\flat),$$

or equivalently,

$$(\operatorname{curl} X) \lrcorner\, dV_g = d(X^\flat). \tag{14.9}$$

The following commutative diagram summarizes the relationships among the gradient, divergence, curl, and exterior derivative operators:

$$\begin{tikzcd}
C^\infty(M) \arrow[r, "\operatorname{grad}"] \arrow[d, "\mathrm{Id}"] & \mathcal{T}(M) \arrow[r, "\operatorname{curl}"] \arrow[d, "\flat"] & \mathcal{T}(M) \arrow[r, "\operatorname{div}"] \arrow[d, "\beta"] & C^\infty(M) \arrow[d, "*"] \\
C^\infty(M) \arrow[r, "d"] & \mathcal{A}^1(M) \arrow[r, "d"] & \mathcal{A}^2(M) \arrow[r, "d"] & \mathcal{A}^3(M),
\end{tikzcd} \tag{14.10}$$

◇ **Exercise 14.5.** Show that $\operatorname{curl} \circ \operatorname{grad} \equiv 0$ and $\operatorname{div} \circ \operatorname{curl} \equiv 0$ on any Riemannian 3-manifold.

Now suppose $S \subset M$ is a compact, embedded, 2-dimensional submanifold with or without boundary in M, and N is a smooth unit normal vector field along S. Let dA denote the induced Riemannian volume form on S with respect to the induced metric $g|_S$ and the orientation determined by N, so that $dA = (N \lrcorner\, dV_g)|_S$ by Proposition 13.24. For any smooth vector field X defined on M, the *surface integral* of X over S (with respect to the given choice of normal field) is defined as

$$\int_S \langle X, N \rangle_g \, dA.$$

(See Figure 14.12.)

The next result, in the special case in which $M = \mathbb{R}^3$, is the original theorem proved by Stokes.

Theorem 14.24 (Stokes's Theorem for Surface Integrals). *Suppose S is a compact, oriented, embedded, 2-dimensional submanifold with boundary in an oriented Riemannian 3-manifold M. For any smooth vector field X on M,*

$$\int_S \langle \operatorname{curl} X, N \rangle_g \, dA = \int_{\partial S} \langle X, T \rangle_g \, ds,$$

where N is the smooth unit normal vector field along S that determines its orientation, ds is the Riemannian volume form for ∂S (with respect to the metric and orientation induced from S), and T is the unique positively oriented unit tangent vector field on ∂S .

Proof. The general version of Stokes's theorem applied to the 1-form $X^\flat$ yields

$$\int_S d(X^\flat) = \int_{\partial S} X^\flat.$$

Thus the theorem will follow from the following two identities:

$$d(X^\flat)|_S = \langle \operatorname{curl} X, N \rangle_g \, dA, \tag{14.11}$$

$$X^\flat|_{\partial S} = \langle X, T \rangle_g \, ds. \tag{14.12}$$

Equation (14.11) is just the defining equation (14.9) for the curl combined with the result of Lemma 13.25. To prove (14.12), we note that $X^\flat|_{\partial S}$ is a smooth 1-form on a 1-manifold, and thus must be equal to $f\,ds$ for some smooth function f on ∂S. To evaluate f, we note that $ds(T) = 1$, and so the definition of $X^\flat$ yields

$$f = f\,ds(T) = X^\flat(T) = \langle X, T \rangle_g.$$

This proves (14.12) and thus the theorem. □

The curl operator is defined only in dimension 3 because it is only in that case that $\Lambda^2 M$ is isomorphic to TM (via the map $\beta\colon X \mapsto X \lrcorner\, dV_g$). In fact, it was largely the desire to generalize the curl and the classical version of Stokes's theorem to higher dimensions that led to the entire theory of differential forms.

Integration on Lie Groups

Let G be a Lie group. A tensor or differential form σ on G is said to be left-invariant if $L_g^*\sigma = \sigma$ for all $g \in G$.

Proposition 14.25. *Let G be a compact Lie group endowed with a left-invariant orientation. Then G has a unique left-invariant orientation form Ω with the property that $\int_G \Omega = 1$.*

Proof. Let $E_1, \dots, E_n$ be a left-invariant global frame on G (i.e., a basis for the Lie algebra of G). By replacing E_1 with $-E_1$ if necessary, we may assume this frame is positively oriented. Let $\varepsilon^1, \dots, \varepsilon^n$ be the dual coframe. Left invariance of E_j implies that

$$(L_g^*\varepsilon^i)(E_j) = \varepsilon^i(L_{g*}E_j) = \varepsilon^i(E_j),$$

which shows that $L_g^*\varepsilon^i = \varepsilon^i$, so ε^i is left-invariant.

Let $\Omega = \varepsilon^1 \wedge \cdots \wedge \varepsilon^n$. Then

$$L_g^*\Omega = L_g^*\varepsilon^1 \wedge \cdots \wedge L_g^*\varepsilon^n = \varepsilon^1 \wedge \cdots \wedge \varepsilon^n = \Omega,$$

so Ω is left-invariant as well. Because $\Omega(E_1, \dots, E_n) = 1 > 0$, Ω is an orientation form for the given orientation. Clearly any positive constant multiple of Ω is also a left-invariant orientation form. Conversely, if $\widetilde{\Omega}$ is any other left-invariant orientation form, we can write $\widetilde{\Omega}_e = c\Omega_e$ for some positive number c. Using left-invariance, we find that

$$\widetilde{\Omega}_g = L_{g^{-1}}^*\widetilde{\Omega}_e = cL_{g^{-1}}^*\Omega_e = c\Omega_g,$$

which proves that $\widetilde{\Omega}$ is a positive constant multiple of Ω.

Since G is compact and oriented, $\int_G \Omega$ is a positive real number, so we can define $\widetilde{\Omega} = (\int_G \Omega)^{-1}\Omega$. Clearly $\widetilde{\Omega}$ is the unique left-invariant orientation form for which G has unit volume. □

Remark. The orientation form whose existence is asserted in this proposition is called the *Haar volume form* on G, and is often denoted by dV. Similarly, the map $f \mapsto \int_G f\,dV$ is called the *Haar integral*. Observe that the proof above did not use the fact that G was compact until the last paragraph; thus every Lie group has a left-invariant orientation form that is uniquely defined up to a constant multiple. It is only in the compact case, however, that we can use the volume normalization to single out a unique one.

Densities

Although differential forms are natural objects to integrate on manifolds, and are essential for use in Stokes's theorem, they have the disadvantage of requiring oriented manifolds in order for their integrals to be defined. There is a way to define integration on nonorientable manifolds as well, which we describe in this section.

As you will recall, the reason an orientation is needed for integrals of differential forms to make sense has to do with transformation laws under changes of coordinates. The transformation law for an n-form on an n-manifold under a change of coordinates involves the Jacobian determinant of the transition map, while the transformation law for integrals involves the absolute value of the determinant. In this section we define objects whose transformation law involves the absolute value of the determinant.

We begin, as always, in the linear-algebraic setting. Let V be an n-dimensional vector space. A *density* on V is a function

$$\mu \colon \underbrace{V \times \cdots \times V}_{n \text{ copies}} \to \mathbb{R}$$

satisfying the following property: If $T \colon V \to V$ is any linear map, then

$$\mu(TX_1, \ldots, TX_n) = |\det T| \, \mu(X_1, \ldots, X_n). \tag{14.13}$$

Observe that a density is *not* a tensor, because it is not linear over $\mathbb{R}$ in any of its arguments. Let $\Omega(V)$ denote the set of all densities on V.

Proposition 14.26 (Properties of Densities). *Let V be a vector space of dimension $n \geq 1$.*

(*a*) *$\Omega(V)$ is a vector space under the obvious vector operations:*

$$\begin{aligned}(c_1\mu_1 + c_2\mu_2)(X_1, \ldots, X_n)& \\ &= c_1\mu_1(X_1, \ldots, X_n) + c_2\mu_2(X_1, \ldots, X_n).\end{aligned}$$

(*b*) *If $\mu_1, \mu_2 \in \Omega(V)$ and $\mu_1(E_1, \ldots, E_n) = \mu_2(E_1, \ldots, E_n)$ for some basis (E_i) of V, then $\mu_1 = \mu_2$.*

(*c*) *If $\omega \in \Lambda^n(V)$, the map $|\omega| : V \times \cdots \times V \to \mathbb{R}$ defined by*

$$|\omega| \, (X_1, \ldots, X_n) = |\omega(X_1, \ldots, X_n)|$$

is a density.

(*d*) *$\Omega(V)$ is 1-dimensional, spanned by $|\omega|$ for any nonzero $\omega \in \Lambda^n(V)$.*

Proof. Part (a) is immediate from the definition. For part (b), suppose μ_1 and μ_2 give the same value when applied to $(E_1, \ldots, E_n)$. If $X_1, \ldots, X_n$ are arbitrary vectors in V, let $T \colon V \to V$ be the unique linear map that takes E_i to X_i for $i = 1, \ldots, n$. It follows that

$$\begin{aligned}\mu_1(X_1, \ldots, X_n) &= \mu_1(TE_1, \ldots, TE_n) \\ &= |\det T| \, \mu_1(E_1, \ldots, E_n) \\ &= |\det T| \, \mu_2(E_1, \ldots, E_n) \\ &= \mu_2(TE_1, \ldots, TE_n) \\ &= \mu_2(X_1, \ldots, X_n).\end{aligned}$$

Part (c) follows from Lemma 12.6:

$$\begin{aligned}|\omega|(TX_1, \ldots, TX_n) &= |\omega(TX_1, \ldots, TX_n)| \\ &= |(\det T)\omega(X_1, \ldots, X_n)| \\ &= |\det T| \; |\omega| \, (X_1, \ldots, X_n).\end{aligned}$$

Finally, to prove (d), suppose ω is any nonzero element of $\Lambda^n(V)$. If μ is an arbitrary element of $\Omega(V)$, it suffices to show that $\mu = c\,|\omega|$ for some $c \in \mathbb{R}$. Let (E_i) be a basis for V, and define $a, b \in \mathbb{R}$ by

$$\begin{aligned} a &= |\omega|\,(E_1, \dots, E_n) = |\omega(E_1, \dots, E_n)|\,, \\ b &= \mu(E_1, \dots, E_n). \end{aligned}$$

Because $\omega \neq 0$, it follows that $a \neq 0$. Thus μ and $(b/a)|\omega|$ give the same result when applied to $(E_1, \dots, E_n)$, so they are equal by part (b). □

A *positive density* on V is a density μ satisfying $\mu(X_1, \dots, X_n) > 0$ whenever $(X_1, \dots, X_n)$ are linearly independent. A *negative density* is defined similarly. If ω is a nonzero element of $\Lambda^n(V)$, then it is clear that $|\omega|$ is a positive density; more generally, a density $c\,|\omega|$ is positive, negative, or zero if and only if c has the same property. Thus every density on V is either positive, negative, or zero, and the set of positive densities is a convex subset of $\Omega(V)$ (namely, a half-line).

Now let M be a smooth manifold. The set

$$\Omega M = \coprod_{p \in M} \Omega(T_pM)$$

is called the *density bundle* of M. Let $\pi\colon \Omega M \to M$ be the natural projection map taking each element of $\Omega(T_pM)$ to p.

Lemma 14.27. *If M is a smooth manifold, its density bundle is a smooth line bundle over M.*

Proof. We will construct local trivializations and use the vector bundle construction lemma (Lemma 5.5). Let $(U, (x^i))$ be any smooth coordinate chart on M, and let $\omega = dx^1 \wedge \dots \wedge dx^n$. Proposition 14.26 shows that $|\omega_p|$ is a basis for $\Omega(T_pM)$ at each point $p \in U$. Therefore the map $\Phi\colon \pi^{-1}(U) \to U \times \mathbb{R}$ given by

$$\Phi(c\,|\omega_p|) = (p, c)$$

is a bijection.

Suppose $(\widetilde{U}, (\widetilde{x}^j))$ is another smooth chart with $U \cap \widetilde{U} \neq \varnothing$. Let $\widetilde{\omega} = d\widetilde{x}^1 \wedge \dots \wedge d\widetilde{x}^n$, and define $\widetilde{\Phi}\colon \pi^{-1}(\widetilde{U}) \to \widetilde{U} \times \mathbb{R}$ correspondingly:

$$\widetilde{\Phi}(c\,|\widetilde{\omega}_p|) = (p, c).$$

It follows from the transformation law (12.13) for n-forms under changes of coordinates that

$$\begin{aligned} \Phi \circ \widetilde{\Phi}^{-1}(p, c) &= \Phi(c\,|\widetilde{\omega}_p|) \\ &= \Phi\left(c \left|\det\left(\frac{\partial \widetilde{x}^j}{\partial x^i}\right)\right| |\omega_p|\right) \\ &= \left(p, c\left|\det\left(\frac{\partial \widetilde{x}^j}{\partial x^i}\right)\right|\right). \end{aligned}$$

Thus the hypotheses of Lemma 5.5 are satisfied, with the transition functions equal to $\left|\det(\partial \widetilde{x}^j/\partial x^i)\right|$. □

If M is a smooth n-manifold, a section of ΩM is called a *density on* M. (One might choose to call such a section a "density field" to distinguish it from a density on a vector space, but we will not do so.) If μ is a density and f is a continuous real-valued function, then $f\mu$ is again a density, which is smooth if both f and μ are. A density on M is said to be positive or negative if its value at each point has that property. Any nonvanishing n-form ω determines a positive density $|\omega|$, defined by $|\omega|_p = |\omega_p|$ for each $p \in M$. If ω is a nonvanishing n-form on an open set $U \subset M$, then any density μ on U can be written $\mu = f|\omega|$ for some real-valued function f.

One important fact about densities is that every manifold admits a global smooth positive density, without any orientability assumptions.

Lemma 14.28. *If M is a smooth manifold, there exists a smooth positive density on M.*

Proof. Because the set of positive elements of ΩM is a convex open subset, the usual partition of unity argument (Problem 11-22) allows us to piece together local densities to obtain a global smooth positive density. □

It is important to understand that this lemma works because positivity of a density is a well-defined property, independent of any choices of coordinates or orientations. There is no corresponding existence result for orientation forms because, without a choice of orientation, there is no way to decide which n-forms are positive.

Under smooth maps, densities pull back in the same way as differential forms. If $F\colon M \to N$ is a smooth map between n-manifolds and μ is a density on N, we define a density $F^*\mu$ on M by

$$F^*\mu(X_1, \dots, X_n) = \mu(F_*X_1, \dots, F_*X_n),$$

or more explicitly, for $p \in M$,

$$(F^*\mu)_p(X_1|_p, \dots, X_n|_p) = \mu_{F(p)}(F_*X_1|_p, \dots, F_*X_n|_p).$$

Lemma 14.29. *Let $F\colon M \to N$ and $G\colon P \to M$ be smooth maps, and let μ be a density on N.*

(*a*) *If μ is smooth, then $F^*\mu$ is a smooth density on M.*

(*b*) $(F \circ G)^*\mu = G^*(F^*\mu)$.

(*c*) *If f is a continuous real-valued function on N, then $F^*(f\mu) = (f \circ F)F^*\mu$.*

(*d*) *If ω is an n-form on N, then $F^*|\omega| = |F^*\omega|$.*

◇ **Exercise 14.6.** Prove the preceding lemma.

The next result shows how to compute the pullback of a density in coordinates. It is an analogue for densities of Proposition 12.12.

Proposition 14.30. *Suppose $F\colon M \to N$ is a smooth map between n-manifolds. If (x^i) and (y^j) are smooth coordinates on open sets $U \subset M$ and $V \subset N$, respectively, and u is a smooth real-valued function on V, then the following holds on $U \cap F^{-1}(V)$:*

$$F^*\left(u\left|dy^1 \wedge \cdots \wedge dy^n\right|\right) = (u \circ F)\left|\det DF\right|\left|dx^1 \wedge \cdots \wedge dx^n\right|, \qquad (14.14)$$

where DF represents the matrix of partial derivatives of F in these coordinates.

Proof. Using Proposition 12.12 and Lemma 14.29, we obtain

$$\begin{aligned} F^*\left(u\left|dy^1 \wedge \cdots \wedge dy^n\right|\right) &= (u \circ F)F^*\left|dy^1 \wedge \cdots \wedge dy^n\right| \\ &= (u \circ F)\left|F^*(dy^1 \wedge \cdots \wedge dy^n)\right| \\ &= (u \circ F)\left|(\det DF)\, dx^1 \wedge \cdots \wedge dx^n\right| \\ &= (u \circ F)\left|\det DF\right|\left|dx^1 \wedge \cdots \wedge dx^n\right|. \end{aligned}$$ □

Integration of Densities

Now we turn to integration. As we did with forms, we begin by defining integrals of densities on subsets of $\mathbb{R}^n$. If $D \subset \mathbb{R}^n$ is a compact domain of integration and μ is a density on U, we can write $\mu = f|dx^1 \wedge \cdots \wedge dx^n|$ for some uniquely determined continuous function $f\colon D \to \mathbb{R}$. We define the *integral of μ over D* by

$$\int_D \mu = \int_D f\, dV,$$

or more suggestively,

$$\int_D f\left|dx^1 \wedge \cdots \wedge dx^n\right| = \int_D f\, dx^1 \cdots dx^n.$$

Similarly, if U is an open subset of $\mathbb{R}^n$ or $\mathbb{H}^n$ and μ is compactly supported in U, we define

$$\int_U \mu = \int_D \mu,$$

where $D \subset U$ is any compact domain of integration containing the support of μ. The key fact is that this is diffeomorphism invariant.

Proposition 14.31. *If D and E are compact domains of integration in $\mathbb{R}^n$, and $G\colon D \to E$ is a smooth map that restricts to a diffeomorphism from $\operatorname{Int} D$ to $\operatorname{Int} E$, then*

$$\int_E \mu = \int_D G^*\mu.$$

for any density μ on E. Similarly, if $U, V \subset \mathbb{R}^n$ are open sets and $G\colon U \to V$ is a diffeomorphism, then

$$\int_V \mu = \int_U G^*\mu.$$

for any compactly supported density μ on V.

Proof. The proof is essentially identical to those of Proposition 14.2 and Corollary 14.3, using (14.14) instead of (12.12). □

Now let M be a smooth n-manifold (with or without boundary). If μ is a density on M whose support is contained in the domain of a single smooth chart (U, φ), the *integral of μ over M* is defined as

$$\int_M \mu = \int_{\varphi(U)} (\varphi^{-1})^*\mu.$$

This is extended to arbitrary densities μ by setting

$$\int_M \mu = \sum_i \int_M \psi_i \mu,$$

where $\{\psi_i\}$ is a smooth partition of unity subordinate to an open cover of M by smooth charts. The fact that this is independent of the choices of coordinates or partition of unity follows just as in the case of forms.

The following proposition is proved in the same way as Proposition 14.6.

Proposition 14.32 (Properties of Integrals of Densities). *Suppose M and N are smooth n-manifolds with or without boundaries, and μ, ν are compactly supported densities on M.*

(*a*) Linearity: *If $a, b \in \mathbb{R}$, then*

$$\int_M a\mu + b\eta = a\int_M \mu + b\int_M \eta.$$

(*b*) Positivity: *If μ is a positive density, then $\int_M \mu > 0$.*

(*c*) Diffeomorphism Invariance: *If $F\colon N \to M$ is a diffeomorphism, then $\int_M \mu = \int_N F^*\mu$.*

◇ **Exercise 14.7.** Prove Proposition 14.32.

Just as for forms, integrals of densities are usually computed by cutting the manifold into pieces and parametrizing each piece, just as in Proposition 14.7. The details are left to the reader.

◇ **Exercise 14.8.** Formulate and prove an analogue of Proposition 14.7 for densities.

The Riemannian Density

Densities are particularly useful on Riemannian manifolds. Throughout the rest of this section, (M,g) will be a Riemannian manifold with or without boundary.

Lemma 14.33 (The Riemannian Density). *Let (M,g) be a Riemannian manifold with or without boundary. There is a unique smooth positive density μ on M, called the Riemannian density, with the property that*

$$\mu(E_1,\dots,E_n)=1 \tag{14.15}$$

for any local orthonormal frame (E_i).

Proof. Uniqueness is obvious, because any two densities that agree on the elements of a basis must be equal. Given any point $p\in M$, let U be a connected smooth coordinate neighborhood of p. Since U is diffeomorphic to an open subset of Euclidean space, it is orientable. Any choice of orientation of U uniquely determines a Riemannian volume form dV_g, with the property that $dV_g(E_1,\dots,E_n)=1$ for any oriented orthonormal frame. If we put $\mu=|dV_g|$, it follows easily that μ is a smooth positive density on U satisfying (14.15). If U and V are two overlapping smooth coordinate neighborhoods, the two definitions of μ agree where they overlap by uniqueness, so this defines μ globally. □

◇ **Exercise 14.9.** Let (M,g) be an oriented Riemannian manifold with or without boundary and let dV_g be its Riemannian volume form.

(a) Show that the Riemannian density of M is equal to $|dV_g|$.
(b) For any compactly supported continuous function $f\colon M\to\mathbb{R}$, show that

$$\int_M f\,|dV_g|=\int_M f\,dV_g.$$

Because of part (b) of this exercise, it is customary to denote the Riemannian density simply by dV_g, and to specify when necessary whether the notation refers to a density or a form. If $f\colon M\to\mathbb{R}$ is a compactly supported continuous function, the *integral of f over M* is defined to be $\int_M f\,dV_g$. Exercise 14.9 shows that when M is oriented, it does not matter whether we interpret dV_g as the Riemannian volume form or the Riemannian density. (If the orientation of M is changed, then both the integral and dV_g change signs, so the result is the same.) When M is not orientable, however, we have no choice but to interpret it as a density.

One of the most useful applications of densities is that they enable us to generalize the divergence theorem to nonorientable manifolds. If X is a smooth vector field on M, it turns out that the divergence of X can be defined even when M is not orientable (see Problem 14-21). The next lemma shows that the divergence theorem holds in that case as well.

Theorem 14.34. *Suppose (M, g) is a Riemannian manifold with boundary, orientable or not. For any compactly supported smooth vector field X on M,*

$$\int_M (\operatorname{div} X)\, dV_g = \int_{\partial M} \langle X, N \rangle_g \, dV_{\widetilde{g}}, \tag{14.16}$$

where N is the outward-pointing unit normal vector field along ∂M, $\widetilde{g}$ is the induced Riemannian metric on ∂M, and dV_g, $dV_{\widetilde{g}}$ are the Riemannian densities of g and $\widetilde{g}$, respectively.

Sketch of proof. One can show that M has a 2-sheeted orientation covering $\widehat{\pi}\colon \widehat{M} \to M$, which satisfies the same properties as in the case of manifolds without boundary. (One way to see this is to use the fact that M can be embedded in a larger manifold M' without boundary of the same dimension, and apply Theorem 13.9) to M'. See Problem 17-13 for the compact case.) Define metrics $\widehat{g} = \widehat{\pi}^* g$ on $\widehat{M}$ and $\overline{g} = \widehat{g}|_{\partial \widehat{M}}$ on $\partial \widehat{M}$. For this proof, we will denote the Riemannian volume element of $\widehat{g}$ by $dV_{\widehat{g}}$ and its Riemannian density by $|dV_{\widehat{g}}|$, with similar notations for $\overline{g}$. It is straightforward to verify the following facts:

- $\widehat{M}$ is a smooth manifold with boundary.
- The restriction of $\widehat{\pi}$ to $\partial \widehat{M}$ is a smooth two-sheeted covering of ∂M.
- $\widehat{g}$ is a Riemannian metric on $\widehat{M}$.
- $\widehat{\pi}^* dV_g = |dV_{\widehat{g}}|$.
- $(\widehat{\pi}|_{\partial \widehat{M}})^* dV_{\widetilde{g}} = |dV_{\overline{g}}|$.
- There is a unique smooth vector field $\widehat{X}$ on $\widehat{M}$ that is $\widehat{\pi}$-related to X.
- The outward unit normal $\widehat{N}$ along $\partial \widehat{M}$ is $\widehat{\pi}$-related to N.
- $(\operatorname{div} X) \circ \widehat{\pi} = \operatorname{div} \widehat{X}$.
- $(\langle X, N \rangle_g) \circ (\widehat{\pi}|_{\partial \widehat{M}}) = \langle \widehat{X}, \widehat{N} \rangle_{\widehat{g}}$.

Using these facts, together with the divergence theorem on $\widehat{M}$ and the result of Problem 14-6, we compute

$$\begin{aligned}
2 \int_M (\operatorname{div} X)\, dV_g &= \int_{\widehat{M}} \widehat{\pi}^* \left((\operatorname{div} X)\, dV_g \right) \\
&= \int_{\widehat{M}} (\operatorname{div} \widehat{X})\ |dV_{\widehat{g}}| \\
&= \int_{\widehat{M}} (\operatorname{div} \widehat{X})\, dV_{\widehat{g}} \\
&= \int_{\partial \widehat{M}} \langle \widehat{X}, \widehat{N} \rangle_{\widehat{g}} \, dV_{\overline{g}}
\end{aligned}$$

$$\begin{aligned}
&= \int_{\partial \widehat{M}} \langle \widehat{X}, \widehat{N} \rangle_{\widehat{g}} \, |dV_{\widehat{g}}| \\
&= \int_{\partial \widehat{M}} (\widehat{\pi}|_{\partial \widehat{M}})^* \left(\langle X, N \rangle_g \, dV_{\widehat{g}} \right) \\
&= 2 \int_{\partial M} \langle X, N \rangle_g \, dV_g.
\end{aligned}$$

Dividing both sides by 2 yields (14.16). □

Problems

14-1. Let $\mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1 \subset \mathbb{R}^4$ denote the 2-torus, defined by $w^2 + x^2 = y^2 + z^2 = 1$, with the orientation determined by its product structure (see Exercise 13.4). Compute $\int_{\mathbb{T}^2} \omega$, where ω is the following 2-form on $\mathbb{R}^4$:

$$\omega = wy \, dx \wedge dz.$$

14-2. Compute $\int_{\mathbb{S}^2} \omega$, where $\mathbb{S}^2$ has its standard orientation and

$$\omega = x^3 \, dy \wedge dz.$$

14-3. Let D denote the torus of revolution in $\mathbb{R}^3$ obtained by revolving the circle $(y-2)^2 + z^2 = 1$ around the z-axis (Example 11.23), with its induced Riemannian metric and with the orientation induced by the outward unit normal.

(a) Compute the surface area of D.
(b) Compute the integral over D of the function $f(x, y, z) = z^2 + 1$.
(c) Compute the integral over D of the 2-form $\omega = z \, dx \wedge dy$.

14-4. Let ω be the $(n-1)$-form on $\mathbb{R}^n \smallsetminus \{0\}$ defined by

$$\omega = |x|^{-n} \sum_{i=1}^{n} (-1)^{i-1} x^i \, dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n. \tag{14.17}$$

(a) Show that $\omega|_{\mathbb{S}^{n-1}}$ is the Riemannian volume element of $\mathbb{S}^{n-1}$ with respect to the round metric.
(b) Show that ω is closed but not exact on $\mathbb{R}^n \smallsetminus \{0\}$.

14-5. Define maps $F_+, F_- : \mathbb{B}^2 \to \mathbb{S}^2$ by

$$F_\pm(u, v) = (u, v, \pm\sqrt{1 - u^2 - v^2}).$$

If ω is a smooth 2-form on $\mathbb{S}^2$, show that

$$\int_{\mathbb{S}^2} \omega = \int_{\mathbb{B}^2} F_+^* \omega - \int_{\mathbb{B}^2} F_-^* \omega,$$

where the integrals on the right-hand side are defined as the limits as $R \nearrow 1$ of the integrals over $\overline{B}_R(0)$. Be sure to justify the limits.

14-6. Suppose $\widetilde{M}$ and M are smooth n-manifolds and $\pi\colon \widetilde{M} \to M$ is a smooth k-sheeted covering map.

(a) If $\widetilde{M}$ and M are oriented and π is orientation-preserving, show that $\int_{\widetilde{M}} \pi^*\omega = k \int_M \omega$ for any compactly supported n-form ω on M.

(b) If μ is any compactly supported density on M, show that $\int_{\widetilde{M}} \pi^*\mu = k \int_M \mu$.

14-7. If M is a compact, smooth, oriented manifold with boundary, show that there does not exist a retraction of M onto its boundary. [Hint: First show that there cannot exist a smooth retraction by considering an orientation form on ∂M.]

14-8. Let (M, g) be a compact, oriented Riemannian manifold with boundary, let $\widetilde{g}$ denote the induced Riemannian metric on ∂M, and let N be the outward unit normal vector field along ∂M.

(a) Show that the divergence operator satisfies the following product rule for $f \in C^\infty(M)$, $X \in \mathfrak{T}(M)$:

$$\operatorname{div}(fX) = f \operatorname{div} X + \langle \operatorname{grad} f, X\rangle_g.$$

(b) Prove the following "integration by parts" formula:

$$\int_M \langle \operatorname{grad} f, X\rangle_g \, dV_g = \int_{\partial M} f\langle X, N\rangle_g \, dV_{\widetilde{g}} - \int_M (f \operatorname{div} X)\, dV_g.$$

(c) Explain what this has to do with integration by parts.

14-9. Let (M, g) be an oriented Riemannian manifold with or without boundary. The linear operator $\Delta\colon C^\infty(M) \to C^\infty(M)$ defined by $\Delta u = -\operatorname{div}(\operatorname{grad} u)$ is called the *Laplace operator* or *Laplacian*. A function $u \in C^\infty(M)$ is said to be *harmonic* if $\Delta u = 0$.

(a) If M is compact, prove *Green's identities*:

$$\int_M u\Delta v \, dV_g = \int_M \langle \operatorname{grad} u, \operatorname{grad} v\rangle_g dV_g - \int_{\partial M} u\, Nv\, dV_{\widetilde{g}},$$

$$\int_M (u\Delta v - v\Delta u)\, dV_g = \int_{\partial M} (v\, Nu - u\, Nv) dV_{\widetilde{g}}.$$

where N and $\widetilde{g}$ are as in Problem 14-8.

(b) If M is compact and connected and $\partial M = \varnothing$, show that the only harmonic functions on M are the constants.

(c) If M is compact and connected, $\partial M \neq \varnothing$, and u, v are harmonic functions on M whose restrictions to ∂M agree, show that $u \equiv v$.

[Remark: There is no general agreement about the sign convention for the Laplacian on a Riemannian manifold, and many authors define the Laplacian to be the negative of the one we have defined here.

Although the definition used here conflicts with the traditional definition of the Laplacian on $\mathbb{R}^n$ (see Problem 14-12), it has two distinct advantages: Our Laplacian has nonnegative eigenvalues (see Problem 14-10), and it agrees with the Laplace-Beltrami operator defined on differential forms (see Problems 15-9 and 15-10). When reading any book or article that mentions the Laplacian, you have to be careful to determine which sign convention the author is using.]

14-10. Let (M, g) be a compact, connected, oriented Riemannian manifold without boundary, and let Δ be its Laplacian. A real number λ is called an *eigenvalue* of Δ if there exists a smooth real-valued function u on M, not identically zero, such that $\Delta u = \lambda u$. In this case, u is called an *eigenfunction* corresponding to λ.

(a) Prove that 0 is an eigenvalue of Δ, and that all other eigenvalues are strictly positive.

(b) If u and v are eigenfunctions corresponding to distinct eigenvalues, show that $\int_M uv\, dV_g = 0$.

14-11. Let M be a compact, connected, oriented Riemannian n-manifold with nonempty boundary. A number $\lambda \in \mathbb{R}$ is called a *Dirichlet eigenvalue* for M if there exists a smooth real-valued function u on M, not identically zero, such that $\Delta u = \lambda u$ and $u|_{\partial M} = 0$. Similarly, λ is called a *Neumann eigenvalue* if there exists such a u satisfying $\Delta u = \lambda u$ and $Nu|_{\partial M} = 0$, where N is the outward unit normal.

(a) Show that every Dirichlet eigenvalue is strictly positive.

(b) Show that 0 is a Neumann eigenvalue, and all other Neumann eigenvalues are strictly positive.

14-12. Let (M, g) be an oriented Riemannian n-manifold with or without boundary.

(a) In any oriented smooth local coordinates (x^i), show that

$$\operatorname{div}\left(X^i \frac{\partial}{\partial x^i}\right) = \frac{1}{\sqrt{\det g}} \frac{\partial}{\partial x^i}\left(X^i \sqrt{\det g}\right),$$

where $\det g = \det(g_{kl})$ is the determinant of the component matrix of g in these coordinates.

(b) Show that the Laplacian is given in any oriented smooth local coordinates by

$$\Delta u = -\frac{1}{\sqrt{\det g}} \frac{\partial}{\partial x^i}\left(g^{ij} \sqrt{\det g} \frac{\partial u}{\partial x^j}\right).$$

(c) Conclude that on $\mathbb{R}^n$ with the Euclidean metric and standard coordinates,

$$\operatorname{div}\left(X^i \frac{\partial}{\partial x^i}\right) = \sum_{i=1}^n \frac{\partial X^i}{\partial x^i}, \qquad \Delta u = -\sum_{i=1}^n \frac{\partial^2 u}{(\partial x^i)^2}.$$

14-13. Let (M, g) be an oriented Riemannian n-manifold. This problem outlines an important generalization of the operator $*\colon C^\infty(M) \to \mathcal{A}^n(M)$ defined in this chapter.

(a) For each $k = 1, \ldots, n$, show that there is a unique inner product on $\Lambda^k(T_pM)$ with the following property: If (E_i) is any orthonormal basis for T_pM and (ε^i) is the dual basis, then $\{\varepsilon^I : I \text{ is increasing}\}$ is an orthonormal basis for $\Lambda^k(T_pM)$.

(b) For each $k = 0, \ldots, n$, show that there is a unique smooth bundle map $*\colon \Lambda^k M \to \Lambda^{n-k} M$ satisfying

$$\omega \wedge * \eta = \langle \omega, \eta \rangle_g dV_g.$$

(For $k = 0$, interpret the inner product as ordinary multiplication.) This map is called the *Hodge star operator*. [Hint: First prove uniqueness, and then define $*$ locally by setting

$$*(\varepsilon^{i_1} \wedge \cdots \wedge \varepsilon^{i_k}) = \pm \varepsilon^{j_1} \wedge \cdots \wedge \varepsilon^{j_{n-k}}$$

in terms of an orthonormal coframe (ε^i), where the indices $j_1, \ldots, j_{n-k}$ are chosen so that $(i_1, \ldots, i_k, j_1, \ldots, j_{n-k})$ is some permutation of $(1, \ldots, n)$.]

(c) Show that $*\colon \Lambda^0 M \to \Lambda^n M$ is given by $* f = f\, dV_g$.

(d) Show that $* * \omega = (-1)^{k(n-k)} \omega$ if ω is a k-form.

14-14. Consider $\mathbb{R}^n$ as a Riemannian manifold with the Euclidean metric and the standard orientation.

(a) Calculate $* dx^i$ for $i = 1, \ldots, n$.

(b) Calculate $*(dx^i \wedge dx^j)$ in the case $n = 4$.

14-15. Let M be an oriented Riemannian 4-manifold. A 2-form ω on M is said to be *self-dual* if $* \omega = \omega$, and *anti-self-dual* if $* \omega = -\omega$.

(a) Show that every 2-form ω on M can be written uniquely as a sum of a self-dual form and an anti-self-dual form.

(b) On $M = \mathbb{R}^4$ with the Euclidean metric, determine the self-dual and anti-self-dual forms in standard coordinates.

14-16. Let (M, g) be a Riemannian manifold and $X \in \mathcal{T}(M)$. Show that

$$X \lrcorner\, dV_g = * X^\flat,$$
$$\operatorname{div} X = * d * X^\flat,$$

and, when $\dim M = 3$,

$$\operatorname{curl} X = (* dX^\flat)^\#.$$

14-17. Let (M, g) be a Riemannian n-manifold. For $1 \le k \le n$, define a map $d^*\colon \mathcal{A}^k(M) \to \mathcal{A}^{k-1}(M)$ by $d^*\omega = (-1)^{n(k+1)+1} * d * \omega$, where $*$ is the Hodge star operator defined in Problem 14-13. Extend this definition to 0-forms by defining $d^*\omega = 0$ for $\omega \in \mathcal{A}^0(M)$.

(a) Show that $d^* \circ d^* = 0$.
(b) If M is compact, show that the formula

$$(\omega, \eta) = \int_M \langle \omega, \eta \rangle_g \, dV_g$$

defines an inner product on $\mathcal{A}^k(M)$ for each k.
(c) Show that $(d^*\omega, \eta) = (\omega, d\eta)$ for all $\omega \in \mathcal{A}^k(M)$ and $\eta \in \mathcal{A}^{k-1}(M)$.

14-18. On $\mathbb{R}^3$ with the Euclidean metric, show that the curl operator we have defined is given by the classical formula:

$$\begin{aligned} &\operatorname{curl}\left(P\frac{\partial}{\partial x} + Q\frac{\partial}{\partial y} + R\frac{\partial}{\partial z}\right) \\ &\quad = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\frac{\partial}{\partial x} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\frac{\partial}{\partial y} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\frac{\partial}{\partial z}. \end{aligned}$$

14-19. Show that any finite product $M_1 \times \cdots \times M_k$ of smooth manifolds with corners is again a smooth manifold with corners. Give a counterexample to show that a finite product of smooth manifolds with boundary need not be a smooth manifold with boundary.

14-20. Suppose M is a smooth manifold with corners, and let $\mathcal{C}$ denote the set of corner points of M. Show that $M \smallsetminus \mathcal{C}$ is a smooth manifold with boundary.

14-21. Show that the divergence operator on an oriented Riemannian manifold does not depend on the choice of orientation, and conclude that it is invariantly defined on all Riemannian manifolds.

14-22. Let M and N be connected, oriented, smooth manifolds, and suppose $F, G\colon M \to N$ are diffeomorphisms. If F and G are homotopic, show that they are both either orientation-preserving or orientation-reversing. [Hint: Use the Whitney approximation theorem and Stokes's theorem on $M \times I$.]

14-23. The Hairy Ball Theorem: *There exists a nowhere-vanishing vector field on $\mathbb{S}^n$ if and only if n is odd.* ("You cannot comb the hair on a ball.") Prove this by showing the following are equivalent:

(a) There exists a nowhere-vanishing vector field on $\mathbb{S}^n$.
(b) There exists a continuous map $V\colon \mathbb{S}^n \to \mathbb{S}^n$ satisfying $V(x) \perp x$ (with respect to the Euclidean dot product on $\mathbb{R}^{n+1}$) for all $x \in \mathbb{S}^n$.
(c) The antipodal map $\alpha\colon \mathbb{S}^n \to \mathbb{S}^n$ is homotopic to $\mathrm{Id}_{\mathbb{S}^n}$.
(d) The antipodal map $\alpha\colon \mathbb{S}^n \to \mathbb{S}^n$ is orientation-preserving.
(e) n is odd.

[Hint: Use Problems 8-7, 13-6, and 14-22.]

15
De Rham Cohomology

In Chapter 12 we defined closed and exact forms: A smooth differential form ω is closed if $d\omega = 0$, and exact if it is of the form $d\eta$. Because $d \circ d = 0$, every exact form is closed. In this chapter, we explore the implications of the converse question: Is every closed form exact? The answer, in general, is no: In Example 6.29 we saw an example of a 1-form on $\mathbb{R}^2 \smallsetminus \{0\}$ that was closed but not exact. In that example, the failure of exactness seemed to be a consequence of the "hole" in the center of the domain. For higher-degree forms, the answer to the question depends on subtle topological properties of the manifold, connected with the existence of "holes" of higher dimensions. Making this dependence quantitative leads to a new set of invariants of smooth manifolds, called the de Rham cohomology groups, which are the subject of this chapter.

There are many situations in which knowledge of which closed forms are exact has important consequences. For example, Stokes's theorem implies that if ω is exact, then the integral of ω over any compact submanifold without boundary is zero. Proposition 6.24 showed that a smooth 1-form is conservative if and only if it is exact.

We begin by defining the de Rham cohomology groups and proving some of their basic properties, including diffeomorphism invariance. Then we prove that they are in fact *homotopy* invariants, which implies in particular that they are topological invariants. Using elementary methods, we compute some de Rham groups, including the zero-dimensional groups of all manifolds, the 1-dimensional groups of simply connected manifolds, the top-dimensional groups of compact manifolds, and all of the de Rham groups of star-shaped open subsets of $\mathbb{R}^n$. Then we prove a general theorem that expresses the de Rham groups of a manifold in terms of those of its open subsets, called the Mayer–Vietoris theorem, and use it to compute all the de Rham groups of spheres.

The de Rham Cohomology Groups

In Chapter 6, we studied the closed 1-form

$$\omega = \frac{x\,dy - y\,dx}{x^2 + y^2}, \tag{15.1}$$

and showed that it is not exact on $\mathbb{R}^2 \smallsetminus \{0\}$, but it is exact on some smaller domains such as the right half-plane $H = \{(x, y) : x > 0\}$, where it is equal to $d\theta$ (see Example 6.29).

As we will see in this chapter, this behavior is typical: Closed p-forms are always *locally* exact, so the question of whether a given closed form is exact depends on the global shape of the domain, not on the local properties of the form.

Let M be a smooth manifold. Because $d\colon \mathcal{A}^p(M) \to \mathcal{A}^{p+1}(M)$ is linear, its kernel and image are linear subspaces. We define

$$\begin{aligned}
\mathcal{Z}^p(M) &= \operatorname{Ker}[d\colon \mathcal{A}^p(M) \to \mathcal{A}^{p+1}(M)] = \{\text{closed } p\text{-forms on } M\},\\
\mathcal{B}^p(M) &= \operatorname{Im}[d\colon \mathcal{A}^{p-1}(M) \to \mathcal{A}^p(M)] = \{\text{exact } p\text{-forms on } M\}.
\end{aligned}$$

By convention, we consider $\mathcal{A}^p(M)$ to be the zero vector space when $p < 0$ or $p > n = \dim M$, so that, for example, $\mathcal{B}^0(M) = 0$ and $\mathcal{Z}^n(M) = \mathcal{A}^n(M)$.

The fact that every exact form is closed implies that $\mathcal{B}^p(M) \subset \mathcal{Z}^p(M)$. Thus it makes sense to define the pth *de Rham cohomology group* (or just *de Rham group*) of M to be the quotient vector space

$$H^p_{dR}(M) = \frac{\mathcal{Z}^p(M)}{\mathcal{B}^p(M)}.$$

(It is a real vector space, and thus in particular a group under vector addition. Perhaps "de Rham cohomology space" would be a more appropriate term, but because most other cohomology theories produce only groups it is traditional to use the term group in this context as well, bearing in mind that these "groups" are actually real vector spaces.) For any closed form ω on M, we let $[\omega]$ denote the equivalence class of ω in this quotient space, called the *cohomology class* of ω. Clearly $H^p_{dR}(M) = 0$ for $p < 0$ or $p > \dim M$, because $\mathcal{A}^p(M) = 0$ in those cases. If $[\omega] = [\omega']$ (that is, if ω and ω' differ by an exact form), we say that ω and ω' are *cohomologous*.

The first order of business is to show that the de Rham groups are diffeomorphism invariants.

Proposition 15.1 (Induced Cohomology Maps). *For any smooth map $G\colon M \to N$, the pullback $G^*\colon \mathcal{A}^p(N) \to \mathcal{A}^p(M)$ carries $\mathcal{Z}^p(N)$ into*

$\mathcal{Z}^p(M)$ and $\mathcal{B}^p(N)$ into $\mathcal{B}^p(M)$. It thus descends to a linear map, still denoted by G^*, from $H^p_{dR}(N)$ to $H^p_{dR}(M)$, called the induced cohomology map. It has the following properties:

(*a*) *If $F\colon N \to P$ is another smooth map, then*
$$(F \circ G)^* = G^* \circ F^* \colon H^p_{dR}(P) \to H^p_{dR}(M).$$

(*b*) *If Id_M denotes the identity map of M, then $(\mathrm{Id}_M)^*$ is the identity map of $H^p_{dR}(M)$.*

Proof. If ω is closed, then
$$d(G^*\omega) = G^*(d\omega) = 0,$$
so $G^*\omega$ is also closed. If $\omega = d\eta$ is exact, then
$$G^*\omega = G^*(d\eta) = d(G^*\eta),$$
which is also exact. Therefore, G^* maps $\mathcal{Z}^p(N)$ into $\mathcal{Z}^p(M)$ and $\mathcal{B}^p(N)$ into $\mathcal{B}^p(M)$. The induced cohomology map $G^*\colon H^p_{dR}(N) \to H^p_{dR}(M)$ is defined in the obvious way: For a closed p-form ω, let
$$G^*[\omega] = [G^*\omega].$$
If $\omega' = \omega + d\eta$, then $[G^*\omega'] = [G^*\omega + d(G^*\eta)] = [G^*\omega]$, so this map is well-defined. Properties (a) and (b) follow immediately from the analogous properties for the pullback map on forms. □

The next two corollaries are immediate.

Corollary 15.2 (Functoriality). *For each integer $p \geq 0$, the assignment $M \mapsto H^p_{dR}(M)$, $F \mapsto F^*$ is a contravariant functor from the category of smooth manifolds and smooth maps to the category of real vector spaces and linear maps.*

Corollary 15.3 (Diffeomorphism Invariance). *Diffeomorphic manifolds have isomorphic de Rham cohomology groups.*

Homotopy Invariance

In this section, we will present a profound generalization of Corollary 15.3, one surprising consequence of which will be that the de Rham cohomology groups are actually *topological* invariants. In fact, they are something much more: They are *homotopy invariants*, which means that homotopy equivalent manifolds have isomorphic de Rham groups. (See page 556 for the definition of homotopy equivalence.)

The underlying fact that will allow us to prove the homotopy invariance of de Rham cohomology is that homotopic smooth maps induce the same cohomology map. To motivate the proof, suppose $F, G\colon M \to N$ are smoothly homotopic maps, and let us think about what needs to be shown. Given a closed p-form ω on N, we need somehow to produce a $(p-1)$-form η on M such that

$$d\eta = G^*\omega - F^*\omega. \tag{15.2}$$

One might hope to construct η in a systematic way, resulting in a map h from closed p-forms on N to $(p-1)$-forms on M that satisfies

$$d(h\omega) = G^*\omega - F^*\omega. \tag{15.3}$$

Instead of defining $h\omega$ only when ω is closed, it turns out to be far simpler to define for each p a map h from the space of *all* smooth p-forms on N to the space of smooth $(p-1)$-forms on M. Such maps cannot satisfy (15.3), but instead we will find maps that satisfy

$$d(h\omega) + h(d\omega) = G^*\omega - F^*\omega. \tag{15.4}$$

This implies (15.3) when ω is closed.

In general, if $F, G\colon M \to N$ are smooth maps, a collection of linear maps $h\colon \mathcal{A}^p(N) \to \mathcal{A}^{p-1}(M)$ such that (15.4) is satisfied for all ω is called a *homotopy operator* between F^* and G^*. (The term *cochain homotopy* is used frequently in the algebraic topology literature.) The key to our proof of homotopy invariance will be to construct a homotopy operator first in the following special case. For each $t \in [0,1]$, let $i_t\colon M \to M \times I$ be the embedding

$$i_t(x) = (x, t).$$

Clearly i_0 is homotopic to i_1. (The homotopy is the identity map of $M \times I$!)

Lemma 15.4 (Existence of a Homotopy Operator). *For any smooth manifold M, there exists a homotopy operator between i_0^* and i_1^*.*

Proof. For each p, we need to define a linear map $h\colon \mathcal{A}^p(M\times I) \to \mathcal{A}^{p-1}(M)$ such that

$$h(d\omega) + d(h\omega) = i_1^*\omega - i_0^*\omega. \tag{15.5}$$

We define h by the formula

$$h\omega = \int_0^1 \left(\frac{\partial}{\partial t} \lrcorner\, \omega\right) dt,$$

where t is the standard coordinate on I. More explicitly, $h\omega$ is the $(p-1)$-form on M whose action on vectors $X_1, \dots, X_{p-1} \in T_qM$ is

$$\begin{aligned}(h\omega)_q(X_1, \dots, X_{p-1}) &= \int_0^1 \left(\frac{\partial}{\partial t} \lrcorner\, \omega_{(q,t)}\right)(X_1, \dots, X_{p-1})\, dt \\ &= \int_0^1 \omega_{(q,t)}(\partial/\partial t, X_1, \dots, X_{p-1})\, dt.\end{aligned}$$

To show that h satisfies (15.5), we choose smooth local coordinates (x^i) on M, and consider separately the cases in which $\omega = f(x,t)\,dt \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_{p-1}}$ and $\omega = f(x,t)\,dx^{i_1} \wedge \cdots \wedge dx^{i_p}$. Since h is linear and every p-form on $M \times I$ can be written locally as a sum of such forms, this suffices.

Case I: $\omega = f(x,t)\,dt \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_{p-1}}$. In this case,

$$\begin{aligned}
d(h\omega) &= d\left(\left(\int_0^1 f(x,t)\,dt\right) dx^{i_1} \wedge \cdots \wedge dx^{i_{p-1}}\right)\\
&= \frac{\partial}{\partial x^j}\left(\int_0^1 f(x,t)\,dt\right) dx^j \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_{p-1}}\\
&= \left(\int_0^1 \frac{\partial f}{\partial x^j}(x,t)\,dt\right) dx^j \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_{p-1}}.
\end{aligned}$$

On the other hand, because $dt \wedge dt = 0$,

$$\begin{aligned}
h(d\omega) &= h\left(\frac{\partial f}{\partial x^j} dx^j \wedge dt \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_{p-1}}\right)\\
&= \int_0^1 \frac{\partial f}{\partial x^j}(x,t) \frac{\partial}{\partial t} \lrcorner \left(dx^j \wedge dt \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_{p-1}}\right) dt\\
&= -\left(\int_0^1 \frac{\partial f}{\partial x^j}(x,t)\,dt\right) dx^j \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_{p-1}}\\
&= -d(h\omega).
\end{aligned}$$

Thus the left-hand side of (15.5) is zero in this case. The right-hand side is zero as well, because $i_0^* dt = i_1^* dt = 0$ (since $t \circ i_0$ and $t \circ i_1$ are constant functions).

Case II: $\omega = f\,dx^{i_1} \wedge \cdots \wedge dx^{i_p}$. Now $\partial/\partial t \lrcorner \omega = 0$, which implies that $d(h\omega) = 0$. On the other hand, by the fundamental theorem of calculus,

$$\begin{aligned}
h(d\omega) &= h\left(\frac{\partial f}{\partial t} dt \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p} + \text{ terms without } dt\right)\\
&= \left(\int_0^1 \frac{\partial f}{\partial t}(x,t) dt\right) dx^{i_1} \wedge \cdots \wedge dx^{i_p}\\
&= (f(x,1) - f(x,0)) dx^{i_1} \wedge \cdots \wedge dx^{i_p}\\
&= i_1^*\omega - i_0^*\omega,
\end{aligned}$$

which proves (15.5) in this case. □

Proposition 15.5. *Let $F, G\colon M \to N$ be homotopic smooth maps. For every p, the induced cohomology maps $F^*, G^*\colon H^p_{dR}(N) \to H^p_{dR}(M)$ are equal.*

Proof. By Proposition 10.22, there is a smooth homotopy $H\colon M \times I \to N$ from F to G. This means that $H \circ i_0 = F$ and $H \circ i_1 = G$, where $i_0, i_1\colon M \to M \times I$ are defined as above (see Figure 15.1). Let $\widetilde{h}$ be the composite map

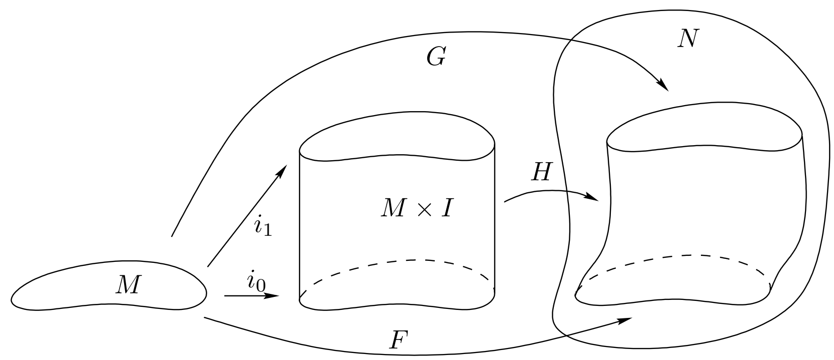


Figure 15.1. Homotopic maps.

$\widetilde{h} = h \circ H^* \colon \mathcal{A}^p(N) \to \mathcal{A}^{p-1}(M)$:

$$\mathcal{A}^p(N) \xrightarrow{H^*} \mathcal{A}^p(M \times I) \xrightarrow{h} \mathcal{A}^{p-1}(M),$$

where h is the homotopy operator constructed in Lemma 15.4.

For any $\omega \in \mathcal{A}^p(N)$, we compute

$$\begin{aligned}
\widetilde{h}(d\omega) + d(\widetilde{h}\omega) &= h(H^* d\omega) + d(hH^*\omega) \\
&= hd(H^*\omega) + dh(H^*\omega) \\
&= i_1^* H^*\omega - i_0^* H^*\omega \\
&= (H \circ i_1)^*\omega - (H \circ i_0)^*\omega \\
&= G^*\omega - F^*\omega.
\end{aligned}$$

Thus if ω is closed,

$$\begin{aligned}
G^*[\omega] - F^*[\omega] &= [G^*\omega - F^*\omega] \\
&= [\widetilde{h}(d\omega) + d(\widetilde{h}\omega)] \\
&= 0,
\end{aligned}$$

where the last line follows from $d\omega = 0$ and the fact that the cohomology class of any exact form is zero. □

The next theorem is the main result of this section.

Theorem 15.6 (Homotopy Invariance). *If M, N are homotopy equivalent smooth manifolds, then $H^p_{dR}(M) \cong H^p_{dR}(N)$ for each p. The isomorphism is induced by any smooth homotopy equivalence $F : M \to N$.*

Proof. Suppose $F \colon M \to N$ is a homotopy equivalence, with homotopy inverse $G \colon N \to M$. By Theorem 10.21, there are smooth maps $\widetilde{F} \colon M \to N$ homotopic to F and $\widetilde{G} \colon N \to M$ homotopic to G. Because homotopy is preserved by composition, it follows that $\widetilde{F} \circ \widetilde{G} \simeq F \circ G \simeq \mathrm{Id}_N$ and $\widetilde{G} \circ \widetilde{F} \simeq G \circ F \simeq \mathrm{Id}_M$, so $\widetilde{F}$ and $\widetilde{G}$ are homotopy inverses of each other.

Now Proposition 15.5 shows that, on cohomology,

$$\widetilde{F}^* \circ \widetilde{G}^* = (\widetilde{G} \circ \widetilde{F})^* = (\mathrm{Id}_M)^* = \mathrm{Id}_{H^p_{dR}(M)}\,.$$

The same argument shows that $\widetilde{G}^* \circ \widetilde{F}^*$ is also the identity, so $\widetilde{F}^*\colon H^p_{dR}(N) \to H^p_{dR}(M)$ is an isomorphism. □

Because every homeomorphism is a homotopy equivalence, the next corollary is immediate.

Corollary 15.7 (Topological Invariance). *The de Rham cohomology groups are topological invariants: If M and N are homeomorphic smooth manifolds, then their de Rham cohomology groups are isomorphic.*

This result is remarkable, because the definition of the de Rham groups of M is intimately tied up with its smooth structure, and we had no reason to expect that different differentiable structures on the same topological manifold should give rise to the same de Rham groups.

Computations

The direct computation of the de Rham groups is not easy in general. However, in this section, we will compute them in several special cases.

We begin with disjoint unions.

Proposition 15.8 (Cohomology of Disjoint Unions). *Let $\{M_j\}$ be a countable collection of smooth manifolds, and let $M = \coprod_j M_j$. For each p, the inclusion maps $\iota_j\colon M_j \to M$ induce an isomorphism from $H^p_{dR}(M)$ to the direct product space $\prod_j H^p_{dR}(M_j)$.*

Proof. The pullback maps $\iota_j^*\colon \mathcal{A}^p(M) \to \mathcal{A}^p(M_j)$ already induce an isomorphism from $\mathcal{A}^p(M)$ to $\prod_j \mathcal{A}^p(M_j)$, namely

$$\omega \mapsto (\iota_1^*\omega, \iota_2^*\omega, \dots) = (\omega|_{M_1}, \omega|_{M_2}, \dots).$$

This map is injective because any smooth p-form whose restriction to each M_j is zero must itself be zero, and it is surjective because giving an arbitrary smooth p-form on each M_j defines one on M. □

Because of this proposition, each de Rham group of a disconnected manifold is just the direct product of the corresponding groups of its components. Thus we can concentrate henceforth on computing the de Rham groups of connected manifolds.

Our next computation gives an explicit characterization of zero-dimensional cohomology.

Proposition 15.9 (Zero-Dimensional Cohomology). *If M is a connected smooth manifold, $H^0_{dR}(M)$ is equal to the space of constant functions and is therefore 1-dimensional.*

Proof. Because there are no (-1)-forms, $\mathcal{B}^0(M) = 0$. A closed 0-form is a smooth real-valued function f such that $df = 0$, and since M is connected this is true if and only if f is constant. Thus $H^0_{dR}(M) = \mathcal{Z}^0(M) = \{\text{constants}\}$. $\square$

Corollary 15.10 (Cohomology of Zero-Manifolds). *If M is a zero-dimensional manifold, the dimension of $H^0_{dR}(M)$ is equal to the cardinality of M, and all other de Rham cohomology groups vanish.*

Proof. By Propositions 15.8 and 15.9, $H^0_{dR}(M)$ is isomorphic to the direct product of one copy of $\mathbb{R}$ for each component of M, which is to say each point. The cohomology groups in dimensions other than zero vanish for dimensional reasons. $\square$

Next we examine the de Rham cohomology of Euclidean space, and more generally of its star-shaped open subsets. (Recall that a subset $V \subset \mathbb{R}^n$ is said to be star-shaped with respect to a point $q \in V$ if for every $x \in V$, the line segment from q to x is entirely contained in V.) In Proposition 6.30, we showed that every closed 1-form on a star-shaped open subset of $\mathbb{R}^n$ is exact. The next theorem is a generalization of that result.

Theorem 15.11 (The Poincaré Lemma). *Let U be a star-shaped open subset of $\mathbb{R}^n$. Then $H^p_{dR}(U) = 0$ for $p \geq 1$.*

Proof. Suppose $U \subset \mathbb{R}^n$ is star-shaped with respect to q. The key feature of star-shaped sets is that they are *contractible*, which means the identity map of U is homotopic to the constant map sending U to q, by the obvious straight-line homotopy:

$$H(x,t) = q + t(x - q).$$

Thus the inclusion of $\{q\}$ into U is a homotopy equivalence. The Poincaré lemma then follows from the homotopy invariance of H^p_{dR} together with the obvious fact that $H^p_{dR}(\{q\}) = 0$ for $p \geq 1$ because $\{q\}$ is a 0-manifold. $\square$

The next two results are easy corollaries of the Poincaré lemma.

Corollary 15.12 (Cohomology of Euclidean Space). *For all $p \geq 1$, $H^p_{dR}(\mathbb{R}^n) = 0$.*

Proof. Euclidean space $\mathbb{R}^n$ is star-shaped. $\square$

Corollary 15.13 (Local Exactness of Closed Forms). *Let M be a smooth manifold, and let ω be a closed p-form on M, $p \geq 1$. For any $q \in M$, there is a neighborhood U of q on which ω is exact.*

Proof. Every $q \in M$ has a neighborhood diffeomorphic to an open ball in $\mathbb{R}^n$, which is star-shaped. The result follows from the diffeomorphism invariance of de Rham cohomology. $\square$

One of the most interesting special cases is that of simply connected manifolds, for which we can compute the first cohomology explicitly.

Theorem 15.14 (First Cohomology, Simply Connected Case). *If M is a simply connected smooth manifold, then $H^1_{dR}(M) = 0$.*

Proof. Let ω be a closed 1-form on M. We need to show that ω is exact. By Theorem 6.24, this is true if and only ω is conservative, that is, if and only if the line integral of ω around any closed curve is zero. Since any closed curve is path homotopic to a constant curve, the result follows from Theorem 14.21 and Proposition 6.18(b). $\square$

Finally, we turn our attention to the top-dimensional cohomology of compact manifolds. We begin with the orientable case. Suppose M is an oriented compact smooth n-manifold. There is a natural linear map $I\colon \mathcal{A}^n(M) \to \mathbb{R}$ given by integration over M:

$$I(\omega) = \int_M \omega.$$

Because the integral of any exact form is zero, I descends to a linear map, still denoted by the same symbol, from $H^n_{dR}(M)$ to $\mathbb{R}$. (Note that every smooth n-form on an n-manifold is closed.)

Theorem 15.15 (Top Cohomology, Orientable Case). *For any compact, connected, oriented, smooth n-manifold M, the integration map $I\colon H^n_{dR}(M) \to \mathbb{R}$ is an isomorphism. Thus $H^n_{dR}(M)$ is 1-dimensional, spanned by the cohomology class of any smooth orientation form.*

Proof. The zero-dimensional case is an immediate consequence of Corollary 15.10, so we may assume that $n \geq 1$. Let Ω_0 be a smooth orientation form for M, and set $b = \int_M \Omega_0$. By Proposition 14.6(c), $b > 0$. Thus $I\colon H^n_{dR}(M) \to \mathbb{R}$ is surjective because $I[a\Omega_0] = ab$ for any $a \in \mathbb{R}$. To complete the proof, we need only show that it is injective. In other words, we have to show the following: If ω is any smooth n-form satisfying $\int_M \omega = 0$, then ω is exact.

Let $\{U_1, \dots, U_m\}$ be a finite cover of M by open sets that are diffeomorphic to $\mathbb{R}^n$, and let $M_k = U_1 \cup \cdots \cup U_k$ for $k = 1, \dots, m$. Since M is connected, by reordering the sets if necessary, we may assume that $M_k \cap U_{k+1} \neq \varnothing$ for each k. We will prove the following claim by induction on k: *If ω is a compactly supported smooth n-form on M_k that satisfies $\int_{M_k} \omega = 0$, then there exists a compactly supported smooth $(n-1)$-form η on M_k such that $d\eta = \omega$.* When $k = m$, this is the statement we are seeking to prove, because every form on a compact manifold is compactly supported.

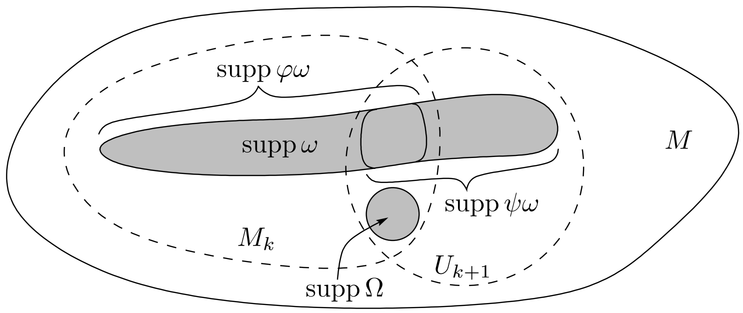


Figure 15.2. Computing the top-dimensional cohomology.

For $k = 1$, since $M_1 = U_1$ is diffeomorphic to $\mathbb{R}^n$, the claim reduces to a statement about compactly supported forms on $\mathbb{R}^n$. This will be proved as a separate lemma at the end of this section (Lemma 15.19). Assuming this for now, we continue with the induction.

Assume the claim is true for some $k \geq 1$, and suppose ω is a compactly supported smooth n-form on $M_{k+1} = M_k \cup U_{k+1}$ that satisfies $\int_{M_{k+1}} \omega = 0$. Choose an auxiliary smooth n-form $\Omega \in \mathcal{A}^n(M_{k+1})$ that is compactly supported in $M_k \cap U_{k+1}$ and satisfies $\int_{M_{k+1}} \Omega = 1$. (Such a form is easily constructed by using a bump function in coordinates.) Let $\{\varphi, \psi\}$ be a smooth partition of unity for M_{k+1} subordinate to the cover $\{M_k, U_{k+1}\}$ (Figure 15.2). Let $c = \int_{M_{k+1}} \varphi\omega$. Observe that $\varphi\omega - c\Omega$ is compactly supported in M_k, and its integral is equal to zero by our choice of c. Therefore, by the induction hypothesis, there is a compactly supported smooth $(n-1)$-form α on M_k such that $d\alpha = \varphi\omega - c\Omega$. Similarly, $\psi\omega + c\Omega$ is compactly supported in U_{k+1}, and its integral is

$$\begin{aligned}
\int_{U_{k+1}} (\psi\omega + c\Omega) &= \int_{M_{k+1}} (1 - \varphi)\omega + c\int_{M_{k+1}} \Omega \\
&= \int_{M_{k+1}} \omega - \int_{M_{k+1}} \varphi\omega + c \\
&= 0.
\end{aligned}$$

Thus by Lemma 15.19, there exists another smooth $(n-1)$-form β, compactly supported in U_{k+1}, such that $d\beta = \psi\omega + c\Omega$. Both α and β can be extended by zero to smooth compactly supported forms on M_{k+1}. We compute

$$d(\alpha + \beta) = (\varphi\omega - c\Omega) + (\psi\omega + c\Omega) = (\varphi + \psi)\omega = \omega,$$

which completes the inductive step. $\square$

We now have enough information to compute the de Rham cohomology of the circle and the punctured plane completely.

Corollary 15.16 (Cohomology of the Circle). *The de Rham cohomology groups of the circle are as follows: $H^0_{dR}(\mathbb{S}^1)$ and $H^1_{dR}(\mathbb{S}^1)$ are both* 1*-dimensional, spanned by the cohomology classes of the function* 1 *and the restriction to $\mathbb{S}^1$ of the form ω defined by* (15.1)*, respectively.*

Proof. Because the restriction of ω to $\mathbb{S}^1$ never vanishes, it is an orientation form. □

Corollary 15.17 (Cohomology of the Punctured Plane). *Let $M = \mathbb{R}^2 \smallsetminus \{0\}$. Then $H^0_{dR}(M)$ and $H^1_{dR}(M)$ are both* 1*-dimensional, spanned by the cohomology classes of the function* 1 *and the form ω defined by* (15.1)*, respectively.*

Proof. Because inclusion $\iota\colon \mathbb{S}^1 \hookrightarrow M$ is a homotopy equivalence, $\iota^*\colon H^p_{dR}(M) \to H^p_{dR}(\mathbb{S}^1)$ is an isomorphism for each p. The result then follows from Corollary 15.16 and the fact that $\iota^*[\omega]$ spans $H^1_{dR}(\mathbb{S}^1)$. □

Next we consider the nonorientable case.

Theorem 15.18 (Top Cohomology, Nonorientable Case). *Let M be a compact, connected, nonorientable, smooth n-manifold. Then $H^n_{dR}(M) = 0$.*

Proof. We have to show that every smooth n-form on M is exact. Let $\widehat{\pi}\colon \widehat{M} \to M$ be the orientation covering of M (Theorem 13.9). Let $\alpha\colon \widehat{M} \to \widehat{M}$ be the unique nontrivial covering transformation of $\widehat{M}$ (see Figure 13.3). Now, α cannot be orientation-preserving—if it were, the entire covering group $\{\mathrm{Id}_{\widehat{M}}, \alpha\}$ would be orientation-preserving, and then M would be orientable by the result of Problem 13-5. By connectedness of $\widehat{M}$ and the fact that α is a diffeomorphism, it follows that α is orientation-reversing.

Suppose ω is any smooth n-form on M, and let $\Omega = \widehat{\pi}^*\omega \in \mathcal{A}^n(\widehat{M})$. Then $\widehat{\pi} \circ \alpha = \widehat{\pi}$ implies

$$\alpha^*\Omega = \alpha^*\widehat{\pi}^*\omega = (\widehat{\pi} \circ \alpha)^*\omega = \widehat{\pi}^*\omega = \Omega.$$

Because α is orientation-reversing, therefore, we conclude from Proposition 14.2 that

$$\int_{\widehat{M}} \Omega = -\int_{\widehat{M}} \alpha^*\Omega = -\int_{\widehat{M}} \Omega.$$

This implies that $\int_{\widehat{M}} \Omega = 0$, so by Theorem 15.15, there exists $\eta \in \mathcal{A}^{n-1}(\widehat{M})$ such that $d\eta = \Omega$. Let $\widetilde{\eta} = \frac{1}{2}(\eta + \alpha^*\eta)$. Using the fact that $\alpha \circ \alpha = \mathrm{Id}_{\widehat{M}}$, we compute

$$\alpha^*\widetilde{\eta} = \tfrac{1}{2}(\alpha^*\eta + (\alpha \circ \alpha)^*\eta) = \widetilde{\eta}$$

and

$$\begin{aligned} d\widetilde{\eta} &= \tfrac{1}{2}(d\eta + d\alpha^*\eta) \\ &= \tfrac{1}{2}(d\eta + \alpha^* d\eta) \\ &= \tfrac{1}{2}(\Omega + \alpha^*\Omega) \\ &= \Omega. \end{aligned}$$

Let $U \subset M$ be any connected, evenly covered open set. There are exactly two smooth local sections $\sigma_1, \sigma_2 \colon U \to \widehat{M}$ over U, which are related by $\sigma_2 = \alpha \circ \sigma_1$. Observe that

$$\sigma_2^* \widetilde{\eta} = (\alpha \circ \sigma_1)^* \widetilde{\eta} = \sigma_1^* \alpha^* \widetilde{\eta} = \sigma_1^* \widetilde{\eta}.$$

Therefore, we can define a smooth global $(n-1)$-form γ on M by setting $\gamma = \sigma^* \widetilde{\eta}$ for any smooth local section σ. To determine its exterior derivative, choose a smooth local section σ in a neighborhood of any point, and compute

$$d\gamma = d\sigma^* \widetilde{\eta} = \sigma^* d\widetilde{\eta} = \sigma^* \Omega = \sigma^* \widehat{\pi}^* \omega = (\widehat{\pi} \circ \sigma)^* \omega = \omega,$$

because $\widehat{\pi} \circ \sigma = \mathrm{Id}_U$. □

Finally, here is the technical lemma that is needed to complete the proof of Theorem 15.15. It can be thought of as a refinement of the Poincaré lemma for compactly supported n-forms.

Lemma 15.19. *Let $n \geq 1$, and suppose ω is a compactly supported smooth n-form on $\mathbb{R}^n$ such that $\int_{\mathbb{R}^n} \omega = 0$. Then there exists a compactly supported smooth $(n-1)$-form η on $\mathbb{R}^n$ such that $d\eta = \omega$.*

Remark. Of course, we know that ω is exact by the Poincaré lemma, so the novelty here is the claim that it is the exterior derivative of a *compactly supported* form.

Proof. When $n = 1$, we can write $\omega = f\,dx$ for some smooth, compactly supported function f. Define $F \colon \mathbb{R} \to \mathbb{R}$ by

$$F(x) = \int_{-\infty}^{x} f(t)\,dt.$$

Then clearly $dF = F'\,dx = f\,dx = \omega$. Choose $R > 0$ such that $\operatorname{supp} f \subset [-R, R]$. When $x < -R$, $F(x) = 0$ by our choice of R. When $x > R$, the fact that $\int_{\mathbb{R}} \omega = 0$ translates to

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{\infty} f(t)\,dt = 0,$$

so in fact $\operatorname{supp} F \subset [-R, R]$. This completes the proof for the case $n = 1$.

Now assume $n \geq 2$, and let $B, B' \subset \mathbb{R}^n$ be open balls centered at the origin such that $\operatorname{supp} \omega \subset B \subset \overline{B} \subset B'$. By the Poincaré lemma, there exists a smooth $(n-1)$-form η_0 on $\mathbb{R}^n$ such that $d\eta_0 = \omega$. Stokes's theorem implies that

$$0 = \int_{\mathbb{R}^n} \omega = \int_{\overline{B'}} \omega = \int_{\overline{B'}} d\eta_0 = \int_{\partial B'} \eta_0. \tag{15.6}$$

Because of (15.6), it follows from Theorem 15.15 that the restriction of η_0 to $\partial B'$ is exact. Since the inclusion $\iota\colon \partial B' \hookrightarrow \mathbb{R}^n \smallsetminus \overline{B}$ is a homotopy equivalence, $\iota^*\colon H^{n-1}(\mathbb{R}^n \smallsetminus \overline{B}) \to H^{n-1}(\partial B')$ is an isomorphism. It follows that η_0 is exact on $\mathbb{R}^n \smallsetminus \overline{B}$. Thus there is a smooth $(n-2)$-form γ on $\mathbb{R}^n \smallsetminus \overline{B}$ such that $d\gamma = \eta_0$ there. If we let ψ be a bump function that is supported in $\mathbb{R}^n \smallsetminus \overline{B}$ and equal to 1 on $\mathbb{R}^n \smallsetminus B'$, then $\eta = \eta_0 - d(\psi\gamma)$ is smooth on all of $\mathbb{R}^n$ and satisfies $d\eta = d\eta_0 = \omega$. Because $d(\psi\gamma) = d\gamma = \eta_0$ on $\mathbb{R}^n \smallsetminus B'$, η is compactly supported. □

For some purposes, it is useful to define a generalization of the de Rham cohomology groups using only compactly supported forms. Let $\mathcal{A}_c^p(M)$ denote the space of compactly-supported smooth p-forms on M. The pth *compactly supported de Rham cohomology group of* M is the quotient space

$$H_c^p(M) = \frac{\operatorname{Ker}[d\colon \mathcal{A}_c^p(M) \to \mathcal{A}_c^{p+1}(M)]}{\operatorname{Im}[d\colon \mathcal{A}_c^{p-1}(M) \to \mathcal{A}_c^p(M)]}.$$

Of course, when M is compact, this just reduces to ordinary de Rham cohomology. But for noncompact manifolds, the two groups can be different, as the next exercise shows.

◇ **Exercise 15.1.** Using Lemma 15.19, show that $H_c^n(\mathbb{R}^n)$ is 1-dimensional.

Compactly supported cohomology has a number of important applications in algebraic topology. One of the most important is the Poincaré duality theorem, which will be outlined in Problem 16-6.

The Mayer–Vietoris Theorem

In this section, we prove a very general theorem that can be used to compute the de Rham cohomology groups of many spaces, by expressing them as unions of open submanifolds with simpler cohomology.

For this purpose, we need to introduce some simple algebraic concepts. More details about the ideas introduced here can be found in [Lee00, Chapter 13] or in any textbook on algebraic topology.

Let $\mathcal{R}$ be a commutative ring, and suppose we are given a sequence of $\mathcal{R}$-modules and $\mathcal{R}$-linear maps:

$$\cdots \to A^{p-1} \xrightarrow{d} A^p \xrightarrow{d} A^{p+1} \to \cdots. \tag{15.7}$$

(In all of our applications, the ring will be either $\mathbb{Z}$, in which case we are looking at abelian groups and homomorphisms, or $\mathbb{R}$, in which case we have vector spaces and linear maps. The terminology of modules is just a convenient way to combine the two cases.) Such a sequence is said to be a *complex* if the composition of any two successive applications of d is the zero map:

$$d \circ d = 0 \colon A^p \to A^{p+2} \quad \text{for each } p.$$

It is called an *exact sequence* if the image of each d is equal to the kernel of the next:

$$\operatorname{Im}[d \colon A^{p-1} \to A^p] = \operatorname{Ker}[d \colon A^p \to A^{p+1}].$$

Clearly every exact sequence is a complex, but the converse need not be true.

Let us denote the sequence (15.7) by A^*. If it is a complex, then the image of each map d is contained in the kernel of the next, so we define the pth *cohomology group* of A^* to be the quotient module

$$H^p(A^*) = \frac{\operatorname{Ker}[d \colon A^p \to A^{p+1}]}{\operatorname{Im}[d \colon A^{p-1} \to A^p]}.$$

It can be thought of as a quantitative measure of the failure of exactness at A^p. The obvious example is the *de Rham complex* of a smooth n-manifold M:

$$0 \to \mathcal{A}^0(M) \xrightarrow{d} \cdots \xrightarrow{d} \mathcal{A}^p(M) \xrightarrow{d} \mathcal{A}^{p+1}(M) \xrightarrow{d} \cdots \xrightarrow{d} \mathcal{A}^n(M) \to 0,$$

whose cohomology groups are the de Rham groups of M. (In algebraic topology, a complex as we have defined it is usually called a *cochain complex*, while a *chain complex* is defined similarly except that the maps go in the direction of decreasing indices:

$$\cdots \to A_{p+1} \xrightarrow{\partial} A_p \xrightarrow{\partial} A_{p-1} \to \cdots .$$

In that case, the term *homology* is used in place of cohomology.)

If A^* and B^* are complexes, a *cochain map* from A^* to B^*, denoted by $F \colon A^* \to B^*$, is a collection of linear maps $F \colon A^p \to B^p$ (it is easiest to use the same symbol for all of the maps) such that the following diagram commutes for each p:

$$\begin{tikzcd}
\cdots \arrow[r] & A^p \arrow[r, "d"] \arrow[d, "F"'] & A^{p+1} \arrow[r] \arrow[d, "F"] & \cdots \\
\cdots \arrow[r] & B^p \arrow[r, "d"'] & B^{p+1} \arrow[r] & \cdots .
\end{tikzcd}$$

The fact that $F \circ d = d \circ F$ means that any cochain map induces a linear map on cohomology $F^* \colon H^p(A^*) \to H^p(B^*)$ for each p, just as in the case of de Rham cohomology. (A map between chain complexes satisfying the analogous relations is called a *chain map*; the same argument shows that a chain map induces a linear map on homology.)

A *short exact sequence of complexes* consists of three complexes A^*, B^*, C^*, together with cochain maps

$$0 \to A^* \xrightarrow{F} B^* \xrightarrow{G} C^* \to 0$$

such that each sequence

$$0 \to A^p \xrightarrow{F} B^p \xrightarrow{G} C^p \to 0$$

is exact. This means F is injective, G is surjective, and $\operatorname{Im} F = \operatorname{Ker} G$.

Lemma 15.20 (The Zigzag Lemma). *Given a short exact sequence of complexes as above, for each p there is a linear map*

$$\delta\colon H^p(C^*) \to H^{p+1}(A^*),$$

called the connecting homomorphism, such that the following sequence is exact:

$$\cdots \xrightarrow{\delta} H^p(A^*) \xrightarrow{F^*} H^p(B^*) \xrightarrow{G^*} H^p(C^*) \xrightarrow{\delta} H^{p+1}(A^*) \xrightarrow{F^*} \cdots. \tag{15.8}$$

Proof. We will sketch only the main idea; you can either carry out the details yourself or look them up.

The hypothesis means that the following diagram commutes and has exact horizontal rows:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & A^p & \xrightarrow{F} & B^p & \xrightarrow{G} & C^p & \longrightarrow & 0 \\
 & & \downarrow d & & \downarrow d & & \downarrow d & & \\
0 & \longrightarrow & A^{p+1} & \xrightarrow{F} & B^{p+1} & \xrightarrow{G} & C^{p+1} & \longrightarrow & 0 \\
 & & \downarrow d & & \downarrow d & & \downarrow d & & \\
0 & \longrightarrow & A^{p+2} & \xrightarrow{F} & B^{p+2} & \xrightarrow{G} & C^{p+2} & \longrightarrow & 0.
\end{array}$$

Suppose $c^p \in C^p$ represents a cohomology class; this means that $dc^p = 0$. Since $G\colon B^p \to C^p$ is surjective, there is some element $b^p \in B^p$ such that $Gb^p = c^p$. Because the diagram commutes, $Gdb^p = dGb^p = dc^p = 0$, and therefore $db^p \in \operatorname{Ker} G = \operatorname{Im} F$. Thus there exists $a^{p+1} \in A^{p+1}$ satisfying $Fa^{p+1} = db^p$. By commutativity of the diagram again, $Fda^{p+1} = dFa^{p+1} = ddb^p = 0$. Since F is injective, this implies $da^{p+1} = 0$, so a^{p+1} represents a cohomology class in $H^{p+1}(A^*)$. The connecting homomorphism δ is defined by setting $\delta[c^p] = [a^{p+1}]$ for any such $a^{p+1} \in A^{p+1}$, that is, provided there exists $b^p \in B^p$ such that

$$Gb^p = c^p,$$
$$Fa^{p+1} = db^p.$$

A number of facts have to be verified: that the cohomology class $[a^{p+1}]$ is well-defined, independently of the choices made along the way; that the resulting map δ is linear; and that the resulting sequence (15.8) is exact. Each of these verifications is a routine "diagram chase" like the one we used to define δ; the details are left as an exercise. $\square$

◇ **Exercise 15.2.** Complete (or look up) the proof of the zigzag lemma.

The situation in which we will apply this lemma is the following. Suppose M is a smooth manifold, and U, V are open subsets of M such that $M = U \cup V$. We have the following diagram of inclusions:

$$\begin{tikzcd} & U \arrow[dr, "k"] & \\ U \cap V \arrow[ur, "i"] \arrow[dr, "j"'] & & M, \\ & V \arrow[ur, "l"'] & \end{tikzcd} \tag{15.9}$$

which induce pullback maps on differential forms:

$$\begin{tikzcd} & \mathcal{A}^p(U) \arrow[dr, "i^*"] & \\ \mathcal{A}^p(M) \arrow[ur, "k^*"] \arrow[dr, "l^*"'] & & \mathcal{A}^p(U \cap V), \\ & \mathcal{A}^p(V) \arrow[ur, "j^*"'] & \end{tikzcd}$$

as well as corresponding induced cohomology maps. Note that these pullback maps are really just restrictions: For example, $k^*\omega = \omega|_U$. We will consider the following sequence:

$$0 \to \mathcal{A}^p(M) \xrightarrow{k^* \oplus l^*} \mathcal{A}^p(U) \oplus \mathcal{A}^p(V) \xrightarrow{i^* - j^*} \mathcal{A}^p(U \cap V) \to 0, \tag{15.10}$$

where

$$\begin{aligned} (k^* \oplus l^*)\omega &= (k^*\omega, l^*\omega), \\ (i^* - j^*)(\omega, \eta) &= i^*\omega - j^*\eta. \end{aligned} \tag{15.11}$$

Because pullbacks commute with d, these maps descend to linear maps on the corresponding de Rham cohomology groups.

Theorem 15.21 (Mayer–Vietoris). *Let M be a smooth manifold, and let U, V be open subsets of M whose union is M. For each p, there is a linear map $\delta\colon H^p_{dR}(U \cap V) \to H^{p+1}_{dR}(M)$ such that the following sequence is* *exact:*

$$\cdots \xrightarrow{\delta} H^p_{dR}(M) \xrightarrow{k^*\oplus l^*} H^p_{dR}(U) \oplus H^p_{dR}(V) \xrightarrow{i^*-j^*} H^p_{dR}(U\cap V) \\ \xrightarrow{\delta} H^{p+1}_{dR}(M) \xrightarrow{k^*\oplus l^*} \cdots \tag{15.12}$$

Remark. The sequence (15.12) is called the *Mayer–Vietoris sequence* for the open cover $\{U,V\}$.

Proof. The heart of the proof will be to show that the sequence (15.10) is exact for each p. Because pullback maps commute with the exterior derivative, (15.10) therefore defines a short exact sequence of chain maps, and the Mayer–Vietoris theorem follows immediately from the zigzag lemma.

We begin by proving exactness at $\mathcal{A}^p(M)$, which just means showing that $k^*\oplus l^*$ is injective. Suppose that $\sigma\in\mathcal{A}^p(M)$ satisfies $(k^*\oplus l^*)\sigma = (\sigma|_U,\sigma|_V)=(0,0)$. This means that the restrictions of σ to U and V are both zero. Since $\{U,V\}$ is an open cover of M, this implies that σ is zero.

To prove exactness at $\mathcal{A}^p(U)\oplus\mathcal{A}^p(V)$, first observe that

$$(i^*-j^*)\circ(k^*\oplus l^*)(\sigma) = (i^*-j^*)(\sigma|_U,\sigma|_V) = \sigma|_{U\cap V} - \sigma|_{U\cap V} = 0,$$

which shows that $\operatorname{Im}(k^*\oplus l^*)\subset\operatorname{Ker}(i^*-j^*)$. Conversely, suppose $(\eta,\eta')\in\mathcal{A}^p(U)\oplus\mathcal{A}^p(V)$ and $(i^*-j^*)(\eta,\eta')=0$. This means $\eta|_{U\cap V}=\eta'|_{U\cap V}$, so there is a global smooth p-form σ on M defined by

$$\sigma = \begin{cases}\eta & \text{on } U,\\ \eta' & \text{on } V.\end{cases}$$

Clearly $(\eta,\eta')=(k^*\oplus l^*)\sigma$, so $\operatorname{Ker}(i^*-j^*)\subset\operatorname{Im}(k^*\oplus l^*)$.

Exactness at $\mathcal{A}^p(U\cap V)$ means that i^*-j^* is surjective. This the only nontrivial part of the proof, and the only part that really uses any properties of smooth manifolds and differential forms.

Let $\omega\in\mathcal{A}^p(U\cap V)$ be arbitrary. We need to show that there exist $\eta\in\mathcal{A}^p(U)$ and $\eta'\in\mathcal{A}^p(V)$ such that

$$\omega = (i^*-j^*)(\eta,\eta') = i^*\eta - j^*\eta' = \eta|_{U\cap V} - \eta'|_{U\cap V}.$$

(See Figure 15.3.) Let $\{\varphi,\psi\}$ be a smooth partition of unity subordinate to the open cover $\{U,V\}$, and define $\eta\in\mathcal{A}^p(U)$ by

$$\eta = \begin{cases}\psi\omega, & \text{on } U\cap V,\\ 0 & \text{on } U\smallsetminus\operatorname{supp}\psi.\end{cases} \tag{15.13}$$

On the set $(U\cap V)\smallsetminus\operatorname{supp}\psi$ where these definitions overlap, they both give zero, so this defines η as a smooth p-form on U. Similarly, define $\eta'\in\mathcal{A}^p(V)$ by

$$\eta' = \begin{cases}-\varphi\omega, & \text{on } U\cap V,\\ 0 & \text{on } V\smallsetminus\operatorname{supp}\varphi.\end{cases} \tag{15.14}$$

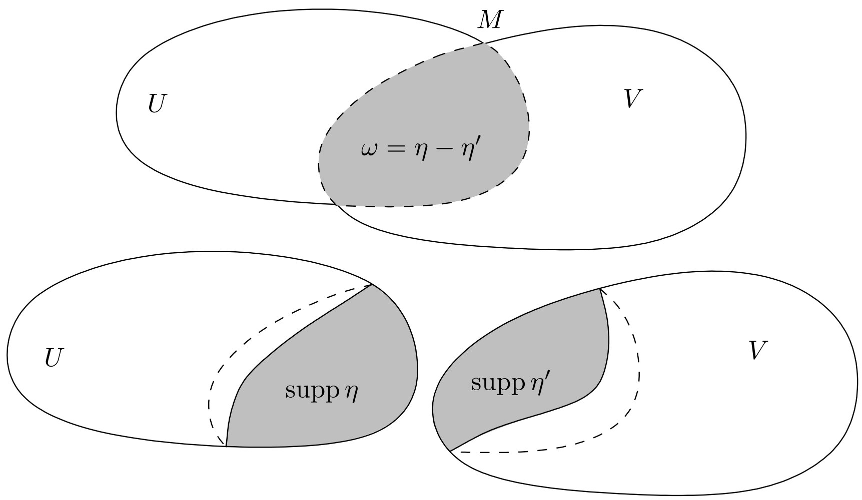


Figure 15.3. Surjectivity of $i^* - j^*$.

Then we have

$$\eta|_{U\cap V} - \eta'|_{U\cap V} = \psi\omega - (-\varphi\omega) = (\psi + \varphi)\omega = \omega,$$

which was to be proved. □

For later use, we record the following corollary to the proof, which explicitly characterizes the connecting homomorphism δ.

Corollary 15.22. *The connecting homomorphism $\delta\colon H^p_{dR}(U \cap V) \to H^{p+1}_{dR}(M)$ is defined as follows. Given $\omega \in \mathcal{Z}^p(U \cap V)$, there are smooth p-forms $\eta \in \mathcal{A}^p(U)$ and $\eta' \in \mathcal{A}^p(V)$ such that $\omega = \eta|_{U\cap V} - \eta'|_{U\cap V}$, and then $\delta[\omega] = [d\eta]$, where $d\eta$ is extended by zero to all of M.*

Proof. A characterization of the connecting homomorphism was given in the proof of the zigzag lemma. Specializing this characterization to the situation of the short exact sequence (15.10), we find that $\delta[\omega] = [\sigma]$ provided there exists $(\eta, \eta') \in \mathcal{A}^p(U) \oplus \mathcal{A}^p(V)$ such that

$$\begin{aligned} &i^*\eta - j^*\eta' = \omega, \\ &(k^*\sigma, l^*\sigma) = (d\eta, d\eta'). \end{aligned} \tag{15.15}$$

Arguing just as in the proof of the Mayer–Vietoris theorem, if $\{\varphi, \psi\}$ is a smooth partition of unity subordinate to $\{U, V\}$, then formulas (15.13) and (15.14) define smooth forms $\eta \in \mathcal{A}^p(U)$ and $\eta' \in \mathcal{A}^p(V)$ satisfying the first equation of (15.15). Let σ be the smooth form on M obtained by extending $d\eta$ to be zero outside of $U \cap V$. Because ω is closed,

$$\sigma|_{U\cap V} = d\eta|_{U\cap V} = d(\omega + \eta')|_{U\cap V} = d\eta'|_{U\cap V},$$

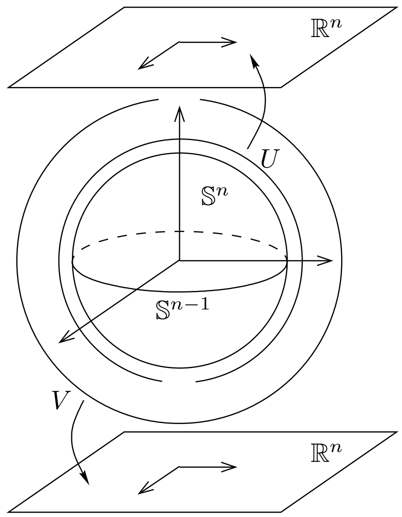


Figure 15.4. Computing the de Rham cohomology of $\mathbb{S}^n$.

and the second equation of (15.15) follows easily. □

Our first application of the Mayer–Vietoris theorem will be to compute all of the de Rham cohomology groups of spheres. In the next chapter, we will use the theorem again as an essential ingredient in the proof of the de Rham theorem.

Theorem 15.23. *For $n \geq 1$, the de Rham cohomology groups of $\mathbb{S}^n$ are*

$$H^p_{dR}(\mathbb{S}^n) \cong \begin{cases} \mathbb{R} & \text{if } p = 0 \text{ or } p = n, \\ 0 & \text{if } 0 < p < n. \end{cases}$$

Proof. We already know $H^0_{dR}(\mathbb{S}^n)$ and $H^n_{dR}(\mathbb{S}^n)$ from the preceding section. For good measure, we give here another proof for $H^n_{dR}(\mathbb{S}^n)$.

Let N and S be the north and south poles in $\mathbb{S}^n$, respectively, and let $U = \mathbb{S}^n \smallsetminus \{S\}$, $V = \mathbb{S}^n \smallsetminus \{N\}$. By stereographic projection, both U and V are diffeomorphic to $\mathbb{R}^n$ (Figure 15.4), and thus $U \cap V$ is diffeomorphic to $\mathbb{R}^n \smallsetminus \{0\}$.

Part of the Mayer–Vietoris sequence for $\{U, V\}$ reads

$$H^{p-1}_{dR}(U) \oplus H^{p-1}_{dR}(V) \to H^{p-1}_{dR}(U \cap V) \to H^p_{dR}(\mathbb{S}^n) \to H^p_{dR}(U) \oplus H^p_{dR}(V).$$

Because U and V are diffeomorphic to $\mathbb{R}^n$, the groups on both ends are trivial when $p > 1$, which implies that $H^p_{dR}(\mathbb{S}^n) \cong H^{p-1}_{dR}(U \cap V)$. Moreover, $U \cap V$ is diffeomorphic to $\mathbb{R}^n \smallsetminus \{0\}$ and therefore homotopy equivalent to $\mathbb{S}^{n-1}$, so in the end we conclude that

$$H^p_{dR}(\mathbb{S}^n) \cong H^{p-1}_{dR}(\mathbb{S}^{n-1}) \quad \text{for } p > 1.$$

We will prove the theorem by induction on n. The case $n = 1$ is taken care of by Corollary 15.16, so suppose $n \geq 2$ and assume the theorem is true for $\mathbb{S}^{n-1}$. Clearly $H^0_{dR}(\mathbb{S}^n) \cong \mathbb{R}$ by Proposition 15.9, and $H^1_{dR}(\mathbb{S}^n) = 0$ because $\mathbb{S}^n$ is simply connected. For $p > 1$, the inductive hypothesis then gives

$$H^p_{dR}(\mathbb{S}^n) \cong H^{p-1}_{dR}(\mathbb{S}^{n-1}) \cong \begin{cases} 0 & \text{if } p < n, \\ \mathbb{R} & \text{if } p = n. \end{cases}$$

This completes the proof. □

Problems

15-1. Compute the de Rham cohomology groups of $\mathbb{R}^n$ minus two points.

15-2. Suppose $U \subset \mathbb{R}^n$ is open and star-shaped with respect to 0. If $\omega = \sum' \omega_I \, dx^I$ is a closed p-form on U, show either directly or by tracing through the proof of the Poincaré lemma that the $(p-1)$-form η given explicitly by the formula

$$\eta = \sum_I{}' \sum_{q=1}^{p} (-1)^{q-1} \left(\int_0^1 t^{p-1} \omega_I(tx) \, dt \right) x^{i_q} \, dx^{i_1} \wedge \cdots \wedge \widehat{dx^{i_q}} \wedge \cdots \wedge dx^{i_p}$$

satisfies $d\eta = \omega$. When ω is a smooth 1-form, show that η is equal to the potential function f defined in Proposition 6.30.

15-3. Let M be a smooth manifold, and let $\omega \in \mathcal{A}^p(M)$, $\eta \in \mathcal{A}^q(M)$ be closed forms. Show that the de Rham cohomology class of $\omega \wedge \eta$ depends only on the de Rham cohomology classes of ω and η, and thus there is a well-defined bilinear map $\smile\colon H^p_{dR}(M) \times H^q_{dR}(M) \to H^{p+q}_{dR}(M)$ given by

$$[\omega] \smile [\eta] = [\omega \wedge \eta].$$

(This bilinear map is called the *cup product*.)

15-4. Let M be a compact, connected, orientable, smooth n-manifold, and let p be any point of M. Let V be a neighborhood of p diffeomorphic to $\mathbb{R}^n$ and let $U = M \smallsetminus \{p\}$.

(a) Show that the connecting homomorphism $\delta\colon H^{n-1}_{dR}(U \cap V) \to H^n_{dR}(M)$ is an isomorphism. [Hint: Show that $\delta[\omega] \neq 0$, where ω is the $(n-1)$-form on $U \cap V \approx \mathbb{R}^n \smallsetminus \{0\}$ defined in coordinates by (14.17) (Problem 14-4).]

(b) Use the Mayer–Vietoris sequence of $\{U, V\}$ to show that $H^n_{dR}(M \smallsetminus \{p\}) = 0$.

15-5. Let M be a compact, connected, smooth manifold of dimension $n \geq 3$. For any $p \in M$ and $0 \leq k < n$, show that the map $H^k_{dR}(M) \to H^k_{dR}(M \smallsetminus \{p\})$ induced by inclusion $M \smallsetminus \{p\} \hookrightarrow M$ is an isomorphism. [Hint: Use a Mayer–Vietoris sequence together with the result of Problem 15-4. The cases $k = 1$ and $k = n - 1$ will require special handling.]

15-6. Suppose M and N are smooth, oriented, compact n-manifolds, and $F\colon M \to N$ is a smooth map. If $\int_M F^*\Omega \neq 0$ for some $\Omega \in \mathcal{A}^n(N)$, show that F is surjective.

15-7. Let M_1, M_2 be smooth, connected, orientable manifolds of dimension $n \geq 2$, and let $M_1 \# M_2$ denote their smooth connected sum (see Problem 7-10). Show that $H^k_{dR}(M_1 \# M_2) \cong H^k_{dR}(M_1) \oplus H^k_{dR}(M_2)$ for $0 < k < n$.

15-8. Suppose (M, ω) is a $2n$-dimensional compact symplectic manifold.

(a) Show that $\omega^n = \omega \wedge \cdots \wedge \omega$ (the n-fold wedge product of ω with itself) is not exact. [Hint: See Problem 13-8.]

(b) Show that $H^{2k}_{dR}(M) \neq 0$ for $k = 1, \ldots, n$.

(c) Show that the only sphere that admits a symplectic structure is $\mathbb{S}^2$.

15-9. Let (M, g) be a compact, oriented Riemannian n-manifold. For $\omega, \eta \in \mathcal{A}^k(M)$, $0 \leq k \leq n$, define $(\omega, \eta) \in \mathbb{R}$ by

$$(\omega, \eta) = \int_M \langle \omega, \eta \rangle \, dV_g, \tag{15.16}$$

where $\langle \omega, \eta \rangle$ is the pointwise inner product on k-forms defined in Problem 14-13. Let $\|\omega\| = (\omega, \omega)^{1/2}$. The *Laplace-Beltrami operator* is the map $\Delta\colon \mathcal{A}^k(M) \to \mathcal{A}^k(M)$ defined by

$$\Delta\omega = dd^*\omega + d^*d\omega,$$

where d^* is the operator defined in Problem 14-17. A smooth form $\omega \in \mathcal{A}^k(M)$ is said to be *harmonic* if $\Delta\omega = 0$.

(a) Show that $(\cdot, \cdot)$ is an inner product and $\|\cdot\|$ is a norm on $\mathcal{A}^k(M)$.

(b) Show that the following are equivalent for any $\omega \in \mathcal{A}^k(M)$.

(i) ω is harmonic.

(ii) $d\omega = 0$ and $d^*\omega = 0$.

(iii) $d\omega = 0$ and ω is the unique smooth k-form in its cohomology class with minimum norm $\|\omega\|$.

[Hint: for (iii), consider $f(t) = \|\omega + d(td^*\omega)\|^2$.]

15-10. Let (M, g) be an oriented Riemannian manifold, and let $\Delta = dd^* + d^*d$ be the Laplace-Beltrami operator on k-forms as in Problem 15-9. When $k = 0$, show that Δ agrees with the Laplacian $\Delta u = -\operatorname{div}(\operatorname{grad} u)$ defined on real-valued functions in Problem 14-9.

15-11. Suppose M is a compact, connected, orientable, smooth manifold with finite fundamental group. Show that $H^1_{dR}(M) = 0$. [Hint: Apply the result of Problem 14-6 to the universal covering of M.]

15-12. Suppose M is a smooth, connected, orientable, compact n-manifold.

(a) Show that there is a one-to-one correspondence between orientations of M and orientations of the vector space $H^n_{dR}(M)$, under which the cohomology class of a smooth orientation form is an oriented basis for $H^n_{dR}(M)$.

(b) If M is given a specific orientation, show that a diffeomorphism $F\colon M \to M$ is orientation preserving if and only if $F^*\colon H^n_{dR}(M) \to H^n_{dR}(M)$ is orientation preserving.

15-13. Show that the compactly supported de Rham cohomology groups $H^p_c(\mathbb{R}^n)$ are all zero for $0 \le p < n$.

16
The de Rham Theorem

The topological invariance of the de Rham groups suggests that there should be some purely topological way of computing them. There is indeed, and the connection between the de Rham groups and topology was first proved by Georges de Rham himself in the 1930s. The theorem that bears his name is a major landmark in the development of smooth manifold theory. The purpose of this chapter is to give a proof of this theorem.

In the category of topological spaces, there are a number of ways of defining cohomology groups that measure the existence of "holes" in different dimensions, but that have nothing to do with differential forms or smooth structures. In the beginning of the chapter, we describe the most straightforward ones, called singular homology and cohomology. Because a complete treatment of singular theory would be far beyond the scope of this book, we can only summarize the basic ideas here. For more details, you can consult a standard textbook on algebraic topology, such as [Bre93, Mun84, Spa89]. (See also [Lee00, Chapter 13] for a more concise treatment.) After introducing the basic definitions, we prove that singular homology can be computed by restricting attention only to smooth simplices.

At the end of the chapter, we turn our attention to the de Rham theorem, which shows that integration of differential forms over smooth simplices induces an isomorphism between the de Rham groups and the singular cohomology groups.

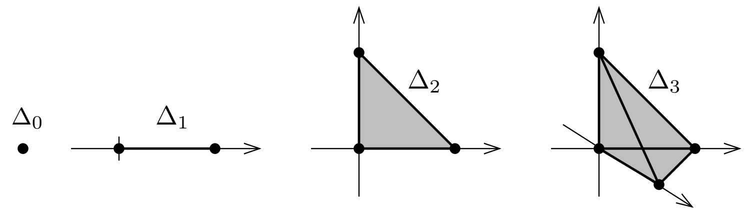


Figure 16.1. Standard p-simplices for $p = 0, 1, 2, 3$.

Singular Homology

We begin the chapter with a very brief summary of singular homology theory. Suppose $v_0, \dots, v_p$ are any $p+1$ points in some Euclidean space $\mathbb{R}^n$. They are said to be in *general position* if they are not contained in any $(p-1)$-dimensional affine subspace. A *geometric p-simplex* is a subset of $\mathbb{R}^n$ of the form

$$\left\{ \sum_{i=0}^{p} t_i v_i : 0 \le t_i \le 1 \text{ and } \sum_{i=0}^{p} t_i = 1 \right\},$$

for some $(p+1)$ points $\{v_0, \dots, v_p\}$ in general position. The integer p (one less than the number of vertices) is called the *dimension* of the simplex. The points v_i are called its *vertices*, and the geometric simplex with vertices $v_0, \dots, v_p$ is denoted by $\langle v_0, \dots, v_p \rangle$. It is a compact convex set, in fact the smallest convex set containing $\{v_0, \dots, v_p\}$. The simplices whose vertices are subsets of $\{v_0, \dots, v_p\}$ are called the *faces* of the simplex. The $(p-1)$-dimensional faces are called its *boundary faces*. There are precisely $p+1$ boundary faces, obtained by omitting each of the vertices in turn; the ith boundary face $\partial_i \langle v_0, \dots, v_p \rangle = \langle v_0, \dots, \widehat{v_i}, \dots, v_p \rangle$ is sometimes called the *face opposite* v_i. (As usual, the hat indicates that v_i is omitted.)

◇ **Exercise 16.1.** Show that a geometric p-simplex is a p-dimensional smooth manifold with corners smoothly embedded in $\mathbb{R}^n$.

The *standard p-simplex* is the simplex $\Delta_p = \langle e_0, e_1, \dots, e_p \rangle \subset \mathbb{R}^p$, where $e_0 = 0$ and e_i is the ith standard basis vector. For example, $\Delta_0 = \{0\}$, $\Delta_1 = [0,1]$, Δ_2 is the triangle with vertices $(0,0)$, $(1,0)$, and $(0,1)$ together with its interior, and Δ_3 is a solid tetrahedron (Figure 16.1).

Let M be a topological space. A continuous map $\sigma\colon \Delta_p \to M$ is called a *singular p-simplex* in M. The *singular chain group* of M in dimension p, denoted by $C_p(M)$, is the free abelian group generated by all singular p-simplices in M. An element of this group, called a *singular p-chain*, is just a finite formal linear combination of singular p-simplices with integer coefficients.

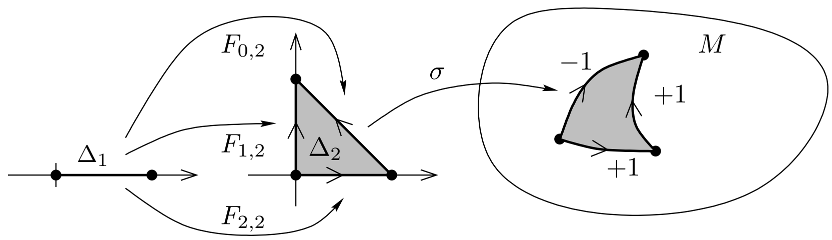


Figure 16.2. The singular boundary operator.

One special case that arises frequently is that in which the space M is a convex subset of some Euclidean space $\mathbb{R}^m$. In that case, for any ordered $(p+1)$-tuple of points $(w_0, \dots, w_p)$ in M, not necessarily in general position, there is a unique affine map from $\mathbb{R}^p$ to $\mathbb{R}^m$ that takes e_i to w_i for $i = 0, \dots, p$. (Such a map is easily constructed by first finding a linear map that takes e_i to $w_i - w_0$ for $i = 1, \dots, p$, and then translating by w_0.) The restriction of this affine map to Δ_p is denoted by $\alpha(w_0, \dots, w_p)$, and is called an *affine singular simplex* in M.

For each $i = 0, \dots, p$, we define the ith *face map* in Δ_p to be the affine singular $(p-1)$-simplex $F_{i,p}\colon \Delta_{p-1} \to \Delta_p$ defined by

$$F_{i,p} = \alpha(e_0, \dots, \widehat{e_i}, \dots, e_p).$$

It maps Δ_{p-1} homeomorphically onto the boundary face $\partial_i \Delta_p$ opposite e_i. More explicitly, it is the unique affine map sending $e_0 \mapsto e_0$, $\dots$, $e_{i-1} \mapsto e_{i-1}$, $e_i \mapsto e_{i+1}$, $\dots$, $e_{p-1} \mapsto e_p$.

The *boundary* of a singular p-simplex $\sigma\colon \Delta_p \to M$ is the singular $(p-1)$-chain $\partial\sigma$ defined by

$$\partial\sigma = \sum_{i=0}^{p} (-1)^i \sigma \circ F_{i,p}.$$

For example, if σ is a singular 2-simplex, its boundary is the formal sum of three singular 1-simplices with coefficients ± 1 as indicated in Figure 16.2. This extends uniquely to a group homomorphism $\partial\colon C_p(M) \to C_{p-1}(M)$, called the *singular boundary operator*. The basic fact about the boundary operator is the following lemma.

Lemma 16.1. *If c is any singular chain, then $\partial(\partial c) = 0$.*

Proof. The starting point is the fact that

$$F_{i,p} \circ F_{j,p-1} = F_{j,p} \circ F_{i-1,p-1} \tag{16.1}$$

when $i > j$, which can be verified by following what both compositions do to each of the vertices of Δ_{p-2}. Using this, the proof of the lemma is just a straightforward computation. $\square$

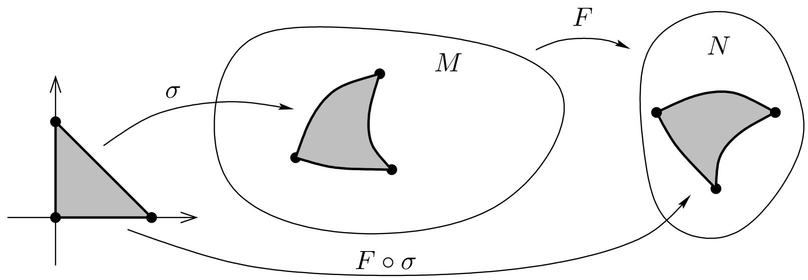


Figure 16.3. The homology homomorphism induced by a continuous map.

A singular p-chain c is called a *cycle* if $\partial c = 0$, and a *boundary* if $c = \partial b$ for some singular $(p+1)$-chain b. Let $Z_p(M)$ denote the set of singular p-cycles in M, and $B_p(M)$ the set of singular p-boundaries. Because ∂ is a homomorphism, $Z_p(M)$ and $B_p(M)$ are subgroups of $C_p(M)$, and because $\partial \circ \partial = 0$, $B_p(M) \subset Z_p(M)$. The pth *singular homology group* of M is the quotient group

$$H_p(M) = \frac{Z_p(M)}{B_p(M)}.$$

To put it another way, the sequence of abelian groups and homomorphisms

$$\cdots \to C_{p+1}(M) \xrightarrow{\partial} C_p(M) \xrightarrow{\partial} C_{p-1}(M) \to \cdots$$

is a complex, called the *singular chain complex*, and $H_p(M)$ is the pth homology group of this complex. The equivalence class in $H_p(M)$ of a singular p-cycle c is called its *homology class*, and is denoted by $[c]$. We say two p-cycles are *homologous* if they differ by a boundary.

Any continuous map $F\colon M \to N$ induces a homomorphism $F_\#\colon C_p(M) \to C_p(N)$ on each singular chain group, defined by $F_\#(\sigma) = F \circ \sigma$ for any singular simplex σ (Figure 16.3), and extended by linearity to chains. An easy computation shows that $F \circ \partial = \partial \circ F$, so F is a chain map, and therefore induces a homomorphism on the singular homology groups, denoted by $F_*\colon H_p(M) \to H_p(N)$. It is immediate that $(G \circ F)_* = G_* \circ F_*$ and $(Id_M)_* = \mathrm{Id}_{H_p(M)}$, so pth singular homology defines a covariant functor from the category of topological spaces and continuous maps to the category of abelian groups and homomorphisms. In particular, homeomorphic spaces have isomorphic singular homology groups.

Proposition 16.2 (Properties of Singular Homology Groups).

(*a*) *For any one-point space* $\{q\}$, $H_0(\{q\})$ *is the infinite cyclic group generated by the homology class of the unique singular* 0*-simplex mapping* Δ_0 *to* q, *and* $H_p(\{q\}) = 0$ *for all* $p \neq 0$.

(*b*) *Let* $\{M_j\}$ *be any collection of topological spaces, and let* $M = \coprod_j M_j$*. The inclusion maps* $\iota_j\colon M_j \hookrightarrow M$ *induce an isomorphism from* $\oplus_j H_p(M_j)$ *to* $H_p(M)$*.*

(*c*) *Homotopy equivalent spaces have isomorphic singular homology groups.*

Sketch of Proof. In a one-point space $\{q\}$, there is exactly one singular p-simplex for each p, namely the constant map. The result of part (a) follows from an analysis of the boundary maps. Part (b) is immediate because the maps ι_j already induce an isomorphism on the chain level: $\oplus_j C_p(M_j) \cong C_p(M)$.

The main step in the proof of homotopy invariance is the construction for any space M of a linear map $h\colon C_p(M) \to C_{p+1}(M \times I)$ satisfying

$$h \circ \partial + \partial \circ h = (i_1)_\# - (i_0)_\#, \tag{16.2}$$

where $i_k\colon M \to M \times I$ is the injection $i_k(x) = (x, k)$. From this it follows just as in the proof of Proposition 15.5 that homotopic maps induce the same homology homomorphism, and then in turn that homotopy equivalent spaces have isomorphic singular homology groups. □

In addition to the properties above, singular homology satisfies the following version of the Mayer–Vietoris theorem. Suppose M is a topological space and $U, V \subset M$ are open subsets whose union is M. The usual diagram (15.9) of inclusions induces homology homomorphisms:

$$\begin{tikzcd} & H_p(U) \arrow[dr, "k_*"] & \\ H_p(U \cap V) \arrow[ur, "i_*"] \arrow[dr, "j_*"'] & & H_p(M). \\ & H_p(V) \arrow[ur, "l_*"'] & \end{tikzcd} \tag{16.3}$$

Theorem 16.3 (Mayer–Vietoris for Singular Homology). *Let* M *be a topological space and let* U, V *be open subsets of* M *whose union is* M*. For each* p *there is a homomorphism* $\partial_*\colon H_p(M) \to H_{p-1}(U \cap V)$ *such that the following sequence is exact:*

$$\cdots \xrightarrow{\partial_*} H_p(U \cap V) \xrightarrow{\alpha} H_p(U) \oplus H_p(V) \xrightarrow{\beta} H_p(M) \xrightarrow{\partial_*} H_{p-1}(U \cap V) \xrightarrow{\alpha} \cdots, \tag{16.4}$$

where

$$\alpha[c] = (i_*[c], -j_*[c]),$$
$$\beta([c], [c']) = k_*[c] + l_*[c'],$$

and $\partial_*[e] = [c]$ *provided there exist* $d \in C_p(U)$ *and* $d' \in C_p(V)$ *such that* $k_\# d + l_\# d'$ *is homologous to* e *and*

$$(i_\# c, -j_\# c) = (\partial d, \partial d').$$

Sketch of Proof. The basic idea, of course, is to construct a short exact sequence of complexes and use the zigzag lemma. The hardest part of the proof is showing that any homology class $[e] \in H_p(M)$ can be represented in the form $[k_\# d + l_\# d']$, where d is a chain in U and d' is a chain in V. □

Note that the maps α and β in this Mayer–Vietoris sequence can be replaced by

$$\widetilde{\alpha}[c] = (i_*[c], j_*[c]),$$
$$\widetilde{\beta}([c],[c']) = k_*[c] - l_*[c'],$$

and the same proof goes through. If you consult various algebraic topology texts, you will find both definitions in use. We have chosen the definition given in the statement of the theorem because it leads to a cohomology exact sequence that is compatible with the Mayer–Vietoris sequence for de Rham cohomology; see the proof of the de Rham theorem below.

Singular Cohomology

In addition to the singular homology groups, for any abelian group G one can define a closely related sequence of groups $H^p(M;G)$ called the *singular cohomology groups* with coefficients in G. The precise definition is unimportant for our purposes; we will only be concerned with the special case $G = \mathbb{R}$, in which case it can be shown that $H^p(M;\mathbb{R})$ is a real vector space that is isomorphic to the space $\operatorname{Hom}(H_p(M),\mathbb{R})$ of group homomorphisms from $H_p(M)$ into $\mathbb{R}$. (We will simply take this as a definition of $H^p(M,\mathbb{R})$.) Any continuous map $F\colon M \to N$ induces a linear map $F^*\colon H^p(N;\mathbb{R}) \to H^p(M;\mathbb{R})$ by $(F^*\gamma)[c] = \gamma(F_*[c])$ for any $\gamma \in H^p(N;\mathbb{R}) \cong \operatorname{Hom}(H_p(M),\mathbb{R})$ and any singular p-chain c. The functorial properties of F_* carry over to cohomology: $(G \circ F)^* = F^* \circ G^*$ and $(\operatorname{Id}_M)^* = \operatorname{Id}_{H^p(M;\mathbb{R})}$.

The following properties of the singular cohomology groups follow easily from the definitions and Proposition 16.2.

Proposition 16.4 (Properties of Singular Cohomology).

(*a*) *For any one-point space* $\{q\}$, $H^p(\{q\};\mathbb{R})$ *is trivial except when* $p = 0$, *in which case it is* 1*-dimensional.*

(*b*) *If* $\{M_j\}$ *is any collection of topological spaces and* $M = \coprod_j M_j$, *then the inclusion maps* $\iota_j\colon M_j \hookrightarrow M$ *induce an isomorphism from* $H^p(M;\mathbb{R})$ *to* $\prod_j H^p(M_j;\mathbb{R})$.

(*c*) *Homotopy equivalent spaces have isomorphic singular cohomology groups.*

The key fact about the singular cohomology groups that we will need is that they too satisfy a Mayer–Vietoris theorem.

Theorem 16.5 (Mayer–Vietoris for Singular Cohomology). *Suppose M, U, and V satisfy the hypotheses of Theorem 16.3. The following sequence is exact:*

$$\cdots \xrightarrow{\partial^*} H^p(M;\mathbb{R}) \xrightarrow{k^*\oplus l^*} H^p(U;\mathbb{R}) \oplus H^p(V;\mathbb{R}) \xrightarrow{i^*-j^*} H^p(U\cap V;\mathbb{R}) \\ \xrightarrow{\partial^*} H^{p+1}(M;\mathbb{R}) \xrightarrow{k^*\oplus l^*} \cdots, \quad (16.5)$$

where the maps $k^*\oplus l^*$ and i^*-j^* are defined as in (15.11)*, and $\partial^*\gamma = \gamma\circ\partial_*$, with ∂_* as in Theorem 16.3.*

Sketch of Proof. For any homomorphism $F\colon A \to B$ between abelian groups, there is a *dual homomorphism* $F^*\colon \operatorname{Hom}(B,\mathbb{R}) \to \operatorname{Hom}(A,\mathbb{R})$ given by $F^*\gamma = \gamma\circ F$. Applying this to the Mayer–Vietoris sequence (16.4) for singular homology, we obtain the cohomology sequence (16.5). The exactness of the resulting sequence is a consequence of the fact that the functor $A \mapsto \operatorname{Hom}(A,\mathbb{R})$ is *exact*, meaning that it takes exact sequences to exact sequences. This in turn follows from the fact that $\mathbb{R}$ is an *injective group*: Whenever H is a subgroup of an abelian group G, every homomorphism from H into $\mathbb{R}$ extends to all of G. □

Smooth Singular Homology

The connection between the singular and de Rham cohomology groups will be established by integrating differential forms over singular chains. More precisely, given a singular p-simplex σ in a manifold M and a p-form ω on M, we would like to pull ω back by σ and integrate the resulting form over Δ_p. However, there is an immediate problem with this approach, because forms can only be pulled back by *smooth* maps, while singular simplices are in general only continuous. (Actually, since only first derivatives of the map appear in the formula for the pullback, it would be sufficient to consider C^1 maps, but merely continuous ones definitely will not do.) In this section we overcome this problem by showing that singular homology can be computed equally well with smooth simplices.

If M is a smooth manifold, a *smooth p-simplex* in M is a smooth map $\sigma\colon \Delta_p \to M$. (Recall that smoothness of a map from a non-open subset of $\mathbb{R}^p$ means that it has a smooth extension to a neighborhood of each point.) The subgroup of $C_p(M)$ generated by smooth simplices is denoted by $C_p^\infty(M)$ and called the *smooth chain group* in dimension p. Elements of this group, which are finite formal linear combinations of smooth simplices, are called *smooth chains*. Because the boundary of a smooth simplex is a smooth chain, we can define the pth *smooth singular homology group* of M to be the quotient group

$$H_p^\infty(M) = \frac{\operatorname{Ker}[\partial\colon C_p^\infty(M) \to C_{p-1}^\infty(M)]}{\operatorname{Im}[\partial\colon C_{p+1}^\infty(M) \to C_p^\infty(M)]}.$$

The inclusion map $\iota\colon C_p^\infty(M) \hookrightarrow C_p(M)$ obviously commutes with the boundary operator, and so induces a map on homology: $\iota_*\colon H_p^\infty(M) \to H_p(M)$ by $\iota_*[c] = [\iota(c)]$.

Theorem 16.6 (Smooth Singular vs. Singular Homology). *For any smooth manifold M, the map $\iota_*\colon H_p^\infty(M) \to H_p(M)$ induced by inclusion is an isomorphism.*

The basic idea of the proof is to construct, with the help of the Whitney approximation theorem, a smoothing operator $s\colon C_p(M) \to C_p^\infty(M)$ such that $s \circ \partial = \partial \circ s$ and $s \circ \iota$ is the identity on $C_p^\infty(M)$, and a homotopy operator that shows that $\iota \circ s$ induces the identity map on $H_p(M)$. The details are rather technical, so unless algebraic topology is your primary interest, you might wish to skim the rest of this section on first reading.

The key to the proof is a systematic construction of a homotopy from each continuous simplex to a smooth one, in a way that respects the restriction to each boundary face of Δ_p. This is summarized in the following lemma.

Lemma 16.7. *Let M be a smooth manifold. For each integer $p \geq 0$ and each singular p-simplex $\sigma\colon \Delta_p \to M$, there exists a continuous map $H_\sigma\colon \Delta_p \times I \to M$, such that the following properties hold:*

- *(i)* *H_σ is a homotopy from $\sigma(x) = H_\sigma(x,0)$ to a smooth p-simplex $\widetilde{\sigma}(x) = H_\sigma(x,1)$.*
- *(ii)* *For each face map $F_{i,p}\colon \Delta_{p-1} \to \Delta_p$,*
$$H_{\sigma \circ F_{i,p}} = H_\sigma \circ (F_{i,p} \times \mathrm{Id}), \tag{16.6}$$
or more explicitly,
$$H_{\sigma \circ F_{i,p}}(x,t) = H_\sigma(F_{i,p}(x),t), \quad (x,t) \in \Delta_{p-1} \times I. \tag{16.7}$$
- *(iii)* *If σ is a smooth p-simplex, then H_σ is the constant homotopy $H_\sigma(x,t) = \sigma(x)$.*

Proof. We will construct the homotopies H_σ (see Figure 16.4) by induction on the dimension of σ. To get started, for each 0-simplex $\sigma\colon \Delta_0 \to M$, we just define $H_\sigma(x,t) = \sigma(x)$. Since every 0-simplex is smooth and there are no face maps, conditions (i)–(iii) are automatically satisfied.

Now suppose that for each $p' < p$ and for every p'-simplex σ' we have defined $H_{\sigma'}$, in such a way that the primed analogues of (i)–(iii) are satisfied. Let $\sigma\colon \Delta_p \to M$ be an arbitrary singular p-simplex in M. If σ is smooth,

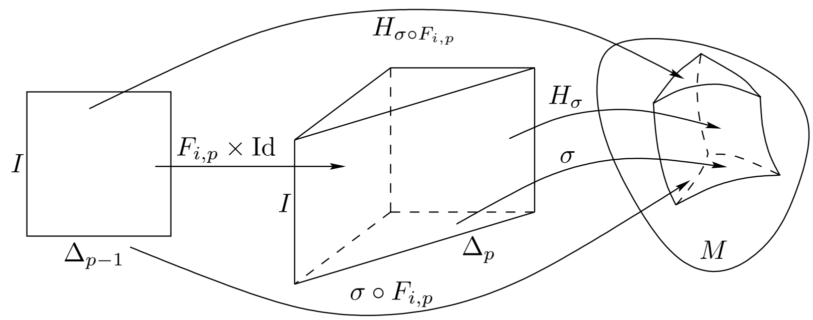


Figure 16.4. The homotopy H_σ.

we just let $H_\sigma(x,t) = \sigma(x)$, and (i)–(iii) are easily verified (using the fact that the restriction of σ to each boundary face is also smooth).

Assume σ is not smooth, and let S be the subset

$$S = (\Delta_p \times \{0\}) \cup (\partial\Delta_p \times I) \subset \Delta_p \times I$$

(the bottom and side faces of the "prism" $\Delta_p \times I$). Observe that $\partial\Delta_p$ is the union of the boundary faces $\partial_i\Delta_p$ for $i = 0, \dots, p$, and recall that for each i, the face map $F_{i,p}\colon \Delta_{p-1} \to \partial_i\Delta_p$ is a homeomorphism onto the ith boundary face. Define a map $H_0\colon S \to M$ by

$$H_0(x,t) = \begin{cases} \sigma(x), & x \in \Delta_p,\ t = 0; \\ H_{\sigma\circ F_{i,p}}(F_{i,p}^{-1}(x), t), & x \in \partial_i\Delta_p,\ t \in I. \end{cases}$$

We need to check that the various definitions agree where they overlap, which will imply that H_0 is continuous by the gluing lemma.

When $t = 0$, the inductive hypothesis (i) applied to the $(p-1)$-simplex $\sigma \circ F_{i,p}$ implies that $H_{\sigma\circ F_{i,p}}(x,0) = \sigma \circ F_{i,p}(x)$. It follows that $H_{\sigma\circ F_{i,p}}(F_{i,p}^{-1}(x), 0) = \sigma(x)$, so the different definitions of H_0 agree at points where $t = 0$.

Suppose now that x is a point in the intersection of two boundary faces $\partial_i\Delta_p$ and $\partial_j\Delta_p$, and assume without loss of generality that $i > j$. Since $F_{i,p} \circ F_{j,p-1}$ is a homeomorphism from Δ_{p-2} onto $\partial_i\Delta_p \cap \partial_j\Delta_p$, we can write $x = F_{i,p} \circ F_{j,p-1}(y)$ for some point $y \in \Delta_{p-2}$. Then (16.7) applied with $\sigma \circ F_{i,p}$ in place of σ and $F_{j-1,p}$ in place of $F_{i,p}$ implies that

$$\begin{aligned} H_{\sigma\circ F_{i,p}}(F_{i,p}^{-1}(x), t) &= H_{\sigma\circ F_{i,p}}(F_{j,p-1}(y), t) \\ &= H_{\sigma\circ F_{i,p}\circ F_{j,p-1}}(y,t). \end{aligned}$$

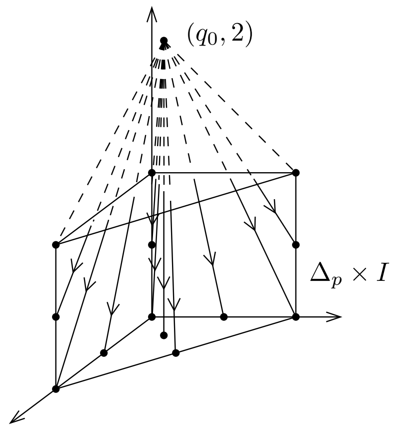


Figure 16.5. A retraction from $\Delta_p \times I$ onto S.

On the other hand, thanks to (16.1), we can also write $x = F_{j,p} \circ F_{i-1,p-1}(y)$, and then the same argument applied to $\sigma \circ F_{j,p}$ yields

$$\begin{aligned} H_{\sigma \circ F_{j,p}}(F_{j,p}^{-1}(x), t) &= H_{\sigma \circ F_{j,p}}(F_{i-1,p-1}(y), t) \\ &= H_{\sigma \circ F_{j,p} \circ F_{i-1,p-1}}(y, t). \end{aligned}$$

Because of (16.1), this shows that the two definitions of $H_0(x,t)$ agree.

To extend H_0 to all of $\Delta_p \times I$, we will use the fact that there is a retraction from $\Delta_p \times I$ onto S. For example, if q_0 is any point in the interior of Δ_p, then the map $R\colon \Delta_p \times I \to S$ obtained by radially projecting from the point $(q_0, 2) \in \mathbb{R}^p \times \mathbb{R}$ is such a retraction (see Figure 16.5). We extend H_0 to a continuous map $H\colon \Delta_p \times I \to M$ by setting

$$H(x,t) = H_0(R(x,t)).$$

Because H agrees with H_0 on S, it is a homotopy from σ to some other (continuous) singular simplex $\sigma'(x) = H(x,1)$, and it satisfies (16.7) by construction. Our only remaining task is to modify H so that it becomes a homotopy from σ to a *smooth* simplex.

Before we do so, we need to observe first that the restriction of H to each boundary face $\partial_i \Delta_p \times \{1\}$ is smooth: Since these faces lie in S, H agrees with H_0 on each of these sets, and hypothesis (i) applied to $\sigma \circ F_{i,p}$ shows that H_0 is smooth there. By virtue of Lemma 16.8 below, this implies that the restriction of H to the entire set $\partial \Delta_p \times \{1\}$ is smooth. Let σ'' be any continuous extension of σ' to an open set $U \subset \mathbb{R}^p$ containing Δ_p. (For example, σ'' could be defined by projecting points outside Δ_p to $\partial \Delta_p$ along radial lines from some point in the interior of Δ_p, and then applying σ'.) By the Whitney approximation theorem for manifolds (Theorem 10.21), σ'' is homotopic relative to $\partial \Delta_p$ to a smooth map, and restricting the homotopy to $\Delta_p \times I$ we obtain a homotopy $G\colon \sigma' \simeq \widetilde{\sigma}$ from σ' to some smooth singular p-simplex $\widetilde{\sigma}$, again relative to $\partial\Delta_p$.

Now let $u\colon \Delta_p \to \mathbb{R}$ be any continuous function that is equal to 1 on $\partial\Delta_p$ and satisfies $0 < u(x) < 1$ for $u \in \operatorname{Int}\Delta_p$. (For example we could take $u(\sum_{0\le i\le p} t_i e_i) = 1 - t_0 t_1 \cdots t_p$, where $\sum_{0\le i\le p} t_i = 1$ and $e_0, \dots, e_p$ are the vertices of Δ_p.) We combine the two homotopies H and G into a single homotopy $H_\sigma\colon \Delta_p \times I \to M$ by

$$H_\sigma(x,t) = \begin{cases} H\left(x, \dfrac{t}{u(x)}\right), & x \in \Delta_p,\ 0 \le t \le u(x), \\ G\left(x, \dfrac{t-u(x)}{1-u(x)}\right), & x \in \operatorname{Int}\Delta_p,\ u(x) \le t \le 1. \end{cases}$$

Because $H(x,1) = \sigma'(x) = G(x,0)$, the gluing lemma shows that H_σ is continuous in $\operatorname{Int}\Delta_p \times I$. Also, $H_\sigma(x,t) = H(x, t/u(x))$ in a neighborhood of $\partial\Delta_p \times [0,1)$, and thus is continuous there. It remains only to show that H_σ is continuous on $\partial\Delta_p \times \{1\}$. Let $x_0 \in \partial\Delta_p$ be arbitrary, and let $U \subset M$ be any neighborhood of $H_\sigma(x_0,1) = H(x_0,1)$. By continuity of H and u, there exists $\delta_1 > 0$ such that $H(x, t/u(x)) \in U$ whenever $|(x,t) - (x_0,1)| < \delta_1$ and $0 \le t \le u(x)$. Since $G(x_0,t) = G(x_0,0) = H(x_0,1) = H_\sigma(x_0,1) \in U$ for all $t \in I$, a simple compactness argument shows that there exists $\delta_2 > 0$ such that $|x - x_0| < \delta_2$ implies $G(x,t) \in U$ for all $t \in I$. Thus if $|(x,t) - (x_0,1)| < \min(\delta_1, \delta_2)$, we have $H_\sigma(x,t) \in U$ in both cases, showing that $(x_0,1)$ has a neighborhood mapped into U by H_σ. Thus H_σ is continuous.

It follows from the definition that $H_\sigma = H$ on $\partial\Delta_p \times I$, so (ii) is satisfied. For any $x \in \Delta_p$, $H_\sigma(x,0) = H(x,0) = \sigma(x)$. Moreover, when $x \in \operatorname{Int}\Delta_p$, $H_\sigma(x,1) = G(x,1) = \widetilde{\sigma}(x)$, and when $x \in \partial\Delta_p$, $H_\sigma(x,1) = H(x,1) = \sigma'(x) = \widetilde{\sigma}(x)$ (because G is a homotopy relative to $\partial\Delta_p$). Thus H_σ is a homotopy from σ to the smooth simplex $\widetilde{\sigma}$, and (i) is satisfied as well. □

Here is the lemma used in the preceding proof.

Lemma 16.8. *Let M be a smooth manifold, let Δ be a geometric p-simplex in $\mathbb{R}^n$, and let $f\colon \partial\Delta \to M$ be a continuous map whose restriction to each individual boundary face of Δ is smooth. Then f is smooth when considered as a map from the entire boundary $\partial\Delta$ to M.*

Proof. Let $(v_0, \dots, v_p)$ denote the vertices of Δ in some order, and for each $i = 0, \dots, p$, let $\partial_i\Delta = \langle v_0, \dots, \widehat{v_i}, \dots, v_p\rangle$ be the boundary face opposite v_i. The hypothesis means that for each i and each $x \in \partial_i\Delta$, there exist an open set $U_x \subset \mathbb{R}^n$ and a smooth map $\widetilde{f}\colon U_x \to M$ whose restriction to $U_x \cap \partial_i\Delta$ agrees with f. We need to show that a single smooth extension can be chosen simultaneously for all the boundary faces containing x.

Suppose $x \in \partial\Delta$. Note that x is in one or more boundary faces of Δ, but cannot be in all of them. By reordering the vertices, we may assume that $x \in \partial_1\Delta \cap \dots \cap \partial_k\Delta$ for some $1 \le k \le p$, but $x \notin \partial_0\Delta$. After composing with an affine diffeomorphism that takes v_i to e_i for $i = 0, \dots, p$, we may assume

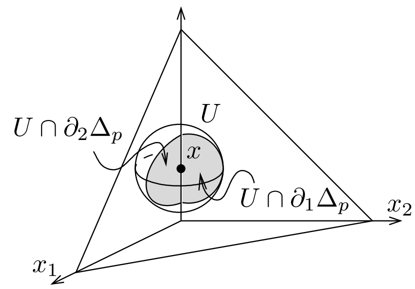


Figure 16.6. Showing that f is smooth on $\partial\Delta_p$.

without loss of generality that $\Delta = \Delta_p$ and $x \notin \partial_0\Delta_p$. Then the boundary faces containing x are precisely the intersections with Δ_p of the coordinate hyperplanes $x^1 = 0, \ldots, x^k = 0$. For each i, there are a neighborhood U_i of x in $\mathbb{R}^n$ (which can be chosen disjoint from $\partial_0\Delta_p$) and a smooth map $\widetilde{f}_i\colon U_i \to M$ whose restriction to $U_i \cap \partial_i\Delta_p$ agrees with f.

Let $U = U_1 \cap \cdots \cap U_k$. We will show by induction on k that there is a smooth map $\widetilde{f}\colon U \to M$ whose restriction to $U \cap \partial_i\Delta_p$ agrees with f for $i = 1, \ldots, k$. (See Figure 16.6.)

For $k = 1$, there is nothing to prove, because $\widetilde{f}_1$ is already such an extension. So suppose $k \geq 2$, and we have shown that there is a smooth map $\widetilde{f}_0\colon U \to M$ whose restriction to $U \cap \partial_i\Delta_p$ agrees with f for $i = 1, \ldots, k-1$. Define $\widetilde{f}\colon U \to M$ by

$$\begin{aligned}\widetilde{f}(x^1, \ldots, x^n) = \widetilde{f}_0&(x^1, \ldots, x^n)\\ &- \widetilde{f}_0(x^1, \ldots, x^{k-1}, 0, x^{k+1}, \ldots, x^n)\\ &+ \widetilde{f}_k(x^1, \ldots, x^{k-1}, 0, x^{k+1}, \ldots, x^n).\end{aligned}$$

For $i = 1, \ldots, k-1$, the restriction of $\widetilde{f}$ to $U \cap \partial_i\Delta_p$ is given by

$$\begin{aligned}&\widetilde{f}(x^1, \ldots, x^{i-1}, 0, x^{i+1}, \ldots, x^n)\\ &\qquad = \widetilde{f}_0(x^1, \ldots, x^{i-1}, 0, x^{i+1}, \ldots, x^n)\\ &\qquad\quad - \widetilde{f}_0(x^1, \ldots, x^{i-1}, 0, x^{i+1}, \ldots, x^{k-1}, 0, x^{k+1}, \ldots, x^n)\\ &\qquad\quad + \widetilde{f}_k(x^1, \ldots, x^{i-1}, 0, x^{i+1}, \ldots, x^{k-1}, 0, x^{k+1}, \ldots, x^n)\\ &\qquad = f(x^1, \ldots, x^{i-1}, 0, x^{i+1}, \ldots, x^n),\end{aligned}$$

since $\widetilde{f}_0$ agrees with f when $x \in \Delta_p$ and $x^i = 0$, as does $\widetilde{f}_k$ when $x \in \Delta_p$ and $x^k = 0$. Similarly, the restriction to $U \cap \partial_k \Delta_p$ is

$$\begin{aligned}
\widetilde{f}(x^1, \dots, &x^{k-1}, 0, x^{k+1}, \dots, x^n) \\
&= \widetilde{f}_0(x^1, \dots, x^{k-1}, 0, x^{k+1}, \dots, x^n) \\
&\quad - \widetilde{f}_0(x^1, \dots, x^{k-1}, 0, x^{k+1}, \dots, x^n) \\
&\quad + \widetilde{f}_k(x^1, \dots, x^{k-1}, 0, x^{k+1}, \dots, x^n) \\
&= f(x^1, \dots, x^{k-1}, 0, x^{k+1}, \dots, x^n).
\end{aligned}$$

This completes the inductive step and thus the proof. □

Proof of Theorem 16.6. Let $i_0, i_1 \colon \Delta_p \to \Delta_p \times I$ be the smooth embeddings $i_0(x) = (x, 0)$, $i_1(x) = (x, 1)$. Define a group homomorphism $s \colon C_p(M) \to C_p^\infty(M)$ by setting

$$s\sigma = H_\sigma \circ i_1$$

for each singular p-simplex σ (where H_σ is the homotopy whose existence is proved in Lemma 16.7), and extending linearly to p-chains. Because of property (i) in Lemma 16.7, $s\sigma$ is a smooth p-simplex homotopic to σ.

Using (16.6), it is easy to verify that s is a chain map: For each singular p-simplex σ,

$$\begin{aligned}
s\partial\sigma &= s \sum_{i=0}^{p} (-1)^i \sigma \circ F_{i,p} \\
&= \sum_{i=0}^{p} (-1)^i H_{\sigma \circ F_{i,p}} \circ i_1 \\
&= \sum_{i=0}^{p} (-1)^i H_\sigma \circ (F_{i,p} \times \mathrm{Id}) \circ i_1 \\
&= \sum_{i=0}^{p} (-1)^i H_\sigma \circ i_1 \circ F_{i,p} \\
&= \partial(H_\sigma \circ i_1) \\
&= \partial s\sigma.
\end{aligned}$$

(In the fourth line, we used the fact that $(F_{i,p} \times \mathrm{Id}) \circ i_1(x) = (F_{i,p}(x), 1) = i_1 \circ F_{i,p}(x)$.) Therefore s descends to a homomorphism $s_* \colon H_p(M) \to H_p^\infty(M)$. We will show that s_* is an inverse for $\iota_* \colon H_p^\infty(M) \to H_p(M)$.

First, observe that condition (iii) in Lemma 16.7 guarantees that $s \circ \iota$ is the identity map of $C_p^\infty(M)$, so clearly $s_* \circ \iota_*$ is the identity on $H_p^\infty(M)$. To show that $\iota_* \circ s_*$ is also the identity, we will construct for each $p \geq 0$ a homotopy operator $h \colon C_p(M) \to C_{p+1}(M)$ satisfying

$$\partial \circ h + h \circ \partial = \iota \circ s - \mathrm{Id}_{C_p(M)}\,. \tag{16.8}$$

Once the existence of such an operator is known, it follows just as in the proof of Proposition 15.5 that $\iota_* \circ s_* = \mathrm{Id}_{H_p(M)}$: For any cycle $c \in C_p(M)$,

$$\iota_* \circ s_*[c] - [c] = [\iota \circ s(c) - c] = [\partial(hc) + h(\partial c)] = 0,$$

because $\partial c = 0$ and $\partial(hc)$ is a boundary.

To define the homotopy operator h, we need to introduce a family of affine singular simplices in the convex set $\Delta_p \times I \subset \mathbb{R}^p \times \mathbb{R}$. For each $i = 0, \dots, p$, let $E_i = (e_i, 0) \in \mathbb{R}^p \times \mathbb{R}$ and $E_i' = (e_1, 1) \in \mathbb{R}^p \times \mathbb{R}$, so that $E_0, \dots, E_p$ are the vertices of the geometric p-simplex $\Delta_p \times \{0\}$, and $E_0' \dots, E_p'$ are those of $\Delta_p \times \{1\}$. For each $i = 0, \dots, p$, let $G_{i,p}\colon \Delta_{p+1} \to \Delta_p \times I$ be the affine singular $(p+1)$-simplex

$$G_{i,p} = \alpha(E_0, \dots, E_i, E_i', \dots, E_p').$$

Thus $G_{i,p}$ is the unique affine map that sends $e_0 \mapsto E_0$, $\dots$, $e_i \mapsto E_i$, $e_{i+1} \mapsto E_i'$, $\dots$, and $e_{p+1} \mapsto E_p'$. A routine computation shows that these maps compose with the face maps as follows:

$$G_{j,p} \circ F_{j,p+1} = G_{j-1,p} \circ F_{j,p+1} = \alpha(E_0, \dots, E_{j-1}, E_j', \dots, E_p'). \tag{16.9}$$

In particular, this implies that

$$\begin{aligned} G_{p,p} \circ F_{p+1,p+1} &= \alpha(E_0, \dots, E_p) = i_0, & (16.10)\\ G_{0,p} \circ F_{0,p+1} &= \alpha(E_0', \dots, E_p') = i_1. & (16.11) \end{aligned}$$

A similar computation shows that

$$(F_{j,p} \times \mathrm{Id}) \circ G_{i,p-1} = \begin{cases} G_{i+1,p} \circ F_{j,p+1}, & i \geq j, \\ G_{i,p} \circ F_{j+1,p+1}, & i < j. \end{cases} \tag{16.12}$$

We define $h\colon C_p(M) \to C_{p+1}(M)$ as follows:

$$h\sigma = \sum_{i=0}^{p} (-1)^i H_\sigma \circ G_{i,p}.$$

The proof that it satisfies the homotopy formula (16.8) is just a laborious computation using (16.7), (16.9), and (16.12):

$$
\begin{aligned}
h(\partial\sigma) &= h\sum_{j=0}^{p}(-1)^j\sigma\circ F_{j,p} \\
&= \sum_{i=0}^{p-1}\sum_{j=0}^{p}(-1)^{i+j}H_{\sigma\circ F_{j,p}}\circ G_{i,p-1} \\
&= \sum_{i=0}^{p-1}\sum_{j=0}^{p}(-1)^{i+j}H_{\sigma}\circ(F_{j,p}\times \mathrm{Id})\circ G_{i,p-1} \\
&= \sum_{0\le j\le i\le p-1}(-1)^{i+j}H_\sigma\circ G_{i+1,p}\circ F_{j,p+1} \\
&\quad + \sum_{0\le i<j\le p}(-1)^{i+j}H_\sigma\circ G_{i,p}\circ F_{j+1,p+1},
\end{aligned}
\tag{16.13}
$$

while

$$
\begin{aligned}
\partial(h\sigma) &= \partial\sum_{i=0}^{p}(-1)^iH_\sigma\circ G_{i,p} \\
&= \sum_{j=0}^{p+1}\sum_{i=0}^{p}(-1)^{i+j}H_\sigma\circ G_{i,p}\circ F_{j,p+1}.
\end{aligned}
$$

Writing separately the terms in $\partial(h\sigma)$ for which $i<j-1$, $i=j-1$, $i=j$, and $i>j$, we get

$$
\begin{aligned}
\partial(h\sigma) = &\sum_{\substack{0\le i<j-1\\ j\le p+1}}(-1)^{i+j}H_\sigma\circ G_{i,p}\circ F_{j,p+1} \\
&- \sum_{1\le j\le p+1}H_\sigma\circ G_{j-1,p}\circ F_{j,p+1} \\
&+ \sum_{0\le j\le p}H_\sigma\circ G_{j,p}\circ F_{j,p+1} \\
&+ \sum_{0\le j<i\le p}(-1)^{i+j}H_\sigma\circ G_{i,p}\circ F_{j,p+1}.
\end{aligned}
$$

After substituting $j=j'+1$ in the first of these four sums and $i=i'+1$ in the last, the first and last sums exactly cancel the two sums in the expression (16.13) for $h(\partial\sigma)$. Using (16.9), all the terms in the middle two sums cancel each other except those in which $j=0$ and $j=p+1$. Thanks to (16.10) and (16.11), these two terms simplify to

$$
\begin{aligned}
h(\partial\sigma)+\partial(h\sigma) &= -H_\sigma\circ G_{p,p}\circ F_{p+1,p+1}+H_\sigma\circ G_{0,p}\circ F_{0,p+1} \\
&= -H_\sigma\circ i_0+H_\sigma\circ i_1 \\
&= -\sigma+s\sigma.
\end{aligned}
$$

Since ι is an inclusion map, $s\sigma = \iota \circ s\sigma$ for any singular p-simplex σ, so this completes the proof. □

The de Rham Theorem

In this section we will state and prove the de Rham theorem. Before getting to the theorem itself, we need one more algebraic lemma. Its proof is another diagram chase like the proof of the zigzag lemma.

Lemma 16.9 (The Five Lemma). *Consider the following commutative diagram of modules and linear maps:*

$$\begin{tikzcd}
A_1 \arrow[r, "\alpha_1"] \arrow[d, "f_1"] & A_2 \arrow[r, "\alpha_2"] \arrow[d, "f_2"] & A_3 \arrow[r, "\alpha_3"] \arrow[d, "f_3"] & A_4 \arrow[r, "\alpha_4"] \arrow[d, "f_4"] & A_5 \arrow[d, "f_5"] \\
B_1 \arrow[r, "\beta_1"] & B_2 \arrow[r, "\beta_2"] & B_3 \arrow[r, "\beta_3"] & B_4 \arrow[r, "\beta_4"] & B_5.
\end{tikzcd}$$

If the horizontal rows are exact and f_1, f_2, f_4, and f_5 are isomorphisms, then f_3 is also an isomorphism.

◇ **Exercise 16.2.** Prove (or look up) the five lemma.

Suppose M is a smooth manifold, ω is a closed p-form on M, and σ is a smooth p-simplex in M. We define the *integral of ω over σ* to be

$$\int_\sigma \omega = \int_{\Delta_p} \sigma^*\omega.$$

This makes sense because Δ_p is a smooth p-submanifold with corners embedded in $\mathbb{R}^p$, and inherits the orientation of $\mathbb{R}^p$. (Or we could just consider Δ_p as a domain of integration in $\mathbb{R}^p$.) Observe that when $p = 1$, this is the same as the line integral of ω over the smooth curve segment $\sigma\colon [0,1] \to M$. If $c = \sum_{i=1}^k c_i\sigma_i$ is a smooth p-chain, the integral of ω over c is defined as

$$\int_c \omega = \sum_{i=1}^k c_i \int_{\sigma_i} \omega.$$

Theorem 16.10 (Stokes's Theorem for Chains). *If c is a smooth p-chain in a smooth manifold M, and ω is a smooth $(p-1)$-form on M, then*

$$\int_{\partial c} \omega = \int_c d\omega.$$

Proof. It suffices to prove the theorem when c is just a smooth simplex σ. Since Δ_p is a manifold with corners, Stokes's theorem says

$$\int_\sigma d\omega = \int_{\Delta_p} \sigma^* d\omega = \int_{\Delta_p} d\sigma^*\omega = \int_{\partial\Delta_p} \sigma^*\omega.$$

The maps $\{F_{i,p} : 0 = 1, \dots, p\}$ are parametrizations of the boundary faces of Δ_p satisfying the conditions of Proposition 14.18, except possibly that they might not be orientation preserving. To check the orientations, note that $F_{i,p}$ is the restriction to $\Delta_p \cap \partial\mathbb{H}^p$ of the affine diffeomorphism sending $\langle e_0, \dots, e_p\rangle$ to $\langle e_0, \dots, \widehat{e_i}, \dots, e_p, e_i\rangle$. This is easily seen to be orientation preserving if and only if $(e_0, \dots, \widehat{e_i}, \dots, e_p, e_i)$ is an even permutation of $(e_0, \dots, e_p)$, which is the case if and only if $p-i$ is even. Since the standard coordinates on $\partial\mathbb{H}^p$ are positively oriented if and only if p is even, the upshot is that $F_{i,p}$ is orientation preserving for $\partial\Delta_p$ if and only if i is even. Thus, by Proposition 14.18.

$$\begin{aligned}\int_{\partial\Delta_p} \sigma^*\omega &= \sum_{i=0}^{p}(-1)^i \int_{\Delta_{p-1}} F_{i,p}^*\sigma^*\omega\\ &= \sum_{i=0}^{p}(-1)^i \int_{\Delta_{p-1}} (\sigma\circ F_{i,p})^*\omega\\ &= \sum_{i=0}^{p}(-1)^i \int_{\sigma\circ F_{i,p}} \omega.\end{aligned}$$

By definition of the singular boundary operator, this is equal to $\int_{\partial\sigma}\omega$. □

Using this theorem, we can define a natural linear map $\mathcal{I}\colon H^p_{dR}(M) \to H^p(M;\mathbb{R})$, called the *de Rham homomorphism*, as follows. For any $[\omega] \in H^p_{dR}(M)$ and $[c] \in H_p(M) \cong H_p^\infty(M)$, we define

$$\mathcal{I}[\omega][c] = \int_{\widetilde{c}} \omega, \tag{16.14}$$

where $\widetilde{c}$ is any smooth p-cycle representing the homology class $[c]$. This is well-defined, because if $\widetilde{c}, \widetilde{c}'$ are smooth cycles representing the same homology class, then Theorem 16.6 guarantees that $\widetilde{c} - \widetilde{c}' = \partial\widetilde{b}$ for some smooth $(p+1)$-chain $\widetilde{b}$, which implies

$$\int_{\widetilde{c}}\omega - \int_{\widetilde{c}'}\omega = \int_{\partial\widetilde{b}}\omega = \int_{\widetilde{b}} d\omega = 0,$$

while if $\omega = d\eta$ is exact, then

$$\int_{\widetilde{c}}\omega = \int_{\widetilde{c}} d\eta = \int_{\partial\widetilde{c}}\omega = 0.$$

(Note that $\partial\widetilde{c} = 0$ because $\widetilde{c}$ represents a homology class, and $d\omega = 0$ because ω represents a cohomology class.) Clearly $\mathcal{I}[\omega][c + c'] = \mathcal{I}[\omega][c] + \mathcal{I}[\omega][c']$, and the resulting homomorphism $\mathcal{I}[\omega]\colon H_p(M) \to \mathbb{R}$ depends linearly on ω. Thus $\mathcal{I}[\omega]$ is a well-defined element of $\mathrm{Hom}(H_p(M), \mathbb{R})$, which we are identifying with $H^p(M;\mathbb{R})$.

Lemma 16.11 (Naturality of the de Rham Homomorphism). *For any smooth manifold M and any nonnegative integer p, let $\mathcal{I}\colon H^p_{dR}(M) \to H^p(M;\mathbb{R})$ denote the de Rham homomorphism.*

(*a*) *If $F\colon M \to N$ is a smooth map, then the following diagram commutes:*

$$\begin{tikzcd} H^p_{dR}(N) \arrow[r, "F^*"] \arrow[d, "\mathcal{I}"'] & H^p_{dR}(M) \arrow[d, "\mathcal{I}"] \\ H^p(N;\mathbb{R}) \arrow[r, "F^*"'] & H^p(M;\mathbb{R}). \end{tikzcd}$$

(*b*) *If M is a smooth manifold and U, V are open subsets of M whose union is M, then the following diagram commutes:*

$$\begin{tikzcd} H^{p-1}_{dR}(U\cap V) \arrow[r, "\delta"] \arrow[d, "\mathcal{I}"'] & H^p_{dR}(M) \arrow[d, "\mathcal{I}"] \\ H^{p-1}(U\cap V;\mathbb{R}) \arrow[r, "\partial^*"'] & H^p(M;\mathbb{R}), \end{tikzcd} \tag{16.15}$$

where δ and ∂^* are the connecting homomorphisms of the Mayer–Vietoris sequences for de Rham and singular cohomology, respectively.

Proof. Directly from the definitions, if σ is a smooth p-simplex in M and ω is a smooth p-form on N,

$$\int_\sigma F^*\omega = \int_{\Delta_p} \sigma^* F^*\omega = \int_{\Delta_p} (F\circ\sigma)^*\omega = \int_{F\circ\sigma} \omega.$$

This implies

$$\begin{aligned} \mathcal{I}(F^*[\omega])[\sigma] &= \mathcal{I}[\omega][F\circ\sigma] \\ &= \mathcal{I}[\omega](F_*[\sigma]) \\ &= F^*(\mathcal{I}[\omega])[\sigma], \end{aligned}$$

which proves (a).

For (b), identifying $H^p(M;\mathbb{R})$ with $\operatorname{Hom}(H_p(M),\mathbb{R}) \cong \operatorname{Hom}(H_p^\infty(M),\mathbb{R})$ as usual, commutativity of (16.15) reduces to the following equation for any $[\omega] \in H^{p-1}_{dR}(U\cap V)$ and any $[e] \in H_p(M)$:

$$\mathcal{I}(\delta[\omega])[e] = (\partial^*\mathcal{I}[\omega])[e] = \mathcal{I}[\omega](\partial_*[e]).$$

If σ is a smooth p-form representing $\delta[\omega]$ and c is a smooth $(p-1)$-chain representing $\partial_*[e]$, this is the same as

$$\int_e \sigma = \int_c \omega.$$

By the characterizations of δ and ∂_* given in Corollary 15.22 and Theorem 16.3, we can choose $\sigma = d\eta$ (extended by zero to all of M), where $\eta \in \mathcal{A}^{p-1}(U)$ and $\eta' \in \mathcal{A}^{p-1}(V)$ are forms such that $\omega = \eta|_{U\cap V} - \eta'|_{U\cap V}$; and

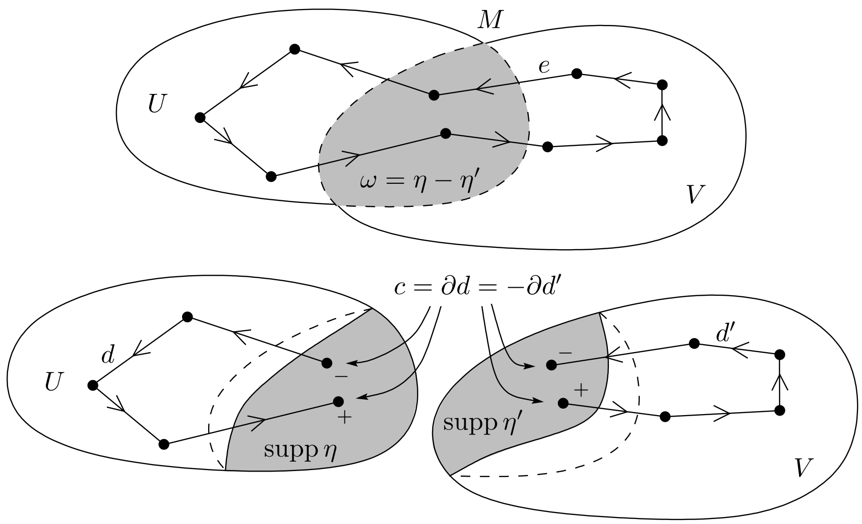


Figure 16.7. Naturality of $\mathcal{I}$ with respect to connecting homomorphisms.

$c = \partial d$, where d, d' are smooth p-chains in U and V, respectively, such that $d+d'$ represents the same homology class as e (Figure 16.7). Then, because $\partial d + \partial d' = \partial e = 0$ and $d\eta|_{U\cap V} - d\eta'|_{U\cap V} = d\omega = 0$, we have

$$\begin{aligned}
\int_c \omega &= \int_{\partial d} \omega \\
&= \int_{\partial d} \eta - \int_{\partial d} \eta' \\
&= \int_{\partial d} \eta + \int_{\partial d'} \eta' \\
&= \int_d d\eta + \int_{d'} d\eta' \\
&= \int_d \sigma + \int_{d'} \sigma \\
&= \int_e \sigma.
\end{aligned}$$

Thus the diagram commutes. □

Theorem 16.12 (de Rham). *For any smooth manifold M and any nonnegative integer p, the de Rham homomorphism $\mathcal{I}\colon H^p_{dR}(M) \to H^p(M;\mathbb{R})$ is an isomorphism.*

Proof. Let us say that a smooth manifold M is a *de Rham manifold* if the de Rham homomorphism $\mathcal{I}\colon H^p_{dR}(M) \to H^p(M;\mathbb{R})$ is an isomorphism for each p. Since the de Rham homomorphism commutes with the cohomology maps induced by smooth maps (Lemma 16.11), any manifold that is diffeomorphic to a de Rham manifold is also de Rham. The theorem will be proved once we show that every smooth manifold is de Rham.

If M is any smooth manifold, an open cover $\{U_i\}$ of M is called a *de Rham cover* if each open set U_i is a de Rham manifold, and every finite intersection $U_{i_1} \cap \dots \cap U_{i_k}$ is de Rham. A de Rham cover that is also a basis for the topology of M is called a *de Rham basis* for M.

Step 1: *If $\{M_j\}$ is any countable collection of de Rham manifolds, then their disjoint union is de Rham.* By Propositions 15.8 and 16.4(b), for both de Rham and singular cohomology, the inclusions $\iota_j \colon M_j \hookrightarrow \coprod_j M_j$ induce isomorphisms between the cohomology groups of the disjoint union and the direct product of the cohomology groups of the manifolds M_j. By Lemma 16.11, $\mathcal{I}$ respects these isomorphisms.

Step 2: *Every convex open subset of $\mathbb{R}^n$ is de Rham.* Let U be such a subset. By the Poincaré lemma, $H^p_{dR}(U)$ is trivial when $p \neq 0$. Since U is homotopy equivalent to a one-point space, Proposition 16.4 implies that the singular cohomology groups of U are also trivial for $p \neq 0$. In the $p = 0$ case, $H^0_{dR}(U)$ is the 1-dimensional space consisting of the constant functions, and $H^0(U;\mathbb{R}) = \operatorname{Hom}(H_0(U), \mathbb{R})$ is also 1-dimensional because $H_0(U)$ is generated by any singular 0-simplex. If $\sigma \colon \Delta_0 \to M$ is a singular 0-simplex (which is smooth because any map from a 0-manifold is smooth), and f is the constant function equal to 1, then

$$\mathcal{I}[f][\sigma] = \int_{\Delta_0} \sigma^* f = (f \circ \sigma)(0) = 1.$$

This shows that $\mathcal{I} \colon H^0_{dR}(U) \to H^0(U;\mathbb{R})$ is not the zero map, so it is an isomorphism.

Step 3: *If M has a finite de Rham cover, then M is de Rham.* This is the heart of the proof. Suppose $M = U_1 \cup \dots \cup U_k$, where the open sets U_i and their finite intersections are de Rham. We will prove the result by induction on k. Suppose first that M has a de Rham cover consisting of two sets $\{U, V\}$. Putting together the Mayer–Vietoris sequences for de Rham and singular cohomology, we obtain the following commutative diagram in which the horizontal rows are exact and the vertical maps are all given by de Rham homomorphisms:

$$\begin{tikzcd}
H^{p-1}_{dR}(U) \oplus H^{p-1}_{dR}(V) \arrow[r] \arrow[d] & H^{p-1}_{dR}(U \cap V) \arrow[r] \arrow[d] & H^p_{dR}(M) \arrow[r] \arrow[d] & {} \\
H^{p-1}(U;\mathbb{R}) \oplus H^{p-1}(V;\mathbb{R}) \arrow[r] & H^{p-1}(U \cap V;\mathbb{R}) \arrow[r] & H^p(M;\mathbb{R}) \arrow[r] & {}
\end{tikzcd}$$

$$\begin{tikzcd}
H^p_{dR}(U) \oplus H^p_{dR}(V) \arrow[r] \arrow[d] & H^p_{dR}(U \cap V) \arrow[d] \\
H^p(U;\mathbb{R}) \oplus H^p(V;\mathbb{R}) \arrow[r] & H^p(U \cap V;\mathbb{R}).
\end{tikzcd}$$

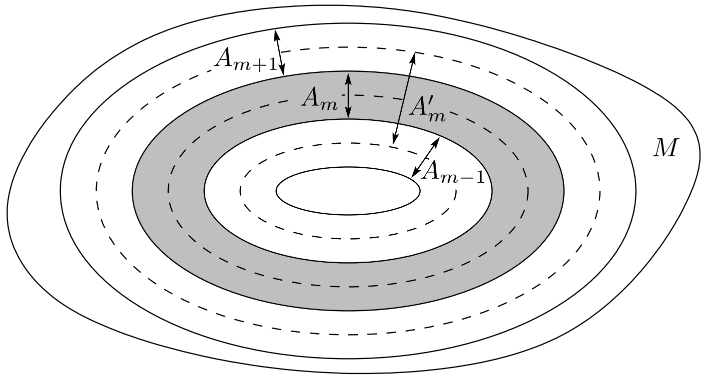


Figure 16.8. Proof of the de Rham theorem, Step 4.

The commutativity of the diagram is an immediate consequence of Lemma 16.11. By hypothesis the first, second, fourth, and fifth vertical maps are all isomorphisms, so by the five lemma the middle map is an isomorphism, which proves that M is de Rham.

Now assume the claim is true for smooth manifolds admitting a de Rham cover with $k \geq 2$ sets, and suppose $\{U_1, \ldots, U_{k+1}\}$ is a de Rham cover of M. Put $U = U_1 \cup \cdots \cup U_k$ and $V = U_{k+1}$. The hypothesis implies that U and V are de Rham, and so is $U \cap V$ because it has a k-fold de Rham cover given by $\{U_1 \cap U_{k+1}, \ldots, U_k \cap U_{k+1}\}$. Therefore, $M = U \cup V$ is also de Rham by the argument above.

Step 4: *If M has a de Rham basis, then M is de Rham.* Suppose $\{U_\alpha\}$ is a de Rham basis for M. Let $f\colon M \to \mathbb{R}$ be an exhaustion function (see Proposition 2.28). For each integer m, define subsets A_m and A'_m of M by

$$A_m = \{q \in M : m \leq f(q) \leq m+1\},$$
$$A'_m = \{q \in M : m - \tfrac{1}{2} < f(q) < m + \tfrac{3}{2}\}.$$

(See Figure 16.8.) For each point $q \in A_m$, there is a basis open set containing q and contained in A'_m. The collection of all such basis sets is an open cover of A_m. Since f is an exhaustion function, A_m is compact, and therefore it is covered by finitely many of these basis sets. Let B_m be the union of this finite collection of sets. This is a finite de Rham cover of B_m, so by Step 3, B_m is de Rham.

Observe that $B_m \subset A'_m$, so B_m can have nonempty intersection with $B_{\widetilde{m}}$ only when $\widetilde{m} = m-1$, m, or $m+1$. Therefore, if we define

$$U = \bigcup_{m \text{ odd}} B_m, \qquad V = \bigcup_{m \text{ even}} B_m,$$

then U and V are disjoint unions of de Rham manifolds, and so they are both de Rham by Step 1. Finally, $U \cap V$ is de Rham because it is the disjoint union of the sets $B_m \cap B_{m+1}$ for m even, each of which has a finite de Rham cover consisting of sets of the form $U_\alpha \cap U_\beta$, where U_α and U_β are basis sets used to define B_m and B_{m+1}, respectively. Thus $M = U \cup V$ is de Rham by Step 3.

STEP 5: *Any open subset of* $\mathbb{R}^n$ *is de Rham.* If $U \subset \mathbb{R}^n$ is such a subset, then U has a basis consisting of Euclidean balls. Because each ball is convex, it is de Rham, and because any finite intersection of balls is again convex, finite intersections are also de Rham. Thus U has a de Rham basis, so it is de Rham by Step 4.

STEP 6: *Every smooth manifold is de Rham.* Any smooth manifold has a basis of smooth coordinate domains. Since every smooth coordinate domain is diffeomorphic to an open subset of $\mathbb{R}^n$, as are their finite intersections, this is a de Rham basis. The claim therefore follows from Step 4. □

This result expresses a deep connection between the topological and analytic properties of a smooth manifold, and plays a central role in differential geometry. If one has some information about the topology of a manifold M, the de Rham theorem can be used to draw conclusions about solutions to differential equations such as $d\eta = \omega$ on M. Conversely, if one can prove that such solutions do or do not exist, then one can draw conclusions about the topology.

As befits so fundamental a theorem, the de Rham theorem has many and varied proofs. The elegant proof given here is due to Glen E. Bredon [Bre93]. Another common approach is via the theory of sheaves; for example, a proof using this technique can be found in [War83]. The sheaf-theoretic proof is extremely powerful and lends itself to countless generalizations, but it has two significant disadvantages for our purposes: It requires the entire technical apparatus of sheaf theory and sheaf cohomology, which would take us too far afield; and, although it produces an isomorphism between de Rham and singular cohomology, it does not make it easy to see that the isomorphism is given specifically by integration. Nonetheless, because the technique leads to other important applications in such fields as differential geometry, algebraic topology, and complex analysis, it is worth taking some time and effort to study it if you get the opportunity.

Problems

16-1. Let M be an orientable smooth manifold and suppose ω is a closed p-form on M.

(a) Show that ω is exact if and only if the integral of ω over every smooth p-cycle is zero.

(b) Now suppose that $H_p(M)$ is generated by the homology classes of finitely many smooth p-cycles $\{c_1, \dots, c_m\}$. The numbers

$P_1(\omega), \ldots, P_m(\omega)$ defined by

$$P_i(\omega) = \int_{c_i} \omega$$

are called the *periods* of ω with respect to this set of generators. Show that ω is exact if and only if all of its periods are zero. [Compare this to Problem 6-14.]

16-2. Let M be a smooth n-manifold and suppose $S \subset M$ is an compact, oriented, embedded p-dimensional submanifold. A *smooth triangulation* of S is a smooth p-cycle $c = \sum_i \sigma_i$ in M with the following properties:

- Each $\sigma_i \colon \Delta_p \to S$ is a smooth orientation-preserving embedding.
- If $i \neq j$, then $\sigma_i(\operatorname{Int} \Delta_p) \cap \sigma_j(\operatorname{Int} \Delta_p) = \varnothing$.
- $S = \bigcup_i \sigma_i(\Delta_p)$.

(It can be shown that every compact smooth orientable submanifold admits a smooth triangulation, but we will not use that fact.) Two p-dimensional submanifolds $S, S' \subset M$ are said to be *homologous* if there exist smooth triangulations c for S and c' for S' such that $c - c'$ is a boundary.

(a) If c is a smooth triangulation of S and ω is any smooth p-form on M, show that $\int_c \omega = \int_S \omega$.
(b) If ω is closed and S, S' are homologous, show that $\int_S \omega = \int_{S'} \omega$.

16-3. Suppose (M, g) is a Riemannian n-manifold. A smooth p-form ω on M is called a *calibration* if ω is closed and $\omega_q(X_1, \ldots, X_p) \leq 1$ whenever $(X_1, \ldots, X_p)$ are orthonormal vectors in some tangent space T_qM. An oriented embedded submanifold $S \subset M$ is said to be *calibrated* if there is a calibration ω such that $\omega|_S$ is the volume form for the induced Riemannian metric on S. If $S \subset M$ is a smoothly triangulated calibrated compact submanifold, show that the volume of S (with respect to the induced Riemannian metric) is less than or equal to that of any other submanifold homologous to S (see Problem 16-2). [Remark: Calibrations were invented in 1982 by Reese Harvey and Blaine Lawson [HL82]; they have become increasingly important in recent years because in many situations a calibration is the only known way of proving that a given submanifold is volume minimizing in its homology class.]

16-4. Let $D \subset \mathbb{R}^3$ be the torus of revolution obtained by revolving the circle $(y-2)^2 + z^2 = 1$ around the z-axis, with the induced Riemannian metric (see Example 11.23), and let $C \subset D$ be the "inner circle" defined by $C = \{(x, y, z) : z = 0,\ x^2 + y^2 = 1\}$. Show that C is calibrated, and therefore has the shortest length in its homology class.

16-5. For any smooth manifold M, let $H_c^p(M)$ denote the pth compactly supported de Rham cohomology group of M (see page 399).

(a) If U is an open subset of M and $\iota\colon U \hookrightarrow M$ is the inclusion map, define a linear map $\iota_\#\colon \mathcal{A}_c^p(U) \to \mathcal{A}_c^p(M)$ by extending each compactly supported form to be zero on $M \smallsetminus U$. Show that $\partial \circ \iota_\# = \iota_\# \circ \partial$, and so $\iota_\#$ induces a linear map on compactly supported cohomology, denoted by $\iota_*\colon H_c^p(U) \to H_c^p(M)$.

(b) Mayer–Vietoris Theorem for Compactly Supported Cohomology: If $U, V \subset M$ are open subsets whose union is M, prove that for each p there is a linear map $\delta_*\colon H_c^p(M) \to H_c^{p+1}(U \cap V)$ such that the following sequence is exact:

$$\cdots \xrightarrow{\delta_*} H_c^p(U \cap V) \xrightarrow{i_* \oplus (-j_*)} H_c^p(U) \oplus H_c^p(V) \xrightarrow{k_* + l_*}$$
$$H_c^p(M) \xrightarrow{\delta_*} H_c^{p+1}(U \cap V) \xrightarrow{i_* \oplus (-j_*)} \cdots,$$

where i, j, k, l are the inclusion maps as in (15.9).

(c) Denoting the dual space to $H_c^p(M)$ by $H_c^p(M)^*$, show that the following sequence is also exact:

$$\cdots \xrightarrow{(\delta_*)^*} H_c^p(M)^* \xrightarrow{(k_*)^* \oplus (l_*)^*} H_c^p(U)^* \oplus H_c^p(V)^* \xrightarrow{(i_*)^* - (j_*)^*}$$
$$H_c^p(U \cap V)^* \xrightarrow{(\delta_*)^*} H_c^{p-1}(M)^* \xrightarrow{(k_*)^* \oplus (l_*)^*} \cdots. \quad (16.16)$$

16-6. The Poincaré Duality Theorem: If M is a smooth, oriented n-manifold, define a map $\mathrm{PD}\colon \mathcal{A}^p(M) \to \mathcal{A}_c^{n-p}(M)^*$ by

$$\mathrm{PD}(\omega)(\eta) = \int_M \omega \wedge \eta.$$

(a) Show that PD descends to a linear map (still denoted by the same symbol) $\mathrm{PD}\colon H^p(M) \to H_c^{n-p}(M)^*$.

(b) Show that PD is an isomorphism for each p. [Hint: Imitate the proof of the de Rham theorem, with "de Rham manifold" replaced by "Poincaré duality manifold." You will need the results of Problems 15-13 and 16-5.]

(c) If M is a compact, orientable smooth n-manifold, show that $\dim H_{dR}^p(M) = \dim H_{dR}^{n-p}(M)$ for each p.

16-7. Let M be a smooth n-manifold all of whose de Rham groups are finite-dimensional. (It can be shown that this is always the case when M is compact.) The *Euler characteristic* of M is the number

$$\chi(M) = \sum_{p=0}^{n} (-1)^p \dim H_{dR}^p(M).$$

Show that $\chi(M)$ is a homotopy invariant of M, and $\chi(M) = 0$ when M is compact, orientable, and odd-dimensional.

17
Integral Curves and Flows

In this chapter, we return to the study of vector fields. The primary geometric objects associated with smooth vector fields are their "integral curves," which are smooth curves whose tangent vector at each point is equal to the value of the vector field there. The collection of all integral curves of a given vector field on a manifold determines a family of diffeomorphisms of (open subsets of) the manifold, called a "flow." Any smooth $\mathbb{R}$-action is a flow, for example; but we will see that there are flows that are not $\mathbb{R}$-actions because the diffeomorphisms may not be defined on the whole manifold for all $t \in \mathbb{R}$.

The main theorem of the chapter, the "fundamental theorem on flows," asserts that every smooth vector field determines a unique maximal integral curve starting at each point, and the collection of all such integral curves determines a unique maximal flow. The proof is an application of the existence, uniqueness, and smoothness theorem for solutions of ordinary differential equations.

After proving the fundamental theorem, we explore some of the properties of vector fields and flows. First we investigate conditions under which a vector field generates a global flow. Then we examine the local behavior of flows, and find that the behavior at points where the vector field vanishes, which correspond to "equilibrium points" of the flow, is very different from the behavior at points where it does not vanish, where the flow looks locally like translation along parallel coordinate lines.

At the end of the chapter, we give a proof of the existence, uniqueness, and smoothness theorem for ordinary differential equations, which is the main ingredient in the proof of the fundamental theorem.

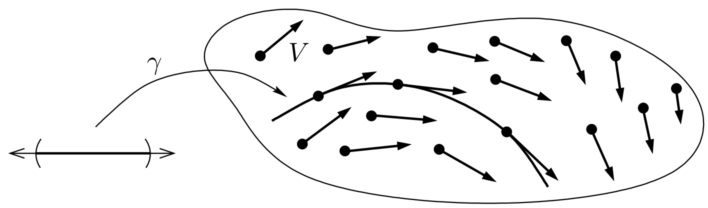


Figure 17.1. An integral curve of a vector field.

Integral Curves

If M is a smooth manifold and $J \subset \mathbb{R}$ is an open interval, a smooth curve $\gamma\colon J \to M$ determines a tangent vector $\gamma'(t) \in T_{\gamma(t)}M$ at each point of the curve. In this section we describe a way to work backwards: Given a tangent vector at each point, we seek a curve that has those tangent vectors.

If V is a smooth vector field on M, an *integral curve* of V is a smooth curve $\gamma\colon J \to M$ such that

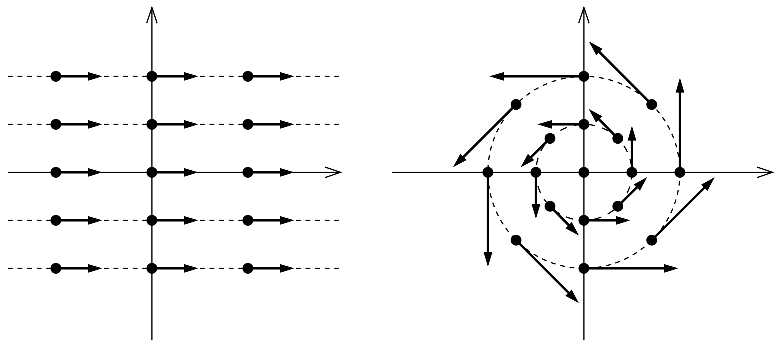
$$\gamma'(t) = V_{\gamma(t)} \quad \text{for all } t \in J.$$

In other words, the tangent vector to γ at each point is equal to the value of V at that point (Figure 17.1). If $0 \in J$, the point $p = \gamma(0)$ is called the *starting point* of γ. (The reason for the term "integral curve" will be explained shortly. Note that this is one definition that requires some differentiability hypothesis, because the definition of an integral curve would make no sense for a curve that is merely continuous. As we will see below, integral curves of smooth vector fields are always smooth, so we lose no generality by including smoothness as part of the definition.)

Example 17.1 (Integral Curves).

(a) Let $V = \partial/\partial x$ be the first coordinate vector field on $\mathbb{R}^2$ (Figure 17.2(a)). It is easy to check that the integral curves of V are precisely the straight lines parallel to the x-axis, with parametrizations of the form $\gamma(t) = (a + t, b)$ for constants a and b. Thus there is a unique integral curve starting at each point of the plane, and the images of different integral curves are either identical or disjoint.

(b) Let $W = x\,\partial/\partial y - y\,\partial/\partial x$ on $\mathbb{R}^2$ (Figure 17.2(b)). If $\gamma\colon \mathbb{R} \to \mathbb{R}^2$ is a smooth curve, written in standard coordinates as $\gamma(t) = (x(t), y(t))$, then the condition $\gamma'(t) = W_{\gamma(t)}$ for γ to be an integral curve



(a) $V = \partial/\partial x$. (b) $W = x\,\partial/\partial y - y\,\partial/\partial x$.

Figure 17.2. Vector fields and their integral curves in the plane.

translates to

$$x'(t)\left.\frac{\partial}{\partial x}\right|_{(x(t),y(t))} + y'(t)\left.\frac{\partial}{\partial y}\right|_{(x(t),y(t))} = x(t)\left.\frac{\partial}{\partial y}\right|_{(x(t),y(t))} - y(t)\left.\frac{\partial}{\partial x}\right|_{(x(t),y(t))}.$$

Comparing the components of these vectors, this is equivalent to the system of ordinary differential equations

$$\begin{aligned} x'(t) &= -y(t), \\ y'(t) &= x(t). \end{aligned}$$

These equations have the solutions

$$\begin{aligned} x(t) &= a\cos t - b\sin t, \\ y(t) &= a\sin t + b\cos t \end{aligned}$$

for arbitrary constants a and b, and thus each curve of the form $\gamma(t) = (a\cos t - b\sin t, a\sin t + b\cos t)$ is an integral curve of W. When $(a,b) = (0,0)$, this is the constant curve $\gamma(t) \equiv (0,0)$; otherwise, it is a circle traversed counterclockwise. Since $\gamma(0) = (a,b)$, we see once again that there is a unique integral curve starting at each point $(a,b) \in \mathbb{R}^2$, and the images of the various integral curves are either identical or disjoint.

As the second example above illustrates, finding integral curves boils down to solving a system of ordinary differential equations in a smooth chart. More generally, let V be a smooth vector field on M and let $\gamma\colon J \to M$ be any smooth curve. Writing γ in smooth local coordinates as $\gamma(t) =$

$(\gamma^1(t), \dots, \gamma^n(t))$, the condition $\gamma'(t) = V_{\gamma(t)}$ for γ to be an integral curve of V can be written on a smooth coordinate domain $U \subset M$ as

$$(\gamma^i)'(t) \left.\frac{\partial}{\partial x^i}\right|_{\gamma(t)} = V^i(\gamma(t)) \left.\frac{\partial}{\partial x^i}\right|_{\gamma(t)},$$

which reduces to the system of ordinary differential equations (ODEs)

$$\begin{aligned}(\gamma^1)'(t) &= V^1(\gamma^1(t), \dots, \gamma^n(t)),\\ &\ \,\vdots\\ (\gamma^n)'(t) &= V^n(\gamma^1(t), \dots, \gamma^n(t)),\end{aligned}$$

where the component functions V^i are smooth on U. The fundamental fact about such systems, which we will state precisely and prove later in the chapter, is that there is a unique solution, at least for t in a small time interval $(-\varepsilon, \varepsilon)$, satisfying any initial condition of the form $(\gamma^1(0), \dots, \gamma^n(0)) = (a^1, \dots, a^n)$ for $(a^1, \dots, a^n) \in U$; and the solution depends smoothly on both t and a. (This is the reason for the terminology "integral curves," because solving a system of ODEs is often referred to as "integrating" the system.) For now, we just note that this implies there is a unique integral curve, at least for a short time, starting at any point in the manifold. Moreover, we will see that up to reparametrization, there is a unique integral curve passing through each point.

The following simple lemma shows how an integral curve can be reparametrized to change its starting point.

Lemma 17.2 (Translation Lemma). *Let V be a smooth vector field on a smooth manifold M, let $J \subset \mathbb{R}$ be an open interval, and let $\gamma\colon J \to M$ be an integral curve of V. For any $a \in \mathbb{R}$, let $J + a$ be the interval*

$$J + a = \{t + a : t \in J\}. \tag{17.1}$$

Then the curve $\widetilde{\gamma}\colon J + a \to M$ defined by $\widetilde{\gamma}(t) = \gamma(t - a)$ is an integral curve of V.

Proof. One way to see this is as a straightforward application of the chain rule in local coordinates. Somewhat more invariantly, we can examine the action of $\widetilde{\gamma}'(t)$ on a smooth real-valued function f defined in a neighborhood of a point $\widetilde{\gamma}(t_0)$. By the chain rule and the fact that γ is an integral curve,

$$\begin{aligned}\widetilde{\gamma}'(t_0)f &= \left.\frac{d}{dt}\right|_{t=t_0} (f \circ \widetilde{\gamma})(t)\\ &= \left.\frac{d}{dt}\right|_{t=t_0} (f \circ \gamma)(t - a)\\ &= (f \circ \gamma)'(t_0 - a)\\ &= \gamma'(t_0 - a)f = V_{\gamma(t_0 - a)}f = V_{\widetilde{\gamma}(t_0)}f.\end{aligned}$$

Thus $\widetilde{\gamma}$ is an integral curve of V. $\square$

Global Flows

Here is another way to visualize the family of integral curves associated with a vector field. Let V be a smooth vector field on a smooth manifold M, and suppose it has the property that for each point $p \in M$ there is a unique integral curve $\theta^{(p)}\colon \mathbb{R} \to M$ starting at p. (It may not always be the case that all of the integral curves are defined for all $t \in \mathbb{R}$, but for purposes of illustration let us assume for the time being that they are.) For each $t \in \mathbb{R}$, we can define a map θ_t from M to itself by sending each point $p \in M$ to the point obtained by following the integral curve starting at p for time t:

$$\theta_t(p) = \theta^{(p)}(t).$$

This defines a family of maps $\theta_t\colon M \to M$ for $t \in \mathbb{R}$. Each such map "slides" the entire manifold along the integral curves for time t. If we set $q = \theta^{(p)}(s)$, the translation lemma implies that $t \mapsto \theta^{(p)}(t+s)$ is an integral curve starting at q; since we are assuming uniqueness of integral curves, we must have $\theta^{(q)}(t) = \theta^{(p)}(t+s)$. When we translate this equality into a statement about the maps θ_t, it becomes

$$\theta_t \circ \theta_s(p) = \theta_{t+s}(p).$$

Together with the equation $\theta_0(p) = \theta^{(p)}(0) = p$, which holds by definition, this implies that the map $\theta\colon \mathbb{R} \times M \to M$ is an action of the additive group $\mathbb{R}$ on M.

Motivated by these considerations, we define a *global flow* on M (sometimes also called a *one-parameter group action*) to be a left action of $\mathbb{R}$ on M; that is, a continuous map $\theta\colon \mathbb{R} \times M \to M$ satisfying the following properties for all $s, t \in \mathbb{R}$ and all $p \in M$:

$$\begin{aligned} \theta(t, \theta(s,p)) &= \theta(t+s, p), \\ \theta(0,p) &= p. \end{aligned} \tag{17.2}$$

Given a global flow θ on M, we define two collections of maps as follows.

- For each $t \in \mathbb{R}$, define $\theta_t\colon M \to M$ by

  $$\theta_t(p) = \theta(t, p).$$

  The defining properties (17.2) are equivalent to the *group laws*:

  $$\begin{aligned} \theta_t \circ \theta_s &= \theta_{t+s}, \\ \theta_0 &= \mathrm{Id}_M . \end{aligned} \tag{17.3}$$

  As is the case for any group action, each map $\theta_t\colon M \to M$ is a homeomorphism, and if the action is smooth, θ_t is a diffeomorphism.

- For each $p \in M$, define a curve $\theta^{(p)}\colon \mathbb{R} \to M$ by

  $$\theta^{(p)}(t) = \theta(t, p).$$

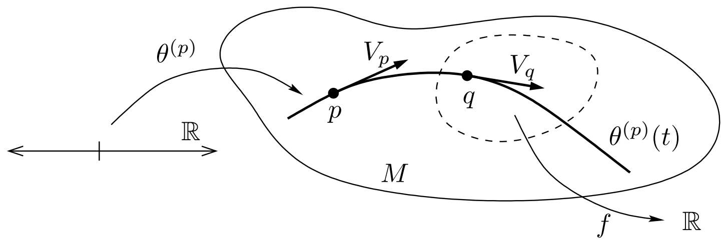


Figure 17.3. The infinitesimal generator of a global flow.

The image of this curve is just the orbit of p under the group action. Because any group action partitions the manifold into disjoint orbits, it follows that M is the disjoint union of the images of these curves.

The next proposition shows that every smooth global flow is derived from the integral curves of some vector field in precisely the way we described above. If $\theta\colon \mathbb{R} \times M \to M$ is a smooth global flow, for each $p \in M$ we define a tangent vector $V_p \in T_pM$ by

$$V_p = \theta^{(p)\prime}(0).$$

The assignment $p \mapsto V_p$ is a (rough) vector field on M, which is called the *infinitesimal generator* of θ, for reasons we will explain below.

Proposition 17.3. *Let $\theta\colon \mathbb{R} \times M \to M$ be a smooth global flow. The infinitesimal generator V of θ is a smooth vector field on M, and each curve $\theta^{(p)}$ is an integral curve of V.*

Proof. To show that V is smooth, it suffices by Lemma 4.2 to show that Vf is smooth for any smooth real-valued function f defined on an open subset $U \subset M$. For any such f and any $p \in U$, just note that

$$Vf(p) = V_pf = \theta^{(p)\prime}(0)f = \left.\frac{d}{dt}\right|_{t=0} f(\theta^{(p)}(t)) = \left.\frac{\partial}{\partial t}\right|_{(0,p)} f(\theta(t,p)).$$

Because $f(\theta(t,p))$ is a smooth function of (t,p) by composition, so is its partial derivative with respect to t. (You can interpret this partial derivative as the action of the smooth vector field $\partial/\partial t$ on the smooth function $f \circ \theta\colon \mathbb{R} \times M \to \mathbb{R}$.) It follows that $Vf(p)$ depends smoothly on p, so V is smooth.

To show that $\theta^{(p)}$ is an integral curve of V, we need to show that

$$\theta^{(p)\prime}(t) = V_{\theta^{(p)}(t)}$$

for all $p \in M$ and all $t \in \mathbb{R}$. Let $t_0 \in \mathbb{R}$ be arbitrary, and set $q = \theta^{(p)}(t_0) = \theta_{t_0}(p)$, so that what we have to show is $\theta^{(p)\prime}(t_0) = V_q$ (see Figure 17.3). By the group law, for all t,

$$\begin{aligned}\theta^{(q)}(t) &= \theta_t(q) \\ &= \theta_t(\theta_{t_0}(p)) \\ &= \theta_{t+t_0}(p) \\ &= \theta^{(p)}(t+t_0).\end{aligned} \tag{17.4}$$

Therefore, for any smooth real-valued function f defined in a neighborhood of q,

$$\begin{aligned}V_q f &= \theta^{(q)\prime}(0) f \\ &= \left.\frac{d}{dt}\right|_{t=0} f(\theta^{(q)}(t)) \\ &= \left.\frac{d}{dt}\right|_{t=0} f(\theta^{(p)}(t+t_0)) \\ &= \theta^{(p)\prime}(t_0) f,\end{aligned} \tag{17.5}$$

which was to be shown. □

Example 17.4 (Global Flows). The two vector fields on the plane described in Example 17.1 both had integral curves defined for all $t \in \mathbb{R}$, so they generate global flows. Using the results of that example, we can write down the flows explicitly.

(a) The flow of $V = \partial/\partial x$ in $\mathbb{R}^2$ is the map $\tau\colon \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$\tau_t(x, y) = (x + t, y).$$

For each $t \in \mathbb{R}$, τ_t translates the plane to the right ($t > 0$) or left ($t < 0$) by a distance $|t|$.

(b) The flow of $W = x\,\partial/\partial y - y\,\partial/\partial x$ is the map $\theta\colon \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$ given by

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

For each $t \in \mathbb{R}$, θ_t rotates the plane through an angle t.

The Fundamental Theorem on Flows

We have seen that every smooth global flow gives rise to a smooth vector field whose integral curves are precisely the curves defined by the flow. Conversely, we would like to be able to say that every smooth vector field is the infinitesimal generator of a global flow. However, it is easy to see that this cannot be the case, because there are smooth vector fields whose integral curves are not defined for all $t \in \mathbb{R}$. Here are two examples.

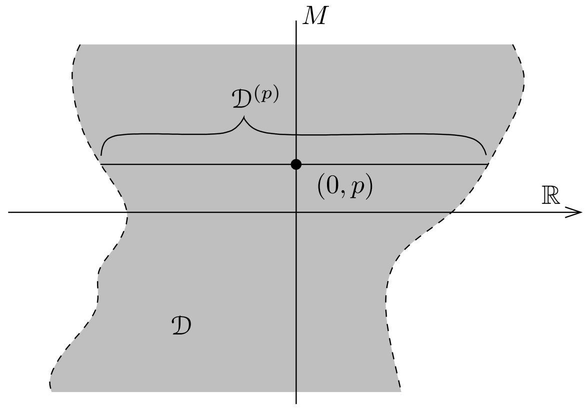


Figure 17.4. A flow domain.

Example 17.5. Let $M = \{(x,y) \in \mathbb{R}^2 : x < 0\}$, and let $V = \partial/\partial x$. The unique integral curve of V starting at $(-1,0) \in M$ is $\gamma(t) = (t-1,0)$. However, in this case, γ cannot be extended past $t = 1$.

Example 17.6. For a somewhat more subtle example, let M be all of $\mathbb{R}^2$ and let $W = x^2\partial/\partial x$. You can check easily that the unique integral curve of W starting at $(1,0)$ is

$$\gamma(t) = \left(\frac{1}{1-t}, 0\right).$$

This curve also cannot be extended past $t = 1$, because it escapes to infinity as $t \nearrow 1$.

For this reason, we make the following definitions. If M is a manifold, a *flow domain* for M is an open subset $\mathcal{D} \subset \mathbb{R} \times M$ with the property that for each $p \in M$, the set $\mathcal{D}^{(p)} = \{t \in \mathbb{R} : (t,p) \in \mathcal{D}\}$ is an open interval containing 0 (Figure 17.4). A *flow* on M is a continuous map $\theta\colon \mathcal{D} \to M$, where $\mathcal{D} \subset \mathbb{R} \times M$ is a flow domain, that satisfies the following group laws: For all $p \in M$,

$$\theta(0,p) = p, \tag{17.6}$$

and for all $s \in \mathcal{D}^{(p)}$ and $t \in \mathcal{D}^{(\theta(s,p))}$ such that $s + t \in \mathcal{D}^{(p)}$,

$$\theta(t, \theta(s,p)) = \theta(t+s,p). \tag{17.7}$$

We sometimes call θ a *local flow* to distinguish it from a global flow as defined earlier. The unwieldy term *local one-parameter group action* is also commonly used.

If θ is a flow, we define $\theta_t(p) = \theta^{(p)}(t) = \theta(t,p)$ whenever $(t,p) \in \mathcal{D}$, just as for a global flow. Similarly, if θ is smooth, the *infinitesimal generator* of θ is defined by $V_p = \theta^{(p)\prime}(0)$.

Proposition 17.7. *If $\theta\colon \mathcal{D} \to M$ is a smooth flow, then the infinitesimal generator V of θ is a smooth vector field, and each curve $\theta^{(p)}$ is an integral curve of V.*

Proof. The proof is essentially identical to the proof of the analogous result for global flows, Proposition 17.3. In the proof that V is smooth, we need only note that for any $p_0 \in M$, $\theta(t,p)$ is defined and smooth for all (t,p) sufficiently close to $(0,p_0)$ because $\mathcal{D}$ is open. In the proof that $\theta^{(p)}$ is an integral curve, we need to verify that all of the expressions in (17.4) and (17.5) make sense. Suppose $t_0 \in \mathcal{D}^{(p)}$. Because both $\mathcal{D}^{(p)}$ and $\mathcal{D}^{(\theta_{t_0}(p))}$ are open intervals containing 0, there is a positive number ε such that $t + t_0 \in \mathcal{D}^{(p)}$ and $t \in \mathcal{D}^{(\theta_{t_0}(p))}$ whenever $|t| < \varepsilon$, and then $\theta_t(\theta_{t_0}(p)) = \theta_{t+t_0}(p)$ by definition of a flow. The rest of the proof goes through just as before. $\square$

The main result of this section is that every smooth vector field gives rise to a smooth flow, which is unique if we require it to be *maximal*, which means that it admits no extension to a flow on a larger flow domain.

Theorem 17.8 (Fundamental Theorem on Flows). *Let V be a smooth vector field on a smooth manifold M. There is a unique maximal smooth flow $\theta\colon \mathcal{D} \to M$ whose infinitesimal generator is V. This flow has the following properties:*

(*a*) *For each $p \in M$, the curve $\theta^{(p)}\colon \mathcal{D}^{(p)} \to M$ is the unique maximal integral curve of V starting at p.*

(*b*) *If $s \in \mathcal{D}^{(p)}$, then $\mathcal{D}^{(\theta(s,p))}$ is the interval $\mathcal{D}^{(p)} - s = \{t - s : t \in \mathcal{D}^{(p)}\}$.*

(*c*) *For each $t \in \mathbb{R}$, the set $M_t = \{p \in M : (t,p) \in \mathcal{D}\}$ is open in M, and $\theta_t\colon M_t \to M_{-t}$ is a diffeomorphism with inverse θ_{-t}.*

(*d*) *For each $(t,p) \in \mathcal{D}$, $(\theta_t)_* V_p = V_{\theta_t(p)}$.*

We will prove the theorem below. First, we make a few comments on what it means.

The flow whose existence and uniqueness are asserted in this theorem is called the *flow generated by V*, or just the *flow of V*. The maximality assertion in (a) means that $\theta^{(p)}$ cannot be extended to an integral curve on any larger open interval. If θ is a smooth flow on M, a vector field W is said to be *invariant under θ* if $(\theta_t)_* W_p = W_{\theta_t(p)}$ for all (t,p) in the domain of θ (or, to put it another way, if W is θ_t-related to itself for each t). Part (d) of the fundamental theorem can be summarized by saying that every smooth vector field is invariant under its own flow.

The term "infinitesimal generator" comes from the following picture. In a smooth chart, a good approximation to an integral curve can be obtained by composing very many small straight-line motions, with the direction and length of each successive motion determined by the value of the vector field at the point arrived at in the previous step. Intuitively, one can think of a flow as being composed of infinitely many infinitesimally small linear steps.

As we saw earlier in this chapter, finding integral curves of V (and therefore finding the flow generated by V) boils down to solving a system of ordinary differential equations, at least locally. A key ingredient in the proof will be the fact that solutions to such systems exist at least locally, are unique, and depend smoothly on both the time variable and the initial conditions. Thus before beginning the proof, we state the following basic theorem about solutions to systems of ODEs. We will give the proof of this theorem in the last section of the chapter.

Theorem 17.9 (ODE Existence, Uniqueness, and Smoothness). *Let $U \subset \mathbb{R}^n$ be open, and let $V : U \to \mathbb{R}^n$ be a smooth map. For $t_0 \in \mathbb{R}$ and $x \in U$, consider the following initial-value problem:*

$$\begin{aligned}(\gamma^i)'(t) &= V^i(\gamma(t)),\\ \gamma^i(t_0) &= x^i.\end{aligned} \tag{17.8}$$

(*a*) Existence: *For any $t_0 \in \mathbb{R}$ and $x_0 \in U$, there exist an open interval J_0 containing t_0 and an open set $U_0 \subset U$ containing x_0 such that for each $x \in U_0$, there is a smooth curve $\gamma\colon J_0 \to U$ that solves* (17.8)*.*

(*b*) Uniqueness: *Any two differentiable solutions to* (17.8) *agree on their common domain.*

(*c*) Smoothness: *Let t_0, x_0, J_0, and U_0 be as in* (a)*, and define a map $\theta\colon J_0 \times U_0 \to U$ by letting $\theta(t,x) = \gamma(t)$, where $\gamma\colon J_0 \to U$ is the unique solution to* (17.8) *with initial condition x. Then θ is smooth.*

Using this result, we now prove the fundamental theorem on flows.

Proof of Theorem 17.8. We begin by noting that the existence assertion of the ODE theorem implies that there exists an integral curve starting at each point $p \in M$, because the equation for an integral curve is a smooth system of ODEs in any smooth local coordinates around p.

Now suppose $\gamma, \widetilde{\gamma}\colon J \to M$ are two integral curves of V defined on the same open interval J such that $\gamma(t_0) = \widetilde{\gamma}(t_0)$ for some $t_0 \in J$. Let $\mathcal{S}$ be the set of $t \in J$ such that $\gamma(t) = \widetilde{\gamma}(t)$. Clearly $\mathcal{S} \neq \varnothing$ because $t_0 \in \mathcal{S}$ by hypothesis, and it is closed in J by continuity. On the other hand, suppose $t_1 \in \mathcal{S}$. Then in a smooth coordinate neighborhood around the point $p = \gamma(t_1)$, γ and $\widetilde{\gamma}$ are both solutions to same ODE with the same initial condition $\gamma(t_1) = \widetilde{\gamma}(t_1) = p$. By the uniqueness part of the ODE theorem, $\gamma \equiv \widetilde{\gamma}$ on an interval containing t_1, which implies that $\mathcal{S}$ is open in J. Since J is connected, $\mathcal{S} = J$, which implies that $\gamma = \widetilde{\gamma}$ on all of J. Thus any two integral curves that agree at one point agree on their common domain.

For each $p \in M$, let $\mathcal{D}^{(p)}$ be the union of all open intervals $J \subset \mathbb{R}$ containing 0 on which an integral curve starting at p is defined. Define $\theta^{(p)}\colon \mathcal{D}^{(p)} \to M$ by letting $\theta^{(p)}(t) = \gamma(t)$, where γ is any integral curve starting at p and defined on an open interval containing 0 and t. Since all such integral curves agree at t by the argument above, $\theta^{(p)}$ is well-defined, and is obviously the unique maximal integral curve starting at p.

Now, let $\mathcal{D} = \{(t,p) \in \mathbb{R} \times M : t \in \mathcal{D}^{(p)}\}$, and define $\theta\colon \mathcal{D} \to M$ by $\theta(t,p) = \theta^{(p)}(t)$. As usual, we also write $\theta_t(p) = \theta(t,p)$. By definition, θ satisfies property (a) in the statement of the fundamental theorem: For each $p \in M$, $\theta^{(p)}$ is the unique maximal integral curve of V starting at p. To verify (b) (the fact that $\mathcal{D}^{(\theta(s,p))} = \mathcal{D}^{(p)} - s$), fix any $p \in M$ and $s \in \mathcal{D}^{(p)}$, and write $q = \theta(s,p) = \theta^{(p)}(s)$. The curve $\gamma\colon \mathcal{D}^{(p)} - s \to M$ defined by

$$\gamma(t) = \theta^{(p)}(t+s)$$

satisfies $\gamma(0) = q$, and the translation lemma shows that γ is an integral curve of V. By maximality of $\theta^{(q)}$, the domain of γ cannot be larger than $\mathcal{D}^{(q)}$, which means that $\mathcal{D}^{(p)} - s \subset \mathcal{D}^{(q)}$. Since $0 \in \mathcal{D}^{(p)}$, this implies in particular that $-s \in \mathcal{D}^{(q)}$. Applying the same argument with $(-s,q)$ in place of (s,p), we find that $\mathcal{D}^{(q)} + s \subset \mathcal{D}^{(p)}$, which is the same as $\mathcal{D}^{(q)} \subset \mathcal{D}^{(p)} - s$. This proves (b). Moreover, the uniqueness of ODE solutions implies that the curve γ defined above agrees with $\theta^{(q)}$, which is equivalent to the second group law (17.7), and the first group law (17.6) is immediate from the definition.

Next we will show that $\mathcal{D}$ is open in $\mathbb{R} \times M$ (so it is a flow domain), and that $\theta\colon \mathcal{D} \to M$ is smooth. Define a subset $W \subset \mathcal{D}$ as the set of all $(t,p) \in \mathcal{D}$ such that θ is defined and smooth on a product neighborhood of (t,p) of the form $J \times U \subset \mathcal{D}$, where $U \subset M$ is a neighborhood of p and $J \subset \mathbb{R}$ is an open interval containing 0 and t. Clearly W is open in $\mathbb{R} \times M$ and the restriction of θ to W is smooth, so it suffices to show that $W = \mathcal{D}$. Suppose this is not the case. Then there exists some point $(t_0, p_0) \in \mathcal{D} \smallsetminus W$. For simplicity, let us assume $t_0 > 0$; the argument for $t_0 < 0$ is similar.

Let $\tau = \sup\{t \in \mathbb{R} : (t, p_0) \in W\}$ (Figure 17.5). By the ODE theorem (applied in smooth coordinates around p_0), θ is defined and smooth in some product neighborhood of $(0, p_0)$, so $\tau > 0$. Since $\tau \le t_0$ and $\mathcal{D}^{(p_0)}$ is an open interval containing 0 and t_0, it follows that $\tau \in \mathcal{D}^{(p_0)}$. Let $q_0 = \theta^{(p_0)}(\tau)$. By the ODE theorem again, there exist $\varepsilon > 0$ and a neighborhood U_0 of q_0 such that $\theta\colon (-\varepsilon, \varepsilon) \times U_0 \to M$ is defined and smooth. We will use the group law to show that θ extends smoothly to a neighborhood of (τ, p_0), which contradicts our choice of τ.

Choose some $t_1 < \tau$ such that $t_1 + \varepsilon > \tau$ and $\theta^{(p_0)}(t_1) \in U_0$. Since $t_1 < \tau$, $(t_1, p_0) \in W$, and so there is a product neighborhood $(-\delta, t_1 + \delta) \times U_1$ of (t_1, p_0) on which θ is defined and smooth. Because $\theta(t_1, p_0) \in U_0$, we can choose U_1 small enough that θ maps $\{t_1\} \times U_1$ into U_0. Because θ satisfies

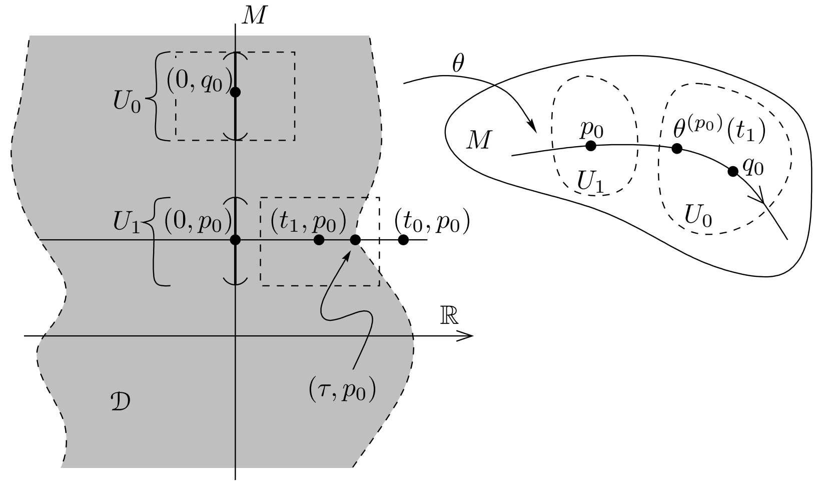


Figure 17.5. Proof that $\mathcal{D}$ is open.

the group law, we have

$$\theta_t(p) = \theta_{t-t_1} \circ \theta_{t_1}(p)$$

whenever $t, t_1 \in \mathcal{D}^{(p)}$. By our choice of t_1, $\theta_{t_1}(p)$ is defined for $p \in U_1$, and depends smoothly on p. Moreover, since $\theta_{t_1}(p) \in U_0$ for all such p, it follows that $\theta_{t-t_1} \circ \theta_{t_1}(p)$ is defined whenever $p \in U_1$ and $|t - t_1| < \varepsilon$, and depends smoothly on (t, p). This gives a smooth extension of θ to the product set $(-\delta, t_1 + \varepsilon) \times U_1$, which contradicts our choice of τ. This completes the proof that $W = \mathcal{D}$.

Next we prove (c): Each set M_t is open and $\theta_t \colon M_t \to M_{-t}$ is a diffeomorphism. The fact that M_t is open is an immediate consequence of the fact that $\mathcal{D}$ is open. From part (b) we deduce

$$\begin{aligned} p \in M_t &\implies t \in \mathcal{D}^{(p)} \\ &\implies \mathcal{D}^{(\theta_t(p))} = \mathcal{D}^{(p)} - t \\ &\implies -t \in \mathcal{D}^{(\theta_t(p))} \\ &\implies \theta_t(p) \in M_{-t}, \end{aligned}$$

which shows that θ_t maps M_t to M_{-t}. Moreover, the group laws then show that $\theta_{-t} \circ \theta_t$ is equal to the identity on M_t. Reversing the roles of t and $-t$ shows that $\theta_t \circ \theta_{-t}$ is the identity on M_{-t}, which completes the proof.

It remains only to prove (d), the fact that V is invariant under its own flow. Let $(t_0, p) \in \mathcal{D}$ be arbitrary, and set $q = \theta_{t_0}(p)$. We need to show that

$$(\theta_{t_0})_* V_p = V_q.$$

Applying the left-hand side to a smooth real-valued function f defined in a neighborhood of q and using the fact that $\theta^{(p)}$ is an integral curve of V, we obtain

$$\begin{aligned}
((\theta_{t_0})_* V_p) f &= V_p(f \circ \theta_{t_0}) \\
&= \left.\frac{d}{dt}\right|_{t=0} f \circ \theta_{t_0} \circ \theta^{(p)}(t) \\
&= \left.\frac{d}{dt}\right|_{t=0} f(\theta_{t_0}(\theta_t(p))) \\
&= \left.\frac{d}{dt}\right|_{t=0} f(\theta_{t_0+t}(p)) \\
&= \left.\frac{d}{dt}\right|_{t=0} f(\theta^{(p)}(t_0 + t)) \\
&= \theta^{(p)\prime}(t_0) f \\
&= V_q f. \qquad \square
\end{aligned}$$

Complete Vector Fields

As we noted above, not every smooth vector field generates a global flow. The ones that do are important enough to deserve a name. We say a smooth vector field is *complete* if it generates a global flow, or equivalently if each of its maximal integral curves is defined for all $t \in \mathbb{R}$. For example, both of the vector fields on the plane whose flows we computed in Example 17.4 are complete.

It is not always easy to determine by looking at a vector field whether it is complete or not. If you can solve the ODE explicitly to find all of the integral curves, and they all exist for all time, then the vector field is complete. On the other hand, if you can find a single integral curve that cannot be extended to all of $\mathbb{R}$, as we did for the vector fields of Examples 17.5 and 17.6, then it is not complete. However, it is often impossible to solve the ODE explicitly, so it is useful to have some general criteria for determining when a vector field is complete.

In this section we will show that all vector fields on a compact manifold are complete. (Problem 17-1 gives a more general sufficient condition.) The proof will be based on the following lemma.

Lemma 17.10 (Escape Lemma). *Let V be a smooth vector field on a smooth manifold M. If γ is an integral curve of V whose maximal domain is not all of $\mathbb{R}$, then the image of γ cannot lie in any compact subset of M.*

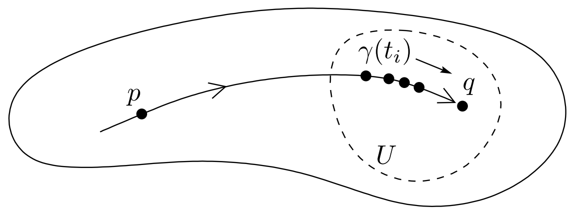


Figure 17.6. Proof of the escape lemma.

Proof. Let (a,b) denote the maximal domain of γ, so that $-\infty \le a < 0 < b \le +\infty$. Let $p = \gamma(0)$, and let θ denote the flow of V, so $\gamma = \theta^{(p)}$ by the uniqueness of integral curves.

Assume that $b < +\infty$ but $\gamma(a,b)$ lies in a compact set $K \subset M$. We will show that γ can be extended past b, which contradicts the maximality of (a,b). (The argument for the case $a > -\infty$ is similar.) If $\{t_i\}$ is any sequence of times approaching b from below, then the sequence $\{\gamma(t_i)\}$ lies in K and therefore has a subsequence converging to a point $q \in M$ (Figure 17.6). Choose a neighborhood U of q and a positive number ε such that θ is defined on $(-\varepsilon, \varepsilon) \times U$. Pick some i large enough that $\gamma(t_i) \in U$ and $t_i > b - \varepsilon$, and define $\sigma\colon (a, t_i + \varepsilon) \to M$ by

$$\sigma(t) = \begin{cases} \gamma(t), & a < t < b, \\ \theta_{t-t_i} \circ \theta_{t_i}(p), & t_i - \varepsilon < t < t_i + \varepsilon. \end{cases}$$

These two definitions agree where they overlap, because $\theta_{t-t_i} \circ \theta_{t_i}(p) = \theta_t(p) = \gamma(t)$ by the group law for θ. Thus σ is an integral curve extending γ, which contradicts the maximality of (a,b). □

Theorem 17.11. *Suppose M is a compact manifold. Then every smooth vector field on M is complete.*

Proof. If M is compact, the escape lemma implies that no integral curve can have a maximal domain that is not all of $\mathbb{R}$, because the image of any integral curve is contained in the compact set M. □

Regular Points and Singular Points

If V is a vector field on M, a point $p \in M$ is said to be a *singular point* for V if $V_p = 0$, and a *regular point* otherwise. The next lemma shows that the integral curves starting at regular and singular points behave very differently from each other.

Lemma 17.12. *Let V be a smooth vector field on a smooth manifold M, and let $\theta\colon \mathcal{D} \to M$ be the flow generated by V. If p is a singular point of V,* *then* $\mathcal{D}^{(p)} = \mathbb{R}$ *and* $\theta^{(p)}$ *is the constant curve* $\theta^{(p)}(t) \equiv p$. *If* p *is a regular point, then* $\theta^{(p)} \colon \mathcal{D}^{(p)} \to M$ *is an immersion.*

Proof. If $V_p = 0$, then the constant curve $\gamma \colon \mathbb{R} \to M$ given by $\gamma(t) \equiv p$ is clearly an integral curve of V, so by uniqueness and maximality it must be equal to $\theta^{(p)}$.

Suppose on the other hand that $V_p \neq 0$. For simplicity, write $\gamma = \theta^{(p)}$. To show that γ is an immersion, it suffices to show that $\gamma'(t) \neq 0$ for any $t \in \mathcal{D}^{(p)}$. Let $t_0 \in \mathcal{D}^{(p)}$ be arbitrary, and put $q = \gamma(t_0)$. Theorem 17.8(d) shows that $V_q = (\theta_{t_0})_* V_p$. Since θ_{t_0} is a diffeomorphism on some open set containing p, $(\theta_{t_0})_* \colon T_pM \to T_qM$ is an isomorphism, so it follows that

$$\gamma'(t_0) = V_q = (\theta_{t_0})_* V_p \neq 0. \qquad \square$$

If $\theta \colon \mathcal{D} \to M$ is a flow, a point $p \in M$ is called an *equilibrium point* for θ if $\theta(t,p) = p$ for all $t \in \mathcal{D}^{(p)}$. The preceding lemma showed that the equilibrium points of a smooth flow are precisely the singular points of its infinitesimal generator.

The next theorem shows that the local structure of a smooth vector field near a regular point is very simple.

Theorem 17.13 (Canonical Form Near a Regular Point). *Let* V *be a smooth vector field on a smooth manifold* M*, and let* $p \in M$ *be a regular point for* V*. There exist smooth coordinates* (u^i) *on some neighborhood of* p *in which* V *has the coordinate representation* $\partial/\partial u^1$*.*

Proof. By the way we have defined coordinate vector fields on a manifold, a smooth chart (U, φ) will satisfy the conclusion of the theorem provided that $(\varphi^{-1})_*(\partial/\partial u^1) = V$. For φ^{-1} to satisfy this, it must take lines parallel to the u^1-axis to integral curves of V. The flow of V is ideally suited to this purpose.

This is a local question, so by choosing smooth coordinates $(x^1, \dots, x^n)$ centered at p, we may replace M by an open subset $U \subset \mathbb{R}^n$, and think of V as a vector field on U. By reordering the coordinates if necessary, we may assume that V has nonzero x^1-component at p.

Let $\theta \colon \mathcal{D} \to U$ be the flow of V. There exist $\varepsilon > 0$ and a neighborhood $U_0 \subset U$ of p such that the product open set $(-\varepsilon, \varepsilon) \times U_0$ is contained in $\mathcal{D}$. Let $S_0 \subset U_0$ be the intersection of U_0 with the coordinate hyperplane where $x^1 = 0$, and define $S \subset \mathbb{R}^{n-1}$ by

$$S = \{(u^2, \dots, u^n) : (0, u^2, \dots, u^n) \in S_0\}.$$

(See Figure 17.7.) Define a smooth map $\psi \colon (-\varepsilon, \varepsilon) \times S \to U$ by

$$\psi(t, u^2, \dots, u^n) = \theta_t(0, u^2, \dots, u^n).$$

Geometrically, for each fixed $(u^2, \dots, u^n)$, ψ maps the set $(-\varepsilon, \varepsilon) \times \{(u^2, \dots, u^n)\}$ to the integral curve starting at $(0, u^2, \dots, u^n)$.

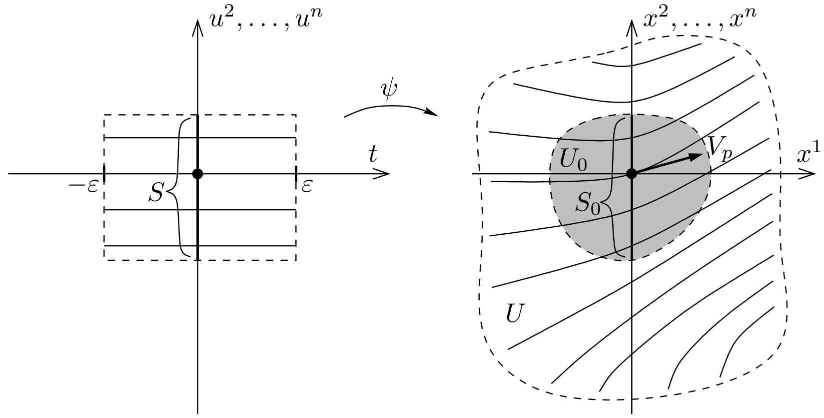


Figure 17.7. Proof of the canonical form theorem.

First we will show that ψ pushes $\partial/\partial t$ forward to V. For any $(t_0, u_0) \in (-\varepsilon, \varepsilon) \times S$, we have

$$\begin{aligned}\left(\psi_* \frac{\partial}{\partial t}\bigg|_{(t_0,u_0)}\right) f &= \frac{\partial}{\partial t}\bigg|_{(t_0,u_0)} (f \circ \psi) \\ &= \frac{\partial}{\partial t}\bigg|_{t=t_0} (f(\theta_t(0, u_0))) \\ &= V_{\psi(t_0,u_0)} f,\end{aligned} \tag{17.9}$$

where we have used the fact that $t \mapsto \theta_t(0, u_0) = \theta^{(0,u_0)}(t)$ is an integral curve of V. On the other hand, when restricted to $\{0\} \times S$, $\psi(0, u^2, \ldots, u^n) = \theta_0(0, u^2, \ldots, u^n) = (0, u^2, \ldots, u^n)$, so

$$\psi_* \frac{\partial}{\partial u^i}\bigg|_{(0,0)} = \frac{\partial}{\partial x^i}\bigg|_p, \qquad i = 2, \ldots, n.$$

Thus at $(0, 0)$, ψ_* takes the basis $(\partial/\partial t|_{(0,0)}, \partial/\partial u^2|_{(0,0)}, \ldots, \partial/\partial u^n|_{(0,0)})$ to $(V_p, \partial/\partial x^2|_p, \ldots, \partial/\partial x^n|_p)$. Since V_p has nonzero x^1-component, this is also a basis, so ψ_* is an isomorphism. Therefore, by the inverse function theorem, there are neighborhoods W of $(0,0)$ and Y of p such that $\psi\colon W \to Y$ is a diffeomorphism.

Let $\varphi = \psi^{-1}\colon Y \to W$, which is a smooth coordinate map on Y. Equation (17.9) says precisely that V is equal to the coordinate vector field $\partial/\partial t$ in these coordinates. Renaming t to u^1, this is what we wanted to prove. □

The proof of this theorem actually provides a technique for finding coordinates that put a given vector field V in canonical form, at least when the corresponding system of ODEs can be explicitly solved. If we begin with smooth coordinates in which $p = 0$ and V_p has nonzero x^1-component, then the proof shows that the inverse of the map ψ constructed above will be a coordinate map of the desired form. Actually, although these normalizations were useful in the proof, they are not really necessary—the requirement that the coordinates be centered at p was only to simplify the notation in the proof; and any coordinate can play the role of x^1, as long as the corresponding component of V_p is not zero. The procedure is best illustrated by an example.

Example 17.14. Let $W = x\,\partial/\partial y - y\,\partial/\partial x$ on $\mathbb{R}^2$. We computed the flow of W in Example 17.4. The point $(1,0) \in \mathbb{R}^2$ is a regular point for W, because $W_{(1,0)} = \partial/\partial y \neq 0$. Because W has nonzero y-coordinate there, we can take S_0 to be coordinate line where $y = 0$ (the x-axis), parametrized by $u \mapsto (u, 0)$. We define $\psi\colon \mathbb{R}^2 \to \mathbb{R}^2$ by

$$\psi(t,u) = \theta_t(u,0) = (u\cos t, u\sin t),$$

and then solve locally for (t,u) in terms of (x,y) to obtain the following coordinate map in a neighborhood of $(1,0)$:

$$(t,u) = \psi^{-1}(x,y) = (\tan^{-1}(y/x), \sqrt{x^2+y^2}).$$

It is easy to check that $W = \partial/\partial t$ in these coordinates. (They are, of course, just polar coordinates with different names.)

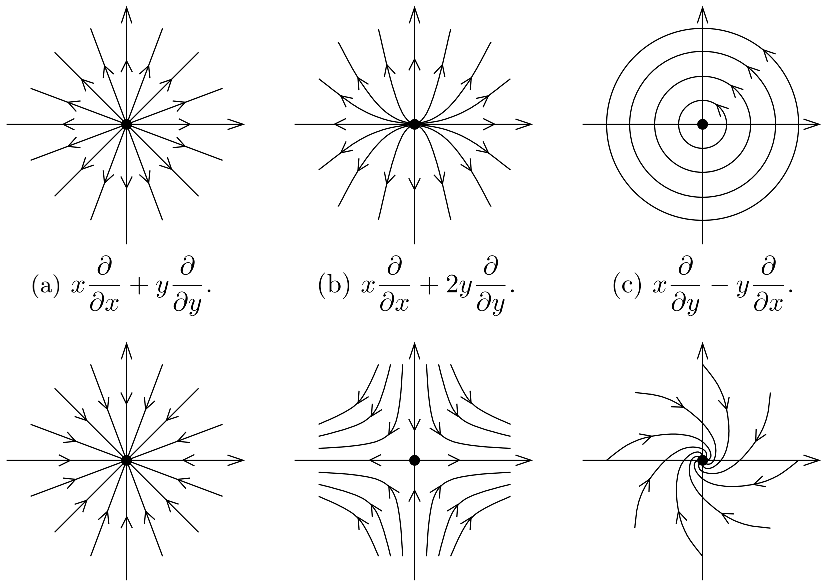
The canonical form theorem implies that the integral curves of V near a regular point behave, up to diffeomorphism, just like parallel coordinate lines in $\mathbb{R}^n$. Thus all of the interesting local behavior of the flow is concentrated near its equilibrium points. The flow near equilibrium points can exhibit a wide variety of behaviors, such as closed orbits surrounding the equilibrium point, orbits converging exponentially or spiraling into the equilibrium point as $t \to +\infty$ or $-\infty$, and many more complicated phenomena. Some typical 2-dimensional examples are illustrated in Figure 17.8. A systematic study of the local behavior of flows near equilibrium points in the plane can be found in many ODE texts, for example [BD65]. The study of global and long-time behaviors of flows on manifolds, called *smooth dynamical systems theory*, is a deep subject with many applications both inside and outside of mathematics.

Time-Dependent Vector Fields

All of the systems of differential equations we have considered so far have been of the form

$$(\gamma^i)'(t) = V^i(\gamma(t)),$$

in which the functions V^i do not depend explicitly on the independent variable t. Such a system is said to be *autonomous*. In contrast, a system



(a) $x\dfrac{\partial}{\partial x} + y\dfrac{\partial}{\partial y}$. (b) $x\dfrac{\partial}{\partial x} + 2y\dfrac{\partial}{\partial y}$. (c) $x\dfrac{\partial}{\partial y} - y\dfrac{\partial}{\partial x}$.

(d) $-x\dfrac{\partial}{\partial x} - y\dfrac{\partial}{\partial y}$. (e) $x\dfrac{\partial}{\partial x} - y\dfrac{\partial}{\partial y}$. (f) $(y - x)\dfrac{\partial}{\partial x} - (y + x)\dfrac{\partial}{\partial y}$.

Figure 17.8. Examples of flows near equilibrium points.

of the form

$$(\gamma^i)'(t) = V^i(t, \gamma(t)),$$

in which V^i is a function defined on some subset of $\mathbb{R} \times \mathbb{R}^n$, is called *nonautonomous*. Many applications of ODEs require the consideration of nonautonomous systems. In this section, we show how the results of this chapter can be extended to cover this case.

Somewhat more generally, if M is a smooth manifold, a *time-dependent vector field* on M is a continuous map $V\colon J \times M \to TM$, where $J \subset \mathbb{R}$ is an open interval, such that $V(t,p) \in T_pM$ for each $(t,p) \in J \times M$. This means that for each $t \in J$, the map $V_t\colon M \to TM$ defined by $V_t(p) = V(t,p)$ is a vector field on M.

Theorem 17.15 (Flows of Time-Dependent Vector Fields). *Let M be a smooth manifold, and let $V\colon J \times M \to TM$ be a smooth time-dependent vector field on M. There exist an open set $\mathcal{E} \subset J \times J \times M$ and a smooth map $\theta\colon \mathcal{E} \to M$ such that for each $s \in J$ and $p \in M$, the set $\mathcal{E}^{(s,p)} = \{t \in J : (t,s,p) \in \mathcal{E}\}$ is an open interval containing s, and the smooth curve $\gamma\colon \mathcal{E}^{(s,p)} \to M$ defined by $\gamma(t) = \theta(t,s,p)$ is the unique* *maximal solution to the initial-value problem*

$$\begin{aligned} \gamma'(t) &= V(t, \gamma(t)), \\ \gamma(s) &= p. \end{aligned} \tag{17.10}$$

Proof. Consider the smooth vector field $\widetilde{V}$ on $J \times M$ defined by

$$\widetilde{V}_{(s,p)} = \left(\left.\frac{\partial}{\partial s}\right|_s, V(s,p) \right),$$

where s is the standard coordinate on $J \subset \mathbb{R}$, and we identify $T_{(s,p)}(J \times M)$ with $T_s J \oplus T_p M$ as usual. Let $\widetilde{\theta}\colon \widetilde{\mathcal{D}} \to J \times M$ denote the flow of $\widetilde{V}$. If we write the component functions of $\widetilde{\theta}$ as

$$\widetilde{\theta}(t, (s,p)) = (\alpha(t, (s,p)), \beta(t, (s,p))),$$

then $\alpha\colon \widetilde{\mathcal{D}} \to J$ and $\beta\colon \widetilde{\mathcal{D}} \to M$ satisfy

$$\begin{aligned} &\frac{\partial \alpha}{\partial t}(t, (s,p)) = 1, && \alpha(0, (s,p)) = s, \\ &\frac{\partial \beta}{\partial t}(t, (s,p)) = V(\alpha(t, (s,p)), \beta(t, (s,p))), && \beta(0, (s,p)) = p. \end{aligned} \tag{17.11}$$

It follows immediately that

$$\alpha(t, (s,p)) = t + s. \tag{17.12}$$

Let $\mathcal{E}$ be the subset of $\mathbb{R} \times J \times M$ defined by

$$\mathcal{E} = \{(t, s, p) : (t - s, (s,p)) \in \widetilde{\mathcal{D}}\}.$$

Clearly $\mathcal{E}$ is open in $\mathbb{R} \times J \times M$ because $\widetilde{\mathcal{D}}$ is. Moreover, since α maps $\widetilde{\mathcal{D}}$ into J, if $(t, s, p) \in \mathcal{E}$, then $t = \alpha(t - s, (s,p)) \in J$, which implies that $\mathcal{E} \subset J \times J \times M$.

Now define $\theta\colon \mathcal{E} \to M$ by

$$\theta(t, s, p) = \beta(t - s, (s,p)).$$

Then θ is smooth because β is, and it follows from (17.11) and (17.12) that θ satisfies (17.10).

To prove uniqueness, suppose that $s \in J$, and $\gamma\colon J_0 \to M$ is any smooth curve satisfying (17.10) on some open interval $J_0 \subset J$ containing s. Define a smooth curve $\widetilde{\gamma}\colon J_0 \to J \times M$ by $\widetilde{\gamma}(t) = (t, \gamma(t))$. Then $\widetilde{\gamma}$ is easily seen to be an integral curve of $\widetilde{V}$, and thus by uniqueness of integral curves, we must have $\widetilde{\gamma}(t) = \widetilde{\theta}(t - s, (s,p))$ on its whole domain, which implies that $\gamma(t) = \theta(t, s, p)$. □

We will call the map θ whose existence is asserted by this theorem a *time-dependent flow.* Some properties of time-dependent flows are described in Problem 17-14.

Proof of the ODE Theorem

In this section we prove the ODE existence, uniqueness, and smoothness theorem (Theorem 17.9). Here is the setting: We are given an open set $U \subset \mathbb{R}^n$ and a map $V\colon U \to \mathbb{R}^n$, and for any $t_0 \in \mathbb{R}$ and any $x \in U$ we will consider the following ODE initial-value problem:

$$\begin{aligned}(\gamma^i)'(t) &= V^i(\gamma(t)),\\ \gamma^i(t_0) &= x^i.\end{aligned} \tag{17.13}$$

In the next three theorems, we will prove existence, uniqueness, and smooth dependence on initial conditions for solutions to (17.13). At the end of the chapter, we will show how Theorem 17.9 follows easily from these results.

The following comparison lemma will be useful in the proofs to follow.

Lemma 17.16 (Comparison Lemma). *Suppose $J_0 \subset \mathbb{R}$ is an open interval containing t_0, and $u\colon J_0 \to \mathbb{R}^n$ is a differentiable map satisfying the following differential inequality for some nonnegative constants A and B and all $t \in J_0$:*

$$|u'(t)| \le A\,|u(t)| + B.$$

Then the following inequality holds for all $t \in J_0$:

$$|u(t)| \le e^{A|t-t_0|}|u(t_0)| + \frac{B}{A}\big(e^{A|t-t_0|} - 1\big). \tag{17.14}$$

Proof. Let $J_0^+ = \{t \in J_0 : t \ge t_0\}$. We will prove that (17.14) holds for all $t \in J_0^+$, and then the analogous statement for $\widetilde{t} \le t_0$ follows easily by substituting $t_0 - \widetilde{t} = t - t_0$.

On the subset of J_0^+ where $|u(t)| > 0$, $|u(t)|$ is a differentiable function of t, and the Schwartz inequality shows that

$$\begin{aligned}\frac{d}{dt}|u(t)| &= \frac{d}{dt}\left(u(t)\cdot u(t)\right)^{1/2}\\ &= \frac{1}{2}\left(u(t)\cdot u(t)\right)^{-1/2}\left(2u(t)\cdot u'(t)\right)\\ &\le \frac{1}{2}|u(t)|^{-1}\left(2\,|u(t)|\;|u'(t)|\right)\\ &= |u'(t)|\\ &\le A|u(t)| + B.\end{aligned}$$

Define a smooth function $g\colon J_0^+ \to \mathbb{R}$ by

$$g(t) = e^{A(t-t_0)}|u(t_0)| + \frac{B}{A}\big(e^{A(t-t_0)} - 1\big).$$

It is easy to check that $g(t_0) = |u(t_0)|$, $g(t) > 0$ for $t > t_0$, and g satisfies the ODE

$$g'(t) = Ag(t) + B.$$

Consider the continuous function $f\colon J_0^+ \to \mathbb{R}$ defined by

$$f(t) = e^{-A(t-t_0)}(|u(t)| - g(t)).$$

Then $f(t_0) = 0$, and (17.14) for $t \in J_0^+$ is equivalent to $f(t) \le 0$. At any $t \in J_0^+$ such that $f(t) > 0$ (and therefore $|u(t)| > 0$), f is differentiable and satisfies

$$\begin{aligned} f'(t) &= -Ae^{-A(t-t_0)}(|u(t)| - g(t)) + e^{-A(t-t_0)}\left(\frac{d}{dt}|u(t)| - g'(t)\right) \\ &\le -Ae^{-A(t-t_0)}(|u(t)| - g(t)) + e^{-A(t-t_0)}\left(A|u(t)| + B - Ag(t) - B\right) \\ &= 0. \end{aligned}$$

Now suppose there is some $t_1 \in J_0^+$ such that $f(t_1) > 0$. Let

$$\tau = \sup\{t \in [t_0, t_1] : f(t) \le 0\}.$$

Then $f(\tau) = 0$ by continuity, and $f(t) > 0$ for $t \in (\tau, t_1]$. Since f is continuous on $[\tau, t_1]$ and differentiable on (τ, t_1), the mean-value theorem implies that there must exist $t \in (\tau, t_1)$ such that $f(t) > 0$ and $f'(t) > 0$. But this contradicts the calculation above, which showed that $f'(t) \le 0$ whenever $f(t) > 0$, thus proving that $f(t) \le 0$ for all $t \in J_0^+$. □

Theorem 17.17 (Existence of ODE Solutions). *Let $U \subset \mathbb{R}^n$ be an open set, and suppose $V\colon U \to \mathbb{R}^n$ is Lipschitz continuous. Let $(t_0, x_0) \in \mathbb{R} \times U$ be given. There exist an open interval $J_0 \subset \mathbb{R}$ containing t_0, an open set $U_0 \subset U$ containing x_0, and for each $x \in U_0$, a C^1 curve $\gamma\colon J_0 \to U$ satisfying the initial-value problem* (17.13).

Proof. Suppose γ is any solution to (17.13) on some interval J_0 containing t_0. Because γ is differentiable, it is continuous, and then the fact that the right-hand side of (17.13) is a continuous function of t implies that γ is of class C^1. Integrating (17.13) with respect to t and applying the fundamental theorem of calculus shows that γ satisfies the following integral equation:

$$\gamma^i(t) = x^i + \int_{t_0}^{t} V^i(\gamma(s))\, ds. \tag{17.15}$$

Conversely, if $\gamma\colon J_0 \to U$ is a continuous map satisfying (17.15), then the fundamental theorem of calculus implies that γ satisfies (17.13) and therefore is actually of class C^1.

This motivates the following definition. Suppose J_0 is an open interval containing t_0. For any continuous curve $\gamma\colon J_0 \to U$, we define a new curve

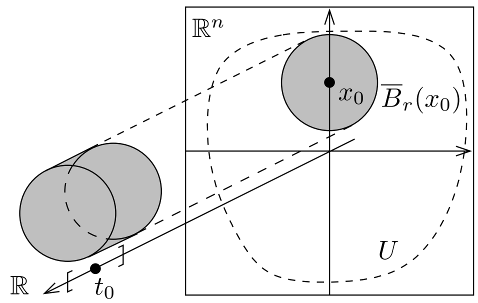


Figure 17.9. Proof of the ODE existence theorem.

$I\gamma\colon J_0 \to \mathbb{R}^n$ by

$$I\gamma(t) = x + \int_{t_0}^{t} V(\gamma(s))\,ds. \tag{17.16}$$

Then we are led to seek a fixed point for I in a suitable metric space of curves.

Let C be a Lipschitz constant for V, so that

$$|V(x) - V(\widetilde{x})| \le C|x - \widetilde{x}|, \qquad x, \widetilde{x} \in U. \tag{17.17}$$

Given $t_0 \in \mathbb{R}$ and $x_0 \in U$, choose $r > 0$ small enough that $\overline{B}_r(x_0) \subset U$ (Figure 17.9). Let M be the supremum of $|V|$ on the compact set $\overline{B}_r(x_0)$. Choose $\delta > 0$ and $\varepsilon > 0$ small enough that

$$\delta < \frac{r}{2}, \qquad \varepsilon < \min\left(r, \frac{r}{2M}, \frac{1}{C}\right),$$

and set $J_0 = (t_0 - \varepsilon, t_0 + \varepsilon) \subset \mathbb{R}$ and $U_0 = B_\delta(x_0) \subset U$. For any $x \in U_0$, let $\mathcal{M}_x$ denote the set of all continuous curves $\gamma\colon J_0 \to \overline{B}_r(x_0)$ satisfying $\gamma(t_0) = x$. We define a metric on $\mathcal{M}_x$ by

$$d(\gamma, \widetilde{\gamma}) = \sup_{t \in J_0} |\gamma(t) - \widetilde{\gamma}(t)|.$$

Any sequence of maps in $\mathcal{M}_x$ that is Cauchy in this metric is uniformly Cauchy, and thus converges to a continuous limit γ. Clearly, the conditions that γ take its values in $\overline{B}_r(x_0)$ and $\gamma(t_0) = x$ are preserved in the limit. Therefore, $\mathcal{M}_x$ is a complete metric space.

We wish to define a map $I\colon \mathcal{M}_x \to \mathcal{M}_x$ by formula (17.16). The first thing we need to verify is that I really does map $\mathcal{M}_x$ into itself. It is clear from the definition that $I\gamma(t_0) = x$ and $I\gamma$ is continuous (in fact, it is differentiable by the fundamental theorem of calculus). Thus we need only check that $I\gamma$ takes its values in $\overline{B}_r(x_0)$. If $\gamma \in \mathcal{M}_x$, then for any $t \in J_0$,

$$\begin{aligned}|I\gamma(t) - x_0| &= \left| x + \int_{t_0}^{t} V(\gamma(s))\, ds - x_0 \right| \\ &\le |x - x_0| + \int_{t_0}^{t} |V(\gamma(s))|\, ds \\ &< \delta + M\varepsilon < r\end{aligned}$$

by our choice of δ and ε.

Next we check that I is a contraction (see page 159). If $\gamma, \widetilde{\gamma} \in \mathcal{M}_x$, then

$$\begin{aligned}d(I\gamma, I\widetilde{\gamma}) &= \sup_{t\in J_0} \left| \int_{t_0}^{t} V(\gamma(s))\, ds - \int_{t_0}^{t} V(\widetilde{\gamma}(s))\, ds \right| \\ &\le \sup_{t\in J_0} \int_{t_0}^{t} |V(\gamma(s)) - V(\widetilde{\gamma}(s))|\, ds \\ &\le \sup_{t\in J_0} \int_{t_0}^{t} C|\gamma(s) - \widetilde{\gamma}(s)|\, ds \\ &\le C\varepsilon d(\gamma, \widetilde{\gamma}).\end{aligned}$$

Because we have chosen ε so that $C\varepsilon < 1$, this shows that I is a contraction. By the contraction lemma (Lemma 7.6), I has a fixed point $\gamma \in \mathcal{M}_x$, which is a solution to (17.15) and thus also (17.13). □

Theorem 17.18 (Uniqueness of ODE Solutions). *Let $U \subset \mathbb{R}^n$ be an open set, and suppose $V\colon U \to \mathbb{R}^n$ is Lipschitz continuous. For any $t_0 \in \mathbb{R}$, any two solutions to* (17.13) *are equal on their common domain.*

Proof. Suppose $\gamma, \widetilde{\gamma}\colon J_0 \to U$ are two differentiable functions that both satisfy the ODE on the same open interval $J_0 \subset \mathbb{R}$, but not necessarily with the same initial conditions. The Lipschitz estimate for V implies

$$\left| \frac{d}{dt}\big(\widetilde{\gamma}(t) - \gamma(t)\big) \right| = |V(\widetilde{\gamma}(t)) - V(\gamma(t)))| \le C\,|\widetilde{\gamma}(t) - \gamma(t)|\,.$$

Applying the comparison lemma (Lemma 17.16) with $A = C$ and $B = 0$, we conclude that

$$|\widetilde{\gamma}(t) - \gamma(t)| \le e^{C|t-t_0|}|\widetilde{\gamma}(t_0) - \gamma(t_0)|. \tag{17.18}$$

Thus $\gamma(t_0) = \widetilde{\gamma}(t_0)$ implies $\gamma \equiv \widetilde{\gamma}$ on all of J_0. □

Theorem 17.19 (Smoothness of ODE Solutions). *Suppose $U \subset \mathbb{R}^n$ is an open set and $V\colon U \to \mathbb{R}^n$ is Lipschitz continuous. Suppose also that $U_0 \subset U$ is an open set, $J_0 \subset \mathbb{R}$ is an open interval containing t_0, and $\theta\colon J_0 \times U_0 \to U$ is any map such that for each $x \in U_0$, $\gamma(t) = \theta(t, x)$ solves the initial-value problem* (17.13). *If V is of class C^k for some $k \ge 0$, then so is θ.*

Proof. We will prove the theorem by induction on k. For the $k = 0$ step, it suffices to choose an arbitrary $(t_1, x_1) \in J_0 \times U_0$ and prove that θ is continuous on some neighborhood of (t_1, x_1). Let J_1 be a bounded open interval containing t_0 and t_1 and such that $\overline{J}_1 \subset J_0$. Choose $r > 0$ such that $\overline{B}_{2r}(x_1) \subset U_0$, and let $U_1 = B_r(x_1)$. Let C be a Lipschitz constant for V as in (17.17), and define constants M and T by

$$M = \sup_{\overline{U}_1} |V|, \qquad T = \sup_{\overline{J}_1} |t - t_0|.$$

We will show that θ is continuous on $\overline{J}_1 \times \overline{U}_1$. First we note that (17.18) implies the following Lipschitz estimate for all $t \in \overline{J}_1$ and all $x, \widetilde{x} \in \overline{U}_1$:

$$|\theta(t, \widetilde{x}) - \theta(t, x)| \le e^{CT} |\widetilde{x} - x|. \tag{17.19}$$

Thus for each t, θ is Lipschitz continuous as a function of x. We need to show it is jointly continuous in (t, x).

Let (t, x), $(\widetilde{t}, \widetilde{x}) \in \overline{J}_1 \times \overline{U}_1$ be arbitrary. Using the fact that every solution to the initial-value problem satisfies the integral equation (17.15), we find that

$$\theta^i(t, x) = x^i + \int_{t_0}^{t} V^i(\theta(s, x))\, ds, \tag{17.20}$$

and therefore (assuming for simplicity that $\widetilde{t} \ge t$),

$$\begin{aligned}
|\theta(\widetilde{t}, \widetilde{x}) - \theta(t, x)| &\le |\widetilde{x} - x| + \left| \int_{t_0}^{\widetilde{t}} V(\theta(s, \widetilde{x}))\, ds - \int_{t_0}^{t} V(\theta(s, x))\, ds \right| \\
&\le |\widetilde{x} - x| + \int_{t_0}^{t} |V(\theta(s, \widetilde{x})) - V(\theta(s, x))|\, ds \\
&\qquad + \int_{t}^{\widetilde{t}} |V(\theta(s, \widetilde{x}))|\, ds \\
&\le |\widetilde{x} - x| + C \int_{t_0}^{t} |\theta(s, \widetilde{x}) - \theta(s, x)|\, ds + \int_{t}^{\widetilde{t}} M\, ds \\
&\le |\widetilde{x} - x| + CTe^{CT} |\widetilde{x} - x| + M|\widetilde{t} - t|.
\end{aligned}$$

It follows that θ is continuous.

Next we tackle the $k = 1$ step, which is the hardest part of the proof. Suppose that V is of class C^1, and let $\overline{J}_1, \overline{U}_1$ be defined as above. Expressed in terms of θ, the initial-value problem (17.13) reads

$$\begin{aligned}
&\frac{\partial \theta^i}{\partial t}(t, x) = V^i(\theta(t, x)), \\
&\theta^i(t_0, x) = x^i.
\end{aligned} \tag{17.21}$$

Because we know θ^i is continuous by the argument above, this shows that $\partial \theta^i / \partial t$ is continuous. We will prove that for each j, $\partial \theta^i / \partial x^j$ exists and is continuous on $\overline{J}_1 \times \overline{U}_1$.

Working directly with the definition of the partial derivative, for any $h \in \overline{B}_r(0) \subset \mathbb{R}^n$ and any indices $i, j \in \{1, \dots, n\}$, we let $(\Delta_h)^i_j : \overline{J}_1 \times \overline{U}_1 \to \mathbb{R}$ be the difference quotient

$$(\Delta_h)^i_j(t,x) = \frac{\theta^i(t, x + he_j) - \theta^i(t,x)}{h}.$$

Then $\partial\theta^i/\partial x^j(t,x) = \lim_{h\to 0}(\Delta_h)^i_j(t,x)$ if the limit exists. In fact, we will show that $(\Delta_h)^i_j$ converges uniformly on $\overline{J}_1 \times \overline{U}_1$ as $h \to 0$, from which it follows that $\partial\theta^i/\partial x^j$ exists and is continuous there, because it is a uniform limit of continuous functions.

Let $\Delta_h : \overline{J}_1 \times \overline{U}_1 \to \mathrm{M}(n,\mathbb{R})$ be the matrix-valued function whose matrix entries are $(\Delta_h)^i_j(t,x)$. Note that (17.19) implies that $|(\Delta_h)^i_j(t,x)| \le e^{CT}$ for each i and j. It follows that Δ_h satisfies the following bound:

$$|\Delta_h(t,x)| \le ne^{CT}, \tag{17.22}$$

where the norm on the left-hand side is the usual Euclidean norm on matrices.

Because V is C^1, Taylor's formula (Theorem A.58 in the Appendix) together with the explicit formula (A.20) for the remainder term give the following formula for all $t \in \overline{J}_1$, $y \in \overline{U}_1$, and $v \in \overline{B}_r(0)$:

$$V^i(y+v) - V^i(y) = v^k \frac{\partial V^i}{\partial y^k}(y) + v^k \int_0^1 \left(\frac{\partial V^i}{\partial y^k}(y+sv) - \frac{\partial V^i}{\partial y^k}(y) \right) ds.$$

(We are using the summation convention as usual.) Letting $G^i_k(y,v)$ denote the integral in the last line above, we have shown that

$$V^i(y+v) = V^i(y) + v^k \frac{\partial V^i}{\partial y^k}(y) + v^k G^i_k(y,v), \tag{17.23}$$

where $G^i_k(y,v)$ is a continuous function of $(y,v) \in \overline{U}_1 \times \overline{B}_r(0)$, which is equal to zero when $v = 0$. Because a continuous function on a compact set is uniformly continuous (see, for example, [Rud76, Theorem 4.19]), for any $\varepsilon > 0$, there exists $\delta > 0$ such that the matrix-valued function G satisfies

$$|G(y,v)| < \varepsilon \text{ for all } y \in \overline{U}_1 \text{ and all } |v| < \delta. \tag{17.24}$$

Note that $\theta^i(t_0, x) = x^i$ implies that Δ_h satisfies the following initial condition:

$$\begin{aligned}(\Delta_h)^i_j(t_0,x) &= \frac{\theta^i(t_0, x + he_j) - \theta^i(t_0,x)}{h}\\ &= \frac{(x^i + h\delta^i_j) - x^i}{h}\\ &= \delta^i_j.\end{aligned} \tag{17.25}$$

Let us compute the derivative of Δ_h with respect to t. Using (17.23) with $y = \theta(t,x)$ and $v = (v^1, \dots, v^k)$ given by

$$v^k = \theta^k(t, x + he_j) - \theta^k(t,x) = h(\Delta_h)^k_j(t,x),$$

we have

$$\begin{aligned}
\frac{\partial}{\partial t}(\Delta_h)^i_j(t,x) &= \frac{1}{h}\left(\frac{\partial \theta^i}{\partial t}(t, x+he_j) - \frac{\partial \theta^i}{\partial t}(t,x)\right)\\
&= \frac{1}{h}\big(V^i(\theta(t,x+he_j)) - V^i(\theta(t,x))\big)\\
&= \frac{1}{h}\left(v^k\frac{\partial V^i}{\partial y^k}(\theta(t,x)) + v^k G^i_k(y,v)\right)\\
&= \left(\frac{\partial V^i}{\partial y^k}(\theta(t,x)) + G^i_k(y,v)\right)(\Delta_h)^k_j(t,x).
\end{aligned}$$

Thus for any nonzero $h, \widetilde{h} \in \overline{B}_r(0)$,

$$\begin{aligned}
&\frac{\partial}{\partial t}\left((\Delta_h)^i_j(t,x) - (\Delta_{\widetilde{h}})^i_j(t,x)\right)\\
&\qquad = \frac{\partial V^i}{\partial y^k}(\theta(t,x))\left((\Delta_h)^k_j(x) - (\Delta_{\widetilde{h}})^k_j(t,x)\right)\\
&\qquad\qquad + G^i_k(y,v)(\Delta_h)^k_j(t,x) - G^i_k(y,\widetilde{v})(\Delta_{\widetilde{h}})^k_j(t,x),
\end{aligned}$$

where $\widetilde{v}$ is defined similarly to v, but with $\widetilde{h}$ in place of h.

Now let $\varepsilon > 0$ be given, and choose $\delta \le r$ satisfying (17.24). Let E denote the supremum of $|DV|$ on $\overline{U}_1$. If $|h|$ and $|\widetilde{h}|$ are both less than $\delta e^{-CT}/n$, then (17.22) implies that $|v|,\ |\widetilde{v}| < \delta$. Therefore,

$$\left|\frac{\partial}{\partial t}\left(\Delta_h(t,x) - \Delta_{\widetilde{h}}(t,x)\right)\right| \le E|\Delta_h(t,x) - \Delta_{\widetilde{h}}(t,x)| + 2\varepsilon n e^{CT}.$$

Since $\Delta_h(t_0,x) - \Delta_{\widetilde{h}}(t_0,x) = 0$ by (17.25), it follows from the comparison lemma that

$$\big|\Delta_h(t,x) - \Delta_{\widetilde{h}}(t,x)\big| \le \frac{2\varepsilon n e^{CT}}{E}\big(e^{E|t-t_0|} - 1\big) \le \frac{2\varepsilon n e^{CT}}{E}\big(e^{ET} - 1\big).$$

Since this can be made as small as desired by choosing h and $\widetilde{h}$ sufficiently small, this shows that for any sequence $h_k \to 0$, $\{(\Delta_{h_k})^i_j(t,x)\}$ is uniformly Cauchy and therefore uniformly convergent for each i and j. The continuous limit function is equal to $\partial\theta^i/\partial x^j$ by definition. This completes the proof of the $k = 1$ case.

Now assume that the theorem is true for some $k \ge 1$, and suppose V is of class C^{k+1}. By the inductive hypothesis, θ is of class C^k, and therefore by (17.21), $\partial\theta^i/\partial t$ is also C^k. We can differentiate under the integral sign in (17.20) to obtain

$$\frac{\partial\theta^i}{\partial x^j}(t,x) = \delta^i_j + \int_{t_0}^t \frac{\partial V^i}{\partial y^k}(\theta(s,x))\frac{\partial\theta^k}{\partial x^j}(s,x)\,ds.$$

By the fundamental theorem of calculus, this implies that $\partial\theta^i/\partial x^j$ satisfies the differential equation

$$\frac{\partial}{\partial t}\frac{\partial\theta^i}{\partial x^j}(t,x) = \frac{\partial V^i}{\partial y^k}(\theta(t,x))\frac{\partial\theta^k}{\partial x^j}(t,x).$$

Consider the following initial-value problem for the $n+n^2$ unknown functions (α^i,β^i_j):

$$\begin{aligned}
(\alpha^i)'(t) &= V^i(\alpha(t)),\\
(\beta^i_j)'(t) &= \frac{\partial V^i}{\partial y^k}(\alpha(t))\beta^k_j(t);\\
\alpha^i(t_0) &= a^i,\\
\beta^i_j(t_0) &= b^i_j.
\end{aligned}$$

The functions on the right-hand side of this system are C^k functions of (α^i,β^i_j), so the inductive hypothesis implies that its solutions are C^k functions of (t,a^i,b^i_j). The discussion in the preceding paragraph shows that $\alpha^i(t) = \theta^i(t,x)$ and $\beta^i_j(t) = \partial\theta^i/\partial x^j(t,x)$ solve this system with initial conditions $a^i = x^i$, $b^i_j = \delta^i_j$. This shows that $\partial\theta^i/\partial x^j$ is a C^k function of (t,x), so θ itself is of class C^{k+1}, thus completing the induction. □

Proof of Theorem 17.9. Suppose $U \subset \mathbb{R}^n$ is open and $V\colon U \to \mathbb{R}^n$ is smooth. Let $t_0 \in \mathbb{R}$ and $x_0 \in U$ be arbitrary. Because V is smooth, replacing U by a precompact convex neighborhood of x_0, we may assume that V is Lipschitz continuous, so the theorems of this section apply. Theorem 17.17 shows that there exist neighborhoods J_0 of t and U_0 of x_0 such that for each $x \in U_0$, there is a C^1 solution $\gamma\colon J_0 \to U$ to (17.8). Because V is smooth, an easy induction using (17.8) then shows that γ is smooth as a function of t. Uniqueness of solutions is an immediate consequence of Theorem 17.18. Finally, Theorem 17.19 shows that the solution is C^k for every k as a function of (t,x), so it is smooth. □

Problems

17-1. Show that every smooth vector field with compact support is complete.

17-2. Compute the flow of each of the following vector fields on $\mathbb{R}^2$:

(a) $V = x\dfrac{\partial}{\partial x} + y\dfrac{\partial}{\partial y}$.

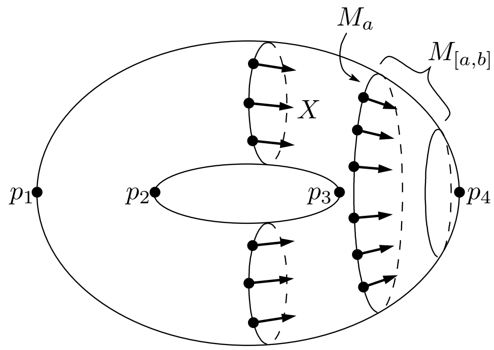


Figure 17.10. The setup for Problem 17-4.

(b) $W = x\dfrac{\partial}{\partial x} + 2y\dfrac{\partial}{\partial y}$.

(c) $X = x\dfrac{\partial}{\partial x} - y\dfrac{\partial}{\partial y}$.

(d) $Y = x\dfrac{\partial}{\partial y} + y\dfrac{\partial}{\partial x}$.

17-3. For each of the vector fields in Problem 17-2, find smooth coordinates in a neighborhood of $(1,0)$ for which the given vector field is a coordinate vector field.

17-4. Let M be a compact Riemannian n-manifold, and let $f \in C^\infty(M)$. Suppose f has only finitely many critical points $\{p_1, \dots, p_k\}$ with corresponding critical values $\{c_1, \dots, c_k\}$. (Assume without loss of generality that $c_1 \le \dots \le c_k$.) For any $a < b \in \mathbb{R}$, define $M_a = f^{-1}(a)$, $M_{[a,b]} = f^{-1}([a,b])$, and $M_{(a,b)} = f^{-1}((a,b))$. If a is a regular value, note that M_a is an embedded hypersurface in M (see Figure 17.10).

(a) Let X be the vector field $X = \operatorname{grad} f / |\operatorname{grad} f|^2$ on $M \smallsetminus \{p_1, \dots, p_k\}$, and let θ denote the flow of X. Show that $f(\theta_t(p)) = f(p) + t$ whenever $\theta_t(p)$ is defined.

(b) Let $[a,b] \subset \mathbb{R}$ be a compact interval containing no critical values of f. Show that

$$\theta : [0, b-a] \times M_a \to M_{[a,b]}$$

is a diffeomorphism, with inverse given by

$$p \mapsto \big(f(p) - a, \theta(a - f(p), p)\big).$$

[Remark: This result shows that M can be decomposed as a union of simpler "building blocks"—the product sets $M_{[c_i+\varepsilon, c_{i+1}-\varepsilon]} \approx I \times M_{c_i+\varepsilon}$, and the neighborhoods $M_{(c_i-\varepsilon, c_i+\varepsilon)}$ of the critical points. This is the starting point of *Morse theory*, which is one of the deepest applications of differential geometry to topology. It is enlightening to think about what this means when M is a torus of revolution in $\mathbb{R}^3$ obtained by revolving a circle around the z-axis, and $f(x, y, z) = x$.]

17-5. Let M be a connected smooth manifold. Show that the group of diffeomorphisms of M acts transitively on M. More precisely, for any two points $p, q \in M$, show that there is a diffeomorphism $F : M \to M$ such that $F(p) = q$. [Hint: First prove that if $p, q \in \mathbb{B}^n$ (the open unit ball in $\mathbb{R}^n$), there is a compactly supported smooth vector field on $\mathbb{B}^n$ whose flow θ satisfies $\theta_1(p) = q$.]

17-6. Let M be a smooth manifold. A curve $\gamma : \mathbb{R} \to M$ is said to be *periodic* if there is a number $T > 0$ such that $\gamma(t) = \gamma(t + kT)$ for all $t \in \mathbb{R}$ and $k \in \mathbb{Z}$. Suppose $X \in \mathcal{T}(M)$ and γ is a maximal integral curve of X.

(a) Show that exactly one of the following holds:
- γ is constant.
- γ is injective.
- γ is periodic and nonconstant.

(b) If γ is periodic and nonconstant, show that there exists $T \in \mathbb{R}$ (called the *period* of γ) such that $\gamma(t) = \gamma(t')$ if and only if $t - t' = kT$ for some $k \in \mathbb{Z}$.

(c) Show that the image of γ is an immersed submanifold of M, diffeomorphic to $\mathbb{R}$, $\mathbb{S}^1$, or $\mathbb{R}^0$.

17-7. Let M be a smooth n-manifold, and suppose V is a smooth vector field on M such that every integral curve of V is periodic with the same period (see Problem 17-6). Define an equivalence relation on M by saying $p \sim q$ if p and q are in the image of the same integral curve of V. Let $M/{\sim}$ be the quotient space, and let $\pi : M \to M/{\sim}$ be the quotient map. Show that $M/{\sim}$ is a topological $(n - 1)$-manifold and has a unique smooth structure such that π is a smooth submersion.

17-8. Show that every connected smooth 1-manifold is diffeomorphic to either $\mathbb{R}$ or $\mathbb{S}^1$. Conclude that the smooth structures on $\mathbb{R}$ and $\mathbb{S}^1$ are unique up to diffeomorphism. [Hint: First show that every connected smooth 1-manifold admits a nowhere-vanishing smooth vector field. The results of Problems 13-1 and 17-6 might be helpful.]

17-9. Let θ be a smooth flow on an oriented smooth manifold. Show that for each $t \in \mathbb{R}$, θ_t is orientation-preserving wherever it is defined.

17-10. Let M be a smooth manifold, and let $S \subset M$ be a compact embedded hypersurface. Suppose $N \in \mathcal{T}(M)$ is a smooth vector field that is nowhere tangent to S. Show that for some $\varepsilon > 0$, the flow of N

restricts to a diffeomorphism from $(-\varepsilon, \varepsilon) \times S$ to a neighborhood of S in M.

17-11. Let S be an embedded hypersurface in the smooth manifold M, and let N be a smooth vector field on M that is nowhere tangent to S. For any $f \in C^\infty(M)$ and $\varphi \in C^\infty(S)$, show that there is a neighborhood U of S in M and a unique function $u \in C^\infty(U)$ that satisfies

$$\begin{aligned} Nu &= f, \\ u|_S &= \varphi. \end{aligned}$$

[Remark: The equation $Nu = f$ is called a *first-order linear partial differential equation* for u—first-order because it involves only first derivatives of u, and linear because the left-hand side depends linearly on u. This problem shows that solving such a PDE reduces to solving a system of ODEs.]

17-12. Let M be a smooth manifold with boundary. An open set $U \subset M$ containing ∂M is called a *collar neighborhood* of ∂M if there exists a diffeomorphism from $[0,1) \times \partial M$ to U that restricts to the obvious identification $\{0\} \times \partial M \to \partial M$. If M is compact, show that ∂M has a collar neighborhood in M. [Hint: Use the flow of an inward-pointing vector field.]

17-13. Suppose M is a smooth n-manifold with boundary. The *double* of M is the quotient space of $M \amalg M$ obtained by identifying each point in the boundary of the first copy of M with the corresponding point in the boundary of the second copy. If M is compact, show that the double of M is a topological n-manifold (without boundary), and has a smooth structure such that M is a smoothly embedded submanifold with boundary. [Hint: Use Problem 17-12.]

17-14. Suppose M is a smooth manifold, $J \subset \mathbb{R}$ is an open interval, and $V \colon J \times M \to TM$ is a smooth time-dependent vector field on M. Let $\theta \colon \mathcal{E} \to M$ be the time-dependent flow of V. For any $(t,s) \in J \times J$, let $M_{t,s}$ denote the set $\{p \in M : (t,s,p) \in \mathcal{E}\}$, and define $\theta_{t,s} \colon M_{t,s} \to M$ by $\theta_{t,s}(p) = \theta(t,s,p)$.

(a) If $(t_1, t_0, p) \in \mathcal{E}$ and $(t_2, t_1, \theta_{t_1,t_0}(p)) \in \mathcal{E}$, show that $(t_2, t_0, p) \in \mathcal{E}$ and

$$\theta_{t_2,t_1} \circ \theta_{t_1,t_0}(p) = \theta_{t_2,t_0}(p).$$

(b) For any $(t,s) \in J \times J$, show that $M_{t,s}$ is open in M, and $\theta_{t,s} \colon M_{t,s} \to M_{s,t}$ is a diffeomorphism with inverse $\theta_{s,t}$.

(c) If M is compact, show that $\mathcal{E} = J \times J \times M$.

18
Lie Derivatives

This chapter is devoted to the study of a particularly important construction involving vector fields, called the Lie derivative. This is a method of computing the "directional derivative" of a vector field with respect to another vector field.

We already know how to make sense of directional derivatives of real-valued functions on a manifold. Indeed, a tangent vector $V \in T_pM$ is by definition an operator that acts on a smooth function f to give a number Vf that we interpret as a directional derivative of f at p. The discussion in Chapter 3 showed that this number can also be interpreted as the ordinary derivative of f along any curve whose initial tangent vector is V.

What about the directional derivative of a vector field? In Euclidean space, we can just differentiate the component functions of the vector field. But as we will see in this chapter, making sense of directional derivatives of a vector field W on a manifold is not as easy as it is in Euclidean space, because the values of W at different points lie in different tangent spaces, and thus cannot be compared directly. This problem can be circumvented if we replace the vector $V \in T_pM$ with a *vector field*. In this case, we can use the flow of the vector field to push values of W back to p and then differentiate. The result is called the Lie derivative of W with respect to the given vector field.

After defining Lie derivatives of vector fields, we show how the definition can be extended to tensor fields and differential forms. Then we explore a few important applications of Lie derivatives to the study of how geometric objects such as Riemannian metrics, volume forms, and symplectic forms behave under flows.

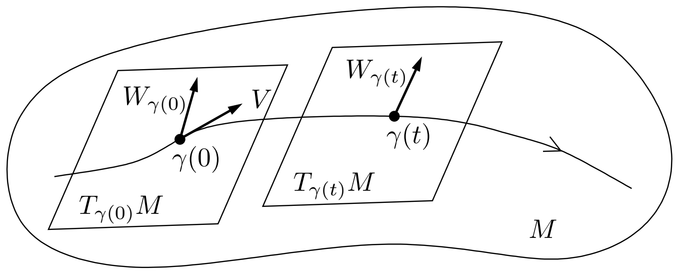


Figure 18.1. The problem with directional derivatives of vector fields.

The Lie Derivative

In Euclidean space, it makes perfectly good sense to define the directional derivative of a smooth vector field W in the direction of a vector $V \in T_p\mathbb{R}^n$—it is the vector

$$D_V W(p) = \left.\frac{d}{dt}\right|_{t=0} W_{p+tV} = \lim_{t\to 0} \frac{W_{p+tV} - W_p}{t}. \tag{18.1}$$

An easy calculation shows that $D_V W(p)$ can be evaluated by applying V to each component of W separately:

$$D_V W(p) = V W^i(p) \left.\frac{\partial}{\partial x^i}\right|_p.$$

Unfortunately, this definition is heavily dependent upon the fact that $\mathbb{R}^n$ is a vector space, so that the tangent vectors W_{p+tV} and W_p can both be viewed as elements of $\mathbb{R}^n$. If we search for a way to make invariant sense of (18.1) on a manifold, we will see very quickly what the problem is. To begin with, we can replace $p+tV$ by a curve $\gamma(t)$ that starts at p and whose initial tangent vector is V. But even with this substitution, the difference quotient still makes no sense because $W_{\gamma(t)}$ and $W_{\gamma(0)}$ are elements of different vector spaces ($T_{\gamma(t)}M$ and $T_{\gamma(0)}M$). (See Figure 18.1.) We got away with it in Euclidean space because there is a canonical identification of each tangent space with $\mathbb{R}^n$ itself; but on a manifold there is no such identification. Thus there is no coordinate-independent way to make sense of the directional derivative of W in the direction of the vector V.

Now suppose that V itself is a smooth vector field instead of a single vector. In this case, we can use the flow of V to push values of W back to p and then differentiate. Thus, for any smooth vector fields V and W on a manifold M, let θ be the flow of V, and define a vector $(\mathcal{L}_V W)_p$ at each $p \in M$, called the *Lie derivative* of W with respect to V at p, by

$$(\mathcal{L}_V W)_p = \left.\frac{d}{dt}\right|_{t=0} (\theta_{-t})_* W_{\theta_t(p)} = \lim_{t\to 0} \frac{(\theta_{-t})_* W_{\theta_t(p)} - W_p}{t}, \tag{18.2}$$

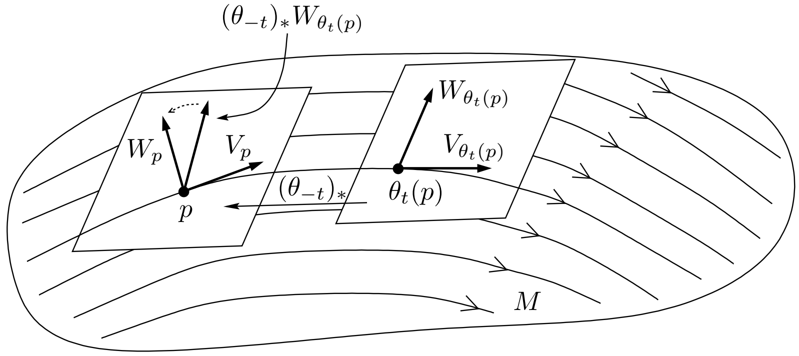


Figure 18.2. The Lie derivative of a vector field.

provided the derivative exists. For small $t \neq 0$, the difference quotient makes sense at least, because θ_t is defined in a neighborhood of p, and both $(\theta_{-t})_* W_{\theta_t(p)}$ and W_p are elements of T_pM (Figure 18.2).

Lemma 18.1. *If V and W are smooth vector fields on a smooth manifold M, then $(\mathcal{L}_V W)_p$ exists for every $p \in M$, and the assignment $p \mapsto (\mathcal{L}_V W)_p$ defines a smooth vector field.*

Proof. Let θ be the flow of V. For arbitrary $p \in M$, let $(U, (x^i))$ be a smooth coordinate chart containing p. Choose an open interval J_0 containing 0 and an open set $U_0 \subset U$ containing p such that θ maps $J_0 \times U_0$ into U. For $(t, x) \in J_0 \times U_0$, we can write the component functions of θ as $(\theta^1(t, x), \dots, \theta^n(t, x))$. Then for any $(t, x) \in J_0 \times U_0$, the matrix of $(\theta_{-t})_* \colon T_{\theta_t(x)}M \to T_xM$ is

$$\left(\frac{\partial \theta^i(-t, \theta(t, x))}{\partial x^j} \right).$$

Therefore,

$$(\theta_{-t})_* W_{\theta_t(x)} = \frac{\partial \theta^i(-t, \theta(t, x)))}{\partial x^j} W^j(\theta(t, x)) \left. \frac{\partial}{\partial x^i} \right|_x .$$

Because θ^i and W^j are smooth, the coefficient of $\partial/\partial x^i|_x$ depends smoothly on (t, x). It follows that $(\mathcal{L}_V W)_x$, which is obtained by taking the derivative of this expression with respect to t and setting $t = 0$, exists for each $x \in U_0$ and depends smoothly on x. □

◇ **Exercise 18.1.** If $V \in \mathbb{R}^n$ and W is a smooth vector field on an open subset of $\mathbb{R}^n$, show that the directional derivative $D_V W(p)$ defined by (18.1) is equal to $(\mathcal{L}_{\widetilde{V}} W)_p$, where $\widetilde{V}$ is the vector field $\widetilde{V} = V^i \partial/\partial x^i$ with constant coefficients in standard coordinates.

Thanks to Lemma 18.1, the assignment $p \mapsto (\mathcal{L}_V W)_p$ is a smooth vector field on M, which we denote by $\mathcal{L}_V W$. The definition of $\mathcal{L}_V W$ is not very useful for computations, however, because for most vector fields the flow is difficult or impossible to write down explicitly. Fortunately, there is a simple formula for computing the Lie derivative without explicitly finding the flow.

Theorem 18.2. *For any smooth vector fields V and W on a smooth manifold M, $\mathcal{L}_V W = [V,W]$.*

Proof. Let $\mathcal{R}(V) \subset M$ be the set of regular points of V (the set of points $p \in M$ such that $V_p \neq 0$). Note that $\mathcal{R}(V)$ is open in M by continuity, and its closure is the support of V.

Step 1: $\mathcal{L}_V W = [V,W]$ *on* $\mathcal{R}(V)$. If $p \in \mathcal{R}(V)$, we can choose smooth coordinates (u^i) on a neighborhood of p in which V has the coordinate representation $V = \partial/\partial u^1$. In these coordinates, the flow of V is

$$\theta_t(u) = (u^1 + t, u^2, \dots, u^n).$$

Therefore, for each fixed t, the matrix of $(\theta_t)_*$ in these coordinates (the Jacobian matrix of θ_t) is the identity at every point. Consequently, for any $u \in U$,

$$\begin{aligned}(\theta_{-t})_* W_{\theta_t(u)} &= (\theta_{-t})_* \left(W^j(u^1+t, u^2, \dots, u^n) \left.\frac{\partial}{\partial u^j}\right|_{\theta_t(u)} \right) \\ &= W^j(u^1+t, u^2, \dots, u^n) \left.\frac{\partial}{\partial u^j}\right|_u .\end{aligned}$$

Using the definition of the Lie derivative,

$$\begin{aligned}(\mathcal{L}_V W)_u &= \left.\frac{d}{dt}\right|_{t=0} W^j(u^1+t, u^2, \dots, u^n) \left.\frac{\partial}{\partial u^j}\right|_u \\ &= \frac{\partial W^j}{\partial u^1}(u^1, \dots, u^n) \left.\frac{\partial}{\partial u^j}\right|_u .\end{aligned}$$

On the other hand, using formula (4.5) for the Lie bracket in coordinates, $[V,W]_u$ is easily seen to be equal to the same expression.

Step 2: $\mathcal{L}_V W = [V,W]$ *on* $\operatorname{supp} V$. Because $\operatorname{supp} V$ is the closure of $\mathcal{R}(V)$, this follows from step 1 by continuity.

Step 3: $\mathcal{L}_V W = [V,W]$ *on* $M \smallsetminus \operatorname{supp} V$. If $p \in M \smallsetminus \operatorname{supp} V$, then $V \equiv 0$ on a neighborhood of p. On the one hand, this implies that θ_t is equal to the identity map in a neighborhood of p for all t, so $(\theta_{-t})_* W_{\theta_t(p)} = W_p$, which implies $(\mathcal{L}_V W)_p = 0$. On the other hand, $[V,W]_p = 0$ by formula (4.5). $\square$

A number of properties of the Lie derivative now follow immediately from things we already know about Lie brackets.

Corollary 18.3. *Suppose $V, W, X \in \mathcal{T}(M)$ and $f \in C^\infty(M)$.*

(*a*) $\mathcal{L}_V W = -\mathcal{L}_W V$.

(*b*) $\mathcal{L}_V [W, X] = [\mathcal{L}_V W, X] + [W, \mathcal{L}_V X]$.

(*c*) $\mathcal{L}_{[V,W]} X = \mathcal{L}_V \mathcal{L}_W X - \mathcal{L}_W \mathcal{L}_V X$.

(*d*) $\mathcal{L}_V (fW) = (Vf)W + f\mathcal{L}_V W$.

(*e*) *If* $F\colon M \to N$ *is a diffeomorphism, then* $F_*(\mathcal{L}_V W) = \mathcal{L}_{F_*V} F_* W$.

◇ **Exercise 18.2.** Prove this corollary.

Theorem 18.2 finally gives us a long-awaited geometric interpretation of the Lie bracket of two vector fields: It is the directional derivative of the second vector field along the flow of the first.

Commuting Vector Fields

Two smooth vector fields are said to *commute* if $[V, W] \equiv 0$, or equivalently if $VWf = WVf$ for every smooth real-valued function f. One simple example of a pair of commuting vector fields is $\partial/\partial x^i$ and $\partial/\partial x^j$ in any smooth coordinate system: Because their component functions are constants, their Lie bracket is identically zero.

Recall that a vector field W is said to be invariant under a flow θ if $(\theta_t)_* W_p = W_{\theta_t(p)}$ for all (t, p) in the domain of θ. We will show that commuting vector fields are invariant under each other's flows. The key is the following somewhat more general result about F-related vector fields.

Lemma 18.4. *Suppose* $F\colon M \to N$ *is a smooth map,* $X \in \mathcal{T}(M)$*, and* $Y \in \mathcal{T}(N)$*, and let* θ *be the flow of* X *and* η *the flow of* Y*. Then* X *and* Y *are* F*-related if and only if for each* $t \in \mathbb{R}$*,* $\eta_t \circ F = F \circ \theta_t$ *on the domain of* θ_t*:*

$$\begin{tikzcd} M \arrow[r, "F"] \arrow[d, "\theta_t"'] & N \arrow[d, "\eta_t"] \\ M \arrow[r, "F"'] & N \end{tikzcd}$$

Proof. The commutativity of the diagram means that the following holds for all (t, p) in the domain of θ:

$$\eta_t \circ F(p) = F \circ \theta_t(p).$$

If we let $\mathcal{D}^{(p)} \subset \mathbb{R}$ denote the domain of $\theta^{(p)}$, this is equivalent to

$$\eta^{(F(p))}(t) = F \circ \theta^{(p)}(t), \quad t \in \mathcal{D}^{(p)}. \tag{18.3}$$

Suppose first that X and Y are F-related. If we define $\gamma\colon \mathcal{D}^{(p)} \to N$ by $\gamma = F \circ \theta^{(p)}$ (see Figure 18.3), then

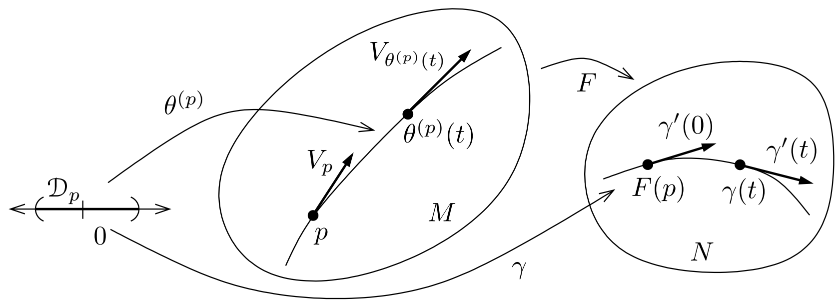


Figure 18.3. Flows of F-related vector fields.

$$\begin{aligned}\gamma'(t) &= (F\circ\theta^{(p)})'(t)\\ &= F_*(\theta^{(p)\prime}(t))\\ &= F_*X_{\theta^{(p)}(t)}\\ &= Y_{F\circ\theta^{(p)}(t)}\\ &= Y_{\gamma(t)},\end{aligned}$$

so γ is an integral curve of Y starting at $F\circ\theta^{(p)}(0) = F(p)$. By uniqueness of integral curves, therefore, the maximal integral curve $\eta^{(F(p))}$ must be defined at least on the interval $\mathcal{D}^{(p)}$, and $\gamma(t) = \eta^{(F(p))}(t)$ on that interval. This proves (18.3).

Conversely, if (18.3) holds, then for each $p \in M$ we have

$$\begin{aligned}F_*X_p &= F_*(\theta^{(p)\prime}(0))\\ &= (F\circ\theta^{(p)})'(0)\\ &= \eta^{(F(p))\prime}(0)\\ &= Y_{F(p)},\end{aligned}$$

which shows that X and Y are F-related. □

The next proposition gives several useful characterizations of what it means for two vector fields to commute.

Proposition 18.5. *Let V and W be smooth vector fields on M, with flows θ and ψ, respectively. The following are equivalent:*

(*a*) $[V, W] = 0$.

(*b*) $\mathcal{L}_V W = 0$.

(*c*) $\mathcal{L}_W V = 0$.

(*d*) *W is invariant under the flow of V.*

(*e*) *V is invariant under the flow of W.*

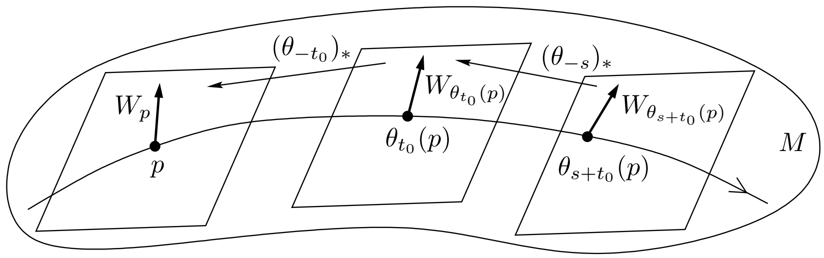


Figure 18.4. Proof of formula (18.5).

(f) $\theta_t \circ \psi_s = \psi_s \circ \theta_t$ *wherever either side is defined.*

Proof. Clearly (a), (b), and (c) are equivalent because $\mathcal{L}_V W = [V, W] = -\mathcal{L}_W V$. Part (d) means by definition that $W_{\theta_t(p)} = (\theta_t)_* W_p$ whenever (t, p) is in the domain of θ. Applying $(\theta_{-t})_*$ to both sides, we conclude that $(\theta_{-t})_* W_{\theta_t(p)} = W_p$, which obviously implies (b) directly from the definition of $\mathcal{L}_V W$. The same argument shows that (e) implies (b).

To prove that (b) implies (d), let $p \in M$ be arbitrary, let $\mathcal{D}^{(p)} \subset \mathbb{R}$ denote the domain of the integral curve $\theta^{(p)}$, and consider the map $X \colon \mathcal{D}^{(p)} \to T_p M$ given by the time-dependent vector

$$X(t) = (\theta_{-t})_* \left(W_{\theta_t(p)} \right) \in T_p M. \tag{18.4}$$

This can be considered as a smooth curve in the vector space $T_p M$. We will show that $X(t)$ is independent of t. Since $X(0) = W_p$, this implies that $X(t) = W_p$ for all $t \in \mathcal{D}^{(p)}$, which says that W is invariant under θ.

The assumption that $\mathcal{L}_V W = 0$ means precisely that the t-derivative of (18.4) is zero when $t = 0$; we need to show that this derivative is zero for all values of t. Making the change of variables $t = t_0 + s$, we obtain

$$\begin{aligned} X'(t_0) &= \left.\frac{d}{dt}\right|_{t=t_0} (\theta_{-t})_* W_{\theta_t(p)} \\ &= \left.\frac{d}{ds}\right|_{s=0} (\theta_{-t_0-s})_* W_{\theta_{s+t_0}(p)} \\ &= \left.\frac{d}{ds}\right|_{s=0} (\theta_{-t_0})_* (\theta_{-s})_* W_{\theta_s(\theta_{t_0}(p))} \\ &= (\theta_{-t_0})_* \left.\frac{d}{ds}\right|_{s=0} (\theta_{-s})_* W_{\theta_s(\theta_{t_0}(p))} \\ &= (\theta_{-t_0})_* (\mathcal{L}_V W)_{\theta_{t_0}(p)} = 0. \end{aligned} \tag{18.5}$$

(See Figure 18.4. The equality on the next-to-last line follows because $(\theta_{-t_0})_* \colon T_{\theta_{t_0}(p)} M \to T_p M$ is a linear map that is independent of s.) The same proof also shows that (b) implies (e).

Finally, we will show that (e) is equivalent to (f). Assume first that (e) holds. For each $s \in \mathbb{R}$, let $M_s \subset M$ denote the domain of ψ_s. The assumption that V is invariant under ψ_s can be rephrased as saying that $V|_{M_s}$ is ψ_s-related to $V|_{M_{-s}}$. Thus Lemma 18.4 applied with $F = \psi_s \colon M_s \to M_{-s}$, $X = V|_{M_s}$, and $Y = V|_{M_{-s}}$ implies that $\theta_t \circ \psi_s = \psi_s \circ \theta_t$ on the set where $\psi_s \circ \theta_t$ is defined. Since (e) implies (d), the same argument with V and W reversed shows that this also holds wherever $\theta_t \circ \psi_s$ is defined. Conversely, if (f) holds, then Lemma 18.4 shows that $V|_{M_s}$ is ψ_s-related to $V|_{M_{-s}}$ for each s, which is the same as saying V is invariant under the flow of W. □

As we mentioned above, the coordinate vector fields in any smooth chart commute with each other. The next theorem shows that any set of smooth, independent, commuting vector fields can be written locally as coordinate vector fields with respect to a suitable choice of coordinates.

Theorem 18.6 (Canonical Form for Commuting Vector Fields). *Let M be a smooth n-manifold, and let $V_1, \dots, V_k$ be smooth independent vector fields on an open subset of M. Then the following are equivalent:*

(*a*) *There exist smooth coordinates $(u^1, \dots, u^n)$ in a neighborhood of each point such that $V_i = \partial/\partial u^i$, $i = 1, \dots, k$.*

(*b*) *$[V_i, V_j] \equiv 0$ for all i and j.*

Proof. The fact that (a) implies (b) is obvious because the coordinate vector fields commute and the Lie bracket is coordinate-independent.

To prove the converse, suppose $V_1, \dots, V_k$ are independent smooth commuting vector fields on an open subset $U \subset M$, and let $p \in U$. The basic outline of the proof is analogous to that of the canonical form theorem for one vector field near a regular point (Theorem 17.13), except that we have to do a bit of extra work to make use of the hypothesis that the vector fields commute.

Choose a smooth chart $(U, (x^i))$ centered at p. By rearranging the coordinates if necessary, we may assume that none of the vectors $V_1|_p, \dots, V_k|_p$ lie in the subspace of T_pM spanned by $\partial/\partial x^{k+1} \dots, \partial/\partial x^n$. Let θ_i denote the flow of V_i for $i = 1, \dots, k$. There exists $\varepsilon > 0$ and a neighborhood W of p such that the composition $(\theta_k)_{t_k} \circ (\theta_{k-1})_{t_{k-1}} \circ \dots \circ (\theta_1)_{t_1}$ is defined on W and maps W into U whenever $|t_1|, \dots, |t_k|$ are all less than ε. (Just choose $\varepsilon_1 > 0$ and $U_1 \subset U$ such that θ_1 maps $(-\varepsilon_1, \varepsilon_1) \times U_1$ into U, and then inductively choose ε_i and U_i such that θ_i maps $(-\varepsilon_i, \varepsilon_i) \times U_i$ into U_{i-1}. Taking $\varepsilon = \min\{\varepsilon_i\}$ and $W = U_k$ does the trick.)

As in the proof of Theorem 17.13, let

$$S = \{(u^{k+1}, \dots, u^n) : (0, \dots, 0, u^{k+1}, \dots, u^n) \in W\},$$

and define $\psi\colon (-\varepsilon,\varepsilon)^k \times S \to U$ by

$$\begin{aligned}\psi(u^1,\ldots,u^k,u^{k+1},\ldots,u^n)&\\ &= (\theta_1)_{u^1}\circ\cdots\circ(\theta_k)_{u^k}(0,\ldots,0,u^{k+1},\ldots,u^n).\end{aligned}$$

We will show first that

$$\psi_*\frac{\partial}{\partial u^i} = V_i, \quad i = 1,\ldots,k.$$

Because all of the flows θ_i commute with each other, for any $i\in\{1,\ldots,k\}$ and any $u_0\in W$ we have

$$\begin{aligned}\left(\psi_*\frac{\partial}{\partial u^i}\bigg|_{u_0}\right)f &= \frac{\partial}{\partial u^i}\bigg|_{u_0} f(\psi(u^1,\ldots,u^n))\\ &= \frac{\partial}{\partial u^i}\bigg|_{u_0} f((\theta_1)_{u^1}\circ\cdots\circ(\theta_k)_{u^k}(0,\ldots,0,u^{k+1},\ldots,u^n))\\ &= \frac{\partial}{\partial u^i}\bigg|_{u_0} f((\theta_i)_{u^i}\circ(\theta_1)_{u^1}\circ\cdots\circ(\theta_{i-1})_{u^{i-1}}\circ(\theta_{i+1})_{u^{i+1}}\\ &\qquad\qquad\circ\cdots\circ(\theta_k)_{u^k}(0,\ldots,0,u^{k+1},\ldots,u^n)).\end{aligned}$$

Now, for any $q\in M$, $t\mapsto(\theta_i)_t(q)$ is an integral curve of V_i, so this last expression is equal to $V_i|_{\psi(u_0)}f$, which proves the claim.

Next we will show that ψ_* is invertible at $u=0$. The computation above shows that for $i=1,\ldots,k$,

$$\psi_*\left.\frac{\partial}{\partial u^i}\right|_0 = V_p.$$

On the other hand, since

$$\psi(0,\ldots,0,u^{k+1},\ldots,u^n) = (0,\ldots,0,u^{k+1},\ldots,u^n),$$

it follows immediately that

$$\psi_*\left.\frac{\partial}{\partial u^i}\right|_0 = \left.\frac{\partial}{\partial x^i}\right|_p$$

for $i=k+1,\ldots,n$. Thus ψ_* takes the basis $(\partial/\partial u^1|_0,\ldots,\partial/\partial u^n|_0)$ to $(V_1|_p,\ldots,V_k|_p,\partial/\partial x^{k+1}|_p,\ldots,\partial/\partial x^n|_p)$, which is also a basis. Thus by the inverse function theorem, ψ is a diffeomorphism in a neighborhood of 0, and $\varphi=\psi^{-1}$ is the desired coordinate map. □

Just as in the case of a single vector field, the proof of this theorem suggests a technique for finding explicit coordinates that put k given independent commuting vector fields into canonical form, as long as their flows can be found explicitly. The method can be summarized as follows: Begin with an $(n-k)$-dimensional coordinate submanifold S to which none of the vectors $V_1,\ldots,V_k$ are tangent. Then define ψ by starting at an arbitrary point in S and following the k flows successively for k arbitrary times. Because the flows commute, it does not matter in which order they are applied. An example will help to clarify the procedure.

Example 18.7. Consider the following two vector fields on $\mathbb{R}^2$:

$$V = x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x},$$
$$W = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}.$$

A simple computation shows that $[V, W] = 0$. We already showed in Example 17.4 that the flow of V is

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

and an easy verification shows that the flow of W is

$$\eta_s(x, y) = (e^s x, e^s y).$$

At $p = (1, 0)$, V_p and W_p are independent. Because $k = n = 2$ in this case, we can take the subset S to be the single point $\{(1, 0)\}$, and define $\psi\colon \mathbb{R}^2 \to \mathbb{R}^2$ by

$$\psi(u, v) = \eta_u \circ \theta_v(1, 0) = (e^u \cos v, e^u \sin v).$$

In this case, we can solve for $(u, v) = \psi^{-1}(x, y)$ explicitly in a neighborhood of $(1, 0)$ to obtain the coordinate map

$$(u, v) = (\log\sqrt{x^2 + y^2}, \tan^{-1}(y/x)).$$

Lie Derivatives of Tensor Fields

The Lie derivative operation can be extended to tensor fields of arbitrary rank. As usual, we focus on covariant tensors; the analogous results for contravariant or mixed tensors require only minor modifications.

Let X be a smooth vector field on a smooth manifold M, and let θ be its flow. For any $p \in M$, if t is sufficiently close to zero, θ_t is a diffeomorphism from a neighborhood of p to a neighborhood of $\theta_t(p)$, so θ_t^* pulls back tensors at $\theta_t(p)$ to ones at p.

Given a smooth covariant tensor field τ on M, we define the *Lie derivative* of τ with respect to X, denoted by $\mathcal{L}_X\tau$, as

$$(\mathcal{L}_X\tau)_p = \left.\frac{d}{dt}\right|_{t=0} (\theta_t^*\tau)_p = \lim_{t\to 0}\frac{\theta_t^*(\tau_{\theta_t(p)}) - \tau_p}{t}, \tag{18.6}$$

provided the derivative exists (Figure 18.5). Because the expression being differentiated lies in $T^k(T_pM)$ for all t, $(\mathcal{L}_X\tau)_p$ makes sense as an element

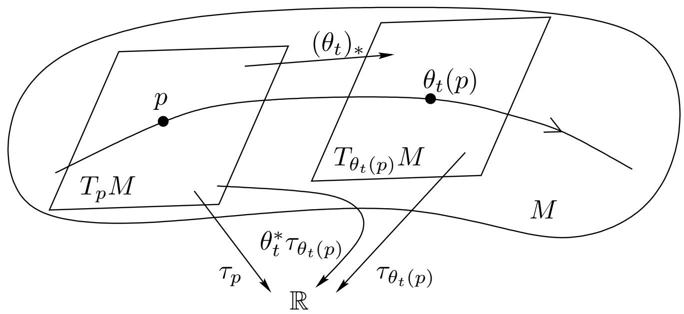


Figure 18.5. The Lie derivative of a tensor field.

of $T^k(T_pM)$. The following lemma is an analogue of Lemma 18.1, and is proved in exactly the same way.

Lemma 18.8. *If X is a smooth vector field and τ is a smooth covariant tensor field on M, then the derivative in* (18.6) *exists for every $p \in M$ and defines $\mathcal{L}_X\tau$ as a smooth tensor field on M.*

◇ **Exercise 18.3.** Prove the preceding lemma.

Proposition 18.9. *Suppose X, Y are smooth vector fields; f is a smooth real-valued function* (*regarded as a* 0*-tensor field*)*; σ, τ are smooth covariant tensor fields; and ω, η are smooth differential forms.*

(*a*) $\mathcal{L}_X f = Xf$.

(*b*) $\mathcal{L}_X(f\sigma) = (\mathcal{L}_X f)\sigma + f\mathcal{L}_X\sigma$.

(*c*) $\mathcal{L}_X(\sigma \otimes \tau) = (\mathcal{L}_X\sigma) \otimes \tau + \sigma \otimes \mathcal{L}_X\tau$.

(*d*) $\mathcal{L}_X(\omega \wedge \eta) = (\mathcal{L}_X\omega) \wedge \eta + \omega \wedge \mathcal{L}_X\eta$.

(*e*) $\mathcal{L}_X(Y \lrcorner\, \omega) = (\mathcal{L}_X Y) \lrcorner\, \omega + Y \lrcorner\, \mathcal{L}_X\omega$.

(*f*) *If $Y_1, \dots, Y_k$ are smooth vector fields and σ is a smooth k-tensor field,*

$$\begin{aligned}\mathcal{L}_X(\sigma(Y_1, \dots, Y_k)) = (\mathcal{L}_X\sigma)(Y_1, \dots, Y_k) &+ \sigma(\mathcal{L}_X Y_1, \dots, Y_k) \\ &+ \dots + \sigma(Y_1, \dots, \mathcal{L}_X Y_k). \quad (18.7)\end{aligned}$$

Proof. The first assertion is just a reinterpretation of the definition in the case of a 0-tensor field: Because $\theta_t^* f = f \circ \theta_t$, the definition implies

$$\mathcal{L}_X f(p) = \left.\frac{d}{dt}\right|_{t=0} f(\theta_t(p)) = Xf(p).$$

The proofs of (b), (c), (d), (e), and (f) are essentially the same, so we will prove (c) and leave the others to you.

$$\begin{aligned}
(\mathcal{L}_X(\sigma\otimes\tau))_p &= \lim_{t\to 0}\frac{\theta_t^*((\sigma\otimes\tau)_{\theta_t(p)})-(\sigma\otimes\tau)_p}{t}\\
&= \lim_{t\to 0}\frac{\theta_t^*(\sigma_{\theta_t(p)})\otimes\theta_t^*(\tau_{\theta_t(p)})-\sigma_p\otimes\tau_p}{t}\\
&= \lim_{t\to 0}\frac{\theta_t^*(\sigma_{\theta_t(p)})\otimes\theta_t^*(\tau_{\theta_t(p)})-\theta_t^*(\sigma_{\theta_t(p)})\otimes\tau_p}{t}\\
&\qquad+\lim_{t\to 0}\frac{\theta_t^*(\sigma_{\theta_t(p)})\otimes\tau_p-\sigma_p\otimes\tau_p}{t}\\
&= \lim_{t\to 0}\theta_t^*(\sigma_{\theta_t(p)})\otimes\frac{\theta_t^*(\tau_{\theta_t(p)})-\tau_p}{t}+\lim_{t\to 0}\frac{\theta_t^*(\sigma_{\theta_t(p)})-\sigma_p}{t}\otimes\tau_p\\
&= \sigma_p\otimes(\mathcal{L}_X\tau)_p+(\mathcal{L}_X\sigma)_p\otimes\tau_p.
\end{aligned}$$

The other parts are similar, and are left as an exercise. □

◇ **Exercise 18.4.** Complete the proof of the preceding proposition.

One consequence of this proposition is the following formula expressing the Lie derivative of any smooth covariant tensor field in terms of Lie brackets and ordinary directional derivatives of functions.

Corollary 18.10. *If X is a smooth vector field and σ is a smooth covariant k-tensor field, then for any smooth vector fields $Y_1,\dots,Y_k$,*

$$\begin{aligned}
(\mathcal{L}_X\sigma)(Y_1,\dots,Y_k) = X(\sigma(Y_1,\dots,Y_k)) - \sigma([X,Y_1],Y_2,\dots,Y_k) - \dots\\
- \sigma(Y_1,\dots,Y_{k-1},[X,Y_k]). \quad (18.8)
\end{aligned}$$

Proof. This formula is obtained simply by solving (18.7) for $\mathcal{L}_X\sigma$, and replacing $\mathcal{L}_Xf$ by Xf and $\mathcal{L}_XY_i$ by $[X,Y_i]$. □

Corollary 18.11. *If $f\in C^\infty(M)$, then $\mathcal{L}_X(df)=d(\mathcal{L}_Xf)$.*

Proof. Using (18.8), we compute

$$\begin{aligned}
(\mathcal{L}_X df)(Y) &= X(df(Y)) - df[X,Y]\\
&= XYf - [X,Y]f\\
&= XYf - (XYf - YXf)\\
&= YXf\\
&= d(Xf)(Y)\\
&= d(\mathcal{L}_Xf)(Y).
\end{aligned}$$
□

Proposition 18.9 and Corollary 18.11 lead to an easy method for computing Lie derivatives of smooth tensor fields in coordinates, since any tensor field can be written locally as a linear combination of functions multiplied by tensor products of exact 1-forms. One drawback of formula (18.8) is that in order to calculate what $\mathcal{L}_X\sigma$ does to vectors $Y_1, \dots, Y_k$ at a point $p \in M$, one must first extend them to vector fields in a neighborhood of p. The next example illustrates a technique that avoids this problem.

Example 18.12. Suppose T is an arbitrary smooth 2-tensor field on a smooth manifold M, and let Y be a smooth vector field. We will compute the Lie derivative $\mathcal{L}_Y T$ in smooth local coordinates (x^i). First, we observe that $\mathcal{L}_Y dx^i = d(\mathcal{L}_Y x^i) = d(Yx^i) = dY^i$. Therefore,

$$\begin{aligned}
\mathcal{L}_Y T &= \mathcal{L}_Y (T_{ij} dx^i \otimes dx^j) \\
&= \mathcal{L}_Y (T_{ij}) dx^i \otimes dx^j + T_{ij} (\mathcal{L}_Y dx^i) \otimes dx^j + T_{ij}\, dx^i \otimes (\mathcal{L}_Y dx^j) \\
&= YT_{ij}\, dx^i \otimes dx^j + T_{ij}\, dY^i \otimes dx^j + T_{ij}\, dx^i \otimes dY^j \\
&= \left(YT_{ij} + T_{kj} \frac{\partial Y^k}{\partial x^i} + T_{ik} \frac{\partial Y^k}{\partial x^j} \right) dx^i \otimes dx^j.
\end{aligned}$$

Differential Forms

In the case of differential forms, the exterior derivative yields a much more powerful formula for computing Lie derivatives, which also has significant theoretical consequences.

Proposition 18.13 (Cartan's Formula). *For any smooth vector field X and any smooth differential form ω,*

$$\mathcal{L}_X \omega = X \lrcorner (d\omega) + d(X \lrcorner \omega). \tag{18.9}$$

Proof. We will prove that (18.9) holds for smooth k-forms by induction on k. We begin with a smooth 0-form f, in which case

$$X \lrcorner (df) + d(X \lrcorner f) = X \lrcorner df = df(X) = Xf = \mathcal{L}_X f,$$

which is (18.9).

Any smooth 1-form can be written locally as a sum of terms of the form $u\, dv$ for smooth functions u and v, so to prove (18.9) for 1-forms, it suffices to consider the case $\omega = u\, dv$. In this case, using Proposition 18.9(a,b) and Corollary 18.11, the left-hand side of (18.9) reduces to

$$\begin{aligned}
\mathcal{L}_X (u\, dv) &= (\mathcal{L}_X u) dv + u(\mathcal{L}_X dv) \\
&= (Xu) dv + u\, d(Xv).
\end{aligned}$$

On the other hand, using the fact that interior multiplication is an antiderivation, the right-hand side is

$$\begin{aligned}
X \lrcorner d(u\, dv) + d(X \lrcorner (u\, dv)) &= X \lrcorner (du \wedge dv) + d(uXv) \\
&= (X \lrcorner du) \wedge dv - du \wedge (X \lrcorner dv) \\
&\qquad + u\, d(Xv) + (Xv) du \\
&= (Xu) dv - (Xv) du + u\, d(Xv) + (Xv) du.
\end{aligned}$$

(Remember that $X \lrcorner du = du(X) = Xu$, and a wedge product with a 0-form is just ordinary multiplication.) After canceling the two $(Xv)du$ terms, this is equal to $\mathcal{L}_X(u\,dv)$.

Now let $k > 1$, and suppose (18.9) has been proved for forms of degree less than k. Let ω be an arbitrary smooth k-form, written in smooth local coordinates as

$$\omega = \sum_I{}' \omega_I \, dx^{i_1} \wedge \cdots \wedge dx^{i_k}.$$

Writing $\alpha = \omega_I \, dx^{i_1}$ and $\beta = dx^{i_2} \wedge \cdots \wedge dx^{i_k}$, we see that ω can be written as a sum of terms of the form $\alpha \wedge \beta$, where α is a smooth 1-form and β is a smooth $(k-1)$-form. For such a term, Proposition 18.9(d) and the induction hypothesis imply

$$\begin{aligned} \mathcal{L}_X(\alpha \wedge \beta) &= (\mathcal{L}_X \alpha) \wedge \beta + \alpha \wedge (\mathcal{L}_X \beta) \\ &= (X \lrcorner d\alpha + d(X \lrcorner \alpha)) \wedge \beta + \alpha \wedge (X \lrcorner d\beta + d(X \lrcorner \beta)). \end{aligned} \tag{18.10}$$

On the other hand, using the fact that both d and interior multiplication by X are antiderivations, we compute

$$\begin{aligned} &X \lrcorner d(\alpha \wedge \beta) + d(X \lrcorner (\alpha \wedge \beta)) \\ &\qquad = X \lrcorner (d\alpha \wedge \beta - \alpha \wedge d\beta) + d((X \lrcorner \alpha) \wedge \beta - \alpha \wedge (X \lrcorner \beta)) \\ &\qquad = (X \lrcorner d\alpha) \wedge \beta + d\alpha \wedge (X \lrcorner \beta) - (X \lrcorner \alpha) \wedge d\beta \\ &\qquad\quad + \alpha \wedge (X \lrcorner d\beta) + d(X \lrcorner \alpha) \wedge \beta + (X \lrcorner \alpha) \wedge d\beta \\ &\qquad\quad - d\alpha \wedge (X \lrcorner \beta) + \alpha \wedge d(X \lrcorner \beta). \end{aligned}$$

After the obvious cancellations are made, this is equal to (18.10). □

Corollary 18.14 (The Lie Derivative Commutes with d). *If X is a smooth vector field and ω is a smooth differential form, then*

$$\mathcal{L}_X(d\omega) = d(\mathcal{L}_X \omega).$$

Proof. This follows from the preceding proposition and the fact that $d \circ d = 0$:

$$\begin{aligned} \mathcal{L}_X d\omega &= X \lrcorner d(d\omega) + d(X \lrcorner d\omega) \\ &= d(X \lrcorner d\omega); \\ d\mathcal{L}_X \omega &= d(X \lrcorner d\omega + d(X \lrcorner \omega)) \\ &= d(X \lrcorner d\omega). \end{aligned}$$

□

Applications to Geometry

What is the geometric meaning of the Lie derivative of a tensor field with respect to a vector field X? We have already seen that the Lie derivative of a vector field Y with respect to X is zero if and only if Y is invariant under the flow of X. It turns out that the Lie derivative of a covariant tensor field has exactly the same interpretation. We say that a tensor field τ is invariant under a flow θ if

$$\theta_t^*(\tau_{\theta_t(p)}) = \tau_p \tag{18.11}$$

for all (t, p) in the domain of θ.

The next lemma shows how the Lie derivative can be used to compute t-derivatives at times other than $t = 0$; it is a generalization to tensor fields of formula (18.5).

Lemma 18.15. *Let M be a smooth manifold, $X \in \mathfrak{T}(M)$, and let θ be the flow of X. For any smooth covariant tensor field τ and any (t_0, p) in the domain of θ,*

$$\left.\frac{d}{dt}\right|_{t=t_0} \theta_t^*(\tau_{\theta_t(p)}) = \theta_{t_0}^*\big((\mathcal{L}_X\tau)_{\theta_{t_0}(p)}\big).$$

Proof. Just as in the proof of Proposition 18.5, the change of variables $t = t_0 + s$ yields

$$\begin{aligned}
\left.\frac{d}{dt}\right|_{t=t_0} \theta_t^*(\tau_{\theta_t(p)}) &= \left.\frac{d}{ds}\right|_{s=0} (\theta_{t_0+s})^*\tau_{\theta_{s+t_0}(p)} \\
&= \left.\frac{d}{ds}\right|_{s=0} (\theta_{t_0})^*(\theta_s)^*\tau_{\theta_s(\theta_{t_0}(p))} \\
&= (\theta_{t_0})^* \left.\frac{d}{ds}\right|_{s=0} (\theta_s)^*\tau_{\theta_s(\theta_{t_0}(p))} \\
&= (\theta_{t_0})^*\big((\mathcal{L}_X\tau)_{\theta_{t_0}(p)}\big).
\end{aligned}$$

□

Proposition 18.16. *Let M be a smooth manifold and let $X \in \mathfrak{T}(M)$. A smooth covariant tensor field τ is invariant under the flow of X if and only if $\mathcal{L}_X\tau = 0$.*

Proof. Let θ denote the flow of X. If τ is invariant under θ, then inserting 18.11 into the definition of the Lie derivative, we see immediately that $\mathcal{L}_X\tau = 0$.

Conversely, suppose $\mathcal{L}_X\tau = 0$. For any $p \in M$, let $\mathcal{D}^{(p)}$ denote the domain of $\theta^{(p)}$, and consider the smooth curve $T\colon \mathcal{D}^{(p)} \to T^k(T_pM)$ defined by

$$T(t) = \theta_t^*(\tau_{\theta_t(p)}).$$

Lemma 18.15 shows that $T'(t) = 0$ for all $t \in \mathcal{D}^{(p)}$. Because $\mathcal{D}^{(p)}$ is a connected interval containing zero, this implies that $T(t) = T(0) = \tau_p$ for all $t \in \mathcal{D}^{(p)}$, which is the same as 18.11. □

Killing Fields

Let (M, g) be a Riemannian manifold. A smooth vector field Y on M is called a *Killing field* for g if g is invariant under the flow of Y. By Proposition 18.16, this is the case if and only if $\mathcal{L}_Y g = 0$.

Example 18.12 applied to the case $T = g$ gives the following coordinate expression for $\mathcal{L}_Y g$:

$$(\mathcal{L}_Y g)\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) = Y g_{ij} + g_{jk}\frac{\partial Y^k}{\partial x^i} + g_{ik}\frac{\partial Y^k}{\partial x^j}. \tag{18.12}$$

Example 18.17 (Euclidean Killing Fields). Let $\overline{g}$ be the Euclidean metric on $\mathbb{R}^n$. In standard coordinates, the condition for a vector field to be a Killing field with respect to $\overline{g}$ reduces to

$$\frac{\partial Y^j}{\partial x^i} + \frac{\partial Y^i}{\partial x^j} = 0.$$

It is easy to check that all constant-coefficient vector fields satisfy this equation, as do the vector fields

$$x^i \frac{\partial}{\partial x^j} - x^j \frac{\partial}{\partial x^i},$$

which generate rotations in the (x^i, x^j)-plane.

The Divergence

For our next application, we let (M, g) be an oriented Riemannian n-manifold. Recall that the *divergence* of a smooth vector field $X \in \mathcal{T}(M)$ is the smooth real-valued function $\operatorname{div} X$ characterized by

$$(\operatorname{div} X)dV_g = d(X \lrcorner\, dV_g).$$

Now we can give a geometric interpretation to the divergence, which explains the origin of the term "divergence." Observe that formula (18.9) for the Lie derivative of a differential form implies

$$\mathcal{L}_X dV_g = X \lrcorner\, d(dV_g) + d(X \lrcorner\, dV_g) = (\operatorname{div} X)dV_g,$$

because the exterior derivative of any smooth n-form on an n-manifold is zero.

A smooth flow θ on M is said to be *volume preserving* if for every compact domain of integration $D \subset M$ and every $t \in \mathbb{R}$ such that D is contained in the domain of θ_t, $\operatorname{Vol}(\theta_t(D)) = \operatorname{Vol}(D)$. It is *volume increasing* if for any such D with positive volume, $\operatorname{Vol}(\theta_t(D))$ is a strictly increasing function of t, and *volume decreasing* if it is strictly decreasing. Note that the properties of flow domains ensure that, if D is contained in the domain of θ_t for some t, then the same is true for all times between 0 and t. The next proposition shows that the divergence of a vector field can be interpreted as a measure of the tendency of its flow to "spread out" or diverge (see Figure 18.6).

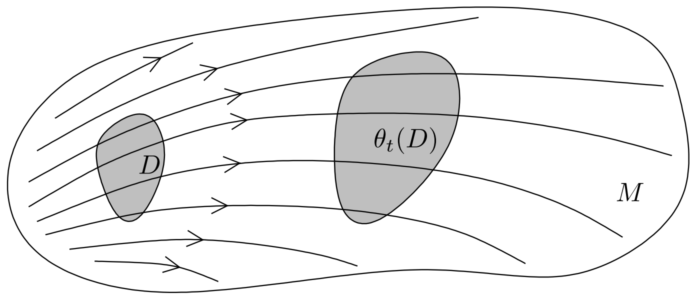


Figure 18.6. Geometric interpretation of the divergence.

Proposition 18.18. *Let M be an oriented Riemannian manifold and let $X \in \mathfrak{T}(M)$.*

(*a*) *The flow of X is volume preserving if and only if* $\operatorname{div} X \equiv 0$.

(*b*) *If* $\operatorname{div} X > 0$, *then the flow of X is volume increasing, and if* $\operatorname{div} X < 0$, *then it is volume decreasing.*

Proof. Let θ be the flow of X, and for each t let M_t be the domain of θ_t. If D is a compact domain of integration contained in M_t, then

$$\operatorname{Vol}(\theta_t(D)) = \int_{\theta_t(D)} dV_g = \int_D \theta_t^* dV_g.$$

Because the integrand is a smooth function of t, we can differentiate this expression with respect to t by differentiating under the integral sign. (Strictly speaking, we should use a partition of unity to express the integral as a sum of integrals over domains in $\mathbb{R}^n$, and then differentiate under the integral signs there. The details are left to you.) Using Lemma 18.15, we obtain

$$\begin{aligned} \left.\frac{d}{dt}\right|_{t=t_0} \operatorname{Vol}(\theta_t(D)) &= \int_D \left.\frac{\partial}{\partial t}\right|_{t=t_0} (\theta_t^* dV_g) \\ &= \int_D \theta_{t_0}^* (\mathcal{L}_X dV_g) \\ &= \int_D \theta_{t_0}^* ((\operatorname{div} X) dV_g) \\ &= \int_{\theta_{t_0}(D)} (\operatorname{div} X) dV_g. \end{aligned}$$

It follows that $\operatorname{div} X \equiv 0$ implies that $\operatorname{Vol}(\theta_t(D))$ is a constant function of t, while $\operatorname{div} X > 0$ or $\operatorname{div} X < 0$ implies that it is strictly increasing or strictly decreasing, respectively.

Now assume that θ is volume preserving. If $\operatorname{div} X \neq 0$ at some point $p \in M$, then there is some open set U containing p on which $\operatorname{div} X$ does not change sign. If $\operatorname{div} X > 0$ on U, then X generates a volume increasing flow on U by the argument above. In particular, for any coordinate ball B such that $\overline{B} \subset U$ and any $t > 0$ sufficiently small that $\theta_t(B) \subset U$, we have $\operatorname{Vol}(\theta_t(B)) > \operatorname{Vol}(B)$, which contradicts the assumption that θ is volume preserving. The argument in the case $\operatorname{div} X < 0$ is exactly analogous. Therefore $\operatorname{div} X \equiv 0$. □

Applications to Symplectic Manifolds

Let (M, ω) be a symplectic manifold. (Recall that this means a smooth manifold M endowed with a symplectic form ω, which is a closed nondegenerate 2-form.) Our next theorem is one of the most fundamental results in the theory of symplectic structures. It is a nonlinear analogue of the canonical form for a symplectic tensor given in Proposition 12.22.

Theorem 18.19 (Darboux). *Let (M, ω) be a $2n$-dimensional symplectic manifold. Near every point $p \in M$, there are smooth coordinates $(x^1, y^1, \dots, x^n, y^n)$ in which ω has the coordinate representation*

$$\omega = \sum_{i=1}^{n} dx^i \wedge dy^i. \tag{18.13}$$

Any coordinates satisfying the conclusion of the Darboux theorem are called *Darboux coordinates*, *symplectic coordinates*, or *canonical coordinates*. Obviously the standard coordinates $(x^1, y^1, \dots, x^n, y^n)$ on $\mathbb{R}^{2n}$ are Darboux coordinates. The proof of Proposition 12.24 showed that the standard coordinates (x^i, ξ_i) are Darboux coordinates for T^*M with its canonical symplectic structure (at least after renaming and reordering the coordinates).

The Darboux theorem was first proved in 1882 by Gaston Darboux, in connection with his work on ordinary differential equations arising in classical mechanics. The proof we will give was discovered in 1965 by Jürgen Moser [Mos65]. It is based on the theory of time-dependent vector fields, discussed in Chapter 17. Recall that a smooth time-dependent vector field on a smooth manifold M is a smooth map $V: J \times M \to TM$, where $J \subset \mathbb{R}$ is an open interval, such that $V(t,p) \in T_pM$ for all $(t,p) \in J \times M$. Thus for each $t \in J$, we get a smooth vector field $V_t: M \to TM$ defined by $V_t|_p = V(t,p)$. Theorem 17.15 shows that any time-dependent vector field generates a time-dependent flow, which is a smooth map $\theta: \mathcal{E} \to M$ defined on some open subset $\mathcal{E} \subset J \times J \times M$ and satisfying

$$\begin{aligned} \frac{\partial \theta}{\partial t}(t,s,p) &= V(t, \theta(t,s,p)), \\ \theta(s,s,p) &= p. \end{aligned}$$

As in Problem 17-14, for each $(t,s) \in J \times J$, we let $M_{t,s}$ denote the set of points $p \in M$ such that $(t,s,p) \in \mathcal{E}$, and let $\theta_{t,s} \colon M_{t,s} \to M$ be the smooth map $\theta_{t,s}(p) = \theta(t,s,p)$. We will need the following generalization of Lemma 18.15 to the case of time-dependent flows.

Proposition 18.20. *Let M be a smooth manifold. Suppose $V \colon J \times M \to TM$ is a smooth time-dependent vector field and $\theta \colon \mathcal{E} \to M$ is its time-dependent flow. For any smooth covariant tensor field $\tau \in \mathcal{T}^k(M)$ and any $(t_1, t_0, p) \in \mathcal{E}$,*

$$\left.\frac{d}{dt}\right|_{t=t_1} (\theta_{t,t_0}^* \tau)_p = \left(\theta_{t_1,t_0}^* (\mathcal{L}_{V_{t_1}} \tau)\right)_p . \tag{18.14}$$

Proof. First assume $t_1 = t_0$. In this case, θ_{t_0,t_0} is the identity map of M, so we need to prove

$$\left.\frac{d}{dt}\right|_{t=t_0} (\theta_{t,t_0}^* \tau)_p = \left(\mathcal{L}_{V_{t_0}} \tau\right)_p . \tag{18.15}$$

We begin with the special case in which $\tau = f$ is a smooth 0-tensor field:

$$\begin{aligned}
\left.\frac{d}{dt}\right|_{t=t_0} (\theta_{t,t_0}^* f)(p) &= \left.\frac{\partial}{\partial t}\right|_{t=t_0} f(\theta(t,t_0,p)) \\
&= V(t_0, \theta(t_0,t_0,p)) f \\
&= (\mathcal{L}_{V_{t_0}} f)(p).
\end{aligned}$$

Next consider an exact 1-form $\tau = df$. In any smooth local coordinates (x^i), $\theta_{t,t_0}^* f(x) = f(\theta(t,t_0,x))$ is a smooth function of all $n+1$ variables $(t, x^1, \dots, x^n)$. Thus the operator d/dt (which is more properly written as $\partial/\partial t$ in this situation) commutes with each of the partial derivatives $\partial/\partial x^i$ when applied to $\theta_{t,t_0}^* f$. In particular, this means that the exterior derivative operator d commutes with $\partial/\partial t$, and so

$$\begin{aligned}
\left.\frac{d}{dt}\right|_{t=t_0} (\theta_{t,t_0}^* df)_p &= \left.\frac{\partial}{\partial t}\right|_{t=t_0} d(\theta_{t,t_0}^* f)_p \\
&= d\left(\left.\frac{\partial}{\partial t}\right|_{t=t_0} (\theta_{t,t_0}^* f) \right)_p \\
&= d\left(\mathcal{L}_{V_{t_0}} f\right)_p \\
&= \left(\mathcal{L}_{V_{t_0}} df\right)_p .
\end{aligned}$$

Thus the proposition is proved for 0-tensors and for exact 1-forms.

Now suppose that $\tau = \alpha \otimes \beta$ for some smooth covariant tensor fields α and β, and assume that the proposition is true for α and β. (We include the possibility that α or β has rank 0, in which case the tensor product is just ordinary multiplication.) By the product rule for Lie derivatives (Lemma 18.9(c)), the right-hand side of (18.15) satisfies

$$\left(\mathcal{L}_{V_{t_0}} (\alpha \otimes \beta)\right)_p = (\mathcal{L}_{V_{t_0}} \alpha)_p \otimes \beta_p + \alpha_p \otimes (\mathcal{L}_{V_{t_0}} \beta)_p .$$

On the other hand, by an argument entirely analogous to that in the proof of Lemma 18.9, the left-hand side satisfies a similar product rule:

$$\begin{aligned}\left.\frac{d}{dt}\right|_{t=t_0} (\theta^*_{t,t_0}(\alpha\otimes\beta))_p &= \left(\left.\frac{d}{dt}\right|_{t=t_0} (\theta^*_{t,t_0}\alpha)\right)_p \otimes \beta_p \\ &\quad + \alpha_p \otimes \left(\left.\frac{d}{dt}\right|_{t=t_0} (\theta^*_{t,t_0}\beta)\right)_p.\end{aligned}$$

This shows that (18.15) holds for $\tau = \alpha\otimes\beta$ provided it holds for α and β. The case of arbitrary tensor fields now follows by induction, using the fact that any smooth covariant tensor field can be written locally as a sum of tensor fields of the form $\tau = f\,dx^1\otimes\cdots\otimes dx^k$.

To handle arbitrary t_1, we use Problem 17-14, which shows that $\theta_{t,t_0} = \theta_{t,t_1}\circ\theta_{t_1,t_0}$ wherever the right-hand side is defined. Therefore, because the linear map $\theta^*_{t_1,t_0}\colon T^k(T_{\theta_{t_1,t_0}(p)}M)\to T^k(T_pM)$ does not depend on t,

$$\begin{aligned}\left.\frac{d}{dt}\right|_{t=t_1} (\theta^*_{t,t_0}\tau)_p &= \left.\frac{d}{dt}\right|_{t=t_1} \theta^*_{t_1,t_0}\theta^*_{t,t_1}(\tau_{\theta_{t,t_0}(p)}) \\ &= \theta^*_{t_1,t_0} \left.\frac{d}{dt}\right|_{t=t_1} \theta^*_{t,t_1}(\tau_{\theta_{t,t_1}\circ\theta_{t_1,t_0}(p)}) \\ &= \left(\theta^*_{t_1,t_0}(\mathcal{L}_{V_{t_1}}\tau)\right)_p. \qquad\square\end{aligned}$$

We will also need the following refinement of the preceding proposition. A *smooth time-dependent tensor field* on a smooth manifold M is a smooth map $\tau\colon J\times M\to T^kM$, where $J\subset\mathbb{R}$ is an open interval, satisfying $\tau(t,p)\in T^k(T_pM)$ for each $(t,p)\in J\times M$.

Corollary 18.21. *Let V and θ be as in Proposition 18.20, and let $\tau\colon J\times M\to T^kM$ be a smooth time-dependent tensor field on M. Then for any $(t_1,t_0,p)\in\mathcal{E}$,*

$$\left.\frac{d}{dt}\right|_{t=t_1} (\theta^*_{t,t_0}\tau_t)_p = \left(\theta^*_{t_1,t_0}\left(\mathcal{L}_{V_{t_1}}\tau_{t_1} + \left.\frac{d}{dt}\right|_{t=t_1}\tau_t\right)\right)_p. \tag{18.16}$$

Proof. For sufficiently small $\varepsilon>0$, consider the smooth map $F\colon(t_1-\varepsilon,t_1+\varepsilon)\times(t_1-\varepsilon,t_1+\varepsilon)\to T^k(T_pM)$ defined by

$$F(u,v) = (\theta^*_{u,t_0}\tau_v)_p.$$

Since F takes its values in the finite-dimensional vector space $T^k(T_pM)$, we can apply the chain rule together with Proposition 18.20 to conclude that

$$\begin{aligned}\left.\frac{d}{dt}\right|_{t=t_1} F(t,t) &= \frac{\partial F}{\partial u}(t_1,t_1) + \frac{\partial F}{\partial v}(t_1,t_1) \\ &= \left(\theta^*_{t_1,t_0}(\mathcal{L}_{V_{t_1}}\tau_{t_1})\right)_p + \left.\frac{d}{dt}\right|_{t=t_1} \left(\theta^*_{t_1,t_0}\tau_t\right)_p.\end{aligned}$$

Just as in the proof of Proposition 18.20, $\theta^*_{t_1,t_0}$ commutes past d/dt, yielding (18.16). □

Proof of the Darboux theorem. Let ω_0 denote the given symplectic form on M, and let $p_0 \in M$ be arbitrary. The theorem will be proved if we can find a smooth coordinate chart (U_0, φ) containing p_0 such that $\varphi^*\omega_1 = \omega_0$, where $\omega_1 = \sum_{i=1}^n dx^i \wedge dy^i$ is the standard symplectic form on $\mathbb{R}^{2n}$. Since this is a local question, by choosing smooth coordinates $(x^1, y^1, \dots, x^n, y^n)$ near p_0, we may replace M with an open ball $U \subset \mathbb{R}^{2n}$. Proposition 12.22 shows that we can arrange by a linear change of coordinates that

$$\omega_0|_{p_0} = \omega_1|_{p_0}.$$

Let $\Omega = \omega_1 - \omega_0$. Because Ω is closed, the Poincaré lemma shows that we can find a smooth 1-form α on U such that

$$d\alpha = -\Omega.$$

By subtracting a constant-coefficient 1-form from α, we may assume without loss of generality that $\alpha_{p_0} = 0$.

For each $t \in \mathbb{R}$, define a closed 2-form ω_t on U by

$$\omega_t = \omega_0 + t\Omega.$$

Let J be a bounded open interval containing $[0,1]$. Because $\omega_t|_{p_0} = \omega_0|_{p_0}$ is nondegenerate for all t, a simple compactness argument shows that there is some neighborhood $U_1 \subset U$ of p_0 such that ω_t is nondegenerate on U_1 for all $t \in \overline{J}$. Because of this nondegeneracy, the smooth bundle map $\widetilde{\omega}_t \colon TU_1 \to T^*U_1$ defined by $\widetilde{\omega}_t(X) = X \lrcorner\, \omega_t$ is an isomorphism for each $t \in \overline{J}$.

Define a smooth time-dependent vector field $V \colon J \times U_1 \to TU_1$ by $V_t = \widetilde{\omega}_t^{-1}\alpha$, or equivalently

$$V_t \lrcorner\, \omega_t = \alpha.$$

Our assumption that $\alpha_{p_0} = 0$ implies that $V_t|_{p_0} = 0$ for each t. If $\theta \colon \mathcal{E} \to U_1$ denotes the time-dependent flow of V, it follows that $\theta(t, 0, p_0) = p_0$ for all $t \in J$, so $J \times \{0\} \times \{p_0\} \subset \mathcal{E}$. Because $\mathcal{E}$ is open in $J \times J \times M$ and $[0,1]$ is compact, there is some neighborhood U_0 of p_0 such that $[0,1] \times \{0\} \times U_0 \subset \mathcal{E}$.

For each $t_1 \in [0,1]$, it follows from Corollary 18.21 that

$$\begin{aligned}
\left.\frac{d}{dt}\right|_{t=t_1} (\theta^*_{t,0}\omega_t) &= \theta^*_{t_1,0}\left(\mathcal{L}_{V_{t_1}}\omega_{t_1} + \left.\frac{d}{dt}\right|_{t=t_1} \omega_t \right) \\
&= \theta^*_{t_1,0}\big(V_{t_1} \lrcorner\, d\omega_{t_1} + d(V_{t_1} \lrcorner\, \omega_{t_1}) + \Omega \big) \\
&= \theta^*_{t_1,0}(0 + d\alpha + \Omega) \\
&= 0.
\end{aligned}$$

Therefore, $\theta^*_{t,0}\omega_t = \theta^*_{0,0}\omega_0 = \omega_0$ for all t. In particular, $\theta^*_{1,0}\omega_1 = \omega_0$. It follows from Problem 12-10 that $(\theta_{1,0})_*$ is bijective at p_0, so after shrinking U_0 further if necessary, $\theta_{1,0}$ is the required coordinate map. □

Hamiltonian Vector Fields

One of the most important constructions on symplectic manifolds is a symplectic analogue of the gradient, defined as follows. Suppose (M,ω) is a symplectic manifold. For any smooth function $f \in C^\infty(M)$, we define the *Hamiltonian vector field* of f to be the smooth vector field X_f defined by

$$X_f = \widetilde{\omega}^{-1}(df),$$

where $\widetilde{\omega}\colon TM \to T^*M$ is the bundle isomorphism determined by ω. Equivalently,

$$X_f \lrcorner\, \omega = df,$$

or for any vector field Y,

$$\omega(X_f, Y) = df(Y) = Yf.$$

In any Darboux coordinates, X_f can be computed explicitly as follows. Writing

$$X_f = \sum_{i=1}^n \left(A^i \frac{\partial}{\partial x^i} + B^i \frac{\partial}{\partial y^i} \right)$$

for some coefficient functions (A^i, B^i) to be determined, we compute

$$\begin{aligned} X_f \lrcorner\, \omega &= \sum_{j=1}^n \left(A^j \frac{\partial}{\partial x^j} + B^j \frac{\partial}{\partial y^j} \right) \lrcorner \sum_{i=1}^n dx^i \wedge dy^i \\ &= \sum_{i=1}^n (A^i\, dy^i - B^i\, dx^i). \end{aligned}$$

On the other hand,

$$df = \sum_{i=1}^n \left(\frac{\partial f}{\partial x^i} dx^i + \frac{\partial f}{\partial y^i} dy^i \right).$$

Setting these two expressions equal to each other, we find that $A^i = \partial f/\partial y^i$ and $B^i = -\partial f/\partial x^i$, which yields the following formula for the Hamiltonian vector field of f:

$$X_f = \sum_{i=1}^n \left(\frac{\partial f}{\partial y^i}\frac{\partial}{\partial x^i} - \frac{\partial f}{\partial x^i}\frac{\partial}{\partial y^i} \right). \tag{18.17}$$

This formula holds, in particular, on $\mathbb{R}^{2n}$ with its standard symplectic form.

Although the definition of the Hamiltonian vector field is formally analogous to that of the gradient on a Riemannian manifold, Hamiltonian vector fields differ from gradients in some very significant ways, as the next lemma shows.

Proposition 18.22 (Properties of Hamiltonian Vector Fields). *Let (M,ω) be a symplectic manifold and let $f \in C^\infty(M)$.*

(*a*) *f is constant along the flow of X_f, i.e., if θ is the flow, then $f(\theta_t(p)) = f(p)$ for all (t,p) in the domain of θ.*

(*b*) *At each regular point of f, the Hamiltonian vector field X_f is tangent to the level set of f.*

Proof. Both assertions follow from the fact that

$$X_f f = df(X_f) = \omega(X_f, X_f) = 0$$

because ω is alternating. □

A smooth vector field X on M is said to be *symplectic* if ω is invariant under the flow of X. It is said to be *Hamiltonian* (or *globally Hamiltonian*) if there exists a function $f \in C^\infty(M)$ such that $X = X_f$, and *locally Hamiltonian* if every point p has a neighborhood on which X is Hamiltonian. Clearly every globally Hamiltonian vector field is locally Hamiltonian.

Proposition 18.23 (Hamiltonian and Symplectic Vector Fields). *Let (M,ω) be a symplectic manifold. A smooth vector field on M is symplectic if and only if it is locally Hamiltonian. Every locally Hamiltonian vector field on M is globally Hamiltonian if and only if $H^1_{dR}(M) = 0$.*

Proof. By Proposition 18.16, a smooth vector field X is symplectic if and only if $\mathcal{L}_X\omega = 0$. Using formula (18.9) for the Lie derivative of a differential form, we compute

$$\mathcal{L}_X\omega = d(X \lrcorner\, \omega) + X \lrcorner\, (d\omega) = d(X \lrcorner\, \omega). \tag{18.18}$$

Therefore X is symplectic if and only if the 1-form $X \lrcorner\, \omega$ is closed. On the one hand, if X is locally Hamiltonian, then in a neighborhood of each point there is a real-valued function f such that $X = X_f$, so $X \lrcorner\, \omega = X_f \lrcorner\, \omega = df$, which is certainly closed. Conversely, if X is symplectic, then by the Poincaré lemma each point $p \in M$ has a neighborhood U on which the closed 1-form $X \lrcorner\, \omega$ is exact. This means there is a smooth real-valued function f defined on U such that $X \lrcorner\, \omega = df$; because ω is nondegenerate, this implies that $X = X_f$ on U.

Now suppose $H^1_{dR}(M) = 0$. Then every closed 1-form is exact, so for any locally Hamiltonian (hence symplectic) vector field X there is a smooth real-valued function f such that $X \lrcorner\, \omega = df$. This means that $X = X_f$, so X is globally Hamiltonian. Conversely, suppose every locally Hamiltonian vector field is globally Hamiltonian. Let η be a closed 1-form, and let X be the vector field $X = \widetilde{\omega}^{-1}\eta$. Then (18.18) shows that $\mathcal{L}_X\omega = 0$, so X is symplectic and therefore locally Hamiltonian. By hypothesis, there is a global smooth real-valued function f such that $X = X_f$, and then unwinding the definitions, we find that $\eta = df$. □

A symplectic manifold (M,ω) together with a smooth function $H \in C^\infty(M)$ is called a *Hamiltonian system*. The function H is called the *Hamiltonian* of the system; the flow of the Hamiltonian vector field X_H is called its *Hamiltonian flow*, and the integral curves of X_H are called the *trajectories* or *orbits* of the system. In Darboux coordinates, formula (18.17) implies that the orbits are those curves $\gamma(t) = (x^i(t), y^i(t))$ that satisfy

$$\begin{aligned}(x^i)'(t) &= \frac{\partial H}{\partial y^i}(x(t), y(t)),\\ (y^i)'(t) &= -\frac{\partial H}{\partial x^i}(x(t), y(t)).\end{aligned} \tag{18.19}$$

These are called *Hamilton's equations*.

Hamiltonian systems play a central role in classical mechanics. We will illustrate how they arise with a simple example.

Example 18.24 (The n-body problem). Consider n physical particles of masses $m_1, \dots, m_n$ moving in $\mathbb{R}^3$. For many purposes, an effective model of such a system is obtained by idealizing the particles as points in $\mathbb{R}^3$, whose coordinates we denote by (q^1, q^2, q^3), $\dots$, $(q^{3n-2}, q^{3n-1}, q^{3n})$. Thus the evolution of the system over time can be represented by a curve $q(t) = (q^1(t), \dots, q^{3n}(t))$ in $\mathbb{R}^{3n}$. The *collision set* is the subset $\mathcal{C} \subset \mathbb{R}^{3n}$ where two or more particles occupy the same position in space:

$$\mathcal{C} = \{q \in \mathbb{R}^{3n} : (q^{3k-2}, q^{3k-1}, q^{3k}) = (q^{3l-2}, q^{3l-1}, q^{3l}) \text{ for some } k \neq l\}.$$

We will only consider motions with no collisions, so we are interested in curves in the open set $Q = \mathbb{R}^{3n} \smallsetminus \mathcal{C}$.

Suppose the particles are acted upon by forces that depend only on their positions. (Typical examples are electromagnetic and gravitational forces.) If we denote the components of the force on the kth particle by $(F_{3k-2}(q), F_{3k-1}(q), F_{3k}(q))$, then Newton's second law of motion asserts that the particles' motion satisfies the $3n \times 3n$ system of second-order ODEs

$$\begin{aligned}m_k(q^{3k-2})''(t) &= F_{3k-2}(q(t)),\\ m_k(q^{3k-1})''(t) &= F_{3k-1}(q(t)),\\ m_k(q^{3k})''(t) &= F_{3k}(q(t)), \qquad k = 1, \dots, n.\end{aligned}$$

(There is no implied summation in these equations.) If we let $M = (M_{ij})$ denote the $3n \times 3n$ diagonal matrix whose diagonal entries are $(m_1, m_1, m_1, m_2, m_2, m_2, \dots, m_n, m_n, m_n)$, then this system can be written in the more compact form

$$M_{ij}(q^j)''(t) = F_i(q(t)). \tag{18.20}$$

(Here the summation convention is in force.) We assume that the forces depend smoothly on q, so we can interpret $(F_1, \dots, F_{3n})$ as the components of a smooth covector field on Q. We assume further that the forces are conservative, which by the results of Chapter 6 is equivalent to the existence of a smooth function $V \in C^\infty(Q)$ (called the *potential energy* of the system) such that $F = -dV$.

Under the physically reasonable assumption that all of the masses are positive, the matrix M is positive definite, and thus can be interpreted as a (constant-coefficient) Riemannian metric on Q. It therefore defines a smooth bundle isomorphism $\widetilde{M}\colon TQ \to T^*Q$. If we denote the standard coordinates on TQ by (q^i, v^i) and those on T^*Q by (q^i, p_i), then $M(v,v) = M_{ij}v^i v^j$, and $\widetilde{M}$ has the coordinate representation

$$p_i = M_{ij}v^j.$$

If $q'(t) = ((q^1)'(t), \dots, (q^{3n})'(t))$ is the velocity vector of the system of particles at time t, then the covector $p(t) = \widetilde{M}(q'(t))$ is given by the formula

$$p_i(t) = M_{ij}(q^j)'(t). \tag{18.21}$$

Physically, $p(t)$ is interpreted as the *momentum* of the system at time t.

Using (18.21), we see that a curve $q(t)$ in Q satisfies Newton's second law (18.20) if and only if the curve $\gamma(t) = (q(t), p(t))$ in T^*Q satisfies the first-order system of ODEs

$$\begin{aligned}(q^i)'(t) &= M^{ij}p_j(t),\\ (p_i)'(t) &= -\frac{\partial V}{\partial q^i}(q(t)),\end{aligned} \tag{18.22}$$

where (M^{ij}) is the inverse of the matrix of (M_{ij}). Define a function $E \in C^\infty(T^*M)$, called the *total energy* of the system, by

$$E(p,q) = V(q) + K(p),$$

where V is the potential energy introduced above, and K is the *kinetic energy*, defined by

$$K(p) = \tfrac{1}{2}M^{ij}p_i p_j.$$

Since (q^i, p_i) are Darboux coordinates on T^*Q, a comparison of (18.22) with (18.19) shows that (18.22) is precisely Hamilton's equations for the Hamiltonian flow of E. The fact that E is constant along the trajectories of its own Hamiltonian flow is known as the *law of conservation of energy.*

An elaboration of the same technique can be applied to virtually any classical dynamical system in which the forces are conservative. For example, if the positions of a system of particles are subject to constraints, as are the constituent particles of a rigid body, for example, then the configuration space will typically be a submanifold of $\mathbb{R}^{3n}$ rather than an open subset. Under very general hypotheses, the equations of motion of such a system can be formulated as a Hamiltonian system on the cotangent bundle of the constraint manifold. For much more on Hamiltonian systems, see [AM78].

Poisson Brackets

Using Hamiltonian vector fields, we define an operation on real-valued functions on a symplectic manifold M similar to the Lie bracket of vector fields. Given $f, g \in C^\infty(M)$, we define their *Poisson bracket* $\{f, g\} \in C^\infty(M)$ by any of the following equivalent formulas:

$$\{f, g\} = \omega(X_f, X_g) = df(X_g) = X_g f. \tag{18.23}$$

Two functions are said to *Poisson commute* if their Poisson bracket is zero.

The geometric interpretation of the Poisson bracket is evident from the characterization $\{f, g\} = X_g f$: It is a measure of the rate of change of f along the Hamiltonian flow of g. In particular, f and g Poisson commute if and only if f is constant along the Hamiltonian flow of g.

Using (18.17), we can readily compute the Poisson bracket of two functions f, g in Darboux coordinates:

$$\{f, g\} = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}\frac{\partial g}{\partial y^i} - \frac{\partial f}{\partial y^i}\frac{\partial g}{\partial x^i}. \tag{18.24}$$

Proposition 18.25 (Properties of the Poisson Bracket). *Suppose (M, ω) is a symplectic manifold, and $f, g, h \in C^\infty(M)$.*

(*a*) Bilinearity: *$\{f, g\}$ is linear over $\mathbb{R}$ in f and in g.*

(*b*) Antisymmetry: *$\{f, g\} = -\{g, f\}$.*

(*c*) Jacobi Identity: *$\{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0$.*

(*d*) *$X_{\{f,g\}} = -[X_f, X_g]$.*

Proof. Parts (a) and (b) are obvious from the characterization $\{f, g\} = \omega(X_f, X_g)$ together with the fact that $X_f = \widetilde{\omega}^{-1}(df)$ depends linearly on f. The proof of (d) is a computation using Proposition 18.9(e) and the fact that Hamiltonian vector fields are symplectic:

$$\begin{aligned}
X_{\{f,g\}} \lrcorner\, \omega &= d\{f, g\} \\
&= d(X_g f) \\
&= d(\mathcal{L}_{X_g} f) \\
&= \mathcal{L}_{X_g} df \\
&= \mathcal{L}_{X_g}(X_f \lrcorner\, \omega) \\
&= (\mathcal{L}_{X_g} X_f) \lrcorner\, \omega + X_f \lrcorner\, \mathcal{L}_{X_g}\omega \\
&= [X_g, X_f] \lrcorner\, \omega + 0 \\
&= -[X_f, X_g] \lrcorner\, \omega,
\end{aligned}$$

which is equivalent to (d) because ω is nondegenerate. Then (c) follows from (d), (b), and (18.23):

$$\begin{aligned}\{f,\{g,h\}\} &= X_{\{g,h\}}f\\ &= -[X_g, X_h]f\\ &= -X_gX_hf + X_hX_gf\\ &= -X_g\{f,h\} + X_h\{f,g\}\\ &= -\{\{f,h\},g\} + \{\{f,g\},h\}\\ &= -\{g,\{h,f\}\} - \{h,\{f,g\}\}.\end{aligned}$$ □

The following corollary is immediate.

Corollary 18.26. *If (M,ω) is a symplectic manifold, the vector space $C^\infty(M)$ is a Lie algebra under the Poisson bracket.*

If (M,ω,H) is a Hamiltonian system, any function $f \in C^\infty(M)$ that is constant on every integral curve of X_H is called a *conserved quantity* of the system. Conserved quantities turn out to be deeply related to symmetries, as we now show.

A smooth vector field V on M is called an *infinitesimal symmetry* of (M,ω,H) if both ω and H are invariant under the flow of V.

◇ **Exercise 18.5.** Let (M,ω,H) be a Hamiltonian system.

(a) Show that $f \in C^\infty(M)$ is a conserved quantity if and only if $\{H,f\} = 0$.
(b) Show that the infinitesimal symmetries are precisely the symplectic vector fields V that satisfy $VH = 0$.
(c) If θ is the flow of an infinitesimal symmetry and γ is a trajectory of the system, show that for each $s \in \mathbb{R}$, $\theta_s \circ \gamma$ is also a trajectory on its domain of definition.

The following theorem, first proved (in a somewhat different form) by Emmy Noether in 1918 [Noe71], has had a profound influence on both physics and mathematics. It shows that for many manifolds (simply connected ones, for example) there is a one-to-one correspondence between conserved quantities (modulo constants) and infinitesimal symmetries.

Theorem 18.27 (Noether's Theorem). *Let (M,ω,H) be a Hamiltonian system. If f is any conserved quantity, then its Hamiltonian vector field is an infinitesimal symmetry. Conversely, if $H^1_{dR}(M) = 0$, then every infinitesimal symmetry is the Hamiltonian vector field of a conserved quantity, which is unique up to addition of a function that is constant on each component of M.*

Proof. Suppose f is a conserved quantity. Exercise 18.5 shows that $\{H,f\} = 0$. This in turn implies that $X_fH = \{H,f\} = 0$, so H is constant along the flow of X_f. Since ω is invariant along the flow of any Hamiltonian vector field by Proposition 18.23, this shows that X_f is an infinitesimal symmetry.

Now suppose that $H^1_{dR}(M) = 0$ and let V be an infinitesimal symmetry. Then V is symplectic by Exercise 18.5, and globally Hamiltonian by Proposition 18.23. Writing $V = X_f$, the fact that H is constant along the flow of V implies that $\{H, f\} = X_f H = VH = 0$, so f is a conserved quantity. If $\widetilde{f}$ is any other function that satisfies $X_{\widetilde{f}} = V = X_f$, then $d(\widetilde{f} - f) = (X_{\widetilde{f}} - X_f) \lrcorner\, \omega = 0$, so $\widetilde{f} - f$ must be constant on each component of M. □

There is one conserved quantity that every Hamiltonian system possesses—the Hamiltonian H itself. The infinitesimal symmetry corresponding to it, of course, generates the Hamiltonian flow of the system, which describes how the system evolves over time. Since H is typically interpreted as the total energy of the system (as in Example 18.24), one usually says that the symmetry corresponding to conservation of energy is "translation in the time variable."

Problems

18-1. Give an example of smooth vector fields V, $\widetilde{V}$, and W on $\mathbb{R}^2$ such that $V = \widetilde{V} = \partial/\partial x$ along the x-axis but $\mathcal{L}_V W \neq \mathcal{L}_{\widetilde{V}} W$ at the origin. [Remark: This shows that it is really necessary to know the vector field V to compute $(\mathcal{L}_V W)_p$; it is not sufficient just to know the vector V_p, or even to know the values of V along an integral curve of V.]

18-2. For each k-tuple of vector fields on $\mathbb{R}^3$ shown below, either find smooth coordinates (u^1, u^2, u^3) in a neighborhood of $(1, 0, 0)$ such that $V_i = \partial/\partial u^i$ for $i = 1, \dots, k$, or explain why there are none.

(a) $k = 2;\ V_1 = \dfrac{\partial}{\partial x},\ V_2 = \dfrac{\partial}{\partial x} + \dfrac{\partial}{\partial y}.$

(b) $k = 2;\ V_1 = x\dfrac{\partial}{\partial y} - y\dfrac{\partial}{\partial x},\ V_2 = x\dfrac{\partial}{\partial x} + y\dfrac{\partial}{\partial y}.$

(c) $k = 3;\ V_1 = x\dfrac{\partial}{\partial y} - y\dfrac{\partial}{\partial x},\ V_2 = y\dfrac{\partial}{\partial z} - z\dfrac{\partial}{\partial y},\ V_3 = z\dfrac{\partial}{\partial x} - x\dfrac{\partial}{\partial z}.$

18-3. This problem generalizes the result of Problem 17-11. Let M be a smooth n-manifold, and let $S \subset M$ be a smooth embedded submanifold of codimension k. Suppose $X_1, \dots, X_k$ are commuting independent smooth vector fields on M that are nowhere tangent to S. If $f_1, \dots, f_k \in C^\infty(M)$ are functions such that $X_i f_j = X_j f_i$ for all $i, j = 1, \dots, k$, and $\varphi \in C^\infty(S)$ is arbitrary, show that there exists a neighborhood U of S in M and a unique function $u \in C^\infty(U)$

satisfying

$$X_i u = f_i, \qquad i = 1, \ldots, k;$$
$$u|_S = \varphi.$$

18-4. Let M, N be smooth manifolds, and suppose $\pi: M \to N$ is a surjective submersion with connected fibers. We say a tangent vector $X \in T_pM$ is *vertical* if $\pi_* X = 0$. Suppose $\omega \in \mathcal{A}^k(M)$. Show that there exists $\eta \in \mathcal{A}^k(N)$ such that $\omega = \pi^*\eta$ if and only if $X \lrcorner\, \omega = 0$ and $X \lrcorner\, d\omega = 0$ for every vertical vector X. [Hint: First do the case in which $\pi: \mathbb{R}^{n+k} \to \mathbb{R}^n$ is projection onto the first n coordinates.]

18-5. Define vector fields V and W on the plane by

$$V = x\frac{\partial}{\partial y} + y\frac{\partial}{\partial x}; \quad W = x\frac{\partial}{\partial x} - y\frac{\partial}{\partial y}.$$

Compute the flows θ, ψ of V and W, and verify that they do not commute by finding explicit times s and t such that $\theta_s \circ \psi_t \neq \psi_t \circ \theta_s$.

18-6. Let (M, ω) be a symplectic manifold.

(a) Show that the set of symplectic vector fields on M is a Lie subalgebra of $\mathcal{T}(M)$.

(b) Show that the set of Hamiltonian vector fields is a Lie subalgebra of the set of symplectic vector fields.

(c) Show that the quotient of the symplectic vector fields modulo the Hamiltonian vector fields is isomorphic (as a vector space) to $H^1_{dR}(M)$.

18-7. Using the same technique as in the proof of the Darboux theorem, prove the following theorem of Moser: If M is an orientable compact smooth manifold, and Ω_0, Ω_1 are smooth orientation forms on M such that $\int_M \Omega_0 = \int_M \Omega_1$, then there exists a diffeomorphism $F: M \to M$ such that $F^*\Omega_1 = \Omega_0$.

18-8. Prove the following global version of the Darboux theorem: Suppose M is a compact symplectic manifold, and ω_0, ω_1 are cohomologous symplectic forms on M. Show that there is a diffeomorphism $F: M \to M$ such that $F^*\omega_1 = \omega_0$.

18-9. This problem outlines a different proof of the Darboux theorem. Let (M, ω) be a $2n$-dimensional symplectic manifold and $p \in M$.

(a) Show that smooth coordinates (x^i, y^i) on an open set $U \subset M$ are Darboux coordinates if and only if their Poisson brackets satisfy

$$\{x^i, y^j\} = \delta^{ij}; \qquad \{x^i, x^j\} = \{y^i, y^j\} = 0. \qquad (18.25)$$

(b) Prove by induction on k that for each $k = 0, \ldots, n$, there are functions $(x^1, y^1, \ldots, x^k, y^k)$ satisfying (18.25) near p such that

$\{dx^1, dy^1, \dots, dx^k, dy^k\}$ are independent at p. When $k = n$, this proves the theorem. [Hint: For the inductive step, assuming that $(x^1, y^1, \dots, x^k, y^k)$ have been found, find smooth coordinates $(u^1, \dots, u^{2n})$ such that

$$\frac{\partial}{\partial u^i} = X_{x^i}, \qquad \frac{\partial}{\partial u^{i+k}} = X_{y^i}, \qquad i = 1, \dots, k,$$

and let $y^{k+1} = u^{2k+1}$. Then find new coordinates $(v^1, \dots, v^{2n})$ such that

$$\frac{\partial}{\partial v^i} = X_{x^i}, \qquad i = 1, \dots, k,$$
$$\frac{\partial}{\partial v^{i+k}} = X_{y^i}, \qquad i = 1, \dots, k+1,$$

and let $x^{k+1} = v^{2k+1}$.]

18-10. Consider the 2-body problem in $\mathbb{R}^3$, i.e., the Hamiltonian system (T^*Q, ω, E) described in Example 18.24 in the special case $n = 2$. Suppose that the potential energy V depends only on the distance between the particles. More precisely, suppose that $V(q) = v(r(q))$ for some smooth function $v\colon (0, \infty) \to \mathbb{R}$, where

$$r(q) = \sqrt{(q^1 - q^4)^2 + (q^2 - q^5)^2 + (q^3 - q^6)^2}.$$

(a) Show that the function $f\colon T^*Q \to \mathbb{R}$ defined by

$$f(p, q) = p^1 + p^4$$

is a conserved quantity (called the linear momentum in the x-direction), and that the corresponding infinitesimal symmetry generates translations in the x-direction:

$$\theta_t(q, p) = (q^1 + t, q^2, q^3, q^4 + t, q^5, q^6, p_1, \dots, p_6).$$

(b) Show that the function $\alpha\colon T^*Q \to \mathbb{R}$ defined by

$$\alpha(p, q) = q^1 p_2 - q^2 p_1 + q^4 p_5 - q^5 p_4$$

is a conserved quantity (called the angular momentum about the z-axis), and find the flow of the corresponding infinitesimal symmetry. Explain what this has to do with rotational symmetry.

19
Integral Manifolds and Foliations

Suppose V is a nonvanishing vector field on a manifold M. The results of Chapter 17 imply that each integral curve of V is an immersion, and that locally the images of the integral curves fit together nicely like parallel lines in Euclidean space. The fundamental theorem on flows tells us that these curves are determined by the knowledge of their tangent vectors.

In this chapter we explore an important generalization of this idea to higher-dimensional submanifolds. The general setup is this: Suppose we are given a k-dimensional subspace of T_pM at each point $p \in M$, varying smoothly from point to point. (Such a collection of subspaces is called a "tangent distribution.") Is there a k-dimensional submanifold (called an "integral manifold" of the tangent distribution) whose tangent space at each point is the given subspace? The answer in this case is more complicated than in the case of vector fields: There is a nontrivial necessary condition, called involutivity, that must be satisfied by the tangent distribution.

In the first section of the chapter, we define involutivity and give examples of both involutive and noninvolutive distributions. Next we show how the involutivity condition can be rephrased in terms of differential forms.

The main theorem of this chapter, the Frobenius theorem, tells us that involutivity is also sufficient for the existence of an integral manifold through each point. We will prove the Frobenius theorem in two forms. First we prove a local form, which says that a neighborhood of every point is filled up with integral manifolds, fitting together nicely like parallel affine subspaces of $\mathbb{R}^n$. After proving this, we give a few applications to the study of partial differential equations. In the last section of the chapter, we prove a global form of the Frobenius theorem, which says that the entire manifold is the disjoint union of immersed integral manifolds.

Tangent Distributions

Let M be a smooth manifold. A choice of k-dimensional linear subspace $D_p \subset T_pM$ at each point $p \in M$ is called a k-dimensional *tangent distribution* on M, or just a *distribution* if there is no opportunity for confusion with the use of the term "distribution" for generalized functions in analysis. A distribution is called *smooth* if the union of all these subspaces forms a smooth subbundle $D = \coprod_{p\in M} D_p \subset TM$. Tangent distributions are also commonly called *k-plane fields* or *tangent subbundles*.

The local frame criterion for subbundles (Lemma 8.39) translates immediately into the following criterion for distribution to be smooth.

Lemma 19.1 (Local Frame Criterion for Distributions). *Let M be a smooth manifold, and suppose $D \subset TM$ is a k-dimensional tangent distribution. Then D is smooth if and only if the following condition is satisfied:*

$$\begin{array}{l}\textit{Each point } p \in M \textit{ has a neighborhood } U \textit{ on which there}\\ \textit{are smooth vector fields } Y_1, \dots, Y_k \colon U \to TM \textit{ such that}\\ Y_1|_q, \dots, Y_k|_q \textit{ form a basis for } D_q \textit{ at each } q \in U.\end{array} \tag{19.1}$$

In the situation of the preceding lemma, we say D is the distribution (locally) spanned by the vector fields $Y_1, \dots, Y_k$.

Integral Manifolds and Involutivity

Suppose $D \subset TM$ is a smooth distribution. An immersed submanifold $N \subset M$ is called an *integral manifold* of D if $T_pN = D_p$ at each point $p \in N$. The main question we want to address in this chapter is that of the existence of integral manifolds.

Before we proceed with the general theory, let us describe some examples of distributions and integral manifolds that you should keep in mind.

Example 19.2 (Tangent Distributions and Integral Manifolds).

(a) If V is any nowhere-vanishing smooth vector field on a manifold M, then V spans a smooth 1-dimensional distribution on M (i.e., $D_p = \operatorname{span}(V_p)$ for each $p \in M$). The image of any integral curve of V is an integral manifold of D.

(b) In $\mathbb{R}^n$, the vector fields $\partial/\partial x^1, \dots, \partial/\partial x^k$ span a smooth k-dimensional distribution. The k-dimensional affine subspaces parallel to $\mathbb{R}^k$ are integral manifolds.

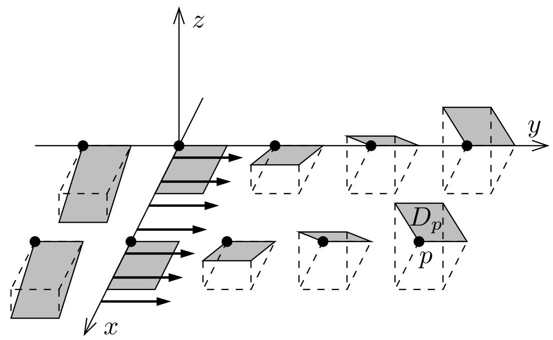


Figure 19.1. A smooth distribution with no integral manifolds.

(c) Define a distribution on $\mathbb{R}^n \smallsetminus \{0\}$ by letting D_p be the subspace of $T_p\mathbb{R}^3$ orthogonal to the radial vector $E_1|_p = x^i\partial/\partial x^i|_p$. If we extend E_1 to a smooth local frame $(E_1, \dots, E_n)$ and apply the Gram-Schmidt algorithm, then D is locally spanned by $(E_2, \dots, E_n)$. Thus D is a smooth $(n-1)$-dimensional distribution on $\mathbb{R}^n \smallsetminus \{0\}$. Through each point $p \in \mathbb{R}^n \smallsetminus \{0\}$, the sphere of radius $|p|$ around 0 is an integral manifold.

(d) Let D be the smooth distribution on $\mathbb{R}^3$ spanned by the following vector fields:

$$X = \frac{\partial}{\partial x} + y\frac{\partial}{\partial z} \qquad Y = \frac{\partial}{\partial y}.$$

(See Figure 19.1). It turns out that D has no integral manifolds. To get an idea why, suppose N is an integral manifold through the origin. Because X and Y are tangent to N, any integral curve of X or Y that starts in N will have to stay in N, at least for a short time. Thus N contains an open subset of the x-axis (which is an integral curve of X). It also contains, for each sufficiently small x, an open subset of the line parallel to the y-axis and passing through $(x, 0, 0)$ (which is an integral curve of Y). Therefore N contains an open subset of the (x, y)-plane. However, the tangent plane to the (x, y)-plane at any point p off of the x-axis is not equal to D_p. Therefore, no such integral manifold exists.

The last example shows that, in general, integral manifolds may fail to exist. The reason for this failure is expressed in the following proposition. Suppose D is a smooth tangent distribution on M. We say that D is *involutive* if given any pair of smooth local sections of D (i.e., smooth vector fields X, Y defined on an open subset of M such that $X_p, Y_p \in D_p$ for each

p), their Lie bracket is also a local section of D. We say D is *integrable* if each point of M is contained in an integral manifold of D.

Proposition 19.3. *Every integrable distribution is involutive.*

Proof. Let $D \subset TM$ be an integrable distribution. Suppose X and Y are smooth local sections of D defined on some open subset $U \subset M$. Let p be any point in U, and let N be an integral manifold of D containing p. The fact that X and Y are sections of D means that X and Y are tangent to N. By Corollary 8.26, $[X, Y]$ is also tangent to N, and therefore $[X, Y]_p \in D_p$. Since this is true at any $p \in U$, D is involutive. □

Note, for example, that the distribution D of Example 19.2(d) is not involutive, because $[X, Y] = -\partial/\partial z$, which is not a section of D.

The next lemma shows that the involutivity condition does not have to be checked for every pair of smooth vector fields, just those of a smooth local frame near each point.

Lemma 19.4. *Let $D \subset TM$ be a distribution. If in a neighborhood of every point of M there exists a smooth local frame $(V_1, \dots, V_k)$ for D such that $[V_i, V_j]$ is a section of D for each $i, j = 1, \dots, k$, then D is involutive.*

Proof. Suppose the hypothesis holds, and suppose X and Y are smooth local sections of D over some open subset $U \subset M$. Given $p \in M$, choose a smooth local frame $(V_1, \dots, V_k)$ satisfying the hypothesis in a neighborhood of p, and write $X = X^i V_i$ and $Y = Y^i V_i$. Then, using formula (4.7),

$$\begin{aligned}
[X, Y] &= \left[X^i V_i, Y^j V_j\right] \\
&= X^i Y^j [V_i, V_j] + X^i (V_i Y^j) V_j - Y^j (V_j X^i) V_i.
\end{aligned}$$

It follows from the hypothesis that this last expression is a section of D. □

Involutivity and Differential Forms

There is an alternative way to characterize involutivity in terms of differential forms. Instead of describing tangent distributions locally by means of smooth vector fields, we can also describe them by means of smooth 1-forms, as the following lemma shows.

Lemma 19.5 (1-Form Criterion for Distributions). *Let M be a smooth n-manifold, and let $D \subset TM$ be a k-dimensional distribution. Then D is smooth if and only if each point $p \in M$ has a neighborhood U on which there are smooth 1-forms $\omega^1, \dots, \omega^{n-k}$ such that for each $q \in U$,*

$$D_q = \operatorname{Ker} \omega^1|_q \cap \cdots \cap \operatorname{Ker} \omega^{n-k}|_q. \tag{19.2}$$

Proof. First suppose that there exist such forms $\omega^1, \dots, \omega^{n-k}$ in a neighborhood of each point. By the result of Problem 5-8, we can extend them to a smooth coframe $(\omega^1, \ldots, \omega^n)$ on a (possibly smaller) neighborhood. If $(E_1, \ldots, E_n)$ is the dual frame, it is easy to check that D is locally spanned by $E_{n-k+1}, \ldots, E_n$, so it is smooth by the local frame criterion.

Conversely, suppose D is smooth. In a neighborhood of any $p \in M$, there are smooth vector fields $Y_1, \ldots, Y_k$ spanning D. By Problem 5-8 again, we can extend these vector fields to a smooth local frame $(Y_1, \ldots, Y_n)$ for M near p. Letting $(\varepsilon^1, \ldots, \varepsilon^n)$ be the dual coframe, it follows easily that D is characterized locally by

$$D_q = \operatorname{Ker} \varepsilon^{k+1}|_q \cap \cdots \cap \operatorname{Ker} \varepsilon^n|_q. \qquad \square$$

Any $(n-k)$ smooth 1-forms $\omega^1, \ldots, \omega^{n-k}$ defined on an open subset $U \subset M$ and satisfying (19.2) for each $q \in U$ will be called *local defining forms* for the distribution D. More generally, we say that a p-form $\omega \in \mathcal{A}^p(M)$ *annihilates* D if $\omega(X_1, \ldots, X_p) = 0$ whenever $X_1, \ldots, X_p$ are local sections of D.

Lemma 19.6. *Suppose M is a smooth manifold and D is a smooth k-dimensional distribution on M. Then a p-form η annihilates D if and only if whenever $\omega^1, \ldots, \omega^{n-k}$ are local defining forms for D over an open subset $U \subset M$, $\eta|_U$ is of the form*

$$\eta|_U = \sum_{i=1}^{n-k} \omega^i \wedge \beta^i \tag{19.3}$$

for some $(p-1)$-forms $\beta^1, \ldots, \beta^{n-k}$ on U.

Proof. It is easy to check that any form η that satisfies (19.3) in a neighborhood of each point annihilates D. Conversely, suppose η annihilates D. Given local defining forms $\omega^1, \ldots, \omega^{n-k}$ for D on $U \subset M$, in a neighborhood of each point we can extend them to a smooth local coframe $(\omega^1, \ldots, \omega^n)$ for M as in the proof of Lemma 19.5. If $(E_1, \ldots, E_n)$ is the dual frame, then D is locally spanned by $E_{n-k+1}, \ldots, E_n$. In terms of this coframe, any $\eta \in \mathcal{A}^p(M)$ can be written uniquely as

$$\eta = \sum_I{}' \eta_I \omega^{i_1} \wedge \cdots \wedge \omega^{i_p},$$

where the coefficients η_I are determined by $\eta_I = \eta(E_{i_1}, \ldots, E_{i_p})$. Thus η annihilates D if and only if $\eta_I = 0$ whenever $n-k+1 \leq i_1 < \cdots < i_p \leq n$, in which case η can be written locally as

$$\begin{aligned}
\eta &= \sum_{I:\ i_1 \leq n-k}{}' \eta_I \omega^{i_1} \wedge \cdots \wedge \omega^{i_p} \\
&= \sum_{i_1=1}^{n-k} \omega^{i_1} \wedge \left(\sum_{I'}{}' \eta_{i_1 I'} \omega^{i_2} \wedge \cdots \wedge \omega^{i_p} \right),
\end{aligned}$$

where we have written $I' = (i_2, \dots, i_p)$. This holds in a neighborhood of each point of U; patching together with a partition of unity, we obtain a similar expression on all of U. □

When expressed in terms of differential forms, the involutivity condition translates into a statement about exterior derivatives.

Proposition 19.7 (1-Form Criterion for Involutivity). *Suppose $D \subset TM$ is a smooth distribution. Then D is involutive if and only if the following condition is satisfied:*

$$\text{If } \eta \text{ is any smooth 1-form that annihilates } D \text{ on an open subset } U \subset M, \text{ then } d\eta \text{ also annihilates } D \text{ on } U. \tag{19.4}$$

Proof. First assume that D is involutive, and suppose η is a smooth 1-form that annihilates D on $U \subset M$. Then for any smooth local sections X, Y of D, formula (12.19) for $d\eta$ gives

$$d\eta(X,Y) = X(\eta(Y)) - Y(\eta(X)) - \eta([X,Y]).$$

The hypothesis implies that each of these terms is zero on U. Conversely, suppose D satisfies (19.4), and suppose X and Y are smooth local sections of D. If $\omega^1, \dots, \omega^{n-k}$ are local defining forms for D, then (12.19) shows that

$$\omega^i([X,Y]) = X(\omega^i(Y)) - Y(\omega^i(X)) - d\omega^i(X,Y) = 0,$$

which implies that $[X,Y]$ takes its values in D. Thus D is involutive. □

Just like the Lie bracket condition for involutivity, the exterior derivative condition need only be checked for a particular set of defining forms in a neighborhood of each point, as the next lemma shows.

Lemma 19.8. *Let D be a smooth k-dimensional distribution on a smooth n-manifold M, and let $\omega^1, \dots, \omega^{n-k}$ be smooth defining forms for D on an open set $U \subset M$. Then D is involutive on U if and only if there exist smooth 1-forms $\{\alpha^i_j : i, j = 1, \dots, n-k\}$ such that*

$$d\omega^i = \sum_{j=1}^{n-k} \omega^j \wedge \alpha^i_j, \qquad \textit{for each } i = 1, \dots, n-k.$$

◇ **Exercise 19.1.** Prove the preceding lemma.

With a bit more algebraic terminology, there is an elegant and concise way to express the involutivity condition in terms of differential forms. Recall that we have defined the graded algebra of smooth differential forms on a smooth n-manifold M as $\mathcal{A}^*(M) = \mathcal{A}^0(M) \oplus \cdots \oplus \mathcal{A}^n(M)$ (see page 302). An *ideal* in $\mathcal{A}^*(M)$ is a linear subspace $\mathcal{I} \subset \mathcal{A}^*(M)$ that is closed under wedge products with arbitrary elements of $\mathcal{A}^*(M)$: That is, $\omega \in \mathcal{I}$ implies that $\eta \wedge \omega \in \mathcal{I}$ for every $\eta \in \mathcal{A}^*(M)$.

If D is a smooth distribution on a smooth n-manifold M, let $\mathcal{I}^p(D) \subset \mathcal{A}^p(M)$ denote the space of smooth p-forms that annihilate D, and let $\mathcal{I}(D) = \mathcal{I}^0(D) \oplus \cdots \oplus \mathcal{I}^n(D) \subset \mathcal{A}^*(M)$.

◇ **Exercise 19.2.** If $D \subset TM$ is a smooth distribution, show that $\mathcal{I}(D)$ is an ideal in $\mathcal{A}^*(M)$.

An ideal $\mathcal{I} \subset \mathcal{A}^*(M)$ is said to be a *differential ideal* if $d(\mathcal{I}) \subset \mathcal{I}$, that is, if $\omega \in \mathcal{I}$ implies $d\omega \in \mathcal{I}$.

Proposition 19.9 (Differential Ideal Criterion for Involutivity). *A smooth distribution $D \subset TM$ is involutive if and only if $\mathcal{I}(D)$ is a differential ideal.*

Proof. Suppose first that $\mathcal{I}(D)$ is a differential ideal. Let η be a smooth 1-form defined on an open set $U \subset M$. Let $q \in U$ be arbitrary, and let $\psi \in C^\infty(M)$ be a smooth bump function that is equal to 1 on a neighborhood of q and supported in U. Then $\psi\eta$ is a smooth global 1-form that annihilates D, so $d(\psi\eta)$ annihilates D by hypothesis. Since $d(\psi\eta)_q = d\eta_q$, it follows that $d\eta$ annihilates D at each point of U, so D is involutive by Proposition 19.7.

Conversely, suppose D is involutive. We need to show that $d\eta$ annihilates D for each $\eta \in \mathcal{I}^p(D)$, $0 \le p \le n$. The $p = 0$ case is trivial, because the only 0-form that annihilates D is the zero form, and the $p = 1$ case is taken care of by Proposition 19.7. Suppose $p \ge 2$, and η is a p-form that annihilates D. If $\omega^1, \dots, \omega^{n-k}$ are local defining forms for D on an open set U, then by Lemma 19.6, $\eta|_U$ can be written in the form (19.3), and thus

$$\begin{aligned} d\eta|_U &= \sum_{i=1}^{n-k} (d\omega^i \wedge \beta^i - \omega^i \wedge d\beta^i) \\ &= \sum_{i,j=1}^{n-k} \omega^j \wedge \alpha^i_j \wedge \beta^i - \sum_{i=1}^{n-k} \omega^i \wedge d\beta^i, \end{aligned}$$

where α^i_j are 1-forms as in Lemma 19.8. It follows from Lemma 19.6 that this form annihilates D. □

The Frobenius Theorem

In Example 19.2, all of the distributions we defined except the last one had the property that there was an integral manifold through each point. Moreover, these submanifolds all "fit together" nicely like parallel affine subspaces of $\mathbb{R}^n$. Given a k-dimensional distribution $D \subset TM$, let us say that a smooth coordinate chart (U, φ) on M is *flat* for D if $\varphi(U)$ is a product of connected open sets $U' \times U'' \subset \mathbb{R}^k \times \mathbb{R}^{n-k}$, and at points of U, D is

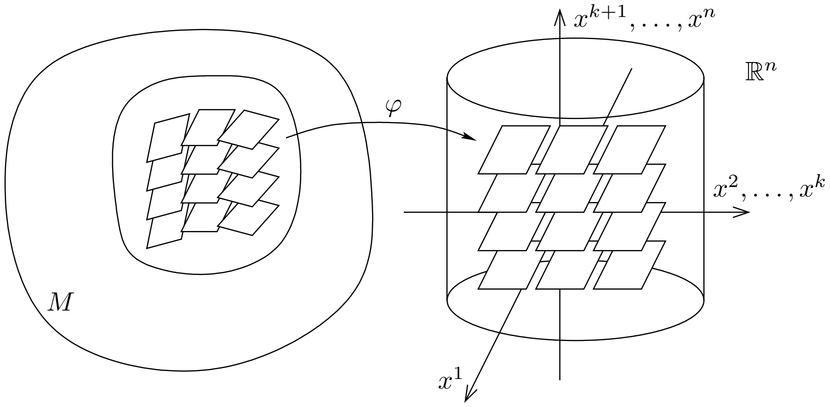


Figure 19.2. A flat chart for a distribution.

spanned by the first k coordinate vector fields $\partial/\partial x^1, \dots, \partial/\partial x^k$ (Figure 19.2). It is obvious that each slice of the form $x^{k+1} = c^{k+1}, \dots, x^n = c^n$ for constants $c^{k+1}, \dots, c^n$ is an integral manifold of D. This is the nicest possible local situation for integral manifolds We say that a distribution $D \subset TM$ is *completely integrable* if there exists a flat chart for D in a neighborhood of every point of M. Obviously every completely integrable distribution is integrable and therefore involutive. In summary:

$$\text{completely integrable} \implies \text{integrable} \implies \text{involutive}.$$

The next theorem is the main result of this chapter, and indeed one of the central theorems in smooth manifold theory. It says that the implications above are actually equivalences:

$$\text{completely integrable} \iff \text{integrable} \iff \text{involutive}.$$

Theorem 19.10 (Frobenius). *Every involutive distribution is completely integrable.*

Proof. The canonical form theorem for commuting vector fields (Theorem 18.6) implies that any distribution locally spanned by smooth *commuting* vector fields is completely integrable. Thus it suffices to show that any involutive distribution is locally spanned by smooth commuting vector fields.

Let D be a k-dimensional involutive distribution on an n-dimensional manifold M, and let $p \in M$. Since complete integrability is a local question, by passing to a smooth coordinate neighborhood of p, we may replace M by an open set $U \subset \mathbb{R}^n$, and choose a smooth local frame $Y_1, \dots, Y_k$ for D. By reordering the coordinates if necessary, we may assume that D_p is complementary to the subspace of $T_p\mathbb{R}^n$ spanned by $(\partial/\partial x^{k+1}|_p, \dots, \partial/\partial x^n|_p)$.

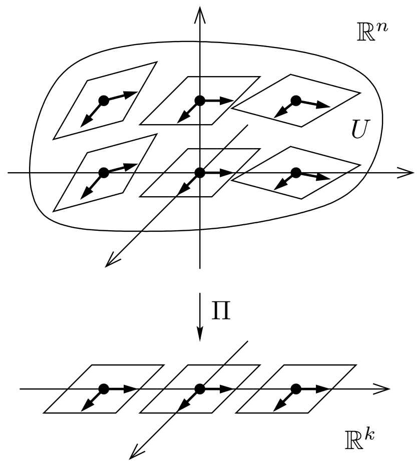


Figure 19.3. Proof of the Frobenius theorem.

Let $\Pi\colon \mathbb{R}^n \to \mathbb{R}^k$ be the projection onto the first k coordinates: $\Pi(x^1,\dots,x^n) = (x^1,\dots,x^k)$ (Figure 19.3). This induces a smooth bundle map $\Pi_*\colon TU \to T\mathbb{R}^k$, which can be written

$$\Pi_*\left(\sum_{i=1}^{n} v^i \left.\frac{\partial}{\partial x^i}\right|_q\right) = \sum_{i=1}^{k} v^i \left.\frac{\partial}{\partial x^i}\right|_{\Pi(q)}.$$

(Notice that the summation is only over $i = 1,\dots,k$ on the right-hand side.) By our choice of coordinates, $D_p \subset T_p\mathbb{R}^n$ is complementary to the kernel of Π_*, so the restriction of Π_* to D_p is bijective. By continuity, the same is true of $\Pi_*\colon D_q \to T_{\Pi(q)}\mathbb{R}^k$ for q in a neighborhood of p. Because $\Pi_*|_D$ is the composition of the inclusion $D \hookrightarrow TM$ followed by Π_*, it is a smooth bundle map. Thus the matrix entries of $\Pi_*|_D$ with respect to the frames $(Y_i|_q)$ and $(\partial/\partial x^j|_{\Pi(q)})$ are smooth functions of q, and thus so are the matrix entries of $(\Pi_*|_D)^{-1}\colon T_{\Pi(q)}\mathbb{R}^k \to D_q$. Define a new smooth local frame $X_1,\dots,X_k$ for D near p by

$$X_i|_q = (\Pi_*|_D)^{-1} \left.\frac{\partial}{\partial x^i}\right|_{\Pi(q)}. \tag{19.5}$$

The theorem will be proved if we can show that $[X_i, X_j] = 0$ for all i, j.

First observe that X_i and $\partial/\partial x^i$ are Π-related, because (19.5) implies that

$$\left.\frac{\partial}{\partial x^i}\right|_{\Pi(q)} = (\Pi_*|_D)\, X_i|_q = \Pi_* X_i|_q.$$

Therefore, by the naturality of Lie brackets,

$$\Pi_*\big([X_i, X_j]_q\big) = \left[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right]_{\Pi(q)} = 0.$$

Since $[X_i, X_j]$ takes its values in D by involutivity, and Π_* is injective on each fiber of D, this implies that $[X_i, X_j]_q = 0$ for each q, thus completing the proof. $\square$

As is often the case, embedded in the proof is a technique for finding integral manifolds. The idea is to use a coordinate projection to find commuting vector fields spanning the same distribution, and then use the technique of Example 18.7 to find a flat chart. Here is an example.

Example 19.11. Let $D \subset T\mathbb{R}^3$ be the distribution spanned by the vector fields

$$\begin{aligned} V &= x\frac{\partial}{\partial x} + \frac{\partial}{\partial y} + x(y+1)\frac{\partial}{\partial z}, \\ W &= \frac{\partial}{\partial x} + y\frac{\partial}{\partial z}. \end{aligned}$$

The computation of Example 4.14 showed that

$$[V, W] = -\frac{\partial}{\partial x} - y\frac{\partial}{\partial z} = -W,$$

so D is involutive. Let us try to find a flat chart in a neighborhood of the origin. Since D is complementary to the span of $\partial/\partial z$, the coordinate projection $\Pi\colon \mathbb{R}^3 \to \mathbb{R}^2$ given by $\Pi(x,y,z) = (x,y)$ induces an isomorphism $\Pi_*|_D\colon D_{(x,y,z)} \to T_{(x,y)}\mathbb{R}^2$ for each $(x,y,z) \in \mathbb{R}^3$. The proof of the Frobenius theorem shows that if we can find smooth local sections X, Y of D that are Π-related to $\partial/\partial x$ and $\partial/\partial y$, respectively, they will be commuting vector fields spanning D. It is easy to check that X, Y have this property if and only if they are of the form

$$\begin{aligned} X &= \frac{\partial}{\partial x} + u(x,y,z)\frac{\partial}{\partial z}, \\ Y &= \frac{\partial}{\partial y} + v(x,y,z)\frac{\partial}{\partial z} \end{aligned}$$

for some smooth real-valued functions u, v. A bit of linear algebra shows that the vector fields

$$\begin{aligned} X &= W = \frac{\partial}{\partial x} + y\frac{\partial}{\partial z}, \\ Y &= V - xW = \frac{\partial}{\partial y} + x\frac{\partial}{\partial z} \end{aligned}$$

are of this form and span D. The flows of these vector fields are easily found by solving the two systems of ODEs. Sparing the details, the flow of X is

$$\alpha_t(x,y,z) = (x+t, y, z+ty), \tag{19.6}$$

and that of Y is

$$\beta_t(x, y, z) = (x, y + t, z + tx). \tag{19.7}$$

Thus, by the procedure of Example 18.7, we can define the inverse ψ of our coordinate map by starting on the z-axis and flowing out along these two flows in succession:

$$\begin{aligned}\psi(u, v, w) &= \alpha_u \circ \beta_v(0, 0, w) \\ &= \alpha_u(0, v, z) \\ &= (u, v, z + uv).\end{aligned}$$

The flat coordinates we seek are given by inverting the map $(x, y, z) = \psi(u, v, w) = (u, v, z + uv)$, to yield

$$(u, v, w) = \psi^{-1}(x, y, z) = (x, y, z - xy).$$

It follows that the integral manifolds of D are the level sets of $w(x, y, z) = z - xy$. (Since the flat chart we have constructed is actually a global chart in this case, this describes all of the integral manifolds, not just the ones near the origin.)

◇ **Exercise 19.3.** Verify that the flows of X and Y are given by (19.6) and (19.7), respectively, and that the level sets of $z - xy$ are integral manifolds of D.

One important consequence of the Frobenius theorem is the following proposition, which will play an important role in the proof of the global version of the Frobenius theorem.

Proposition 19.12 (Local Structure of Integral Manifolds). *Let D be an involutive k-dimensional distribution on a smooth manifold M, and let (U, φ) be a flat chart for D. If N is any integral manifold of D, then $N \cap U$ is a countable disjoint union of open subsets of k-dimensional slices of U, each of which is open in N and embedded in M.*

Proof. Because the inclusion map $\iota\colon N \hookrightarrow M$ is continuous, $N \cap U = \iota^{-1}(U)$ is open in N, and thus consists of a countable disjoint union of connected components, each of which is open in N.

Let V be any component of $N \cap U$ (Figure 19.4). We will show first that V is contained in a single slice. Choosing some $p \in V$, it suffices to show that any $q \in V$ satisfies $x^i(q) = x^i(p)$ for $i = k + 1, \ldots, n$. Since V is connected, there exists a path γ in V from p to q, which we may take to be smooth by Problem 10-1. Because γ lies in V and V is an integral manifold of D, we have $\gamma'(t) \in T_{\gamma(t)}V = D_{\gamma(t)}$. Because D is spanned by $\partial/\partial x^1, \ldots, \partial/\partial x^k$ in U, this implies that the last $n - k$ components of $\gamma'(t)$ are zero: $(\gamma^i)'(t) \equiv 0$ for $i \geq k + 1$. Since the domain of γ is connected, this means γ^i is constant for these values of i, so p and q lie in a single slice S.

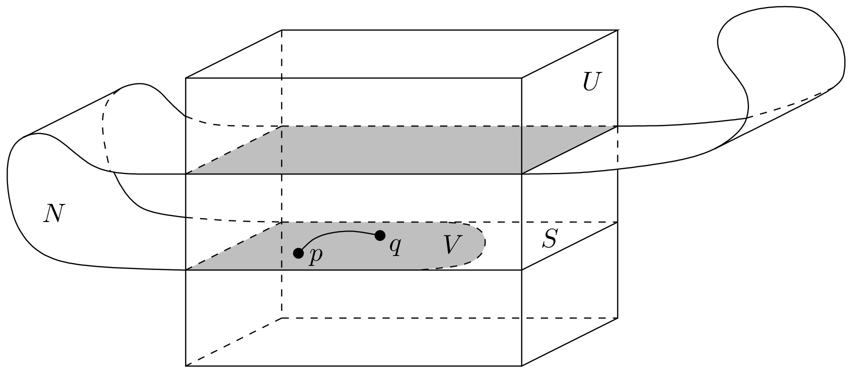


Figure 19.4. The local structure of an integral manifold.

Because S is embedded in M, the inclusion map $V \hookrightarrow M$ is also smooth as a map into S by Corollary 8.23. The inclusion $V \hookrightarrow S$ is thus an injective immersion between manifolds of the same dimension, and therefore a local diffeomorphism, an open map, and a homeomorphism onto an open subset of S. The inclusion map $V \hookrightarrow M$ is a composition of the smooth embeddings $V \hookrightarrow S \hookrightarrow M$, so it is a smooth embedding. $\square$

This local characterization of integral manifolds implies the following strengthening of our theorem about restricting the range of a smooth map. (This result will be used in Chapter 20.)

Proposition 19.13. *Suppose $H \subset N$ is an integral manifold of an involutive distribution D on N. If $F\colon M \to N$ is a smooth map such that $F(M) \subset H$, then F is smooth as a map from M to H.*

Proof. Let $p \in M$ be arbitrary, and set $q = F(p) \in H$. Let $(y^1, \dots, y^n)$ be flat coordinates for D on a neighborhood U of q. Choose smooth coordinates (x^i) for M on a connected neighborhood B of p such that $F(B) \subset U$. Writing the coordinate representation of F as

$$(y^1, \dots, y^n) = (F^1(x), \dots, F^n(x)),$$

the fact that $F(B) \subset H \cap U$ means that the coordinate functions $F^{k+1}, \dots, F^n$ take on only countably many values. Because B is connected, the intermediate value theorem implies that these coordinate functions are constant, and thus $F(B)$ lies in a single slice. On this slice, $(y^1, \dots, y^k)$ are smooth coordinates for H, so $F\colon N \to H$ has the local coordinate representation

$$(y^1, \dots, y^k) = (F^1(x), \dots, F^k(x)),$$

which is smooth. $\square$

Applications to Partial Differential Equations

In the next chapter, the Frobenius theorem will play a key role in our study of Lie subgroups. In this chapter, we concentrate on some applications to the study of partial differential equations (PDEs).

Our first application is not really so much an application as a simple rephrasing of the theorem. Because explicitly finding integral manifolds boils down to solving a system of PDEs, we can interpret the Frobenius theorem as an existence and uniqueness theorem for such systems.

Suppose $1 \le k \le n-1$, and (Y_i^j) is an $n \times k$ matrix of smooth real-valued functions defined on an open subset $U \subset \mathbb{R}^n$. Consider the following system of equations for a function $u \in C^\infty(U)$:

$$\begin{aligned} Y_1^1 \frac{\partial u}{\partial x^1} + \cdots + Y_1^n \frac{\partial u}{\partial x^n} &= 0, \\ &\vdots \\ Y_k^1 \frac{\partial u}{\partial x^1} + \cdots + Y_k^n \frac{\partial u}{\partial x^n} &= 0. \end{aligned} \tag{19.8}$$

This a homogeneous linear first-order system of PDEs. (Homogeneous means that the right-hand sides are all zero.) If $k > 1$, it is also *overdetermined*, meaning that there are more equations than unknown functions. In general, overdetermined systems have solutions only if they satisfy certain compatibility conditions; in this case, involutivity is the key. (Note that the case $k = 1$ was treated in Problem 17-11.)

Letting Y_i denote the vector field $Y_i^j \partial/\partial x^j$, (19.8) can be written more succinctly as

$$Y_1 u = \cdots = Y_k u = 0.$$

To avoid equations that are either redundant or self-contradictory, we will assume that the matrix (Y_i^j) has rank k at each point of U, or equivalently that the vector fields $Y_1, \dots, Y_k$ are independent. Clearly any constant function is a solution to (19.8), so the interesting question is whether there are nonconstant solutions, and if so how many there are.

Proposition 19.14. *Suppose $Y_1, \dots, Y_k$ are smooth independent vector fields on an open subset $U \subset \mathbb{R}^n$. Then the following are equivalent:*

(*a*) *For each $x_0 \in U$, there exist a neighborhood V of x_0 in U and $n-k$ smooth solutions to* (19.8) *whose differentials are independent at each point of V.*

(*b*) *The distribution spanned by $Y_1, \dots, Y_k$ is involutive.*

Proof. Let D denote the distribution spanned by $Y_1, \dots, Y_k$. If there are $(n-k)$ smooth solutions $u^1, \dots, u^{n-k}$ with independent differentials on any open set V, then the map $(u^1, \dots, u^{n-k})\colon V \to \mathbb{R}^{n-k}$ is a submersion whose level sets are integral manifolds of D, which implies that D is involutive on V. Conversely, suppose D is involutive, and let $(v^1, \dots, v^n)$ be any flat coordinates for D on an open set $V \subset U$. Then the coordinate functions $(v^{k+1}, \dots, v^n)$ are solutions to (19.8) on V whose differentials are independent at each point. $\square$

The next proposition gives a form of local uniqueness for (19.8).

Proposition 19.15. *Let $Y_1, \dots, Y_k$ be as above, and suppose $(u^1, \dots, u^{n-k})$ are any smooth solutions to* (19.8) *on an open subset $V \subset U$ with independent differentials at each point of V. Then each point of V has a neighborhood on which the most general solution to* (19.8) *is any function of the form $u(x) = g(u^1(x), \dots, u^{n-k}(x))$ for an arbitrary smooth real-valued function g of $n-k$ variables.*

Proof. Near any point $x_0 \in V$, we can choose functions $y^1, \dots, y^k$ on such that $(y^1, \dots, y^k, u^1, \dots, u^{n-k})$ are smooth coordinates for $\mathbb{R}^n$ on a neighborhood W of x_0. By shrinking W if necessary, we may assume its image is an open rectangle in $\mathbb{R}^n$. Since both D and the span of $(\partial/\partial y^i)$ are given by the simultaneous vanishing of $du^1, \dots, du^{n-k}$ at each point of W, it follows that this coordinate chart is flat for D. In these coordinates, if g is any smooth real-valued function of $(y^1, \dots, y^k, u^1, \dots, u^{n-k})$ defined on W,

$$\begin{aligned} & g \text{ is a solution to (19.8)} \\ & \quad \Longleftrightarrow dg_p|_{D_p} = 0 \text{ for each } p \in W \\ & \quad \Longleftrightarrow \partial g/\partial y^i = 0,\ i = 1, \dots, k \\ & \quad \Longleftrightarrow g \text{ is independent of } (y^1, \dots, y^k). \end{aligned}$$ $\square$

Example 19.16. Consider the system

$$\begin{aligned} x\frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} + x(y+1)\frac{\partial u}{\partial z} &= 0, \\ \frac{\partial u}{\partial x} + y\frac{\partial u}{\partial z} &= 0 \end{aligned}$$

of PDEs for a real-valued function $u(x, y, z)$. This can be rewritten as $Vu = Wu = 0$, where V and W are the vector fields of Example 19.11. Because the distribution spanned by V and W is involutive, Proposition 19.14 implies that near each point of $\mathbb{R}^n$ there is a smooth solution with nonvanishing differential. In fact, the computation in that example shows that the solutions are precisely the functions of the form $u(x, y, z) = g(z - xy)$ for any smooth real-valued function g of one variable.

Our next application of the Frobenius theorem to PDE theory is more substantial. We introduce it first in the case of a single real-valued function $f(x, y)$ of two independent variables, in which case the notation is considerably simpler.

Suppose we seek a solution f to the system

$$\begin{aligned}\frac{\partial f}{\partial x}(x,y) &= \alpha(x,y,f(x,y)),\\ \frac{\partial f}{\partial y}(x,y) &= \beta(x,y,f(x,y)),\end{aligned} \tag{19.9}$$

where α and β are smooth real-valued functions defined on some open subset $U \subset \mathbb{R}^3$. This is an overdetermined system of (possibly nonlinear) first-order partial differential equations. (In fact, almost any pair of smooth first-order partial differential equations for one unknown function can be put into this form, at least locally, simply by solving the two equations for $\partial f/\partial x$ and $\partial f/\partial y$. Whether this can be done in principle is a question that is completely answered by the implicit function theorem; whether it can be done in practice depends on the specific equations and how clever you are.)

To determine the necessary conditions that α and β must satisfy for solvability of (19.9), assume f is a smooth solution on some open set in $\mathbb{R}^2$. Because $\partial^2 f/\partial x\partial y = \partial^2 f/\partial y\partial x$, (19.9) implies

$$\frac{\partial}{\partial y}\big(\alpha(x,y,f(x,y))\big) = \frac{\partial}{\partial x}\big(\beta(x,y,f(x,y))\big)$$

and therefore by the chain rule

$$\frac{\partial \alpha}{\partial y} + \beta\frac{\partial \alpha}{\partial z} = \frac{\partial \beta}{\partial x} + \alpha\frac{\partial \beta}{\partial z}. \tag{19.10}$$

This is true at any point $(x,y,z) \in U$ provided there is a smooth solution f with $f(x,y) = z$. In particular, (19.10) is a necessary condition for (19.9) to have a solution in a neighborhood of any point (x_0,y_0) with arbitrary initial value $f(x_0,y_0) = z_0$. Using the Frobenius theorem, we can show that this condition is sufficient.

Proposition 19.17. *Suppose α and β are smooth real-valued functions defined on some open set $U \subset \mathbb{R}^3$ and satisfying* (19.10) *there. For any $(x_0,y_0,z_0) \in U$, there is a neighborhood V of (x_0,y_0) in $\mathbb{R}^2$ and a unique smooth function $f\colon V \to \mathbb{R}$ satisfying* (19.9) *and $f(x_0,y_0) = z_0$.*

Proof. The idea of the proof is that the system (19.9) determines the partial derivatives of f in terms of its values, and therefore determines the tangent plane to the graph of f at each point in terms of the coordinates of the point on the graph. This collection of tangent planes defines a smooth 2-dimensional distribution on U (Figure 19.5), and (19.10) is equivalent to the involutivity condition for this distribution.

If there were a solution f on an open set $V \subset \mathbb{R}^2$, the map $F\colon V \to U$ given by

$$F(x,y) = (x,y,f(x,y))$$

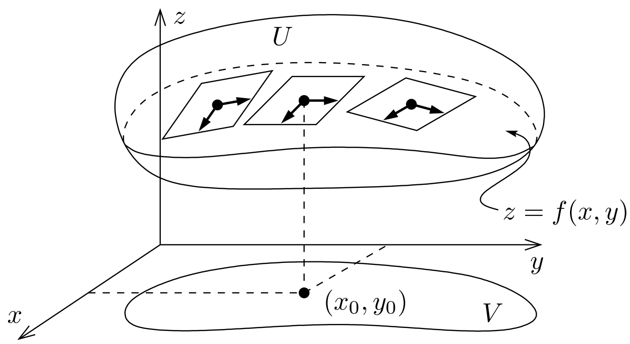


Figure 19.5. Solving for f by finding its graph.

would be a diffeomorphism onto the graph $\Gamma(f) \subset V \times \mathbb{R}$. At any point $p = F(x,y)$, the tangent space $T_p\Gamma(f)$ is spanned by the vectors

$$\begin{aligned} F_* \left.\frac{\partial}{\partial x}\right|_{(x,y)} &= \left.\frac{\partial}{\partial x}\right|_p + \frac{\partial f}{\partial x}(x,y) \left.\frac{\partial}{\partial z}\right|_p, \\ F_* \left.\frac{\partial}{\partial y}\right|_{(x,y)} &= \left.\frac{\partial}{\partial y}\right|_p + \frac{\partial f}{\partial y}(x,y) \left.\frac{\partial}{\partial z}\right|_p. \end{aligned} \tag{19.11}$$

The system (19.9) is satisfied if and only if

$$\begin{aligned} F_* \left.\frac{\partial}{\partial x}\right|_{(x,y)} &= \left.\frac{\partial}{\partial x}\right|_p + \alpha(x,y,f(x,y)) \left.\frac{\partial}{\partial z}\right|_p, \\ F_* \left.\frac{\partial}{\partial y}\right|_{(x,y)} &= \left.\frac{\partial}{\partial y}\right|_p + \beta(x,y,f(x,y)) \left.\frac{\partial}{\partial z}\right|_p. \end{aligned} \tag{19.12}$$

Let X and Y be the vector fields

$$\begin{aligned} X &= \frac{\partial}{\partial x} + \alpha(x,y,z)\frac{\partial}{\partial z}, \\ Y &= \frac{\partial}{\partial y} + \beta(x,y,z)\frac{\partial}{\partial z} \end{aligned} \tag{19.13}$$

on U, and let D be the distribution on U spanned by X and Y. A little linear algebra will convince you that (19.12) holds if and only if $\Gamma(f)$ is an integral manifold of D. A straightforward computation using (19.10) shows that $[X,Y] \equiv 0$, so given any point $p = (x_0, y_0, z_0) \in U$, there is an integral manifold N of D containing p. Let $\Phi\colon W \to \mathbb{R}$ be a defining function for N on some neighborhood W of p; for example, we could take Φ to be the third coordinate function in a flat chart. The tangent space to N at each point $p \in N$ (namely D_p) is equal to the kernel of $\Phi_*\colon T_p\mathbb{R}^3 \to T_p\mathbb{R}$. Since $\partial/\partial z|_p \notin D_p$ at any point p, this implies that $\partial\Phi/\partial z \neq 0$ at p, so by the implicit function theorem N is the graph of a smooth function $z = f(x,y)$

in some neighborhood of p. You can verify easily that f is a solution to the problem. Uniqueness follows immediately from Proposition 19.12. □

As in several cases we have seen before, the proof of Proposition 19.17 actually contains a procedure for finding solutions to (19.9): Find flat coordinates (u, v, w) for the distribution spanned by the vector fields X and Y defined by (19.13), and solve the equation $w =$ constant for $z = f(x, y)$. Some examples are given in Problem 19-10.

There is a straightforward generalization of this result to higher dimensions. The general statement of the theorem is a bit complicated, but verifying the necessary conditions in specific examples usually just amounts to computing mixed partial derivatives and applying the chain rule.

Proposition 19.18. *Suppose U is an open subset of $\mathbb{R}^n \times \mathbb{R}^m$, and $\alpha = (\alpha^i_j)\colon U \to \mathrm{M}(m \times n, \mathbb{R})$ is a smooth matrix-valued function satisfying*

$$\frac{\partial \alpha^i_j}{\partial x^k} + \alpha^l_k \frac{\partial \alpha^i_j}{\partial z^l} = \frac{\partial \alpha^i_k}{\partial x^j} + \alpha^l_j \frac{\partial \alpha^i_k}{\partial z^l} \qquad \text{for all } i, j, k,$$

where we denote a point in $\mathbb{R}^n \times \mathbb{R}^m$ by $(x, z) = (x^1, \ldots, x^n, z^1, \ldots, z^m)$. For any $(x_0, z_0) \in U$, there is a neighborhood V of x_0 in $\mathbb{R}^n$ and a unique smooth map $f\colon V \to \mathbb{R}^m$ such that $f(x_0) = z_0$ and the Jacobian of f satisfies

$$\frac{\partial f^i}{\partial x^j}(x^1, \ldots, x^n) = \alpha^i_j(x^1, \ldots, x^n, f^1(x), \ldots, f^m(x)).$$

◇ **Exercise 19.4.** Prove Proposition 19.18.

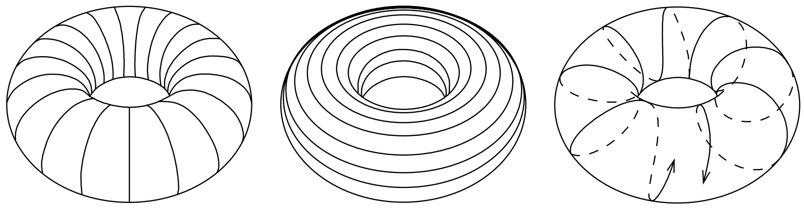
Foliations

When we put together all the maximal integral manifolds of a k-dimensional involutive distribution D, we obtain a decomposition of M into k-dimensional submanifolds that "fit together" locally like the slices in a flat chart.

We define a *foliation* of dimension k on an n-manifold M to be a collection of disjoint, connected, immersed k-dimensional submanifolds of M (called the *leaves* of the foliation), whose union is M, and such that in a neighborhood of each point $p \in M$ there is a smooth chart (U, φ) with the property that $\varphi(U)$ is a product of connected open sets $U' \times U'' \subset \mathbb{R}^k \times \mathbb{R}^{n-k}$, and each leaf of the foliation intersects U in either the empty set or a countable union of k-dimensional slices of the form $x^{k+1} = c^{k+1}, \ldots, x^n = c^n$. (Such a chart is called a *flat chart* for the foliation.)

Example 19.19 (Foliations).

(a) The collection of all k-dimensional affine subspaces of $\mathbb{R}^n$ parallel to $\mathbb{R}^k$ is a k-dimensional foliation of $\mathbb{R}^n$.



(a) (b) (c)

Figure 19.6. Foliations of the torus.

(b) The collection of all open rays of the form $\{\lambda x_0 : \lambda > 0\}$ is a 1-dimensional foliation of $\mathbb{R}^n \smallsetminus \{0\}$.

(c) The collection of all spheres centered at 0 is an $(n-1)$-dimensional foliation of $\mathbb{R}^n \smallsetminus \{0\}$.

(d) If M and N are connected smooth manifolds, the collection of subsets of the form $\{q\} \times N$ as q ranges over M forms a foliation of $M \times N$, each of whose leaves is diffeomorphic to N. For example, the collection of all circles of the form $\mathbb{S}^1 \times \{q\} \subset \mathbb{T}^2$ for $q \in \mathbb{S}^1$ yields a foliation of the torus $\mathbb{T}^2$ (Figure 19.6(a)). A different foliation of $\mathbb{T}^2$ is given by the collection of circles of the form $\{p\} \times \mathbb{S}^1$ (Figure 19.6(b)).

(e) If α is a fixed irrational number, the images of all curves of the form

$$\gamma_\theta(t) = (e^{it}, e^{i(\alpha t + \theta)})$$

as θ ranges over $[0, 2\pi)$ form a 1-dimensional foliation of the torus in which each leaf is dense (Figure 19.6(c)). (See Example 7.3 and Problem 7-3).

(f) The collection of connected components of the curves in the (y,z)-plane defined by the following equations is a foliation of $\mathbb{R}^2$ (Figure 19.7(a)):

$$\begin{aligned} z &= \sec y + c, && c \in \mathbb{R}; \\ y &= (k + \tfrac{1}{2})\pi, && k \in \mathbb{Z}. \end{aligned}$$

(g) If we rotate the curves of the previous example around the z-axis, we obtain a 2-dimensional foliation of $\mathbb{R}^3$ in which some of the leaves are diffeomorphic to disks and some are diffeomorphic to cylinders (Figure 19.7(b)).

The main fact about foliations is that they are in one-to-one correspondence with involutive distributions. One direction, expressed in the next lemma, is an easy consequence of the definition.

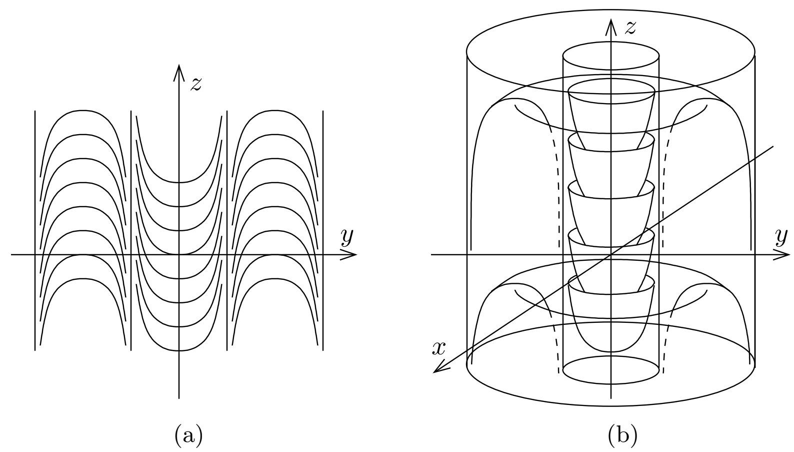


Figure 19.7. Foliations of $\mathbb{R}^2$ and $\mathbb{R}^3$.

Lemma 19.20. *Let $\mathcal{F}$ be a foliation on a smooth manifold M. The collection of tangent spaces to the leaves of $\mathcal{F}$ forms an involutive distribution on M.*

◇ **Exercise 19.5.** Prove Lemma 19.20.

The Frobenius theorem allows us to conclude the following converse, which is much more profound. By the way, it is worth noting that this result is one of the two primary reasons why the notion of immersed submanifold has been defined. (The other is for the study of Lie subgroups.)

Theorem 19.21 (Global Frobenius Theorem). *Let D be an involutive distribution on a smooth manifold M. The collection of all maximal connected integral manifolds of D forms a foliation of M.*

The theorem will be an easy consequence of the following lemma.

Lemma 19.22. *Suppose $D \subset TM$ is an involutive distribution, and let $\{N_\alpha\}_{\alpha \in A}$ be any collection of connected integral manifolds of D with a point in common. Then $N = \bigcup_\alpha N_\alpha$ has a unique smooth manifold structure making it into a connected integral manifold of D in which each N_α is an open submanifold.*

Proof. Define a topology on N by declaring a subset $U \subset N$ to be open if and only if $U \cap N_\alpha$ is open in N_α for each α. It is easy to check that this is a topology. To prove that each N_α is open in N, we need to show that $N_\alpha \cap N_\beta$ is open in N_β for each β. Let $q \in N_\alpha \cap N_\beta$ be arbitrary, and choose a flat chart for D on a neighborhood W of q (Figure 19.8). Let V_α, V_β denote the components of $N_\alpha \cap W$ and $N_\beta \cap W$, respectively, containing

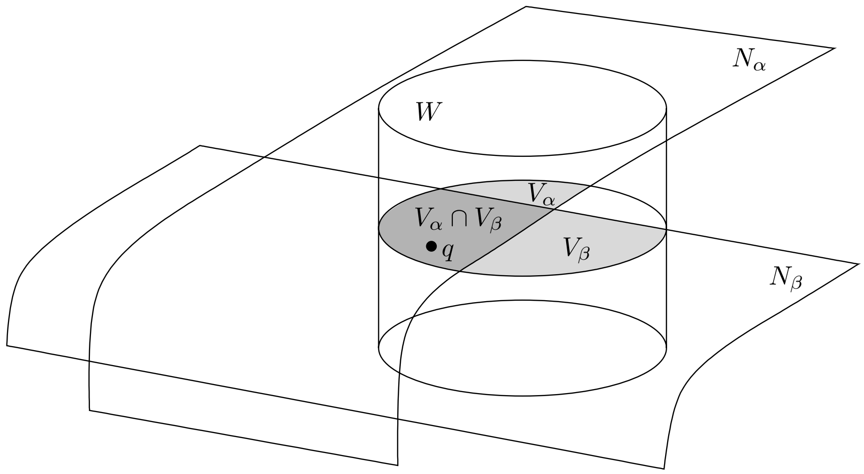


Figure 19.8. A union of integral manifolds.

q. By Lemma 19.12, V_α and V_β are open subsets of single slices with the subspace topology, and since both contain q, they both must lie in the same slice S. Thus $V_\alpha \cap V_\beta$ is open in S and also in both N_α and N_β. Since each $q \in N_\alpha \cap N_\beta$ has a neighborhood in N_β contained in the intersection, it follows that $N_\alpha \cap N_\beta$ is open in N_β as claimed. Clearly this is the unique topology on N with the property that each N_α is a subspace of N.

With this topology, N is locally Euclidean of dimension k, because each point $q \in N$ has a coordinate neighborhood V in some N_α, and V is an open subset of N because N_α is open in N. Moreover, the inclusion map $N \hookrightarrow M$ is continuous: For any open subset $U \subset M$, $U \cap N$ is open in N because $U \cap N_\alpha$ is open in N_α for each α.

To see that N is Hausdorff, let $q, q' \in N$ be given. There are disjoint open sets $U, U' \subset M$ containing q and q', respectively, and then (because inclusion $N \hookrightarrow M$ is continuous) $N \cap U$ and $N \cap U'$ are disjoint open subsets of N containing q.

Next we show that N is second countable. Let p be a point that lies on N_α for each α. We can cover M with countably many flat charts for D, say $\{W_i\}$. It suffices to show that $N \cap W_i$ is contained in a countable union of slices for each i, because any open subset of a single slice is second countable.

Suppose W_k is one of these flat charts and $S \subset W_k$ is a slice containing a point $q \in N$. There is some connected integral manifold N_α containing p and q. Because connected manifolds are path connected, there is a continuous path $\gamma\colon [0,1] \to N_\alpha$ connecting p and q. Since $\gamma[0,1]$ is compact, there exist finitely many numbers $0 = t_0 < t_1 < \cdots < t_m = 1$ such that $\gamma[t_{j-1}, t_j]$ is contained in one of the flat charts W_{i_j} for each j. Since $\gamma[t_{j-1}, t_j]$ is con-

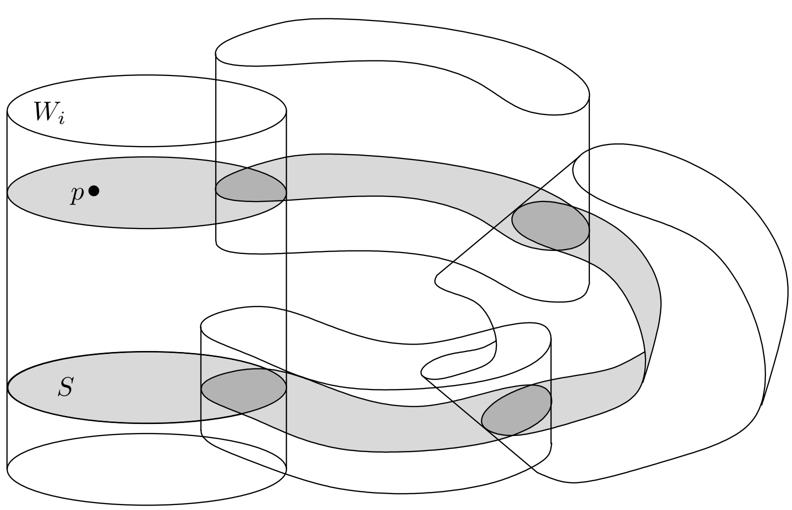


Figure 19.9. A slice S accessible from p.

nected, it is contained in a single component of $W_{i_j} \cap N_\alpha$ and therefore in a single slice $S_{i_j} \subset W_{i_j}$.

Let us say that a slice S of some W_k is *accessible from* p if there is a finite sequence of indices $i_0, \dots, i_m$ and for each i_j a slice $S_{i_j} \subset W_{i_j}$, with the properties that $p \in S_{i_0}$, $S_{i_m} = S$, and $S_{i_j} \cap S_{i_{j+1}} \neq \varnothing$ for each j (Figure 19.9). The discussion in the preceding paragraph showed that every slice that contains a point of N is accessible from p. To complete the proof of second countability, we just note that each S_{i_j} is itself an integral manifold, and therefore it meets at most countably many slices of $W_{i_{j+1}}$ by Proposition 19.12; thus there are only countably many slices accessible from p. Therefore, N is a topological manifold of dimension k. It is connected because it is a union of connected subspaces with a point in common.

To construct a smooth structure on N, we define an atlas consisting of all charts of the form $(S \cap N, \psi)$, where S is a single slice of some flat chart, and $\psi\colon S \to \mathbb{R}^k$ is the map whose coordinate representation is projection onto the first k coordinates: $\psi(x^1, \dots, x^n) = (x^1, \dots, x^k)$. Because any slice is an embedded submanifold, its smooth structure is uniquely determined, and thus whenever two such slices S, S' overlap the transition map $\psi' \circ \psi$ is smooth.

With respect to this smooth structure, the inclusion map $N \hookrightarrow M$ is an immersion (because it is a smooth embedding on each slice), and the tangent space to N at each point $q \in N$ is equal to D_q (because this is true for slices). The smooth structure is the unique one such that the inclusion $N \hookrightarrow M$ is an immersion, by the result of Problem 8-12(b). $\square$

Proof of the global Frobenius theorem. For each $p \in M$, let L_p be the union of all connected integral manifolds of D containing p. By the preceding lemma, L_p is a connected integral manifold of D containing p, and it is clearly maximal. If any two such maximal integral manifolds L_p and $L_{p'}$ intersect, their union $L_p \cup L_{p'}$ is an integral manifold containing both p and p', so by maximality $L_p = L_{p'}$. Thus the various maximal integral manifolds are either disjoint or identical.

If (U, φ) is any flat chart for D, then $L_p \cap U$ is a countable union of open subsets of slices by Proposition 19.12. For any such slice S, if $L_p \cap S$ is neither empty nor all of S, then $L_p \cup S$ is a connected integral manifold properly containing L_p, which contradicts the maximality of L_p. Therefore, $L_p \cap U$ is precisely a countable union of slices, so the collection $\{L_p : p \in M\}$ is the desired foliation. □

Problems

19-1. Let M be a smooth n-manifold. An ideal $\mathcal{I} \subset \mathcal{A}^*(M)$ is called a *Pfaffian system* if for some $k \leq n$, $\mathcal{I}$ is locally generated by $n-k$ independent 1-forms, in the following sense: There exist open sets $\{U_\alpha\}_{\alpha \in A}$ covering M, and for each $\alpha \in A$ there exist $n-k$ independent smooth 1-forms $\omega_\alpha^1, \dots, \omega_\alpha^{n-k}$ on U_α, such that an element $\eta \in \mathcal{A}^*(M)$ is in $\mathcal{I}$ if and only if for each $\alpha \in A$, the restriction of η to U_α is of the form

$$\eta|_{U_\alpha} = \sum_{i=1}^{n-k} \omega_\alpha^i \wedge \beta_\alpha^i$$

for some $\beta_\alpha^1, \dots, \beta_\alpha^{n-k} \in \mathcal{A}^*(U_\alpha)$. Show that an ideal $\mathcal{I} \subset \mathcal{A}^*(M)$ is a Pfaffian system if and only if there is a smooth distribution D on M such that $\mathcal{I} = \mathcal{I}(D)$.

19-2. Let D be a smooth k-dimensional distribution on a smooth n-manifold M, and suppose $\omega^1, \dots, \omega^{n-k}$ are smooth local defining forms for D on an open set $U \subset M$. Show that D is involutive on U if and only if the following identity holds for each $i = 1, \dots, n-k$:

$$d\omega^i \wedge \omega^1 \wedge \cdots \wedge \omega^{n-k} = 0.$$

19-3. Let $D \subset TM$ be a smooth distribution, and let $\Gamma(D) \subset \mathcal{T}(M)$ denote the space of smooth global sections of D. Show that D is involutive if and only if $\Gamma(D)$ is a Lie subalgebra of $\mathcal{T}(M)$.

19-4. If H is an integral manifold of an involutive distribution on N, and $\gamma\colon J \to N$ is a smooth curve whose image lies in H, show that $\gamma'(t)$ is in $T_{\gamma(t)}H \subset T_{\gamma(t)}N$ for all $t \in J$. (Compare this to the result of Problem 8-14.)

19-5. Let ω be a smooth 1-form on a smooth manifold M. A smooth real-valued function μ on some open subset $U \subset M$ is called an *integrating factor* for ω if $\mu\omega$ is exact on U.

(a) If ω is nowhere-vanishing, show that ω admits an integrating factor in a neighborhood of each point if and only if $d\omega \wedge \omega = 0$.
(b) If $\dim M = 2$, show that every nonvanishing smooth 1-form admits an integrating factor in a neighborhood of each point.

19-6. Let $U \subset \mathbb{R}^3$ be the subset where all three coordinates are positive, and let D be the distribution on U spanned by the vector fields

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

Find an explicit (global) flat chart for D on U.

19-7. Let D be the distribution on $\mathbb{R}^3$ spanned by

$$X = \frac{\partial}{\partial x} + yz\frac{\partial}{\partial z}, \qquad Y = \frac{\partial}{\partial y}.$$

(a) Find an integral submanifold of D passing through the origin.
(b) Is D involutive? Explain your answer in light of part (a).

19-8. Consider the following system of PDEs for $u \in C^\infty(\mathbb{R}^3)$:

$$-2z^2\frac{\partial u}{\partial x} + 2x\frac{\partial u}{\partial z} = 0,$$
$$-3z^2\frac{\partial u}{\partial y} + 2y\frac{\partial u}{\partial z} = 0.$$

(a) Suppose (x_0, y_0, z_0) is a point in $\mathbb{R}^3$ with $z_0 \neq 0$. Determine whether this system of equations has a solution with nonvanishing differential in a neighborhood of (x_0, y_0, z_0).
(b) Can you say anything about the existence or nonexistence of nonconstant solutions in a neighborhood of a point $(x_0, y_0, 0)$?

19-9. Consider the following system of PDEs for $f \in C^\infty(\mathbb{R}^4)$:

$$\frac{\partial f}{\partial x^1} + x^2\frac{\partial f}{\partial x^3} + 2(x^1 + x^2x^3)\frac{\partial f}{\partial x^4} = 0,$$
$$\frac{\partial f}{\partial x^2} - x^1\frac{\partial f}{\partial x^3} + 2(x^2 - x^1x^3)\frac{\partial f}{\partial x^4} = 0.$$

(a) Show that there do not exist two solutions f^1, f^2 with independent differentials on any open subset of $\mathbb{R}^4$.
(b) Show that there exists a solution with nonvanishing differential in a neighborhood of any point.

19-10. Of the systems of partial differential equations below, determine which ones have solutions $z(x, y)$ (or, for part (c), $z(x, y)$ and $w(x, y)$)

in a neighborhood of the origin for arbitrary positive values of $z(0,0)$ (respectively, $z(0,0)$ and $w(0,0)$).

(a) $\dfrac{\partial z}{\partial x} = z\cos y; \qquad \dfrac{\partial z}{\partial y} = -z\log z\tan y.$

(b) $\dfrac{\partial z}{\partial x} = e^{xz}; \qquad \dfrac{\partial z}{\partial y} = xe^{yz}.$

(c) $\dfrac{\partial z}{\partial x} = z; \qquad \dfrac{\partial z}{\partial y} = w; \qquad \dfrac{\partial w}{\partial x} = w; \qquad \dfrac{\partial w}{\partial y} = z.$

19-11. Show that the local uniqueness property of Proposition 19.15 cannot be strengthened to global uniqueness in general, by considering the partial differential equation $\partial u/\partial x = 0$ on the set $U = \{(x,y) \in \mathbb{R}^2 : x > 0,\ 1 < x^2 + y^2 < 4\}$.

(a) Show that each point $p \in U$ has a neighborhood on which every smooth solution is a function of y alone.

(b) Show that there is a solution on all of U that is not a function of y alone.

19-12. Let D be an involutive distribution on a smooth manifold M, and let N be a connected integral manifold of D. If N is a closed subset of M, show that N is a maximal connected integral manifold and is therefore a leaf of the foliation determined by D.

19-13. If $F\colon M \to N$ is a submersion, show that the connected components of the level sets of F form a foliation of M.

19-14. If G is a connected Lie group acting smoothly, freely, and properly on a smooth manifold M, show that the orbits of G form a foliation of M. Give a counterexample to show that the conclusion is false if the action is not free.

19-15. Suppose (M,ω) is a symplectic manifold and $S \subset M$ is a coisotropic submanifold. An immersed submanifold $N \subset S$ is said to be *characteristic* if $T_pN = (T_pS)^{\perp}$ for each $p \in S$. Show that the set of connected characteristic submanifolds of S forms a foliation of S, whose dimension is equal to the codimension of S in M.

20
Lie Groups and Their Lie Algebras

In this chapter, we apply the tools developed in Chapters 17–19 (flows, Lie derivatives, and foliations) to delve deeper into the relationships between Lie groups and Lie algebras.

In the first section, we define one-parameter subgroups, which are just Lie group homomorphisms from $\mathbb{R}$, and show that there is a one-to-one correspondence between elements of $\operatorname{Lie}(G)$ and one-parameter subgroups of G. Next we introduce the exponential map, which is a canonical smooth map from the Lie algebra into the group. We give two applications of the exponential map: First, we use it to prove the closed subgroup theorem, which says that every topologically closed subgroup of a Lie group is actually an embedded Lie subgroup; and second, we use it to analyze the relationship between the adjoint representation of a Lie group and that of its Lie algebra.

Then we use the Frobenius theorem to show that every Lie subalgebra of $\operatorname{Lie}(G)$ corresponds to some Lie subgroup of G, and apply this to show that every Lie algebra homomorphism from $\operatorname{Lie}(G)$ to $\operatorname{Lie}(H)$ is induced by some Lie group homomorphism.

The culmination of the chapter is a description of the fundamental correspondence between Lie groups and Lie algebras: There is a one-to-one correspondence (up to isomorphism) between finite-dimensional Lie algebras and simply-connected Lie groups, and all of the connected Lie groups with a given Lie algebra are quotients of the simply connected one by discrete normal subgroups.

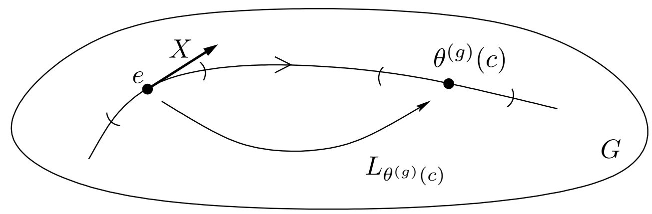


Figure 20.1. Proof that left-invariant vector fields are complete.

One-Parameter Subgroups

Let G be a Lie group. We define a *one-parameter subgroup* of G to be a Lie group homomorphism $F\colon \mathbb{R} \to G$. Notice that, by this definition, a one-parameter subgroup is *not* a Lie subgroup of G, but rather a homomorphism into G. (However, the *image* of a one-parameter subgroup is a Lie subgroup; see Problem 20-5.)

We will see shortly that the one-parameter subgroups are precisely the integral curves of left-invariant vector fields starting at the identity. Before we do so, however, we need the following lemma.

Lemma 20.1. *Every left-invariant vector field on a Lie group is complete.*

Proof. Let G be a Lie group, let $X \in \operatorname{Lie}(G)$, and let θ denote the flow of X. Suppose some maximal integral curve $\theta^{(g)}$ is defined on an interval $(a, b) \subset \mathbb{R}$, and assume that $b < \infty$. (The case $a > -\infty$ is handled similarly.) We will use left-invariance to define an integral curve on a slightly larger interval.

Left-invariance of X means that X is L_g-related to itself, and so it follows from Lemma 18.4 that

$$L_g \circ \theta_t = \theta_t \circ L_g \tag{20.1}$$

on the domain of θ_t. Observe that the integral curve $\theta^{(e)}$ starting at the identity is defined at least on some interval $(-\varepsilon, \varepsilon)$ for $\varepsilon > 0$. Choose some $c \in (b - \varepsilon, b)$, and define a new curve $\gamma\colon (a, c + \varepsilon) \to G$ by

$$\gamma(t) = \begin{cases} \theta^{(g)}(t) & t \in (a, b) \\ L_{\theta^{(g)}(c)}(\theta^{(e)}(t - c)) & t \in (c - \varepsilon, c + \varepsilon). \end{cases}$$

(See Figure 20.1.) By (20.1), when $t \in (a, b) \cap (c - \varepsilon, c + \varepsilon)$, we have

$$\begin{aligned}
L_{\theta^{(g)}(c)}(\theta^{(e)}(t-c)) &= L_{\theta^{(g)}(c)}(\theta_{t-c}(e)) \\
&= \theta_{t-c}(L_{\theta^{(g)}(c)}(e)) \\
&= \theta_{t-c}(\theta_c(g)) \\
&= \theta_t(g) \\
&= \theta^{(g)}(t),
\end{aligned}$$

so the two definitions of γ agree where they overlap.

Now, γ is clearly an integral curve of X on (a,b), and for $t_0 \in (c-\varepsilon, c+\varepsilon)$ we use left-invariance of X to compute

$$\begin{aligned}
\gamma'(t_0) &= \left.\frac{d}{dt}\right|_{t=t_0} L_{\theta^{(g)}(c)}(\theta^{(e)}(t-c)) \\
&= (L_{\theta^{(g)}(c)})_* \left.\frac{d}{dt}\right|_{t=t_0} \theta^{(e)}(t-c) \\
&= (L_{\theta^{(g)}(c)})_* X_{\theta^{(e)}(t_0-c)} \\
&= X_{\gamma(t_0)}.
\end{aligned}$$

Thus γ is an integral curve of X defined for $t \in (a, c+\varepsilon)$. Since $c+\varepsilon > b$, this contradicts the assumption that (a,b) was the maximal domain of $\theta^{(g)}$. □

Proposition 20.2. *Let G be a Lie group, and let $X \in \mathrm{Lie}(G)$. The integral curve of X starting at e is a one-parameter subgroup of G.*

Proof. Let θ be the flow of X, so that $\theta^{(e)}\colon \mathbb{R} \to G$ is the integral curve in question. Clearly $\theta^{(e)}$ is smooth, so we need only show that it is a group homomorphism, i.e., that $\theta^{(e)}(s+t) = \theta^{(e)}(s)\theta^{(e)}(t)$ for all $s, t \in \mathbb{R}$. Using (20.1) once again, we compute

$$\begin{aligned}
\theta^{(e)}(s)\theta^{(e)}(t) &= L_{\theta^{(e)}(s)}\theta_t(e) \\
&= \theta_t(L_{\theta^{(e)}(s)}(e)) \\
&= \theta_t(\theta^{(e)}(s)) \\
&= \theta_t(\theta_s(e)) \\
&= \theta_{t+s}(e) \\
&= \theta^{(e)}(t+s).
\end{aligned}$$ □

The main result of this section is that all one-parameter subgroups are obtained in this way.

Theorem 20.3. *Every one-parameter subgroup of a Lie group is an integral curve of a left-invariant vector field. Thus there are one-to-one correspondences*

$$\{\textit{one-parameter subgroups of } G\} \longleftrightarrow \mathrm{Lie}(G) \longleftrightarrow T_eG.$$

In particular, a one-parameter subgroup is uniquely determined by its initial tangent vector in T_eG.

Proof. Let $F\colon \mathbb{R} \to G$ be a one-parameter subgroup, and let $X = F_*(d/dt) \in \operatorname{Lie}(G)$, where we think of d/dt as a left-invariant vector field on $\mathbb{R}$. To prove the theorem, it suffices to show that F is an integral curve of X. Recall that $F_*(d/dt)$ is defined as the unique left-invariant vector field on G that is F-related to d/dt (see Theorem 4.25). Therefore, for any $t_0 \in \mathbb{R}$,

$$F'(t_0) = F_* \left.\frac{d}{dt}\right|_{t_0} = X_{F(t_0)},$$

so F is an integral curve of X. □

Given $X \in \operatorname{Lie}(G)$, we will call the one-parameter subgroup determined in this way the *one-parameter subgroup generated by* X.

The one-parameter subgroups of the general linear group are not hard to compute explicitly.

Proposition 20.4. *For any $A \in \mathfrak{gl}(n,\mathbb{R})$, let*

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k = I_n + A + \frac{1}{2}A^2 + \cdots . \tag{20.2}$$

This series converges to an invertible matrix $e^A \in \mathrm{GL}(n,\mathbb{R})$, and the one-parameter subgroup of $\mathrm{GL}(n,\mathbb{R})$ generated by $A \in \mathfrak{gl}(n,\mathbb{R})$ is $F(t) = e^{tA}$.

Proof. First we verify convergence. From Exercise A.56 in the Appendix, matrix multiplication satisfies $|AB| \le |A|\,|B|$, where the norm is the Euclidean norm on $\mathfrak{gl}(n,\mathbb{R})$ under its obvious identification with $\mathbb{R}^{n^2}$. It follows by induction that $|A^k| \le |A|^k$. The Weierstrass M-test shows that (20.2) converges uniformly on any bounded subset of $\mathfrak{gl}(n,\mathbb{R})$ (by comparison with the series $\sum_k (1/k!)c^k = e^c$).

Fix $A \in \mathfrak{gl}(n,\mathbb{R})$. The one-parameter subgroup generated by A is an integral curve of the left-invariant vector field $\widetilde{A}$ on $\mathrm{GL}(n,\mathbb{R})$, and therefore satisfies the ODE initial value problem

$$\begin{aligned} F'(t) &= \widetilde{A}_{F(t)},\\ F(0) &= I_n. \end{aligned}$$

Using formula (4.11) for $\widetilde{A}$, the condition for F to be an integral curve can be rewritten as

$$(F^i_k)'(t) = F^i_j(t) A^j_k,$$

or in matrix notation

$$F'(t) = F(t)A.$$

We will show that $F(t) = e^{tA}$ satisfies this equation. Since $F(0) = I_n$, this implies that F is the unique integral curve of $\widetilde{A}$ starting at the identity and is therefore the desired one-parameter subgroup.

To see that F is differentiable, we note that differentiating the series (20.2) formally term-by-term yields the result

$$\begin{aligned} F'(t) &= \sum_{k=1}^{\infty} \frac{k}{k!} t^{k-1} A^k \\ &= \left(\sum_{k=1}^{\infty} \frac{1}{(k-1)!} t^{k-1} A^{k-1} \right) A \\ &= F(t)A. \end{aligned}$$

Since the differentiated series converges uniformly on bounded sets (because, apart from the additional factor of A, it is the same series!), the term-by-term differentiation is justified. A similar computation shows that $F'(t) = AF(t)$. By smoothness of solutions to ODEs, F is a smooth curve.

It remains only to show that $F(t)$ is invertible for all t, so that F actually takes its values in $\mathrm{GL}(n, \mathbb{R})$. If we let $\sigma(t) = F(t)F(-t) = e^{tA}e^{-tA}$, then σ is a smooth curve in $\mathfrak{gl}(n, \mathbb{R})$, and by the previous computation and the product rule it satisfies

$$\sigma'(t) = (F(t)A)F(-t) - F(t)(AF(-t)) = 0.$$

Therefore σ is the constant curve $\sigma(t) \equiv \sigma(0) = I_n$, which is to say that $F(t)F(-t) = I_n$. Substituting $-t$ for t, we obtain $F(-t)F(t) = I_n$, which shows that $F(t)$ is invertible and $F(t)^{-1} = F(-t)$. □

Next we would like to compute the one-parameter subgroups of subgroups of $\mathrm{GL}(n, \mathbb{R})$, such as $\mathrm{O}(n)$. To do so, we need the following result.

Proposition 20.5. *Suppose $H \subset G$ is a Lie subgroup. The one-parameter subgroups of H are precisely those one-parameter subgroups of G whose initial tangent vectors lie in T_eH.*

Proof. Let $F \colon \mathbb{R} \to H$ be a one-parameter subgroup. Then the composite map

$$\mathbb{R} \xrightarrow{F} H \hookrightarrow G$$

is a Lie group homomorphism and thus a one-parameter subgroup of G, which clearly satisfies $F'(0) \in T_eH$.

Conversely, suppose $F \colon \mathbb{R} \to G$ is a one-parameter subgroup whose initial tangent vector lies in T_eH. Let $\widetilde{F} \colon \mathbb{R} \to H$ be the one-parameter subgroup of H with the same initial tangent vector $\widetilde{F}'(0) = F'(0) \in T_eH \subset T_eG$. As in the preceding paragraph, by composing with the inclusion map, we can also consider $\widetilde{F}$ as a one-parameter subgroup of G. Since F and $\widetilde{F}$ are both one-parameter subgroups of G with the same initial tangent vector, they must be equal. □

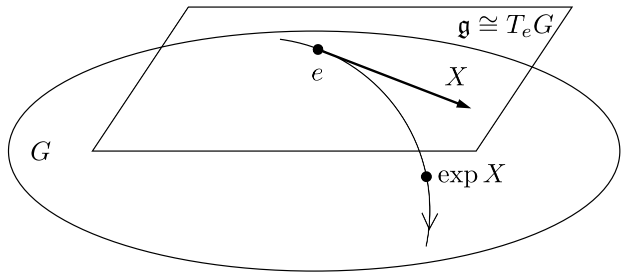


Figure 20.2. The exponential map.

Example 20.6. If H is a Lie subgroup of $\mathrm{GL}(n,\mathbb{R})$, the preceding proposition shows that the one-parameter subgroups of H are precisely the maps of the form $F(t) = e^{tA}$ for $A \in \mathfrak{h}$, where $\mathfrak{h} \subset \mathfrak{gl}(n,\mathbb{R})$ is the subalgebra corresponding to $\mathrm{Lie}(H)$ as in Proposition 8.36. For example, taking $H = \mathrm{O}(n)$, this shows that the exponential of any skew-symmetric matrix is orthogonal.

The Exponential Map

In the preceding section, we saw that the matrix exponential maps $\mathfrak{gl}(n,\mathbb{R})$ to $\mathrm{GL}(n,\mathbb{R})$ and takes each line through the origin to a one-parameter subgroup. This has a powerful generalization to arbitrary Lie groups.

Given a Lie group G with Lie algebra $\mathfrak{g}$, define a map $\exp\colon \mathfrak{g} \to G$, called the *exponential map* of G, by letting $\exp X = F(1)$, where F is the one-parameter subgroup generated by X, or equivalently the integral curve of X starting at the identity (Figure 20.2).

Example 20.7. The results of the preceding section show that the exponential map of $\mathrm{GL}(n,\mathbb{R})$ (or any Lie subgroup of it) is given by $\exp A = e^A$. This, obviously, is the reason for the term exponential map.

Example 20.8. If V is a finite-dimensional real vector space, a choice of basis for V yields isomorphisms $\mathrm{GL}(V) \cong \mathrm{GL}(n,\mathbb{R})$ and $\mathfrak{gl}(V) \cong \mathfrak{gl}(n,\mathbb{R})$. The analysis of the $\mathrm{GL}(n,\mathbb{R})$ case then shows that the exponential map of $\mathrm{GL}(V)$ can be written in the form

$$\exp A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k, \tag{20.3}$$

where we consider $A \in \mathfrak{gl}(V)$ as a linear map from V to itself, and $A^k = A \circ \cdots \circ A$ is the k-fold composition of A with itself.

Proposition 20.9 (Properties of the Exponential Map). *Let G be a Lie group and let $\mathfrak{g}$ be its Lie algebra.*

(*a*) *The exponential map is a smooth map from $\mathfrak{g}$ to G.*

(*b*) *For any $X \in \mathfrak{g}$, $F(t) = \exp tX$ is the one-parameter subgroup of G generated by X.*

(*c*) *For any $X \in \mathfrak{g}$, $\exp(s+t)X = \exp sX \exp tX$.*

(*d*) *The push-forward $\exp_*\colon T_0\mathfrak{g} \to T_eG$ is the identity map, under the canonical identifications of both $T_0\mathfrak{g}$ and T_eG with $\mathfrak{g}$ itself.*

(*e*) *The exponential map restricts to a diffeomorphism from some neighborhood of 0 in $\mathfrak{g}$ to a neighborhood of e in G.*

(*f*) *If H is another Lie group and $\mathfrak{h}$ is its Lie algebra, for any Lie group homomorphism $F\colon G \to H$, the following diagram commutes:*

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "F_*"] \arrow[d, "\exp"'] & \mathfrak{h} \arrow[d, "\exp"] \\ G \arrow[r, "F"'] & H \end{tikzcd} \tag{20.4}$$

(*g*) *The flow θ of a left-invariant vector field X is given by $\theta_t = R_{\exp tX}$ (right multiplication by $\exp tX$).*

Proof. For this proof, for any $X \in \mathfrak{g}$ we let $\theta_{(X)}$ denote the flow of X. To prove (a), we need to show that the expression $\theta^{(e)}_{(X)}(1)$ depends smoothly on X, which amounts to showing that the flow varies smoothly as the vector field varies. This is a situation not covered by the fundamental theorem on flows, but we can reduce it to that theorem by the following simple trick. Define a smooth vector field Ξ on the product manifold $G \times \mathfrak{g}$ by

$$\Xi_{(g,X)} = (X_g, 0) \in T_gG \oplus T_X\mathfrak{g} \cong T_{(g,X)}(G \times \mathfrak{g}).$$

(See Figure 20.3.) It is easy to verify that the flow Θ of Ξ is given by

$$\Theta_t(g, X) = (\theta_{(X)}(t, g), X).$$

By the fundamental theorem on flows, Θ is a smooth map. Since $\exp X = \pi_1(\Theta_1(e, X))$, where $\pi_1\colon G \times \mathfrak{g} \to G$ is the projection, it follows that exp is smooth.

Since the one-parameter subgroup generated by X is equal to the integral curve of X starting at e, to prove (b) it suffices to show that $\exp tX = \theta^{(e)}_{(X)}(t)$, or in other words that

$$\theta^{(e)}_{(tX)}(1) = \theta^{(e)}_{(X)}(t). \tag{20.5}$$

In fact, we will prove that for all $s, t \in \mathbb{R}$,

$$\theta^{(e)}_{(tX)}(s) = \theta^{(e)}_{(X)}(st), \tag{20.6}$$

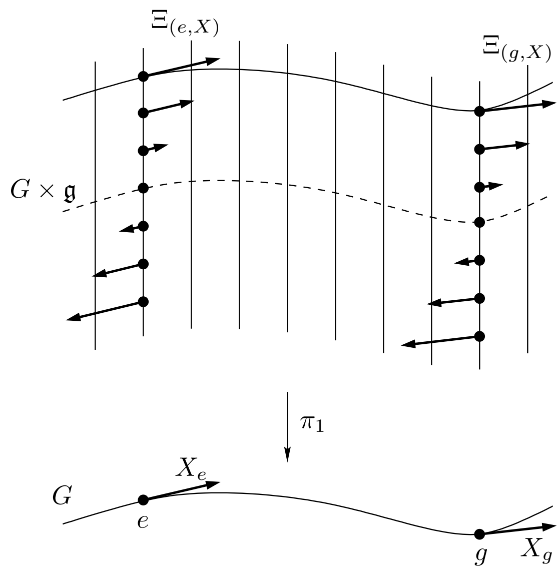


Figure 20.3. Proof that the exponential map is smooth.

which clearly implies (20.5).

To prove (20.6), fix $t \in \mathbb{R}$ and define a smooth curve $\gamma\colon \mathbb{R} \to G$ by

$$\gamma(s) = \theta^{(e)}_{(X)}(st).$$

By the chain rule,

$$\gamma'(s) = t(\theta^{(e)}_{(X)})'(st) = tX_{\gamma(s)},$$

so γ is an integral curve of the vector field tX. Since $\gamma(0) = e$, by uniqueness of integral curves we must have $\gamma(s) = \theta^{(e)}_{(tX)}(s)$, which is (20.6). This proves (b).

Next, (c) follows immediately from (b), because $t \mapsto \exp tX$ is a group homomorphism.

To prove (d), let $X \in \mathfrak{g}$ be arbitrary, and let $\sigma\colon \mathbb{R} \to \mathfrak{g}$ be the curve $\sigma(t) = tX$. Then $\sigma'(0) = X$, and (b) implies

$$\begin{aligned}
\exp_* X &= \exp_* \sigma'(0) \\
&= (\exp \circ \sigma)'(0) \\
&= \left.\frac{d}{dt}\right|_{t=0} \exp tX \\
&= X.
\end{aligned}$$

Part (e) then follows immediately from (d) and the inverse function theorem.

Next, to prove (f) we need to show that $\exp(F_*X) = F(\exp X)$ for any $X \in \mathfrak{g}$. In fact, we will show that for all $t \in \mathbb{R}$,

$$\exp(tF_*X) = F(\exp tX).$$

The left-hand side is, by (b), the one-parameter subgroup generated by F_*X. Thus if we put $\sigma(t) = F(\exp tX)$, it suffices to show that $\sigma\colon \mathbb{R} \to G$ is a group homomorphism satisfying $\sigma'(0) = F_*X$. We compute

$$\sigma'(0) = \left.\frac{d}{dt}\right|_{t=0} F(\exp tX) = F_* \left.\frac{d}{dt}\right|_{t=0} \exp tX = F_*X$$

and

$$\begin{aligned}
\sigma(s+t) &= F(\exp(s+t)X) \\
&= F(\exp sX \exp tX) && \text{(by (c))} \\
&= F(\exp sX)F(\exp tX) && \text{(since } F \text{ is a homomorphism)} \\
&= \sigma(s)\sigma(t).
\end{aligned}$$

Finally, to show that $\theta_{(X)t} = R_{\exp tX}$, we use part (b) and (20.1) to show that for any $g \in G$,

$$\begin{aligned}
R_{\exp tX}(g) &= g \exp tX \\
&= L_g(\exp tX) \\
&= L_g(\theta_{(X)t}(e)) \\
&= \theta_{(X)t}(L_g(e)) \\
&= \theta_{(X)t}(g).
\end{aligned}$$

□

The Closed Subgroup Theorem

One of the most powerful applications of the exponential map is to show that every closed subgroup of a Lie group is actually an embedded Lie subgroup. For example, this theorem allows us to strengthen the homogeneous space construction theorem (Theorem 9.18), because we need only assume that the subgroup H is topologically closed in G, not that it is a closed Lie subgroup. A similar remark applies to Proposition 9.24 about sets with transitive group actions.

We begin with the a simple result that shows how group multiplication in G is reflected "to first order" in the vector space structure of $\mathfrak{g}$.

Proposition 20.10. *Let G be a Lie group and let $\mathfrak{g}$ be its Lie algebra. For any $X, Y \in \mathfrak{g}$, there is a smooth function $Z\colon (-\varepsilon, \varepsilon) \to \mathbb{R}$ satisfying $Z(0) = 0$, and such that the following identity holds for all $t \in (-\varepsilon, \varepsilon)$:*

$$(\exp tX)(\exp tY) = \exp t(X + Y + Z(t)). \tag{20.7}$$

Proof. Since the exponential map is a diffeomorphism on some neighborhood of the origin in $\mathfrak{g}$, there is some $\varepsilon > 0$ such that the map $\gamma\colon (-\varepsilon, \varepsilon) \to \mathfrak{g}$ defined by

$$\gamma(t) = \exp^{-1}(\exp tX \exp tY)$$

is smooth. It obviously satisfies $\gamma(0) = 0$ and

$$\exp tX \exp tY = \exp \gamma(t).$$

Observe that we can write γ as the composition

$$\mathbb{R} \xrightarrow{e_X \times e_Y} G \times G \xrightarrow{m} G \xrightarrow{\exp^{-1}} G,$$

where $e_X(t) = \exp tX$ and $e_Y(t) = \exp tY$. Problem 3-6 shows that $m_*(X, Y) = X + Y$ for $X, Y \in T_eG$, which implies

$$\gamma'(0) = \exp_*^{-1}(e'_X(0) + e'_Y(0)) = X + Y.$$

Therefore, the first-order Taylor formula for γ reads

$$\gamma(t) = t(X + Y) + tZ(t)$$

for some smooth function Z satisfying $Z(0) = 0$. □

Theorem 20.11 (Closed Subgroup Theorem). *Suppose G is a Lie group and $H \subset G$ is a subgroup that is also a closed subset of G. Then H is an embedded Lie subgroup.*

Proof. By Proposition 8.28, it suffices to show that H is an embedded submanifold of G. We begin by identifying a subspace of $\mathrm{Lie}(G)$ that will turn out to be the Lie algebra of H.

Let $\mathfrak{g} = \mathrm{Lie}(G)$, and define a subset $\mathfrak{h} \subset \mathfrak{g}$ by

$$\mathfrak{h} = \{X \in \mathfrak{g} : \exp tX \in H \text{ for all } t \in \mathbb{R}\}.$$

We need to show that $\mathfrak{h}$ is a vector subspace of $\mathfrak{g}$. It is obvious from the definition that if $X \in \mathfrak{h}$, then $tX \in \mathfrak{h}$ for all $t \in \mathbb{R}$. To see that $\mathfrak{h}$ is closed under vector addition, let $X, Y \in \mathfrak{h}$ be arbitrary. Observe that formula (20.7) implies that for any $t \in \mathbb{R}$ and any sufficiently large integer n,

$$\left(\exp \frac{t}{n} X\right)\left(\exp \frac{t}{n} Y\right) = \exp \frac{t}{n}\left(X + Y + Z\left(\frac{t}{n}\right)\right)$$

with $Z(0) = 0$, and a simple induction using Proposition 20.9(c) yields

$$\begin{aligned}\left(\left(\exp \frac{t}{n} X\right)\left(\exp \frac{t}{n} Y\right)\right)^n &= \left(\exp \frac{t}{n}\left(X + Y + Z\left(\frac{t}{n}\right)\right)\right)^n \\ &= \exp t\left(X + Y + Z\left(\frac{t}{n}\right)\right).\end{aligned}$$

Fixing t and taking the limit as $n \to \infty$, we obtain

$$\lim_{n\to\infty}\left(\left(\exp \frac{t}{n} X\right)\left(\exp \frac{t}{n} Y\right)\right)^n = \exp t(X + Y).$$

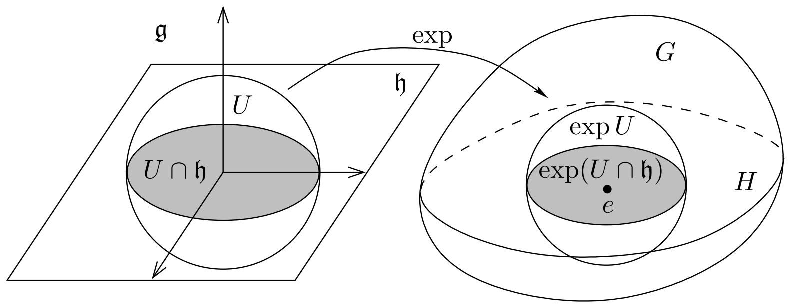


Figure 20.4. A neighborhood used to construct a slice chart for H.

Since $\exp((t/n)X)$ and $\exp((t/n)Y)$ are in H by assumption, and H is a closed subgroup of G, this shows that $\exp t(X+Y)$ is in H for each t. Thus $X + Y \in \mathfrak{h}$, and so $\mathfrak{h}$ is a subspace.

Next we will show that there is a neighborhood U of the origin in $\mathfrak{g}$ on which the exponential map of G is a diffeomorphism, and which has the property that

$$\exp(U \cap \mathfrak{h}) = (\exp U) \cap H. \tag{20.8}$$

(See Figure 20.4.) This will enable us to construct a slice chart for H near the identity, and we will then use left translation to get a slice chart in a neighborhood of any point of H.

If U is any neighborhood of $0 \in \mathfrak{g}$ on which exp is a diffeomorphism, then $\exp(U \cap \mathfrak{h}) \subset (\exp U) \cap H$ by definition of $\mathfrak{h}$. So to find a neighborhood satisfying (20.8), all we need to do is to show that U can be chosen small enough that $(\exp U) \cap H \subset \exp(U \cap \mathfrak{h})$. Assume this is not possible. Let $\{U_i\}$ be any countable neighborhood basis for $\mathfrak{g}$ at 0 (for example, a countable sequence of balls whose radii approach zero). The assumption implies that for each i, there exists $h_i \in (\exp U_i) \cap H$ such that $h_i \notin \exp(U_i \cap \mathfrak{h})$.

Choose a basis $E_1, \dots, E_k$ for $\mathfrak{h}$ and extend it to a basis $E_1, \dots, E_m$ for $\mathfrak{g}$. Let $\mathfrak{b}$ be the subspace spanned by $E_{k+1}, \dots, E_m$, so that $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{b}$ as vector spaces. As soon as i is large enough, the map from $\mathfrak{h} \oplus \mathfrak{b}$ to G given by $X + Y \mapsto \exp X \exp Y$ is a diffeomorphism from U_i to a neighborhood of e in G (see Problem 20-2). Therefore we can write

$$h_i = \exp X_i \exp Y_i$$

for some $X_i \in U_i \cap \mathfrak{h}$ and $Y_i \in U_i \cap \mathfrak{b}$, with $Y_i \neq 0$ because $h_i \notin \exp(U_i \cap \mathfrak{h})$ (Figure 20.5). Since $\{U_i\}$ is a neighborhood basis, $Y_i \to 0$ as $i \to \infty$. Observe that $\exp X_i \in H$ by definition of $\mathfrak{h}$, so it follows that $\exp Y_i = (\exp X_i)^{-1} h_i \in H$ as well.

The basis $\{E_j\}$ determines an inner product on $\mathfrak{g}$ for which $\{E_j\}$ is orthonormal. Let $|\cdot|$ denote the norm associated with this inner product

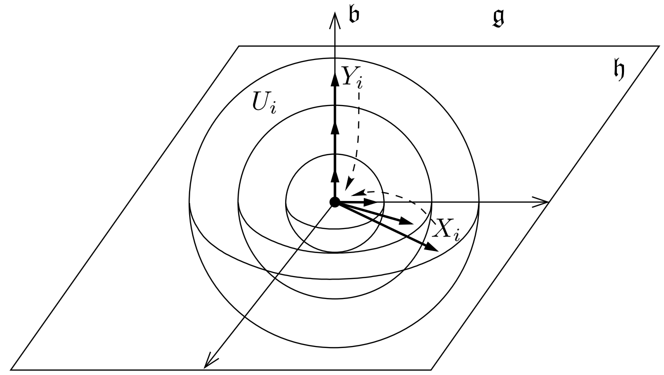


Figure 20.5. Proof of the closed subgroup theorem.

and define $c_i = |Y_i|$, so that $c_i \to 0$ as $i \to \infty$. The sequence $\{c_i^{-1}Y_i\}$ lies in the unit sphere in $\mathfrak{b}$ with respect to this norm, so replacing it by a subsequence we may assume that $c_i^{-1}Y_i \to Y \in \mathfrak{b}$, with $|Y| = 1$ by continuity. In particular, $Y \neq 0$. We will show that $\exp tY \in H$ for all $t \in \mathbb{R}$, which implies that $Y \in \mathfrak{h}$. Since $\mathfrak{h} \cap \mathfrak{b} = \{0\}$, this is a contradiction.

Let $t \in \mathbb{R}$ be arbitrary, and for each i, let n_i be the greatest integer less than or equal to t/c_i. Then

$$\left| n_i - \frac{t}{c_i} \right| \leq 1,$$

which implies

$$|n_i c_i - t| \leq c_i \to 0,$$

so $n_i c_i \to t$. Thus

$$n_i Y_i = (n_i c_i)(c_i^{-1} Y_i) \to tY,$$

which implies $\exp n_i Y_i \to \exp tY$ by continuity. But $\exp n_i Y_i = (\exp Y_i)^{n_i} \in H$, so the fact that H is closed implies $\exp tY \in H$. This completes the proof of the existence of U satisfying (20.8).

Let $E\colon \mathbb{R}^m \to \mathfrak{g}$ be the basis isomorphism determined by the basis $(E_1, \ldots, E_m)$. The composite map $\varphi = E^{-1} \circ \exp^{-1}\colon \exp U \to \mathbb{R}^m$ is easily seen to be a smooth chart for G, and by our choice of basis, $\varphi((\exp U) \cap H) = E^{-1}(U \cap \mathfrak{h})$ is the slice obtained by setting the last $m-k$ coordinates equal to zero. Moreover, if $h \in H$ is arbitrary, the left translation map L_h is a diffeomorphism from $\exp U$ to a neighborhood of h. Since H is a subgroup, $L_h(H) = H$, and so

$$L_h((\exp U) \cap H) = L_h(\exp U) \cap H,$$

and $\varphi \circ L_h^{-1}$ is easily seen to be a slice chart for H in a neighborhood of h. Thus H is an embedded submanifold of G, hence a Lie subgroup. □

The following corollary summarizes the closed subgroup theorem and Proposition 8.28.

Corollary 20.12. *If G is a Lie group and H is any subgroup of G, the following are equivalent:*

(*a*) *H is closed in G.*

(*b*) *H is an embedded submanifold of G.*

(*c*) *H is an embedded Lie subgroup of G.*

The Adjoint Representation

Suppose G is a Lie group and $\mathfrak{g}$ is its Lie algebra. In Chapter 9, we introduced the adjoint representation of G—this is the representation $\mathrm{Ad}\colon G \to \mathrm{GL}(\mathfrak{g})$ defined by $\mathrm{Ad}(g) = (C_g)_*\colon \mathfrak{g} \to \mathfrak{g}$, where $C_g(h) = ghg^{-1}$ (see Example 9.3). In problem 9-28, we also introduced the adjoint representation of a Lie algebra $\mathfrak{g}$, which is the representation $\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ defined by $\mathrm{ad}(X)Y = [X, Y]$. Using the exponential map, we can show that these two representations are intimately related.

Theorem 20.13. *Let G be a Lie group, let $\mathfrak{g}$ be its Lie algebra, and let $\mathrm{Ad}\colon G \to \mathrm{GL}(\mathfrak{g})$ be the adjoint representation of G. The induced Lie algebra representation $\mathrm{Ad}_*\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is given by $\mathrm{Ad}_* = \mathrm{ad}$.*

Proof. Let $X \in \mathfrak{g}$ be arbitrary. Because $t \mapsto \exp tX$ is a smooth curve in G whose tangent vector at $t = 0$ is X, we can compute $\mathrm{Ad}_*(X)$ by

$$\mathrm{Ad}_*(X) = \left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}(\exp tX).$$

Since $\mathrm{Ad}_*(X)$ is an element of the Lie algebra of $\mathrm{GL}(\mathfrak{g})$, which we canonically identify with $\mathfrak{gl}(\mathfrak{g})$ (Corollary 4.24), we can apply both sides to an arbitrary element $Y \in \mathfrak{g}$ to obtain

$$\mathrm{Ad}_*(X)Y = \left(\left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}(\exp tX) \right) Y = \left.\frac{d}{dt}\right|_{t=0} \left(\mathrm{Ad}(\exp tX)Y \right).$$

As an element of $\mathfrak{g}$, $\mathrm{Ad}(\exp tX)Y$ is a left-invariant vector field on G, and thus is determined by its value at the identity. Using the fact that $\mathrm{Ad}(g) = (C_g)_* = (R_{g^{-1}})_* \circ (L_g)_*$, its value at $e \in G$ can be computed as

$$\begin{aligned} (\mathrm{Ad}(\exp tX)Y)_e &= (R_{\exp(-tX)})_*(L_{\exp tX})_* Y_e \\ &= (R_{\exp(-tX)})_* Y_{\exp tX}. \end{aligned} \tag{20.9}$$

Recall from Proposition 20.9(g) that the flow of X is given by $\theta_t(g) = R_{\exp tX}(g)$. Therefore, (20.9) can be rewritten as

$$(\mathrm{Ad}(\exp tX)Y)_e = (\theta_{-t})_* Y_{\theta_t(e)}.$$

Taking the derivative with respect to t and setting $t = 0$, we obtain

$$(\mathrm{Ad}_*(X)Y)_e = \left.\frac{d}{dt}\right|_{t=0} (\theta_{-t})_* Y_{\theta_t(e)} = (\mathcal{L}_X Y)_e = [X, Y]_e.$$

Since $\mathrm{Ad}_*(X)Y$ is determined by its value at e, this completes the proof. □

Lie Subalgebras and Lie Subgroups

Suppose G is a Lie group and $\mathfrak{g}$ is its Lie algebra. Recall from Proposition 8.36 that for any Lie subgroup $H \subset G$, there is a Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$ that is canonically isomorphic to the Lie algebra of H, namely:

$$\mathfrak{h} = \{X \in \mathfrak{g} : X_e \in T_e H\}.$$

In this section, using the Frobenius theorem, we will show that the converse is true: Every Lie subalgebra corresponds to some Lie subgroup. This result has important consequences that we will explore in the remainder of the chapter.

Theorem 20.14. *Suppose G is a Lie group with Lie algebra $\mathfrak{g}$. If $\mathfrak{h}$ is any Lie subalgebra of $\mathfrak{g}$, then there is a unique connected Lie subgroup of G whose Lie algebra is $\mathfrak{h}$.*

Proof. Define a distribution $D \subset TG$ by

$$D_g = \{X_g \in T_g G : X \in \mathfrak{h}\}.$$

If $(X_1, \dots, X_k)$ is any basis for $\mathfrak{h}$, then clearly D_g is spanned by $X_1|_g, \dots, X_k|_g$ at any $g \in G$. Thus D is locally (in fact, globally) spanned by smooth vector fields, so it is smooth. Moreover, because $[X_i, X_j] \in \mathfrak{h}$ for each i, j, D is involutive. Let $\mathcal{H}$ denote the foliation determined by D, and for any $g \in G$, let $\mathcal{H}_g$ denote the leaf of $\mathcal{H}$ containing g (Figure 20.6).

If g, g' are arbitrary elements of G, then

$$\begin{aligned} L_{g*}(D_{g'}) &= \mathrm{span}(L_{g*}X_1|_{g'}, \dots, L_{g*}X_k|_{g'}) \\ &= \mathrm{span}(X_1|_{gg'}, \dots, X_k|_{gg'}) \\ &= D_{gg'}, \end{aligned}$$

so D is invariant under all left translations. It follows that if M is any connected integral manifold of D, then so is $L_g(M)$. If M is maximal, it is easy to see that $L_g(M)$ is as well. Therefore, left translation takes leaves to leaves: $L_g(\mathcal{H}_{g'}) = \mathcal{H}_{gg'}$ for any $g, g' \in G$.

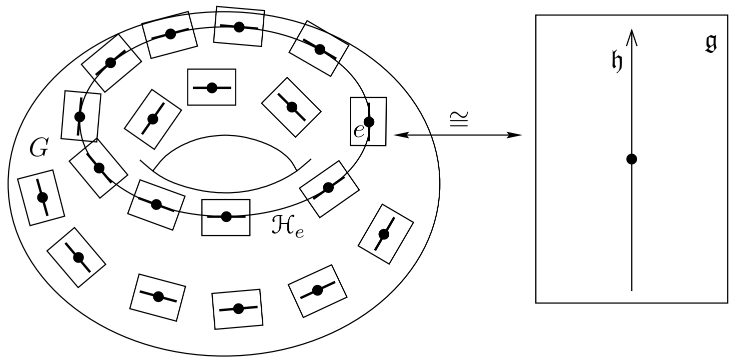


Figure 20.6. Finding a subgroup whose Lie algebra is $\mathfrak{h}$.

Define $H = \mathcal{H}_e$, the leaf containing the identity. We will show that H is the desired Lie subgroup.

First, to see that H is a subgroup, observe that for any $h, h' \in H$,

$$hh' = L_h(h') \in L_h(H) = L_h(\mathcal{H}_e) = \mathcal{H}_h = H.$$

Similarly,

$$h^{-1} = h^{-1}e \in L_{h^{-1}}(\mathcal{H}_e) = L_{h^{-1}}(\mathcal{H}_h) = \mathcal{H}_{h^{-1}h} = H.$$

To show that H is a Lie group, we need to show that the map $\mu\colon (h, h') \mapsto hh'^{-1}$ is smooth as a map from $H \times H$ to H. Because $H \times H$ is a submanifold of $G \times G$, it is immediate that $\mu\colon H \times H \to G$ is smooth. Since H is an integral manifold of an involutive distribution, Proposition 19.13 shows that μ is also smooth as a map into H.

The fact that H is a leaf of $\mathcal{H}$ implies that the Lie algebra of H is $\mathfrak{h}$ (identified with a subalgebra of $\mathfrak{g}$ in the usual way), because the tangent space to H at the identity is $D_e = \{X_e : X \in \mathfrak{h}\}$. To see that H is the unique connected subgroup with Lie algebra $\mathfrak{h}$, suppose $\widetilde{H}$ is any other connected subgroup with the same Lie algebra. Any such Lie subgroup is easily seen to be an integral manifold of D, so by maximality of H, we must have $\widetilde{H} \subset H$. On the other hand, if U is the domain of a flat chart for D near the identity, then by Proposition 19.12, $\widetilde{H} \cap U$ is a union of open subsets of slices. Since the slice containing e is an open subset of H, this implies that $\widetilde{H}$ contains a neighborhood V of the identity in H. Since any neighborhood of the identity generates H, this implies that $\widetilde{H} = H$. $\square$

An interesting corollary to this proof is the following characterization of the Lie subalgebra of a subgroup in terms of the exponential map. We will use this in the next section when we study normal subgroups.

Corollary 20.15. *Let G be a Lie group, and let $H \subset G$ be a Lie subgroup. Then the exponential map of H is the restriction to* $\mathrm{Lie}(H)$ *of the* *exponential map of G, and*

$$\operatorname{Lie}(H) = \{X \in \operatorname{Lie}(G) : \exp tX \in H \text{ for all } t \in \mathbb{R}\}.$$

Proof. The fact that the exponential map of H is the restriction of that of G is an immediate consequence of Proposition 20.5. To prove the second assertion, by the way we have identified $\operatorname{Lie}(H)$ as a subalgebra of $\operatorname{Lie}(G)$, we need to show that the following conditions are equivalent for every $X \in \operatorname{Lie}(G)$:

$$\exp tX \in H \text{ for all } t \in \mathbb{R} \iff X_e \in T_eH.$$

Assume first that $\exp tX \in H$ for all t, and let $\gamma\colon \mathbb{R} \to G$ be the smooth curve defined by $\gamma(t) = \exp tX$. By connectedness, $\gamma(t)$ lies in the identity component of H for all t, which the proof of Theorem 20.14 showed is a leaf of a foliation on G. By the result of Problem 19-4, this implies that $X_e = \gamma'(0) \in T_eH$. Conversely, if $X_e \in T_eH$, then Proposition 20.5 implies that $\exp tX \in H$ for all t. □

The most important application of Theorem 20.14 is the next theorem.

Theorem 20.16. *Suppose G and H are Lie groups with G simply connected, and let $\mathfrak{g}$ and $\mathfrak{h}$ be their Lie algebras. For any Lie algebra homomorphism $\varphi\colon \mathfrak{g} \to \mathfrak{h}$, there is a unique Lie group homomorphism $\Phi\colon G \to H$ such that $\Phi_* = \varphi$.*

Proof. The Lie algebra of $G \times H$ is isomorphic to the product Lie algebra $\mathfrak{g} \times \mathfrak{h}$ by Problem 4-16. Let $\mathfrak{k} \subset \mathfrak{g} \times \mathfrak{h}$ be the graph of φ:

$$\mathfrak{k} = \{(X, \varphi X) : X \in \mathfrak{g}\}.$$

Then $\mathfrak{k}$ is a vector subspace of $\mathfrak{g} \times \mathfrak{h}$ because φ is linear, and in fact it is a Lie subalgebra because φ is a homomorphism:

$$[(X, \varphi X), (X', \varphi X')] = ([X, X'], [\varphi X, \varphi X']) = ([X, X'], \varphi[X, X']) \in \mathfrak{k}.$$

Therefore, by the preceding theorem, there is a unique connected Lie subgroup $K \subset G \times H$ whose Lie algebra is $\mathfrak{k}$ (Figure 20.7).

The restrictions to K of the projections

$$\pi_1\colon G \times H \to G, \qquad \pi_2\colon G \times H \to H$$

are Lie group homomorphisms because π_1 and π_2 are. Let $\Pi = \pi_1|_K\colon K \to G$. We will show that Π is a smooth covering map. Since G is simply connected, this will imply that Π is bijective by Proposition A.28, and thus is a Lie group isomorphism.

To show that Π is a smooth covering map, it suffices by Proposition 9.36 to show that its induced Lie algebra homomorphism Π_* is an isomorphism. Consider the sequence of maps

$$K \hookrightarrow G \times H \xrightarrow{\pi_1} G$$

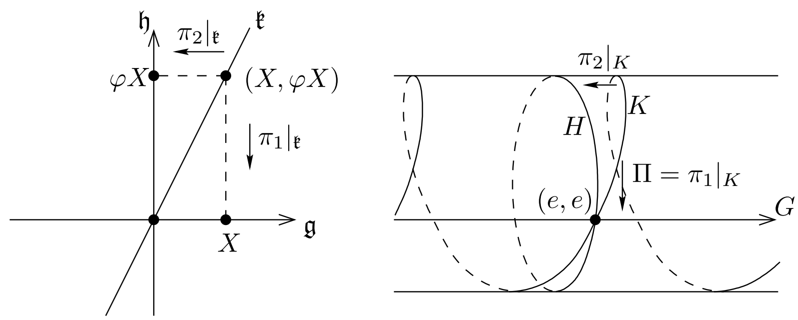


Figure 20.7. Constructing a Lie group homomorphism.

whose composition is Π. The induced Lie algebra homomorphism Π_* is just inclusion followed by projection on the algebra level:

$$\mathfrak{k} \hookrightarrow \mathfrak{g} \times \mathfrak{h} \xrightarrow{\pi_1} \mathfrak{g}.$$

This last composition is nothing more than the restriction to $\mathfrak{k}$ of the projection $\mathfrak{g} \times \mathfrak{h} \to \mathfrak{g}$. Because $\mathfrak{k} \cap \mathfrak{h} = \{0\}$ (since $\mathfrak{k}$ is a graph), it follows that $\Pi_* \colon \mathfrak{k} \to \mathfrak{g}$ is an isomorphism. This completes the proof that Π is a smooth covering map and thus a Lie group isomorphism.

Define a Lie group homomorphism $\Phi \colon G \to H$ by $\Phi = \pi_2|_K \circ \Pi^{-1}$. Note that the definition of Φ implies that

$$\pi_2|_K = \Phi \circ \pi_1|_K.$$

Because the Lie algebra homomorphism induced by the projection $\pi_1 \colon G \times H \to H$ is just the linear projection $\pi_1 \colon \mathfrak{g} \times \mathfrak{h} \to \mathfrak{h}$, this implies

$$\pi_2|_{\mathfrak{k}} = \Phi_* \circ \pi_1|_{\mathfrak{k}} \colon \mathfrak{k} \to \mathfrak{h}.$$

Thus if $X \in \mathfrak{g}$ is arbitrary,

$$\begin{aligned} \varphi X &= \pi_2|_{\mathfrak{k}}(X, \varphi X) \\ &= \Phi_* \circ \pi_1|_{\mathfrak{k}}(X, \varphi X) \\ &= \Phi_* X, \end{aligned}$$

which shows that $\Phi_* = \varphi$.

The proof is completed by invoking Problem 20-3, which shows that Φ is the unique homomorphism with this property. □

◇ **Exercise 20.1.** Let G be a simply connected Lie group and let $\mathfrak{g}$ be its Lie algebra. Show that every representation of $\mathfrak{g}$ is of the form $\rho_* \colon \mathfrak{g} \to \mathfrak{gl}(V)$ for some representation $\rho \colon G \to \mathrm{GL}(V)$ of G.

Corollary 20.17. *If G and H are simply connected Lie groups with isomorphic Lie algebras, then G and H are isomorphic.*

Proof. Let $\mathfrak{g}$, $\mathfrak{h}$ be the Lie algebras, and let $\varphi\colon \mathfrak{g} \to \mathfrak{h}$ be a Lie algebra isomorphism between them. By the preceding theorem, there are Lie group homomorphisms $\Phi\colon G \to H$ and $\Psi\colon H \to G$ satisfying $\Phi_* = \varphi$ and $\Psi_* = \varphi^{-1}$. Both the identity map of G and the composition $\Psi \circ \Phi$ are maps from G to itself whose induced homomorphisms are equal to the identity, so the uniqueness part of Theorem 20.16 implies that $\Psi \circ \Phi = \mathrm{Id}_G$. Similarly, $\Phi \circ \Psi = \mathrm{Id}_H$, so Φ is a Lie group isomorphism. □

A version of this theorem was proved in the nineteenth century by Sophus Lie. However, since global topological notions such as simple connectedness (or even manifolds!) had not yet been formulated, what he was able to prove was essentially a local version of this corollary. Two Lie groups G and H are said to be *locally isomorphic* if there exist neighborhoods of the identity $U \subset G$ and $V \subset H$ and a diffeomorphism $F\colon U \to V$ such that $F(g_1g_2) = F(g_1)F(g_2)$ whenever g_1, g_2, and g_1g_2 are all in U.

Theorem 20.18 (Fundamental Theorem of Sophus Lie). *Two Lie groups are locally isomorphic if and only if they have isomorphic Lie algebras.*

The proof in one direction is essentially to follow the arguments in Theorem 20.16 and Corollary 20.17, except that one just uses the inverse function theorem to show that Φ is a local isomorphism instead of appealing to the theory of covering spaces. The details are left as an exercise.

◇ **Exercise 20.2.** Carry out the details of the proof of Lie's fundamental theorem.

Normal Subgroups

Recall that a subgroup $H \subset G$ is said to be normal in G if $ghg^{-1} \in H$ for every $g \in G$ and $h \in H$.

Lemma 20.19. *Let G be a connected Lie group, and let $H \subset G$ be a connected Lie subgroup. Let $\mathfrak{g}$ and $\mathfrak{h}$ denote the Lie algebras of G and H, respectively. Then H is normal in G if and only if*

$$(\exp X)(\exp Y)(\exp(-X)) \in H \textit{ for all } X \in \mathfrak{g} \textit{ and } Y \in \mathfrak{h}. \tag{20.10}$$

Proof. Note that $\exp(-X) = (\exp X)^{-1}$. Thus if H is normal, then (20.10) holds by definition. Conversely, suppose (20.10) holds, and choose open sets $U \subset \mathfrak{g}$ containing 0 and $V \subset G$ containing the identity such that $\exp\colon U \to V$ is a diffeomorphism. Since the exponential map of H is the restriction of that of G, after shrinking U if necessary, we may also assume that the restriction of exp to $U \cap \mathfrak{h}$ is a diffeomorphism from $U \cap \mathfrak{h}$ to a neighborhood V_0 of the identity in H. Then (20.10) implies that $ghg^{-1} \in H$ whenever $g \in V$ and $h \in V_0$. Since every element $g \in G$ can be written as a finite product $g = g_1 \cdots g_k$ of elements $g_1, \dots, g_k \in V$ (Problem 2-11), it follows by induction that the same is true for all $g \in G$. Similarly, any $h \in H$ can be written $h = h_1 \cdots h_m$ for $h_1, \dots, h_m \in V_0$, and then for any $g \in G$ we have

$$ghg^{-1} = gh_1 \cdots h_m g^{-1} = (gh_1 g^{-1}) \cdots (gh_m g^{-1}) \in H. \qquad \square$$

The next theorem expresses one of the deepest relationships between Lie groups and their Lie algebras. If $\mathfrak{g}$ is a Lie algebra, a linear subspace $\mathfrak{h} \subset \mathfrak{g}$ is called an *ideal* in $\mathfrak{g}$ if $[X, Y] \in \mathfrak{h}$ whenever $X \in \mathfrak{g}$ and $Y \in \mathfrak{h}$ (see Problem 4-15).

Theorem 20.20 (Ideals and Normal Subgroups). *Let G be a connected Lie group, and let $H \subset G$ be a connected Lie subgroup. Then H is a normal subgroup of G if and only if* $\operatorname{Lie}(H)$ *is an ideal in* $\operatorname{Lie}(G)$*.*

Proof. Write $\mathfrak{g} = \operatorname{Lie}(G)$ and $\mathfrak{h} = \operatorname{Lie}(H)$, considering $\mathfrak{h}$ as a Lie subalgebra of $\mathfrak{g}$. For any $g \in G$, the commutative diagram (20.4) applied to the Lie group homomorphism $C_g(h) = ghg^{-1}$ yields

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{\operatorname{Ad}(g)} & \mathfrak{g} \\ {\scriptstyle \exp}\downarrow & & \downarrow{\scriptstyle \exp} \\ G & \xrightarrow[C_g]{} & G. \end{array} \tag{20.11}$$

Applying this to $Y \in \mathfrak{g}$ with $g = \exp X$, we obtain

$$\begin{aligned} (\exp X)(\exp Y)(\exp(-X)) &= C_{\exp X}(\exp Y) \\ &= \exp(\operatorname{Ad}(\exp X)Y). \end{aligned} \tag{20.12}$$

On the other hand, applying (20.4) to the homomorphism $\operatorname{Ad}\colon G \to \operatorname{GL}(\mathfrak{g})$ and noting that $\operatorname{Ad}_* = \operatorname{ad}$ by Theorem 20.13, we obtain

$$\operatorname{Ad}(\exp X) = \exp(\operatorname{ad} X). \tag{20.13}$$

Using formula (20.3) for the exponential map of the group $\operatorname{GL}(\mathfrak{g})$,

$$\operatorname{Ad}(\exp X)Y = (\exp(\operatorname{ad} X))Y = \sum_{k=0}^{\infty} \frac{1}{k!} (\operatorname{ad} X)^k Y. \tag{20.14}$$

Now suppose that $\mathfrak{h}$ is an ideal. Then whenever $X \in \mathfrak{g}$ and $Y \in \mathfrak{h}$, $(\operatorname{ad} X)Y = [X, Y] \in \mathfrak{h}$, and by induction $(\operatorname{ad} X)^k Y \in \mathfrak{h}$ for all k. Therefore (20.14) implies that $\operatorname{Ad}(\exp X)Y \in \mathfrak{h}$, and so (20.12) implies that $(\exp X)(\exp Y)(\exp(-X)) \in \exp \mathfrak{h} \subset H$. By Lemma 20.19, H is normal.

Conversely, suppose H is normal. If $X \in \mathfrak{g}$ and $Y \in \mathfrak{h}$, (20.12) implies

$$\begin{aligned} \exp t(\operatorname{Ad}(\exp X)Y) &= \exp(\operatorname{Ad}(\exp X)(tY)) \\ &= (\exp X)(\exp tY)(\exp(-X)), \end{aligned}$$

which is in H for all t because H is normal. Therefore, by Corollary 20.15,

$$\operatorname{Ad}(\exp X)Y \in \mathfrak{h} \quad \text{whenever } X \in \mathfrak{g} \text{ and } Y \in \mathfrak{h}. \tag{20.15}$$

Let $X \in \mathfrak{g}$ and $Y \in \mathfrak{h}$ be arbitrary, and let $\gamma\colon \mathbb{R} \to \mathfrak{g}$ be the smooth curve $\gamma(t) = \operatorname{Ad}(\exp tX)Y$. Using (20.14),

$$\begin{aligned}\gamma'(0) &= \frac{d}{dt}\bigg|_{t=0} \operatorname{Ad}(\exp tX)Y \\ &= \frac{d}{dt}\bigg|_{t=0} (\exp(\operatorname{ad} tX))Y \\ &= \left(\frac{d}{dt}\bigg|_{t=0} (\exp t(\operatorname{ad} X))\right)Y \\ &= (\operatorname{ad} X)Y \\ &= [X,Y].\end{aligned}$$

Now, (20.15) shows that $\gamma(t) \in \mathfrak{h}$ for all t. This means, in particular, that $[X,Y] = \gamma'(0) \in \mathfrak{h}$, which proves that $\mathfrak{h}$ is an ideal. □

The Fundamental Correspondence Between Lie Algebras and Lie Groups

Many of the results of this chapter show how essential properties of a Lie group are reflected in its Lie algebra, and vice versa. This raises a natural question: To what extent is the correspondence between Lie groups and Lie algebras (or at least between their isomorphism classes) one-to-one? We have already seen in Proposition 4.26(c) that isomorphic Lie groups have isomorphic Lie algebras. The converse is easily seen to be false: Both $\mathbb{R}^n$ and $\mathbb{T}^n$ have n-dimensional abelian Lie algebras, which are obviously isomorphic, but $\mathbb{R}^n$ and $\mathbb{T}^n$ are certainly not isomorphic Lie groups. However, if we restrict our attention to simply connected Lie groups, then we do obtain a one-to-one correspondence. The central result is the following theorem.

Theorem 20.21 (Lie Group–Lie Algebra Correspondence). *There is a one-to-one correspondence between isomorphism classes of finite-dimensional Lie algebras and isomorphism classes of simply connected Lie groups, given by associating each simply connected Lie group with its Lie algebra.*

Proof. We need to show that the association that sends a simply connected Lie group to its Lie algebra is both surjective and injective up to isomorphism. Injectivity is precisely the content of Corollary 20.17.

To prove surjectivity, suppose $\mathfrak{g}$ is any finite-dimensional Lie algebra. By Ado's theorem (see page 211), $\mathfrak{g}$ has a faithful representation $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ for some finite-dimensional vector space V. Replacing $\mathfrak{g}$ with its isomorphic image under this representation, we may assume that $\mathfrak{g}$ is a Lie subalgebra of $\mathfrak{gl}(V) \cong \mathrm{Lie}(\mathrm{GL}(V))$. By Theorem 20.14, there is a connected Lie subgroup $G \subset \mathrm{GL}(V)$ that has $\mathfrak{g}$ as its Lie algebra. Letting $\widetilde{G}$ be the universal covering group of G, Proposition 9.36 shows that $\mathrm{Lie}(\widetilde{G}) \cong \mathrm{Lie}(G) \cong \mathfrak{g}$. □

What happens when we remove the restriction to simply-connected groups? The connected case is easy to describe, because every connected Lie group has a simply connected covering group. (For the disconnected case, see Problem 20-12.)

Theorem 20.22. *Let $\mathfrak{g}$ be a finite-dimensional Lie algebra. The connected Lie groups whose Lie algebras are isomorphic to $\mathfrak{g}$ are (up to isomorphism) precisely those of the form G/Γ, where G is the simply connected Lie group with Lie algebra $\mathfrak{g}$, and Γ is a discrete central subgroup of G.*

Proof. Given $\mathfrak{g}$, let G be a simply connected Lie group with Lie algebra isomorphic to $\mathfrak{g}$. Suppose H is any other Lie group whose Lie algebra is isomorphic to $\mathfrak{g}$, and let $\varphi\colon \mathrm{Lie}(G) \to \mathrm{Lie}(H)$ be a Lie algebra isomorphism. Theorem 20.16 guarantees that there is a Lie group homomorphism $\Phi\colon G \to H$ such that $\Phi_* = \varphi$. Because φ is an isomorphism, Proposition 9.36 implies that Φ is surjective and has discrete kernel. It follows that H is isomorphic to $G/\operatorname{Ker}\Phi$ by Proposition 9.34. By the result of problem 9-26, $\operatorname{Ker}\Phi$ is contained in the center of G. □

◇ **Exercise 20.3.** Show that two connected Lie groups are locally isomorphic if and only if they have isomorphic universal covering groups.

Problems

20-1. Compute the exponential maps of the abelian Lie groups $\mathbb{R}^n$ and $\mathbb{T}^n$.

20-2. Let G be a Lie group and let $\mathfrak{g}$ be its Lie algebra. If $A, B \subset \mathfrak{g}$ are complementary linear subspaces of $\mathfrak{g}$, show that the map $A \oplus B \to G$ given by $(X, Y) \mapsto \exp X \exp Y$ is a diffeomorphism from some neighborhood of $(0, 0)$ in $A \oplus B$ to a neighborhood of e in G.

20-3. Suppose G is a connected Lie group and H is any Lie group. If $\Phi, \Psi\colon G \to H$ are Lie group homomorphisms such that $\Phi_* = \Psi_*\colon \mathrm{Lie}(G) \to \mathrm{Lie}(H)$, show that $\Phi = \Psi$.

20-4. Show that the matrix exponential satisfies the identity

$$\det e^A = e^{\operatorname{tr} A}.$$

[Hint: Apply Proposition 20.9(f) to $\det\colon \mathrm{GL}(n, \mathbb{R}) \to \mathbb{R}^*$.]

20-5. Let G be a Lie group.

(a) Show that the images of one-parameter subgroups in G are precisely the connected Lie subgroups of dimension less than or equal to 1.

(b) If $H \subset G$ is the image of a one-parameter subgroup, show that H is isomorphic as a Lie group to one of the following: the trivial group $\{e\}$, $\mathbb{R}$, or $\mathbb{S}^1$.

20-6. Let A and B be the following elements of $\mathfrak{gl}(2,\mathbb{R})$:

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}; \qquad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Give explicit formulas for the one-parameter subgroups of $\mathrm{GL}(2,\mathbb{R})$ generated by A and B.

20-7. Let $\mathrm{GL}^+(n,\mathbb{R})$ be the subgroup of $\mathrm{GL}(n,\mathbb{R})$ consisting of matrices with positive determinant. (By Proposition 9.29, it is the identity component of $\mathrm{GL}(n,\mathbb{R})$.)

(a) Suppose $A \in \mathrm{GL}^+(n,\mathbb{R})$ is of the form e^B for some $B \in \mathfrak{gl}(n,\mathbb{R})$. Show that A has a square root, i.e., a matrix $C \in \mathrm{GL}^+(n,\mathbb{R})$ such that $C^2 = A$.

(b) Let

$$A = \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix}.$$

Show that the exponential map $\exp\colon \mathfrak{gl}(2,\mathbb{R}) \to \mathrm{GL}^+(2,\mathbb{R})$ is not surjective, by showing that A is not in its image.

20-8. Let G be a connected Lie group and let $\mathfrak{g}$ be its Lie algebra.

(a) If $X, Y \in \mathfrak{g}$, show that $[X,Y] = 0$ if and only if

$$\exp tX \exp sY = \exp sY \exp tX \text{ for all } s,\ t \in \mathbb{R}.$$

(b) Show that G is abelian if and only if $\mathfrak{g}$ is abelian.

(c) Give a counterexample when G is not connected.

20-9. Show that every connected abelian Lie group is isomorphic to $\mathbb{R}^k \times \mathbb{T}^l$ for some nonnegative integers k and l.

20-10. By the result of Problem 9-24, $\mathrm{SL}(2,\mathbb{R})$ is diffeomorphic to $\mathrm{SO}(2) \times \mathbb{R}^2 \approx \mathbb{S}^1 \times \mathbb{R}^2$, and therefore its fundamental group is isomorphic to $\mathbb{Z}$. Let $\widetilde{\mathrm{SL}(2,\mathbb{R})}$ denote the universal covering group of $\mathrm{SL}(2,\mathbb{R})$ (see Theorem 2.13). Show that $\widetilde{\mathrm{SL}(2,\mathbb{R})}$ does not admit a faithful representation, as follows. Suppose $\rho\colon \widetilde{\mathrm{SL}(2,\mathbb{R})} \to \mathrm{GL}(V)$ is any representation. By choosing a basis for V over $\mathbb{R}$, we might as well replace $\mathrm{GL}(V)$ with $\mathrm{GL}(n,\mathbb{R})$ for some n. Then ρ induces a Lie algebra homomorphism $\rho_*\colon \mathfrak{sl}(2,\mathbb{R}) \to \mathfrak{gl}(n,\mathbb{R})$. Define a map

$\varphi\colon \mathfrak{sl}(2,\mathbb{C}) \to \mathfrak{gl}(n,\mathbb{C})$ by

$$\varphi(A+iB) = \rho_*A + i\rho_*B, \qquad A, B \in \mathfrak{sl}(2,\mathbb{R}).$$

(a) Show that φ is a Lie algebra homomorphism.
(b) Show that there is a Lie group homomorphism $\Phi\colon \mathrm{SL}(2,\mathbb{C}) \to \mathrm{GL}(n,\mathbb{C})$ such $\Phi_* = \varphi$ and the following diagram commutes:

$$\begin{tikzcd}
\widetilde{\mathrm{SL}(2,\mathbb{R})} \arrow[r, "\rho"] \arrow[d, "\pi"'] & \mathrm{GL}(n,\mathbb{R}) \arrow[dd, "j"] \\
\mathrm{SL}(2,\mathbb{R}) \arrow[d, "i"'] & \\
\mathrm{SL}(2,\mathbb{C}) \arrow[r, "\Phi"] & \mathrm{GL}(n,\mathbb{C}),
\end{tikzcd}$$

where π is the universal covering map, and i and j are inclusions. [Hint: Use the fact that $\mathrm{SL}(2,\mathbb{C})$ is simply connected by Problems 9-24 and 9-8.]
(c) Show that ρ is not faithful.

20-11. If $F\colon G \to H$ is a Lie group homomorphism, show that the kernel of $F_*\colon \mathrm{Lie}(G) \to \mathrm{Lie}(H)$ is the Lie algebra of $\operatorname{Ker} F$ (under our usual identification of the Lie algebra of a subgroup with a Lie subalgebra).

20-12. If G and H are Lie groups, and there exists a surjective Lie group homomorphism from G to H with kernel G_0, we say that G is an *extension of* G_0 *by* H. If $\mathfrak{g}$ is any finite-dimensional Lie algebra, show that the disconnected Lie groups whose Lie algebras are isomorphic to $\mathfrak{g}$ are precisely the extensions of the connected ones by discrete groups. [Hint: Begin by showing that the identity component of any Lie group is a normal subgroup.]

Appendix
Review of Prerequisites

The essential prerequisites for reading this book are a thorough acquaintance with basic topology, abstract linear algebra, and advanced multivariable calculus. In this appendix, we summarize the most important facts from these subjects that are used throughout the book.

Topology

This book is written for readers who have already completed a rigorous course in basic topology, including an introduction to the fundamental group and covering maps. A convenient source for this material is [Lee00], which covers all the topological ideas you will need, and uses notations and conventions that are compatible with those in the present book. But almost any other good topology text would do as well, such as [Mun00, Sie92, Mas89]. In this section, we state the most important definitions and results, with most of the proofs left as exercises. If you have had sufficient exposure to topology, these exercises should be straightforward, although you might want to look a few of them up in the topology texts listed above.

We begin with the definitions. A *topology* on a set X is a collection $\mathcal{T}$ of subsets of X, called *open sets*, satisfying

(i) X and $\varnothing$ are open.

(ii) The union of any family of open sets is open.

(iii) The intersection of any finite family of open sets is open.

A pair $(X, \mathcal{T})$ consisting of a set X together with a topology $\mathcal{T}$ on X is called a *topological space*. Ordinarily, when the topology is understood, one omits mention of it and simply says "X is a topological space."

There are myriad constructions and definitions associated with topological spaces. Here we summarize the ones that will be most important for our purposes.

Suppose X is a topological space. Let p be a point of X, and let S be any subset of X.

- A *neighborhood* of p is an open set containing p. Similarly, a neighborhood of the set S is an open set containing S. (Be warned that some authors use the word "neighborhood" in the more general sense of a set containing an open set containing p or S.)
- S is said to be *closed* if $X \smallsetminus S$ is open (where $X \smallsetminus S$ denotes the *set difference* $\{x \in X : x \notin S\}$).
- The *interior* of S, denoted by $\operatorname{Int} S$, is the union of all open subsets of X contained in S.
- The *exterior* of S, denoted by $\operatorname{Ext} S$, is the union of all open subsets of X contained in $X \smallsetminus S$.
- The *closure* of S, denoted by $\overline{S}$, is the intersection of all closed subsets of X containing S.
- The *boundary* of S, denoted by ∂S, is the set of all points of X that are in neither $\operatorname{Int} S$ nor $\operatorname{Ext} S$.
- The point p is said to be a *limit point* of S if every neighborhood of p contains at least one point of S other than p.
- S is said to be *dense* in X if $\overline{S} = X$, or equivalently if every open subset of X contains at least one point of S.

Continuity and Convergence

The most important concepts of topology are continuous maps and convergent sequences, which we define next. Let X and Y be topological spaces.

- A map $F \colon X \to Y$ is said to be *continuous* if for every open set $U \subset Y$, the inverse image $F^{-1}(U)$ is open in X.
- A continuous bijective map $F \colon X \to Y$ with continuous inverse is called a *homeomorphism*. If there exists a homeomorphism from X to Y, we say that X and Y are *homeomorphic*.

- A continuous map $F\colon X \to Y$ is said to be a *local homeomorphism* if every point $p \in X$ has a neighborhood $U \subset X$ such that F restricts to a homeomorphism from U to an open subset $V \subset Y$.
- A sequence $\{p_i\}$ of points in X is said to *converge* to $p \in X$ if for every neighborhood U of p, there exists an integer N such that $p_i \in U$ for all $i \geq N$. In this case, we write $p_i \to p$ or $\lim_{i\to\infty} p_i = p$.

◇ **Exercise A.1.** Show that every bijective local homeomorphism is a homeomorphism.

◇ **Exercise A.2.** Let X, Y, and Z be topological spaces. Show that the following maps are continuous:

(a) The *identity map* $\mathrm{Id}_X : X \to X$, defined by $\mathrm{Id}_X(x) = x$ for all $x \in X$.
(b) Any constant map $f\colon X \to Y$. (A map $f\colon X \to Y$ is *constant* if there exists some $c \in Y$ such that $f(x) = c$ for all $x \in X$.)
(c) The composition $g \circ f$ of any two continuous maps $f\colon X \to Y$ and $g\colon Y \to Z$.

The most important examples of topological spaces, from which most of our examples of manifolds will be built in one way or another, are described below.

Example A.1 (Discrete Spaces). If X is any set, the *discrete topology* on X is the topology defined by declaring every subset of X to be open. Any space that has the discrete topology is called a *discrete space*.

Example A.2 (Metric Spaces). A *metric space* is a set M endowed with a *distance function* (also called a *metric*) $d\colon M \times M \to \mathbb{R}$ (where $\mathbb{R}$ denotes the set of real numbers) satisfying the following properties for all $x, y, z \in M$:

(i) Positivity: $d(x, y) > 0$ when $x \neq y$, and $d(x, x) = 0$.

(ii) Symmetry: $d(x, y) = d(y, x)$.

(iii) Triangle Inequality: $d(x, z) \leq d(x, y) + d(y, z)$.

If M is a metric space, $x \in M$, and $r > 0$, we define the *open ball* of radius r around x as the set

$$B_r(x) = \{y \in M : d(x, y) < r\},$$

and the *closed ball* of radius r as

$$\overline{B}_r(x) = \{y \in M : d(x, y) \leq r\}.$$

A subset $S \subset M$ is said to be open if for every point $x \in S$, there is some $r > 0$ such that the open ball $B_r(x)$ is contained in S. The collection of all open subsets of M is a topology on M, called the *metric topology*. If M is a metric space and S is any subset of M, the restriction of the distance function to points in S is easily seen to turn S into a metric space and thus also a topological space. We use the following standard terminology for metric spaces:

- A subset $S \subset M$ is *bounded* if there exists a positive number R such that $d(x, y) \leq R$ for all $x, y \in S$.
- A sequence of points $\{x_i\}$ in M is *Cauchy* if for every $\varepsilon > 0$, there exists an integer N such that $i, j \geq N$ implies $d(x_i, x_j) < \varepsilon$.
- A metric space M is said to be *complete* if every Cauchy sequence in M converges to a point of M.

Example A.3 (Euclidean Spaces). For any integer $n \geq 1$, n-dimensional *Euclidean space* is the set $\mathbb{R}^n$ of ordered n-tuples of real numbers. We denote a point in $\mathbb{R}^n$ by any of the notations $(x^1, \dots, x^n)$ or (x^i) or x; the numbers x^i are called the *coordinates* of x. (When n is small, we often use more traditional names such as (x, y, z) for the coordinates.) Notice that we write the coordinates of a point $(x^1, \dots, x^n) \in \mathbb{R}^n$ with superscripts, not subscripts as is usually done in linear algebra and calculus books, so as to be consistent with the Einstein summation convention, which is explained in Chapter 1. Endowed with the *Euclidean distance function* defined by

$$d(x, y) = \sqrt{(x^1 - y^1)^2 + \cdots + (x^n - y^n)^2}, \tag{A.1}$$

$\mathbb{R}^n$ is a metric space, as is any subset of $\mathbb{R}^n$. The resulting metric topology is called the *Euclidean topology*. By convention, $\mathbb{R}^0$ is defined to be the one-element set $\{0\}$, with the discrete topology. Whenever we work with subsets of $\mathbb{R}^n$, we will always consider them to be endowed with the Euclidean topology unless otherwise specified. We will also have occasion to work with the space $\mathbb{C}^n$ of ordered n-tuples of complex numbers; as a topological space, we identify $\mathbb{C}^n$ with $\mathbb{R}^{2n}$ via the correspondence

$$(x^1 + iy^1, \dots, x^n + iy^n) \leftrightarrow (x^1, y^1, \dots, x^n, y^n).$$

Hausdorff Spaces

The definition of topological spaces is wonderfully flexible, and can be used to describe a rich assortment of concepts of "space." However, without further qualification, arbitrary topological spaces are far too general for most purposes, because they can have some unpleasant properties, as the following exercise illustrates.

◇ **Exercise A.3.** Let X be any set. Show that $\{X, \varnothing\}$ is a topology on X, called the *trivial topology*. Show that every sequence in X converges to every point of X in this topology.

To avoid pathological cases like this, which result when X does not have sufficiently many open sets, we will usually restrict our attention to topological spaces satisfying the following special condition. A topological space X is said to be a *Hausdorff space* if for every pair of distinct points $p, q \in X$, there exist disjoint open subsets $U \subset X$ containing p and $V \subset X$ containing q.

◇ **Exercise A.4.** Show that the metric topology on any metric space is Hausdorff.

◇ **Exercise A.5.** If X is a Hausdorff space, show that every finite subset of X is closed, and that each convergent sequence in X has a unique limit.

Bases and Countability

Suppose X is a set. A *basis* for a topology on X is a collection $\mathcal{B}$ of subsets of X such that

(i) $X = \bigcup_{B \in \mathcal{B}} B$.

(ii) If $B_1, B_2 \in \mathcal{B}$ and $x \in B_1 \cap B_2$, there exists $B_3 \in \mathcal{B}$ such that $x \in B_3 \subset B_1 \cap B_2$.

◇ **Exercise A.6.** Suppose X is a set and $\mathcal{B}$ is a basis for a topology on X. Show that the collection of all unions of elements of $\mathcal{B}$ is a topology on X.

The topology described in the preceding exercise is called the *topology generated by* $\mathcal{B}$. It is easy to check that a subset $S \subset X$ is open in this topology if and only if for each $p \in S$, there is a basis element $B \in \mathcal{B}$ such that $p \in B \subset S$.

If X is a topological space and $p \in X$, a *neighborhood basis* at p is a collection $\mathcal{B}_p$ of neighborhoods of p such that every neighborhood of p contains at least one $B \in \mathcal{B}_p$.

A set is said to be *countably infinite* if it admits a bijection with the natural numbers, and *countable* if it is finite or countably infinite. A topological space X is said to be *first countable* if there is a countable neighborhood basis at each point, and *second countable* if there is a countable basis for its topology. Since a countable basis for X contains a countable neighborhood basis at each point, second countability implies first countability.

◇ **Exercise A.7.** Let $U \subset \mathbb{R}^n$ be an open subset with the Euclidean topology. Show that the set of all open balls contained in U with rational radii and rational center points is a countable basis for U, and thus every open subset of $\mathbb{R}^n$ is second countable.

◇ **Exercise A.8.** Let X be a first countable space, and let $A \subset X$ be any subset. Show that a point $p \in X$ is in $\overline{A}$ if and only if there exists a sequence $\{p_i\}$ of points in A such that $p_i \to p$.

One of the most important properties of second countable spaces is expressed the following lemma. An *open cover* of a topological space X is a collection $\mathcal{U}$ of open subsets of X whose union is X. A *subcover* is a subcollection of $\mathcal{U}$ that is still an open cover.

Lemma A.4. *Let X be a second countable topological space. Every open cover of X has a countable subcover.*

Proof. Let $\mathcal{B}$ be a countable basis for X, and let $\mathcal{U}$ be an arbitrary open cover of X. Let $\mathcal{B}' \subset \mathcal{B}$ be the collection of basis open sets $B \in \mathcal{B}$ such that $B \subset U$ for some $U \in \mathcal{U}$. For each $B \in \mathcal{B}'$, choose a particular set $U_B \in \mathcal{U}$ containing B. The collection $\{U_B : B \in \mathcal{B}'\}$ is countable, so it suffices to show that it covers X. Given any point $x \in X$, there is some $U_0 \in \mathcal{U}$ containing x, and because $\mathcal{B}$ is a basis there exists $B \in \mathcal{B}$ such that $x \in B \subset U_0$. This implies, in particular, that $B \in \mathcal{B}'$, and therefore $x \in B \subset U_B$. □

Subspaces, Product Spaces, and Disjoint Unions

Probably the simplest way to obtain new topological spaces from old ones is by taking subsets of other spaces. If X is a topological space and $S \subset X$ is any subset, we define the *subspace topology* (sometimes called the *relative topology*) on S by declaring a subset $U \subset S$ to be open in S if and only if there exists an open set $V \subset X$ such that $U = V \cap S$. A subset of S that is open or closed in the subspace topology is sometimes said to be *relatively open* or *relatively closed* in S, to make it clear that we do not mean open or closed as a subset of X. Any subset of X endowed with the subspace topology is said to be a *subspace* of X. Whenever we treat a subset of a topological space as a space in its own right, we will always assume that it has the subspace topology unless otherwise specified.

If X and Y are topological spaces, a continuous injective map $f \colon X \to Y$ is called a *topological embedding* if it is a homeomorphism onto its image $f(X) \subset Y$ in the subspace topology.

The most important properties of the subspace topology are summarized in the next lemma.

Lemma A.5 (Properties of the Subspace Topology). *Let X be a topological space and let S be a subspace of X.*

(*a*) Characteristic Property: *For any topological space Y, a map $f \colon Y \to S$ is continuous if and only if the composition $\iota_S \circ f \colon Y \to X$ is continuous, where $\iota_S \colon S \hookrightarrow X$ is the inclusion map (the restriction of the identity map of X to S).*

(*b*) *The subspace topology is the unique topology on S for which the characteristic property holds.*

(*c*) *A subset* $K \subset S$ *is closed in* S *if and only if there exists a closed subset* $L \subset X$ *such that* $K = L \cap S$*.*

(*d*) *The inclusion map* $\iota_S \colon S \hookrightarrow X$ *is a topological embedding.*

(*e*) *If* Y *is a topological space and* $f \colon X \to Y$ *is continuous, then* $f|_S \colon S \to Y$ *(the restriction of* f *to* S*) is continuous.*

(*f*) *If* $\mathcal{B}$ *is a basis for the topology of* X*, then*

$$\mathcal{B}_S = \{B \cap S : B \in \mathcal{B}\}$$

is a basis for the subspace topology on S*.*

(*g*) *If* X *is Hausdorff, then so is* S*.*

(*h*) *If* X *is second countable, then so is* S*.*

◇ **Exercise A.9.** Prove the preceding lemma.

If X and Y are topological spaces and $F \colon X \to Y$ is a continuous map, part (e) of the preceding lemma guarantees that the restriction of F to every subspace of X is continuous (in the subspace topology). We can also ask the converse question: If we know that the restriction of F to certain subspaces of X is continuous, is F itself continuous? The next two lemmas express two somewhat different answers to this question.

Lemma A.6 (Continuity is Local). *Continuity is a local property, in the following sense: If* $F \colon X \to Y$ *is a map between topological spaces such that every point* $p \in X$ *has a neighborhood* U *on which the restriction* $F|_U$ *is continuous, then* F *is continuous.*

Lemma A.7 (Gluing Lemma). *Let* X *and* Y *be topological spaces, and let* $K_1, \dots, K_n$ *be finitely many closed subsets of* X *whose union is* X*. Suppose that we are given continuous maps* $f_i \colon K_i \to Y$*,* $i = 1, \dots, n$*, that agree on overlaps:* $f_i|_{K_i \cap K_j} = f_j|_{K_i \cap K_j}$*. Then there exists a unique continuous map* $f \colon X \to Y$ *whose restriction to each* K_i *is equal to* f_i*.*

◇ **Exercise A.10.** Prove the two preceding lemmas.

◇ **Exercise A.11.** Let X be a topological space, and suppose X admits a countable open cover $\{U_i\}$ such that each set U_i is second countable in the subspace topology. Show that X is second countable.

Next we consider finite products of topological spaces. If $X_1, \dots, X_k$ are (finitely many) sets, their *Cartesian product* is the set $X_1 \times \cdots \times X_k$ consisting of all ordered k-tuples of the form $(x_1, \dots, x_k)$ with $x_i \in X_i$ for each i. The ith *projection map* is the map $\pi_i \colon X_1 \times \cdots \times X_k \to X_i$ defined by $\pi_i(x_1, \dots, x_k) = x_i$.

Suppose $X_1, \dots, X_k$ are topological spaces. The collection of all subsets of $X_1 \times \cdots \times X_k$ of the form $U_1 \times \cdots \times U_k$, where each U_i is open in X_i, forms a basis for a topology on $X_1 \times \cdots \times X_k$, called the *product topology.* Endowed with this topology, a finite product of topological spaces is called a *product space.* Any open subset of the form $U_1 \times \cdots \times U_k \subset X_1 \times \cdots \times X_k$, where each U_i is open in X_i, is called a *product open set.* (A slightly different definition is required for products of infinitely many spaces, but we will need only the finite case. See [Mun00] or [Sie92] for details on infinite product spaces.)

Lemma A.8 (Properties of the Product Topology). *Let $X_1 \times \cdots \times X_k$ be a finite product space.*

(*a*) Characteristic Property: *For any topological space B, a map $f \colon B \to X_1 \times \cdots \times X_k$ is continuous if and only if each of its component functions $f_i = \pi_i \circ f \colon B \to X_i$ is continuous.*

(*b*) *The product topology is the unique topology on $X_1 \times \cdots \times X_k$ for which the characteristic property holds.*

(*c*) *Each projection map $\pi_i \colon X_1 \times \cdots \times X_k \to X_i$ is continuous.*

(*d*) *Given continuous maps $F_i \colon X_i \to Y_i$ for $i = 1, \dots, k$, the product map $F_1 \times \cdots \times F_k \colon X_1 \times \cdots \times X_k \to Y_1 \times \cdots \times Y_k$, defined by*

$$F_1 \times \cdots \times F_k(x_1, \dots, x_k) = (F_1(x_1), \dots, F_k(x_k)),$$

is continuous.

(*e*) *If $\mathcal{B}_i$ is a basis for the topology of X_i for $i = 1, \dots, k$, then the collection*

$$\mathcal{B} = \{B_1 \times \cdots \times B_k : B_i \in \mathcal{B}_i\}$$

is a basis for the topology of $X_1 \times \cdots \times X_k$.

(*f*) *Any finite product of Hausdorff spaces is Hausdorff.*

(*g*) *Any finite product of second countable spaces is second countable.*

◇ **Exercise A.12.** Prove the preceding lemma.

Another simple way of building new topological spaces is by taking disjoint unions of other spaces. From a set-theoretic point of view, the disjoint union is defined as follows. If $\{X_\alpha\}_{\alpha \in A}$ is any indexed collection of sets, their *disjoint union* is defined to be the set

$$\coprod_{\alpha \in A} X_\alpha = \{(x, \alpha) : \alpha \in A,\ x \in X_\alpha\}.$$

For each α, there is a canonical injective map $\iota_\alpha \colon X_\alpha \to \coprod_{\alpha \in A} X_\alpha$ given by $\iota_\alpha(x) = (x, \alpha)$, and the images of these maps for different values of α are disjoint. Typically, we implicitly identify X_α with its image in the disjoint union, thereby viewing X_α as a subset of $\coprod_{\alpha \in A} X_\alpha$. The α in the notation (x, α) should be thought of as a "tag" to indicate which set x comes from, so that the subsets corresponding to different values of α are disjoint, even if some or all of the original sets X_α were identical.

Given any indexed collection of topological spaces $\{X_\alpha\}_{\alpha\in A}$, we define the *disjoint union topology* on $\coprod_{\alpha\in A} X_\alpha$ by declaring a subset of $\coprod_{\alpha\in A} X_\alpha$ to be open if and only if its intersection with each X_α is open in X_α.

Lemma A.9 (Properties of the Disjoint Union Topology). *Let $\{X_\alpha\}_{\alpha\in A}$ be an indexed collection of topological spaces, and suppose $\coprod_{\alpha\in A} X_\alpha$ is endowed with the disjoint union topology.*

(*a*) Characteristic Property: *For any topological space Y, a map $f\colon \coprod_{\alpha\in A} X_\alpha \to Y$ is continuous if and only if $f\circ\iota_\alpha\colon X_\alpha\to Y$ is continuous for each $\alpha\in A$.*

(*b*) *The disjoint union topology is the unique topology on $\coprod_{\alpha\in A} X_\alpha$ for which the characteristic property holds.*

(*c*) *A subset of $\{X_\alpha\}_{\alpha\in A}$ is closed if and only if its intersection with each X_α is closed.*

◇ **Exercise A.13.** Prove the preceding lemma.

Quotient Spaces and Quotient Maps

In addition to the subspace, product, and disjoint union topologies, there is a fourth important way of constructing new topological spaces from old ones: the quotient topology.

If X is a topological space, Y is a set, and $\pi\colon X\to Y$ is any surjective map, the *quotient topology* on Y determined by π is defined by declaring a subset $U\subset Y$ to be open if and only if $\pi^{-1}(U)$ is open in X. If X and Y are topological spaces, a map $\pi\colon X\to Y$ is called a *quotient map* if it is surjective and continuous and Y has the quotient topology determined by π.

The following construction is the most common way of producing quotient maps. A relation $\sim$ on a set X is called an *equivalence relation* if it is *reflexive* ($x\sim x$ for all $x\in X$), *symmetric* ($x\sim y$ implies $y\sim x$), and *transitive* ($x\sim y$ and $y\sim z$ imply $x\sim z$). For each $x\in X$, the *equivalence class* of x, denoted by $[x]$, is the set of all $y\in X$ such that $y\sim x$.

Suppose X is a topological space and $\sim$ is an equivalence relation on X. Let $X/{\sim}$ denote the set of equivalence classes in X, and let $\pi\colon X\to X/{\sim}$ be the natural projection sending each point to its equivalence class. Endowed with the quotient topology determined by π, the space $X/{\sim}$ is called the *quotient space* of X determined by the given equivalence relation.

If $\pi\colon X\to Y$ is a map, a subset $U\subset X$ is said to be *saturated* with respect to π if U is the entire inverse image of its image: $U=\pi^{-1}(\pi(U))$. The *fiber* of π over a point $y\in Y$ is the set $\pi^{-1}(y)$. Thus a subset of X is saturated if and only if it is a union of fibers.

Lemma A.10 (Properties of Quotient Maps). *Suppose $\pi\colon X \to Y$ is a quotient map.*

(*a*) Characteristic Property: *For any topological space B, a map $f\colon Y \to B$ is continuous if and only if $f \circ \pi\colon X \to B$ is continuous.*

(*b*) *The quotient topology is the unique topology on Y for which the characteristic property holds.*

(*c*) *A subset $K \subset Y$ is closed if and only if $\pi^{-1}(K)$ is closed in X.*

(*d*) *The restriction of π to any saturated open or closed subset of X is a quotient map.*

(*e*) *The composition of π with any quotient map is again a quotient map.*

◇ **Exercise A.14.** Prove the preceding lemma.

◇ **Exercise A.15.** Let X and Y be topological spaces, and suppose $f\colon X \to Y$ is a surjective continuous map.

(a) Show that f is a quotient map if and only if it takes saturated open sets to open sets.
(b) Show that f is a quotient map if and only if it takes saturated closed sets to closed sets.

The next two properties of quotient maps play an important role in topology, and have equally important generalizations in smooth manifold theory (see Chapter 7).

Lemma A.11 (Passing to the Quotient). *Suppose $\pi\colon X \to Y$ is a quotient map, B is a topological space, and $f\colon X \to B$ is a continuous map that is constant on the fibers of π (i.e., $\pi(p) = \pi(q)$ implies $f(p) = f(q)$). Then there exists a unique continuous map $\widetilde{f}\colon Y \to B$ such that $f = \widetilde{f} \circ \pi$.*

Proof. The existence and uniqueness of $\widetilde{f}$ follow from set theoretic considerations, and its continuity is an immediate consequence of the characteristic property of the quotient topology. □

Lemma A.12 (Uniqueness of Quotient Spaces). *If $\pi_1\colon X \to Y_1$ and $\pi_2\colon X \to Y_2$ are quotient maps that are constant on each other's fibers (i.e., $\pi_1(p) = \pi_1(q)$ if and only if $\pi_2(p) = \pi_2(q)$), then there exists a unique homeomorphism $\varphi\colon Y_1 \to Y_2$ such that $\varphi \circ \pi_1 = \pi_2$.*

Proof. Applying the preceding lemma to the quotient map $\pi_2\colon X \to Y_2$, we see that π_1 passes to the quotient, yielding a continuous map $\widetilde{\pi}_1\colon Y_2 \to Y_1$ satisfying $\widetilde{\pi}_1 \circ \pi_2 = \pi_1$. Applying the same argument with the roles of π_1 and π_2 reversed, there is a continuous map $\widetilde{\pi}_2\colon Y_1 \to Y_2$ satisfying $\widetilde{\pi}_2 \circ \pi_1 = \pi_2$. Together, these identities imply that $\widetilde{\pi}_1 \circ \widetilde{\pi}_2 \circ \pi_1 = \pi_1$. Applying Lemma A.11 again with π_1 playing the roles of both π and f, we see that both $\widetilde{\pi}_1 \circ \widetilde{\pi}_2$ and Id_{Y_1} are obtained from π_1 by passing to the quotient. The uniqueness assertion of Lemma A.11 therefore implies that $\widetilde{\pi}_1 \circ \widetilde{\pi}_2 = \mathrm{Id}_{Y_1}$. A similar argument shows that $\widetilde{\pi}_2 \circ \widetilde{\pi}_1 = \mathrm{Id}_{Y_2}$, so that $\widetilde{\pi}_2$ is the desired homeomorphism. □

Open and Closed Maps

A map $F\colon X \to Y$ (continuous or not) is said to be an *open map* if for every open subset $U \subset X$, $F(U)$ is open in Y, and a *closed map* if for every closed subset $K \subset X$, $F(K)$ is closed in Y. Continuous maps may be open, closed, both, or neither, as can be seen by examining simple examples involving subsets of the plane.

◇ **Exercise A.16.** Show that every local homeomorphism is an open map.

The most important classes of continuous maps in topology are the homeomorphisms, quotient maps, and topological embeddings. Obviously, it is necessary for a map to be bijective in order for it to be a homeomorphism, surjective for it to be a quotient map, and injective for it to be a topological embedding. However, even when a continuous map is known to satisfy one of these obvious necessary conditions, it is not always easy to tell whether it has the desired topological property. One simple sufficient condition is that it be either an open or a closed map, as the next lemma shows.

Lemma A.13. *Suppose X and Y are topological spaces, and $F\colon X \to Y$ is a continuous map that is either open or closed.*

(*a*) *If F is surjective, it is a quotient map.*

(*b*) *If F is injective, it is a topological embedding.*

(*c*) *If F is bijective, it is a homeomorphism.*

Proof. Suppose first that F is surjective. If it is open, it certainly takes saturated open sets to open sets. Similarly, if it is closed, it takes saturated closed sets to closed sets. Thus it is a quotient map by Exercise A.15.

Now suppose F is open and injective. Then $F\colon X \to F(X)$ is bijective, so $F^{-1}\colon F(X) \to X$ exists by elementary set-theoretic considerations. If $U \subset X$ is open, then $(F^{-1})^{-1}(U) = F(U)$ is open in Y by hypothesis, and therefore is also open in $F(X)$ by definition of the subspace topology on $F(X)$. This proves that F^{-1} is continuous, so that F is a homeomorphism onto its image. If F is closed, the same argument goes through with "open" replaced by "closed" throughout (using the characterization of closed subsets of $F(X)$ given in Lemma A.5(c)). This proves part (b), and part (c) is just the special case of (b) in which $F(X) = Y$. □

Connectedness

Let X be a topological space. A *separation* of X is a pair of disjoint, nonempty, open subsets $U, V \subset X$ whose union is X. If there exists a separation of X, then X is said to be *disconnected*; if there does not, then it is *connected*. Equivalently, X is connected if and only if the only subsets of X that are both open and closed are $\varnothing$ and X itself. For example, the nonempty connected subsets of $\mathbb{R}$ are precisely the *intervals*, which are the nonempty subsets $J \subset \mathbb{R}$ with the property that whenever $a, b \in J$ and $a < c < b$, then $c \in J$ as well.

If X is any topological space, we can define an equivalence relation on X called the *connectivity relation* by declaring $p \sim q$ if and only if there exists a connected subset of X containing both p and q. The equivalence classes in X under this equivalence relation are called the *components* (or *connected components*) of X; they are the maximal connected subsets of X.

Lemma A.14 (Properties of Connected Spaces). *Let X and Y be topological spaces.*

(*a*) *If $f\colon X \to Y$ is continuous and X is connected, then $f(X)$ is connected.*

(*b*) *Any connected subset of X is contained in a single component of X.*

(*c*) *Any union of connected subspaces of X with a point in common is connected.*

(*d*) *Any finite product of connected spaces is connected.*

(*e*) *Any quotient space of a connected space is connected.*

◇ **Exercise A.17.** Prove the preceding lemma.

Closely related to connectedness is path connectedness. If X is a topological space and p, q are points in X, a *path* in X from p to q is a continuous map $f\colon [0,1] \to X$ such that $f(0) = p$ and $f(1) = q$. If for every pair of points $p, q \in X$, there exists a path in X from p to q, then X is said to be *path connected*.

Associated with path connectedness is an equivalence relation analogous to the connectivity relation. If X is any space, define the *path connectivity relation* on X by declaring $p \underset{\mathrm{p}}{\sim} q$ if and only if there is a path in X from p to q. This is an equivalence relation whose equivalence classes are called the *path components* of X.

Lemma A.15 (Properties of Path Connected Spaces). *Let X and Y be topological spaces.*

(*a*) *If $f\colon X \to Y$ is continuous and X is path connected, then $f(X)$ is path connected.*

(*b*) *If X is path connected, then it is connected.*

(*c*) *Any path connected subset of X is contained in a single path component.*

(*d*) *Any finite product of path connected spaces is path connected.*

(*e*) *Any quotient space of a path connected space is path connected.*

◇ **Exercise A.18.** Prove the preceding lemma.

For most topological spaces we treat in this book, including all manifolds, connectedness and path connectedness turn out to be equivalent. The link between the two concepts is provided by the following notion. A topological space is said to be *locally path connected* if it admits a basis of path connected open sets.

Lemma A.16 (Properties of Locally Path Connected Spaces). *Let X be a locally path connected topological space.*

(*a*) *The components of X are open in X.*

(*b*) *The path components of X are equal to its components.*

(*c*) *X is connected if and only if it is path connected.*

◇ **Exercise A.19.** Prove the preceding lemma.

Compactness

A topological space X is said to be *compact* if every open cover of X has a finite subcover. For example, a subset of $\mathbb{R}^n$ is compact if and only if it is closed and bounded.

Lemma A.17 (Properties of Compact Spaces). *Let X and Y be topological spaces.*

(*a*) *If $f\colon X \to Y$ is continuous and X is compact, then $f(X)$ is compact.*

(*b*) *If X is compact and $f\colon X \to \mathbb{R}$ is a continuous real-valued function, then f is bounded and attains its maximum and minimum values on X.*

(*c*) *If X is compact, then every closed subset of X is compact.*

(*d*) *If X is Hausdorff, then every compact subset of X is closed in X.*

(*e*) *If X is Hausdorff and K and L are compact subsets of X, then there exist disjoint open subsets $U, V \subset X$ such that $K \subset U$ and $L \subset V$.*

(*f*) *Every finite product of compact spaces is compact.*

(*g*) *Every quotient of a compact space is compact.*

◇ **Exercise A.20.** Prove the preceding lemma.

For manifolds, subsets of manifolds, and most other spaces we will be working with, there are two other equivalent formulations of compactness that are frequently useful. Proofs of the next proposition can be found in [Lee00, Chapter 4] or [Sie92, Chapter 7].

Proposition A.18 (Equivalent Formulations of Compactness). *Let M be a second countable Hausdorff space or a metric space. Then the following are equivalent.*

(*a*) *M is compact.*

(*b*) *Every infinite subset of M has a limit point in M.*

(*c*) *Every sequence in M has a convergent subsequence in M.*

◇ **Exercise A.21.** Show that every compact metric space is complete.

One of the most useful properties of compact spaces is expressed in the following lemma.

Lemma A.19 (Closed Map Lemma). *Suppose X is a compact space, Y is a Hausdorff space, and $F\colon X \to Y$ is a continuous map.*

(*a*) *F is a closed map.*

(*b*) *If F is surjective, it is a quotient map.*

(*c*) *If F is injective, it is a topological embedding.*

(*d*) *If F is bijective, it is a homeomorphism.*

Proof. By virtue of Lemma A.13, the last three assertions follow from the first, so we need only prove that F is closed. Suppose $K \subset X$ is a closed set. Then part (c) of Lemma A.17 implies that K is compact; part (a) of that lemma implies that $F(K)$ is compact; and part (d) implies that $F(K)$ is closed in Y. □

Homotopy and the Fundamental Group

If X and Y are topological spaces and $F_0, F_1\colon X \to Y$ are continuous maps, a *homotopy* from F_0 to F_1 is a continuous map $H\colon X \times I \to Y$ (where $I = [0,1]$ is the closed unit interval in $\mathbb{R}$) satisfying

$$H(x,0) = F_0(x),$$
$$H(x,1) = F_1(x)$$

for all $x \in X$. If there exists a homotopy from F_0 to F_1, we say F_0 and F_1 are *homotopic*, and write $F_0 \simeq F_1$. If the homotopy satisfies $H(x,t) = F_0(x) =$

$F_1(x)$ for all $t \in I$ and all x in some subset $A \subset X$, the maps F_0 and F_1 are said to be *homotopic relative to* A. Both homotopy and homotopy relative to A are equivalence relations on the set of all continuous maps from X to Y.

The most important application of homotopies is to paths. Suppose X is a topological space. Two paths $f_0, f_1 \colon I \to X$ are said to be *path homotopic*, denoted symbolically by $f_0 \sim f_1$, if they are homotopic relative to $\{0,1\}$. Explicitly, this means that there is a continuous map $H \colon I \times I \to X$ satisfying

$$\begin{aligned} H(s,0) &= f_0(s), && s \in I; \\ H(s,1) &= f_1(s), && s \in I; \\ H(0,t) &= f_0(0) = f_1(0), && t \in I; \\ H(1,t) &= f_0(1) = f_1(1), && t \in I. \end{aligned}$$

For any given points $p, q \in X$, path homotopy is an equivalence relation on the set of all paths from p to q. The equivalence class of a path f is called its *path class*, and is denoted by $[f]$.

Given two paths f, g such that $f(1) = g(0)$, their *product* is the path $f \cdot g$ defined by

$$f \cdot g(s) = \begin{cases} f(2s), & 0 \le s \le \frac{1}{2}; \\ g(2s-1), & \frac{1}{2} \le s \le 1. \end{cases}$$

If $f \sim f'$ and $g \sim g'$, it is not hard to show that $f \cdot g \sim f' \cdot g'$. Therefore it makes sense to define the product of the path classes $[f]$ and $[g]$ by $[f] \cdot [g] = [f \cdot g]$. Although multiplication of paths is not associative, it is associative up to path homotopy: $([f] \cdot [g]) \cdot [h] = [f] \cdot ([g] \cdot [h])$. When we need to consider products of three or more actual paths (as opposed to path classes), we adopt the convention that such products are to be evaluated from left to right: $f \cdot g \cdot h = (f \cdot g) \cdot h$.

If X is a topological space and q is a point in X, a *loop* in X based at q is a path in X from q to q; i.e., a continuous map $f \colon I \to X$ such that $f(0) = f(1) = q$. The set of path classes of loops based at q is denoted by $\pi_1(X, q)$. Equipped with the product described above, it is a group, called the *fundamental group of* X *based at* q. The identity element of this group is the path class of the *constant loop* $c_q(s) \equiv q$, and the inverse of $[f]$ is the path class of the *reverse loop* $f^{-1}(s) = f(1-s)$.

It can be shown that for path connected spaces, the fundamental groups based at different points are isomorphic. If X is path connected and for some (hence any) $q \in X$, $\pi_1(X, q)$ is the trivial group consisting of $[c_q]$ alone, we say that X is *simply connected*. This means that every loop is path homotopic to a constant loop.

◇ **Exercise A.22.** Let X be a path connected topological space. Show that X is simply connected if and only if any two paths in X with the same starting and ending points are path homotopic.

A key feature of the path homotopy relation is that it is preserved by continuous maps, as the next lemma shows.

Lemma A.20. *If $f_0, f_1 \colon I \to X$ are path homotopic and $F \colon X \to Y$ is a continuous map, then $F \circ f_0 \sim F \circ f_1$.*

◇ **Exercise A.23.** Prove the preceding lemma.

Thus if $F \colon X \to Y$ is a continuous map, for each $q \in X$ we obtain a well-defined map $F_* \colon \pi_1(X, q) \to \pi_1(Y, F(q))$ by setting

$$F_*[f] = [F \circ f].$$

Lemma A.21. *If $F \colon X \to Y$ is a continuous map, then $F_* \colon \pi_1(X, q) \to \pi_1(Y, F(q))$ is a group homomorphism, known as the homomorphism induced by F.*

Lemma A.22 (Properties of the Induced Homomorphism).

(*a*) *Let $F \colon X \to Y$ and $G \colon Y \to Z$ be continuous maps. Then for any $q \in X$, $(G \circ F)_* = G_* \circ F_* \colon \pi_1(X, q) \to \pi_1(Z, G(F(q)))$.*

(*b*) *For any space X and any $q \in X$, the homomorphism induced by the identity map $\mathrm{Id}_X \colon X \to X$ is the identity map of $\pi_1(X, q)$.*

(*c*) *If $F \colon X \to Y$ is a homeomorphism, then $F_* \colon \pi_1(X, q) \to \pi_1(Y, F(q))$ is an isomorphism. Thus homeomorphic spaces have isomorphic fundamental groups.*

◇ **Exercise A.24.** Prove the two preceding lemmas.

◇ **Exercise A.25.** A subset $U \subset \mathbb{R}^n$ is said to be *star-shaped* with respect to a point $p \in U$ if for every $q \in U$, the line segment from p to q is contained in U. Show that any star-shaped set is simply connected.

For $n \geq 0$, we define the (*unit*) *n-sphere* to be the set

$$\mathbb{S}^n = \{x \in \mathbb{R}^{n+1} : |x| = 1\},$$

with the Euclidean topology. In the special case $n = 1$, we often think of $\mathbb{S}^1$ as a subset of $\mathbb{C}$ under our usual identification of $\mathbb{R}^2$ with $\mathbb{C}$.

Proposition A.23 (Fundamental Groups of Spheres).

(*a*) *$\pi_1(\mathbb{S}^1, 1)$ is the infinite cyclic group generated by the path class of the loop $\alpha \colon [0, 1] \to \mathbb{S}^1$ given (in complex notation) by $\alpha(s) = e^{2\pi i s}$.*

(*b*) *If $n > 1$, $\mathbb{S}^n$ is simply connected.*

Proposition A.24 (Fundamental Groups of Product Spaces). *Let $X_1, \dots, X_k$ be topological spaces, and let $\pi_i \colon X_1 \times \cdots \times X_k \to X_i$ denote the ith projection map. For any points $q_i \in X_i$, $i = 1, \dots, k$, define a map $P \colon \pi_1(X_1 \times \cdots \times X_k, (q_1, \dots, q_k)) \to \pi_1(X_1, q_1) \times \cdots \times \pi_1(X_k, q_k)$ by*

$$P[f] = (\pi_{1*}[f], \dots, \pi_{k*}[f]).$$

Then P is an isomorphism.

◇ **Exercise A.26.** Prove the two preceding propositions.

A continuous map $F \colon X \to Y$ between topological spaces is said to be a *homotopy equivalence* if there is a continuous map $G \colon Y \to X$ such that $F \circ G \simeq \mathrm{Id}_Y$ and $G \circ F \simeq \mathrm{Id}_X$. Such a map G is called a *homotopy inverse* for F. If there exists a homotopy equivalence between X and Y, the two spaces are said to be *homotopy equivalent*. For example, the inclusion map $\iota \colon \mathbb{S}^{n-1} \hookrightarrow \mathbb{R}^n \smallsetminus \{0\}$ is a homotopy equivalence with homotopy inverse $r(x) = x/|x|$, because $r \circ \iota = \mathrm{Id}_{\mathbb{S}^{n-1}}$ and $\iota \circ r$ is homotopic to the identity map of $\mathbb{R}^n \smallsetminus \{0\}$ via the straight-line homotopy $H(x,t) = tx + (1-t)x/|x|$.

Theorem A.25 (Homotopy Invariance). *If $F \colon X \to Y$ is a homotopy equivalence, then for each $p \in X$, $F_* \colon \pi_1(X, p) \to \pi_1(Y, F(p))$ is an isomorphism.*

For a proof, see any of the topology texts mentioned at the beginning of this section.

Covering Maps

Suppose $\widetilde{X}$ and X are topological spaces. A map $\pi \colon \widetilde{X} \to X$ is called a *covering map* if $\widetilde{X}$ is path connected and locally path connected, π is surjective and continuous, and each point $p \in X$ has a neighborhood U that is *evenly covered* by π, meaning that U is connected and each component of $\pi^{-1}(U)$ is mapped homeomorphically onto U by π. In this case, X is called the *base* of the covering, and $\widetilde{X}$ is called a *covering space* of X.

◇ **Exercise A.27.** Show that a covering map is a local homeomorphism, an open map, and a quotient map.

◇ **Exercise A.28.** Show that an injective covering map is a homeomorphism.

◇ **Exercise A.29.** If $\pi \colon \widetilde{X} \to X$ is a covering map, show that all fibers of π have the same cardinality, called the *number of sheets* of the covering.

◇ **Exercise A.30.** Show that any finite product of covering maps is a covering map.

The main properties of covering maps that we will need are summarized in the next four propositions. For proofs, you can consult [Lee00, Chapters 11 and 12] or [Sie92, Chapter 14].

If $\pi\colon \widetilde{X} \to X$ is a covering map and $F\colon Y \to X$ is any continuous map, a *lift* of F is a continuous map $\widetilde{F}\colon Y \to \widetilde{X}$ such that $\pi \circ \widetilde{F} = F$:

$$\begin{tikzcd} & \widetilde{X} \arrow[d, "\pi"] \\ Y \arrow[ur, "\widetilde{F}"] \arrow[r, "F"'] & X. \end{tikzcd}$$

Proposition A.26 (Lifting Properties of Covering Maps). *Suppose $\pi\colon \widetilde{X} \to X$ is a covering map.*

(*a*) Unique Lifting Property: *If Y is a connected space and $F\colon Y \to X$ is a continuous map, then any two lifts of F that agree at one point are identical.*

(*b*) Path Lifting Property: *If $f\colon I \to X$ is a path and $\widetilde{q}_0 \in \widetilde{X}$ is a point such that $\pi(\widetilde{q}_0) = f(0)$, then there exists a unique lift $\widetilde{f}\colon I \to \widetilde{X}$ of f such that $\widetilde{f}(0) = \widetilde{q}_0$.*

(*c*) Homotopy Lifting Property: *If $f_0, f_1\colon I \to X$ are path homotopic and $\widetilde{f}_0, \widetilde{f}_1\colon I \to \widetilde{X}$ are lifts of f_0 and f_1 such that $\widetilde{f}_0(0) = \widetilde{f}_1(0)$, then $\widetilde{f}_0$ and $\widetilde{f}_1$ are path homotopic.*

Proposition A.27 (Lifting Criterion). *Suppose $\pi\colon \widetilde{X} \to X$ is a covering map, Y is a connected and locally path connected space, and $F\colon Y \to X$ is a continuous map. Let $y \in Y$ and $\widetilde{x} \in \widetilde{X}$ be such that $\pi(\widetilde{x}) = F(y)$. Then there exists a lift $\widetilde{F}\colon Y \to \widetilde{X}$ of F satisfying $\widetilde{F}(y) = \widetilde{x}$ if and only if $F_*\pi_1(Y, y) \subset \pi_*\pi_1(\widetilde{X}, \widetilde{x})$.*

Proposition A.28 (Coverings of Simply Connected Spaces). *If X is a simply connected space, then any covering map $\pi\colon \widetilde{X} \to X$ is a homeomorphism.*

A topological space is said to be *locally simply connected* if it admits a basis of simply connected open sets.

Proposition A.29 (Existence of a Universal Covering Space). *If X is a connected and locally simply connected topological space, there exists a simply connected topological space $\widetilde{X}$ and a covering map $\pi\colon \widetilde{X} \to X$. If $\widehat{\pi}\colon \widehat{X} \to X$ is any other simply connected covering of X, there is a homeomorphism $\varphi\colon \widetilde{X} \to \widehat{X}$ such that $\widehat{\pi} \circ \varphi = \pi$.*

The simply connected covering space $\widetilde{X}$ whose existence and uniqueness (up to homeomorphism) are guaranteed by this proposition is called the *universal covering space* of X.

Linear Algebra

For the basic properties of vector spaces and linear maps, you can consult almost any linear algebra book that treats vector spaces abstractly, such as [FIS97]. Here we just summarize the main points, with emphasis on those aspects that will prove most important in the study of smooth manifolds.

Vector Spaces

Let $\mathbb{R}$ denote the field of real numbers. A *vector space* over $\mathbb{R}$ (or *real vector space*) is a set V endowed with two operations: *vector addition* $V \times V \to V$, denoted by $(X, Y) \mapsto X + Y$, and *scalar multiplication* $\mathbb{R} \times V \to V$, denoted by $(a, X) \mapsto aX$; the operations are required to satisfy

(i) V is an abelian group under vector addition.

(ii) Scalar multiplication satisfies the following identities:

$$\begin{aligned} a(bX) &= (ab)X && \text{for all } X \in V \text{ and } a, b \in \mathbb{R}; \\ 1X &= X && \text{for all } X \in V. \end{aligned}$$

(iii) Scalar multiplication and vector addition are related by the following distributive laws:

$$\begin{aligned} (a+b)X &= aX + bX && \text{for all } X \in V \text{ and } a, b \in \mathbb{R}; \\ a(X+Y) &= aX + aY && \text{for all } X, Y \in V \text{ and } a \in \mathbb{R}. \end{aligned}$$

The elements of V are usually called *vectors*. When necessary to distinguish them from vectors, real numbers are sometimes called *scalars*.

This definition can be generalized in several directions. First, replacing $\mathbb{R}$ by an arbitrary field $\mathbb{F}$ everywhere, we obtain the definition of a vector space over $\mathbb{F}$. Second, if $\mathbb{R}$ is replaced by a commutative ring $\mathcal{R}$, this becomes the definition of a *module* over $\mathcal{R}$. For example, it is straightforward to check that modules over the ring $\mathbb{Z}$ of integers are just abelian groups under addition. We will be concerned almost exclusively with real vector spaces, but it is useful to be aware of these more general definitions. Unless we specify otherwise, all vector spaces will be assumed to be real.

If V is a vector space, a subset $W \subset V$ that is closed under vector addition and scalar multiplication is itself a vector space, and is called a *subspace* of V.

Let V be a vector space. A finite sum of the form $\sum_{i=1}^{k} a^i X_i$, where $a^i \in \mathbb{R}$ and $X_i \in V$, is called a *linear combination* of the vectors $X_1, \dots, X_k$. (The reason we write the coefficients a^i with superscripts instead of subscripts is to be consistent with the Einstein summation convention, explained in Chapter 1.) If S is an arbitrary subset of V, the set of all linear combinations of elements of S is called the *span* of S and is denoted by $\operatorname{span}(S)$; it is easily seen to be the smallest subspace of V containing S. If $V = \operatorname{span}(S)$, we say S spans V. By convention, a linear combination of no elements is considered to sum to zero, and the span of the empty set is $\{0\}$.

If p, q are points of V, the *line segment* from p to q is the set $\{tp+(1-t)q : 0 \le t \le 1\}$. A subset $B \subset V$ is said to be *convex* if for every two points $p, q \in B$, the line segment from p to q is contained in B.

Bases and Dimension

Suppose V is a real vector space. A subset $S \subset V$ is said to be *linearly dependent* if there exists a linear relation of the form $\sum_{i=1}^{k} a^i X_i = 0$, where $X_1, \dots, X_k$ are distinct elements of S and at least one of the coefficients a^i is nonzero; it is said to be *linearly independent* otherwise. In other words, S is linearly independent if and only if the only linear combination of distinct elements of S that sums to zero is the one in which all the scalar coefficients are zero. Note that any set containing the zero vector is linearly dependent. By convention, the empty set is considered to be linearly independent.

◇ **Exercise A.31.** Let V be a vector space and $S \subset V$.

(a) If S is linearly independent, show that any subset of S is linearly independent.
(b) If S is linearly dependent or spans V, show that any subset of V that properly contains S is linearly dependent.
(c) Show that S is linearly dependent if and only if some element $X \in S$ can be expressed as a linear combination of elements of $S \smallsetminus \{X\}$.
(d) If $(X_1, \dots, X_m)$ is a finite, ordered, linearly dependent subset of V, show that some X_i can be written as a linear combination of the *preceding* vectors $(X_1, \dots, X_{i-1})$.

A *basis* for V is a subset $S \subset V$ that is linearly independent and spans V. If S is a basis for V, every element of V has a *unique* expression as a linear combination of elements of S. If V has a finite basis, then V is said to be *finite-dimensional*, and otherwise it is *infinite-dimensional*. The trivial vector space $\{0\}$ (which we denote by $\mathbb{R}^0$) is finite-dimensional, because it has the empty set as a basis.

Lemma A.30. *Let V be a finite-dimensional vector space. If V is spanned by n vectors, then every subset of V containing more than n vectors is linearly dependent.*

Proof. Suppose the vectors $\{X_1, \dots, X_n\}$ span V. To prove the lemma, it clearly suffices to show that any $n+1$ vectors $\{Y_1, \dots, Y_{n+1}\}$ are dependent. Suppose not. By Exercise A.31(b), the set $\{Y_1, X_1, \dots, X_n\}$ is dependent. By Exercise A.31(d), one of the vectors X_j can be written as a linear combination of $\{Y_1, \dots, X_{j-1}\}$. Renumbering the X_is if necessary, we may assume that $j = 1$, so the set $\{Y_1, X_2, \dots, X_n\}$ still spans V.

Now suppose by induction that $\{Y_1, Y_2, \dots, Y_{k-1}, X_k, \dots, X_n\}$ spans V. As before, the set $\{Y_1, Y_2, \dots, Y_{k-1}, Y_k, X_k, \dots, X_n\}$ is dependent, so one of the vectors in this list can be written as a linear combination of the preceding ones. Because the Y_is are independent, it must be one of the X_js that can be so written, and after reordering we may assume it is X_k. Thus the set $\{Y_1, Y_2, \dots, Y_k, X_{k+1}, \dots, X_n\}$ still spans V. Continuing by induction, we conclude that the vectors $\{Y_1, \dots, Y_n\}$ span V. But this means that $\{Y_1, \dots, Y_{n+1}\}$ are dependent by Exercise A.31(b). □

Proposition A.31. *If V is a finite-dimensional vector space, all bases for V contain the same number of elements.*

Proof. If $\{E_1, \dots, E_n\}$ is a basis for V, then Lemma A.30 implies that any set containing more than n elements is dependent, so no basis can have more than n elements. Conversely, if there were a basis containing fewer than n elements, then Lemma A.30 would imply that $\{E_1, \dots, E_n\}$ is dependent, which is a contradiction. □

Because of the preceding proposition, it makes sense to define the *dimension* of a finite-dimensional vector space V, denoted by $\dim V$, to be the number of elements in any basis.

◇ **Exercise A.32.** Suppose V is a finite-dimensional vector space.

(a) Show that every set that spans V contains a basis, and every linearly independent subset of V is contained in a basis.
(b) If $S \subset V$ is a subspace, show that S is finite-dimensional and $\dim S \leq \dim V$, with equality if and only if $S = V$.
(c) Show that $\dim V = 0$ if and only if $V = \{0\}$.

If S is a subspace of a finite-dimensional vector space V, we define the *codimension of S in V* to be $\dim V - \dim S$. By virtue of part (b) of the preceding exercise, the codimension of S is always nonnegative, and is zero if and only if $S = V$. A *hyperplane* is a subspace of codimension 1.

An *ordered basis* of a finite-dimensional vector space is a basis endowed with a specific ordering of the basis vectors. For most purposes, ordered bases are more useful than bases, so we will assume, usually without comment, that each basis comes with a given ordering. We will denote an ordered basis by a notation such as $(E_1, \dots, E_n)$ or (E_i).

If $(E_1, \dots, E_n)$ is an (ordered) basis for V, each vector $X \in V$ has a unique expression as a linear combination of basis vectors:

$$X = \sum_{i=1}^{n} X^i E_i.$$

The numbers X^i are called the *components* of X with respect to this basis, and the ordered n-tuple $(X^1, \dots, X^n)$ is called its *basis representation.* (Here is an example of a definition that requires an ordered basis.)

The fundamental example of a finite-dimensional vector space is of course n-dimensional Euclidean space $\mathbb{R}^n$. It is a vector space under the usual operations of vector addition and scalar multiplication. It has a natural basis $(e_1, \dots, e_n)$, called the *standard basis*, where $e_i = (0, \dots, 1, \dots, 0)$ is the vector with a 1 in the ith place and zeros elsewhere. Any point $x \in \mathbb{R}^n$ can be written $(x^1, \dots, x^n) = \sum_{i=1}^n x^i e_i$, so its components with respect to the standard basis are just its coordinates $(x^1, \dots, x^n)$.

If S and T are subspaces of a vector space V, the notation $S+T$ denotes the set of all vectors of the form $X+Y$, where $X \in S$ and $Y \in T$. It is easily seen to be a subspace, and in fact is the subspace spanned by $S \cup T$. If $S+T = V$ and $S \cap T = \{0\}$, V is said to be the *direct sum* of S and T, and we write $V = S \oplus T$.

If S is any subspace of V, another subspace $T \subset V$ is said to be *complementary* to S if $V = S \oplus T$. In this case, it is easy to check that every vector in V has a *unique* expression as a sum of an element of S plus an element of T.

◇ **Exercise A.33.** Suppose S and T are subspaces of a finite-dimensional vector space V.

(a) Show that $S \cap T$ is a subspace of V.
(b) Show that $\dim(S+T) = \dim S + \dim T - \dim(S \cap T)$.
(c) If $V = S+T$, show that $V = S \oplus T$ if and only if $\dim V = \dim S + \dim T$.

◇ **Exercise A.34.** Let V be a finite-dimensional vector space. Show that every subspace $S \subset V$ has a complementary subspace in V. In fact, if $(E_1, \dots, E_n)$ is any basis for V, show that there is some subset $\{i_1, \dots, i_k\}$ of the integers $\{1, \dots, n\}$ such that $\operatorname{span}(E_{i_1}, \dots, E_{i_k})$ is a complement to S. [Hint: Choose a basis $(F_1, \dots, F_m)$ for S, and apply Exercise A.31(d) to the ordered $(m+n)$-tuple $(F_1, \dots, F_m, E_1, \dots, E_n)$.]

Suppose $S \subset V$ is a subspace. For any vector $x \in V$, the *coset* of S determined by x is the set

$$x + S = \{x + y : y \in S\}.$$

A coset is also sometimes called an *affine subspace* of V parallel to S. The set V/S of cosets of S is called the *quotient* of V by S.

◇ **Exercise A.35.** Suppose V is a vector space and S is a subspace of V. Define vector addition and scalar multiplication of cosets by

$$\begin{aligned}(x+S) + (y+S) &= (x+y) + S,\\ c(x+S) &= (cx) + S.\end{aligned}$$

(a) Show that the quotient V/S is a vector space under these operations.
(b) If V is finite-dimensional, show that $\dim V/S = \dim V - \dim S$.

Linear Maps

Let V and W be vector spaces. A map $T\colon V \to W$ is *linear* if $T(aX+bY) = aTX + bTY$ for all vectors $X, Y \in V$ and all scalars a, b. (Because of the close connection between linear maps and matrix multiplication described below, we generally write the action of a linear map T on a vector X as TX without parentheses, unless parentheses are needed for grouping.) The *kernel* or *nullspace* of T, denoted by $\operatorname{Ker} T$, is the set $\{X \in V : TX = 0\}$, and the *image* of T, denoted by $\operatorname{Im} T$ or $T(V)$, is the set $\{Y \in W : Y = TX \text{ for some } X \in V\}$.

One simple but important example of a linear map arises in the following way. Given a subspace $S \subset V$ and a complementary subspace T, there is a unique linear map $\pi\colon V \to S$ defined by

$$\pi(X+Y) = X \text{ for } X \in S,\ Y \in T.$$

This map is called the *projection onto S with kernel T*.

A bijective linear map $T\colon V \to W$ is called an *isomorphism*. In this case, there is a unique inverse map $T^{-1}\colon W \to V$, and the following computation shows that T^{-1} is also linear:

$$\begin{aligned} aT^{-1}X + bT^{-1}Y &= T^{-1}T(aT^{-1}X + bT^{-1}Y) \\ &= T^{-1}(aTT^{-1}X + bTT^{-1}Y) \qquad \text{(by linearity of } T\text{)} \\ &= T^{-1}(aX + bY). \end{aligned}$$

For this reason, a bijective linear map is also said to be *invertible*. If there exists an isomorphism $T\colon V \to W$, then V and W are said to be *isomorphic*. Isomorphism is easily seen to be an equivalence relation.

Example A.32. Let V be any n-dimensional vector space, and let $(E_1, \dots, E_n)$ be any ordered basis for V. Define a map $E\colon \mathbb{R}^n \to V$ by

$$E(x^1, \dots, x^n) = x^1E_1 + \cdots + x^nE_n.$$

Then E is bijective, so it is an isomorphism, called the *basis isomorphism* determined by this basis. Thus every n-dimensional vector space is isomorphic to $\mathbb{R}^n$.

◇ **Exercise A.36.** Let V and W be vector spaces, and suppose $(E_1, \dots, E_n)$ is a basis for V. For any n elements $X_1, \dots, X_n \in W$, show that there is a unique linear map $T\colon V \to W$ satisfying $T(E_i) = X_i$ for $i = 1, \dots, n$.

◇ **Exercise A.37.** Let $S\colon V \to W$ and $T\colon W \to X$ be linear maps.

(a) Show that $\operatorname{Ker} S$ and $\operatorname{Im} S$ are subspaces of V and W, respectively.
(b) Show that S is injective if and only if $\operatorname{Ker} S = \{0\}$.
(c) If S is an isomorphism, show that $\dim V = \dim W$ (in the sense that these dimensions are either both infinite or both finite and equal).

(d) If S and T are both injective or both surjective, show that $T \circ S$ has the same property.
(e) If $T \circ S$ is surjective, show that T is surjective; give an example to show that S may not be.
(f) If $T \circ S$ is injective, show that S is injective; give an example to show that T may not be.

◇ **Exercise A.38.** Suppose V is a vector space and S is a subspace of V, and let $\pi\colon V \to V/S$ denote the projection defined by $\pi(x) = x + S$.

(a) Show that π is a surjective linear map with kernel equal to S.
(b) If $T\colon V \to W$ is any linear map, show that there exists a linear map $\widetilde{T}\colon V/S \to W$ such that $\widetilde{T} \circ \pi = T$ if and only if $S \subset \operatorname{Ker} T$.

Now suppose V and W are finite-dimensional vector spaces with ordered bases $(E_1, \dots, E_n)$ and $(F_1, \dots, F_m)$, respectively. If $T\colon V \to W$ is a linear map, the *matrix of* T with respect to these bases is the $m \times n$ matrix

$$A = (A^i_j) = \begin{pmatrix} A^1_1 & \dots & A^1_n \\ \vdots & \ddots & \vdots \\ A^m_1 & \dots & A^m_n \end{pmatrix}$$

whose jth column consists of the components of TE_j with respect to the basis (F_i):

$$TE_j = \sum_{i=1}^m A^i_j F_i.$$

By linearity, the action of T on any vector $X = \sum_j X^j E_j$ is then given by

$$T\left(\sum_{j=1}^n X^j E_j\right) = \sum_{i=1}^m \sum_{j=1}^n A^i_j X^j F_i.$$

If we write the components of a vector with respect to a basis as a column matrix, then the matrix representation of $Y = TX$ is given by matrix multiplication:

$$\begin{pmatrix} Y^1 \\ \vdots \\ Y^m \end{pmatrix} = \begin{pmatrix} A^1_1 & \dots & A^1_n \\ \vdots & \ddots & \vdots \\ A^m_1 & \dots & A^m_n \end{pmatrix} \begin{pmatrix} X^1 \\ \vdots \\ X^n \end{pmatrix},$$

or, more succinctly,

$$Y^i = \sum_{j=1}^n A^i_j X^j.$$

Insofar as possible, we will denote the row index of a matrix by a superscript and the column index by a subscript, so that A^i_j represents the element in the ith row and jth column. Thus the entry in the ith row and jth column of a matrix product AB is given by

$$(AB)^i_j = \sum_{k=1}^{n} A^i_k B^k_j.$$

It is straightforward to check that the composition of two linear maps is represented by the product of their matrices, and the identity map on any n-dimensional vector space V is represented with respect to any basis by the $n \times n$ *identity matrix*, which we denote by I_n; it is the matrix with ones on the main diagonal and zeros elsewhere.

The set $\mathrm{M}(m \times n, \mathbb{R})$ of all $m \times n$ real matrices is easily seen to be a real vector space of dimension mn. (In fact, by stringing out the matrix entries in a single row, we can identify it in a natural way with $\mathbb{R}^{mn}$.) Similarly, because $\mathbb{C}$ is a real vector space of dimension 2, the set $\mathrm{M}(m \times n, \mathbb{C})$ of $m \times n$ complex matrices is a real vector space of dimension $2mn$. When $m = n$, we will abbreviate the spaces of $n \times n$ square real and complex matrices by $\mathrm{M}(n, \mathbb{R})$ and $\mathrm{M}(n, \mathbb{C})$, respectively. In this case, matrix multiplication gives these spaces additional algebraic structure. If V, W, and Z are vector spaces, a map $B\colon V \times W \to Z$ is said to be *bilinear* if it is linear in each variable separately when the other is held fixed:

$$\begin{aligned} B(a_1X_1 + a_2X_2, Y) &= a_1B(X_1, Y) + a_2B(X_2, Y), \\ B(X, a_1Y_1 + a_2Y_2) &= a_1B(X, Y_1) + a_2B(X, Y_2). \end{aligned}$$

An *algebra* (over $\mathbb{R}$) is a real vector space V endowed with a bilinear product map $V \times V \to V$. The algebra is said to be commutative or associative if the bilinear product has that property.

◇ **Exercise A.39.** Show that matrix multiplication turns $\mathrm{M}(n, \mathbb{R})$ and $\mathrm{M}(n, \mathbb{C})$ into associative algebras over $\mathbb{R}$. Show that they are noncommutative unless $n = 1$.

Suppose A is an $n \times n$ matrix. If there is a matrix B such that $AB = BA = I_n$, then A is said to be *invertible* or *nonsingular*; it is *singular* otherwise.

◇ **Exercise A.40.** Suppose A is an $n \times n$ matrix.

(a) If A is nonsingular, show that there is a *unique* $n \times n$ matrix B such that $AB = BA = I_n$. This matrix is denoted by A^{-1} and is called the *inverse* of A.
(b) If A is the matrix of a linear map $T\colon V \to W$ with respect to some bases for V and W, show that T is invertible if and only if A is invertible, in which case A^{-1} is the matrix of T^{-1} with respect to the same bases.
(c) If B is an $n \times n$ matrix such that either $AB = I_n$ or $BA = I_n$, show that A is nonsingular and $B = A^{-1}$.

Because $\mathbb{R}^n$ comes equipped with the canonical basis (e_i), we can unambiguously identify linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$ with $m \times n$ matrices, and we will often do so without further comment.

In this book, we often need to be concerned with how various objects transform when we change bases. Suppose (E_i) and $(\widetilde{E}_j)$ are two bases for a finite-dimensional vector space V. Then each basis can be written uniquely in terms of the other, so there is an invertible matrix B, called the *transition matrix* between the two bases, such that

$$\begin{aligned} E_i &= \sum_{j=1}^{n} B_i^j \widetilde{E}_j, \\ \widetilde{E}_j &= \sum_{i=1}^{n} (B^{-1})_j^i E_i. \end{aligned} \tag{A.2}$$

Now suppose V and W are finite-dimensional vector spaces and $T\colon V \to W$ is a linear map. With respect to bases (E_i) for V and (F_j) for W, T is represented by some matrix $A = (A_j^i)$. If $(\widetilde{E}_i)$ and $(\widetilde{F}_j)$ are any other choices of bases for V and W, respectively, let B and C denote the transition matrices satisfying (A.2) and

$$\begin{aligned} F_i &= \sum_{j=1}^{m} C_i^j \widetilde{F}_j, \\ \widetilde{F}_j &= \sum_{i=1}^{m} (C^{-1})_j^i F_i. \end{aligned}$$

Then a straightforward computation shows that the matrix $\widetilde{A}$ representing T with respect to the new bases is related to A by

$$\widetilde{A}_j^i = \sum_{k,l} C_l^i A_k^l (B^{-1})_j^k,$$

or, in matrix form,

$$\widetilde{A} = CAB^{-1}.$$

In particular, if T is a map from V to itself, we usually use the same basis in the domain and the range. In this case, if A denotes the matrix of T with respect to (E_i), and $\widetilde{A}$ is its matrix with respect to $(\widetilde{E}_i)$, we have

$$\widetilde{A} = BAB^{-1}. \tag{A.3}$$

If V and W are vector spaces, the set $\operatorname{Hom}(V, W)$ of linear maps from V to W is a vector space under the operations

$$(S + T)X = SX + TX; \qquad (cT)X = c(TX).$$

If $\dim V = n$ and $\dim W = m$, then any choices of bases for V and W give us a map $\operatorname{Hom}(V, W) \to \mathrm{M}(m \times n, \mathbb{R})$, by sending each linear map to its matrix with respect to the chosen bases. This map is easily seen to be linear and bijective, so $\dim \operatorname{Hom}(V, W) = \dim \mathrm{M}(m \times n, \mathbb{R}) = mn$.

If $T\colon V \to W$ is a linear map between finite-dimensional spaces, the dimension of $\operatorname{Im} T$ is called the *rank* of T, and the dimension of $\operatorname{Ker} T$ is called its *nullity.* The following theorem shows that, up to choices of bases, a linear map is completely determined by its rank together with the dimensions of its domain and range.

Theorem A.33 (Canonical Form for a Linear Map). *Suppose V and W are finite-dimensional vector spaces, and $T\colon V \to W$ is a linear map of rank r. Then there are bases for V and W with respect to which T has the following matrix representation (in block form):*

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Proof. Choose any bases $(F_1, \dots, F_r)$ for $\operatorname{Im} T$ and $(K_1, \dots, K_k)$ for $\operatorname{Ker} T$. Extend (F_j) arbitrarily to a basis $(F_1, \dots, F_m)$ for W. By definition of the image, there are vectors $E_1, \dots, E_r \in V$ such that $TE_i = F_i$ for $i = 1, \dots, r$. We will show that $(E_1, \dots, E_r, K_1, \dots, K_k)$ is a basis for V; once we know this, it follows easily that T has the desired matrix representation.

Suppose first that $\sum_i a^i E_i + \sum_j b^j K_j = 0$. Applying T to this equation yields $\sum_{i=1}^r a^i F_i = 0$, which implies that all the coefficients a^i are zero. Then it follows also that all the b^js are zero because the K_js are independent. Therefore, the vectors $(E_1, \dots, E_r, K_1, \dots, K_k)$ are independent.

To show that they span V, let $X \in V$ be arbitrary. We can express $TX \in \operatorname{Im} T$ as a linear combination of $(F_1, \dots, F_r)$:

$$TX = \sum_{i=1}^r c^i F_i.$$

If we put $Y = \sum_i c^i E_i \in V$, it follows that $TY = TX$, so $Z = X - Y \in \operatorname{Ker} T$. Writing $Z = \sum_j d^j K_j$, we obtain

$$X = Y + Z = \sum_{i=1}^r c^i E_i + \sum_{j=1}^k d^j K_j,$$

so $(E_1, \dots, E_r, K_1, \dots, K_k)$ do indeed span V. □

This theorem says that any linear map can be put into a particularly nice diagonal form by appropriate choices of bases in the domain and range. However, it is important to be aware of what the theorem does *not* say: If $T\colon V \to V$ is a linear map from a finite-dimensional vector space to itself, it may not be possible to choose a *single* basis for V with respect to which the matrix of T is diagonal.

The next result is central in applications of linear algebra to smooth manifold theory; it is a corollary to the proof of the preceding theorem.

Corollary A.34 (Rank-Nullity Law). *Suppose $T: V \to W$ is a linear map between finite-dimensional vector spaces. Then*

$$\begin{aligned}\dim V &= \operatorname{rank} T + \operatorname{nullity} T \\ &= \dim(\operatorname{Im} T) + \dim(\operatorname{Ker} T).\end{aligned}$$

Proof. The preceding proof showed that V has a basis consisting of $k + r$ elements, where $k = \dim \operatorname{Ker} T$ and $r = \dim \operatorname{Im} T$. □

◇ **Exercise A.41.** Suppose V, W, X are finite-dimensional vector spaces, and let $S: V \to W$ and $T: W \to X$ be linear maps.

(a) Show that $\operatorname{rank} S \leq \dim V$, with equality if and only if S is injective.
(b) Show that $\operatorname{rank} S \leq \dim W$, with equality if and only if S is surjective.
(c) If $\dim V = \dim W$ and S is either injective or surjective, show that it is an isomorphism.
(d) Show that $\operatorname{rank}(T \circ S) \leq \operatorname{rank} S$, with equality if and only if $\operatorname{Im} S \cap \operatorname{Ker} T = \{0\}$.
(e) Show that $\operatorname{rank}(T \circ S) \leq \operatorname{rank} T$, with equality if and only if $\operatorname{Im} S + \operatorname{Ker} T = W$.
(f) If S is an isomorphism, show that $\operatorname{rank}(T \circ S) = \operatorname{rank} T$, and if T is an isomorphism, show that $\operatorname{rank}(T \circ S) = \operatorname{rank} S$.

Suppose A is an $m \times n$ matrix. The *transpose* of A is the $n \times m$ matrix A^T obtained by interchanging the rows and columns of A: $(A^T)_i^j = A_j^i$. The matrix A is said to be *symmetric* if $A = A^T$ and *skew-symmetric* if $A = -A^T$.

◇ **Exercise A.42.** If A and B are matrices of dimensions $m \times n$ and $n \times k$, respectively, show that $(AB)^T = B^T A^T$.

The *rank* of an $m \times n$ matrix A is defined to be its rank as a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$. Because the columns of A, thought of as vectors in $\mathbb{R}^m$, are the images of the standard basis vectors under this linear map, the rank of A can also be thought of as the dimension of the span of its columns, and is sometimes called its *column rank*. Analogously, we define the *row rank* of A to be the dimension of the span of its rows, thought of similarly as vectors in $\mathbb{R}^n$.

Proposition A.35. *The row rank of any matrix is equal to its column rank.*

Proof. Let A be an $m \times n$ matrix. Because the row rank of A is equal to the column rank of A^T, we must show that $\operatorname{rank} A = \operatorname{rank} A^T$.

Suppose the (column) rank of A is k. Thought of as a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$, A factors through $\operatorname{Im} A$ as follows:

$$\begin{tikzcd} \mathbb{R}^n \arrow[rr, "A"] \arrow[dr, "\widetilde{A}"'] & & \mathbb{R}^m \\ & \operatorname{Im} A, \arrow[ur, "\iota"'] & \end{tikzcd}$$

where $\widetilde{A}$ is just the map A with its range restricted to $\operatorname{Im} A$, and ι is the inclusion of $\operatorname{Im} A$ into $\mathbb{R}^m$. Choosing a basis for the k-dimensional subspace $\operatorname{Im} A$, we can write this as a matrix equation $A = BC$, where B and C are the matrices of ι and $\widetilde{A}$ with respect to the standard bases in $\mathbb{R}^n$ and $\mathbb{R}^m$ and the chosen basis in $\operatorname{Im} A$. Taking transposes, we find $A^T = C^T B^T$, from which it follows that $\operatorname{rank} A^T \leq \operatorname{rank} B^T$. Since B^T is a $k \times m$ matrix, its column rank is at most k, which shows that $\operatorname{rank} A^T \leq \operatorname{rank} A$. Reversing the roles of A and A^T and using the fact that $(A^T)^T = A$, we conclude that $\operatorname{rank} A = \operatorname{rank} A^T$. □

Suppose $A = (A^i_j)$ is an $m \times n$ matrix. By choosing nonempty subsets $\{i_1, \dots, i_k\} \subset \{1, \dots, m\}$ and $\{j_1, \dots, j_l\} \subset \{1, \dots, n\}$, we obtain a $k \times l$ matrix whose entry in the pth row and qth column is $A^{i_p}_{j_q}$:

$$\begin{pmatrix} A^{i_1}_{j_1} & \dots & A^{i_1}_{j_l} \\ \vdots & \ddots & \vdots \\ A^{i_k}_{j_1} & \dots & A^{i_k}_{j_l} \end{pmatrix}.$$

Such a matrix is called a $k \times l$ *minor* of A. Looking at minors gives a convenient criterion for checking the rank of a matrix.

Proposition A.36. *Suppose A is an $m \times n$ matrix. Then* $\operatorname{rank} A \geq k$ *if and only if some $k \times k$ minor of A is nonsingular.*

Proof. By definition, $\operatorname{rank} A \geq k$ if and only if A has at least k independent columns, which is equivalent to A having some $m \times k$ minor with rank k. But by Proposition A.35, any such $m \times k$ minor has rank k if and only if it has k independent rows. Thus A has rank at least k if and only if it has an $m \times k$ minor with k independent rows, if and only if it has a $k \times k$ minor that is nonsingular. □

The Determinant

There are a number of ways of defining the determinant of a square matrix, each of which has advantages in different contexts. The definition we will give here is the simplest to state and fits nicely with our treatment of alternating tensors in Chapter 12.

If X is a set, a *permutation* of X is a bijective map from X to itself. The set of all permutations of X is a group under composition. We let S_n denote the group of permutations of the set $\{1, \ldots, n\}$, called the *symmetric group* on n elements. The properties of S_n that we will need are summarized in the following lemma; proofs can be found in any good undergraduate algebra text such as [Hun90] or [Her75]. A *transposition* is a permutation that interchanges two elements and leaves all the others fixed. A permutation that can be written as a product of an even number of transpositions is called *even*, and one that can be written as a product of an odd number of transpositions is called *odd*.

Lemma A.37 (Properties of the Symmetric Group).

(*a*) *Every element of S_n can be written as a finite product of transpositions.*

(*b*) *For any $\sigma \in S_n$, the parity (evenness or oddness) of the number of factors in any decomposition of σ as a product of transpositions is independent of the choice of decomposition.*

(*c*) *The map* $\operatorname{sgn}\colon S_n \to \{\pm 1\}$ *given by*

$$\operatorname{sgn}\sigma = \begin{cases} 1 & \text{if } \sigma \text{ is even,} \\ -1 & \text{if } \sigma \text{ is odd} \end{cases}$$

is a surjective group homomorphism, where we consider $\{\pm 1\}$ as a group under multiplication.

◇ **Exercise A.43.** Prove (or look up) Lemma A.37.

If $A = (A^i_j)$ is an $n \times n$ (real or complex) matrix, the *determinant* of A is defined by the expression

$$\det A = \sum_{\sigma \in S_n} (\operatorname{sgn}\sigma) A_1^{\sigma(1)} \cdots A_n^{\sigma(n)}. \tag{A.4}$$

For simplicity, we assume throughout this section that our matrices are real. The statements and proofs, however, hold equally well in the complex case. In our study of Lie groups we also have occasion to consider determinants of complex matrices.

Although the determinant is defined as a function of matrices, it is also useful to think of it as a function of n vectors in $\mathbb{R}^n$: If $A_1, \ldots, A_n \in \mathbb{R}^n$, we interpret $\det(A_1, \ldots, A_n)$ to mean the determinant of the matrix whose columns are $(A_1, \ldots, A_n)$:

$$\det(A_1, \ldots, A_n) = \det \begin{pmatrix} A_1^1 & \ldots & A_n^1 \\ \vdots & \ddots & \vdots \\ A_1^n & \ldots & A_n^n \end{pmatrix}.$$

It is obvious from the defining formula (A.4) that the function $\det\colon \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$ so defined is *multilinear*, which means that it is linear as a function of each vector when all the other vectors are held fixed.

Proposition A.38 (Properties of the Determinant). *Let A be an $n \times n$ matrix.*

(*a*) *If one column of A is multiplied by a scalar c, the determinant is multiplied by the same scalar:*

$$\det(A_1, \dots, cA_i, \dots, A_n) = c \det(A_1, \dots, A_i, \dots, A_n).$$

(*b*) *The determinant changes sign when two columns are interchanged:*

$$\begin{aligned} &\det(A_1, \dots, A_q, \dots, A_p, \dots, A_n) \\ &\qquad = -\det(A_1, \dots, A_p, \dots, A_q, \dots, A_n). \end{aligned} \tag{A.5}$$

(*c*) *The determinant is unchanged by adding a scalar multiple of one column to any other column:*

$$\begin{aligned} &\det(A_1, \dots, A_i, \dots, A_j + cA_i, \dots, A_n) \\ &\qquad = \det(A_1, \dots, A_i, \dots, A_j \dots, A_n). \end{aligned}$$

(*d*) *For any scalar c,* $\det(cA) = c^n \det A$.

(*e*) *If any two columns of A are identical, then* $\det A = 0$.

(*f*) $\det A^T = \det A$.

(*g*) $\det I_n = 1$.

(*h*) *If A is singular, then* $\det A = 0$.

Proof. Part (a) is part of the definition of multilinearity, and (d) follows immediately from (a). To prove (b), suppose $p < q$ and let $\tau \in S_n$ be the transposition that interchanges p and q, leaving all other indices fixed. Then the left-hand side of (A.5) is equal to

$$\begin{aligned} &\det(A_1, \dots, A_q, \dots, A_p, \dots, A_n) \\ &\quad = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) A_1^{\sigma(1)} \cdots A_q^{\sigma(p)} \cdots A_p^{\sigma(q)} \cdots A_n^{\sigma(n)} \\ &\quad = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) A_1^{\sigma(1)} \cdots A_p^{\sigma(q)} \cdots A_q^{\sigma(p)} \cdots A_n^{\sigma(n)} \\ &\quad = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) A_1^{\sigma(\tau(1))} \cdots A_n^{\sigma(\tau(n))} \\ &\quad = -\sum_{\sigma \in S_n} (\operatorname{sgn}(\sigma\tau)) A_1^{\sigma(\tau(1))} \cdots A_n^{\sigma(\tau(n))} \\ &\quad = -\sum_{\eta \in S_n} (\operatorname{sgn} \eta) A_1^{\eta(1)} \cdots A_n^{\eta(n)} \\ &\quad = -\det(A_1, \dots, A_n), \end{aligned}$$

where the next-to-last line follows by substituting $\eta = \sigma\tau$ and noting that η runs over all elements of S_n as σ does. Part (e) is then an immediate consequence of (b), and (c) follows by multilinearity:

$$\begin{aligned}
&\det(A_1, \dots, A_i, \dots, A_j + cA_i, \dots, A_n)\\
&\quad = \det(A_1, \dots, A_i, \dots, A_j \dots, A_n) + c\det(A_1, \dots, A_i, \dots, A_i \dots, A_n)\\
&\quad = \det(A_1, \dots, A_i, \dots, A_j \dots, A_n) + 0.
\end{aligned}$$

Part (f) follows directly from the definition of the determinant:

$$\begin{aligned}
\det A^T &= \sum_{\sigma \in S_n} (\operatorname{sgn}\sigma) A^1_{\sigma(1)} \cdots A^n_{\sigma(n)}\\
&= \sum_{\sigma \in S_n} (\operatorname{sgn}\sigma) A^{\sigma^{-1}(\sigma(1))}_{\sigma(1)} \cdots A^{\sigma^{-1}(\sigma(n))}_{\sigma(n)}\\
&= \sum_{\sigma \in S_n} (\operatorname{sgn}\sigma) A^{\sigma^{-1}(1)}_{1} \cdots A^{\sigma^{-1}(n)}_{n}\\
&= \sum_{\eta \in S_n} (\operatorname{sgn}\eta) A^{\eta(1)}_{1} \cdots A^{\eta(n)}_{n}\\
&= \det A.
\end{aligned}$$

In the third line, we have used the fact that multiplication is commutative, and the numbers $\{A^{\sigma^{-1}(\sigma(1))}_{\sigma(1)}, \dots, A^{\sigma^{-1}(\sigma(n))}_{\sigma(n)}\}$ are just $\{A^{\sigma^{-1}(1)}_{1}, \dots, A^{\sigma^{-1}(n)}_{n}\}$ in a different order; and the fourth line follows by substituting $\eta = \sigma^{-1}$ and noting that $\operatorname{sgn}\sigma^{-1} = \operatorname{sgn}\sigma$. Similarly, (g) follows from the definition, because when A is the identity matrix, for each σ except the identity permutation there is some j such that $A^{\sigma(j)}_j = 0$.

Finally, to prove (h), suppose A is singular. Then, as a linear map from $\mathbb{R}^n$ to $\mathbb{R}^n$, A as rank less than n by parts (a) and (b) of Exercise A.41. Thus the columns of A are dependent, so at least one column can be written as a linear combination of the others: $A_j = \sum_{i \neq j} c^i A_i$. The result then follows from the multilinearity of det and (e). □

The operations on matrices described in parts (a), (b), and (c) of the preceding proposition (multiplying one column by a scalar, interchanging two columns, and adding a multiple of one column to another) are called *elementary column operations*. Part of the proposition, therefore, describes precisely how a determinant is affected by elementary column operations. If we define *elementary row operations* analogously, the fact that the determinant of A^T is equal to that of A implies that the determinant behaves similarly under elementary row operations.

Since the columns of an $n \times n$ matrix A are the images of the standard basis vectors under the linear map from $\mathbb{R}^n$ to itself that A defines, elementary column operations correspond to changes of basis in the domain. Thus each elementary column operation on a matrix A can be realized by multiplying A on the right by a suitable matrix, called an *elementary matrix*. For example, multiplying the ith column by c is achieved by multiplying A

by the matrix E_c that is equal to the identity matrix except for a c in the (i,i) position:

$$\begin{pmatrix} A_1^1 & \dots & A_i^1 & \dots & A_n^1 \\ \vdots & & \vdots & & \vdots \\ A_1^j & \dots & A_i^j & \dots & A_n^j \\ \vdots & & \vdots & & \vdots \\ A_1^n & \dots & A_i^n & \dots & A_n^n \end{pmatrix} \begin{pmatrix} 1 & \dots & 0 & \dots & 0 \\ & \ddots & & & \\ & & c & & \\ & & & \ddots & \\ 0 & \dots & 0 & \dots & 1 \end{pmatrix}$$
$$= \begin{pmatrix} A_1^1 & \dots & cA_i^1 & \dots & A_n^1 \\ \vdots & & \vdots & & \vdots \\ A_1^j & \dots & cA_i^j & \dots & A_n^j \\ \vdots & & \vdots & & \vdots \\ A_1^n & \dots & cA_i^n & \dots & A_n^n \end{pmatrix}.$$

Observe that $\det E_c = c \det I_n = c$.

◇ **Exercise A.44.** Show that interchanging two columns of a matrix is equivalent to multiplying on the right by a matrix whose determinant is -1, and adding a multiple of one column to another is equivalent to multiplying on the right by a matrix of determinant 1.

◇ **Exercise A.45.** Suppose A is a nonsingular $n \times n$ matrix.

(a) Show that A can be reduced to the identity I_n by a sequence of elementary column operations.
(b) Show that A is equal to a product of elementary matrices.

Elementary matrices form a key ingredient in the proof of the following theorem, which is arguably the deepest and most important property of the determinant.

Theorem A.39. *If A and B are $n \times n$ matrices, then*

$$\det(AB) = (\det A)(\det B).$$

Proof. If B is singular, then $\operatorname{rank} B < n$, which implies that $\operatorname{rank} AB < n$. Therefore both $\det B$ and $\det AB$ are zero by Proposition A.38(h). On the other hand, parts (a), (b), and (c) of Proposition A.38 combined with Exercise A.44 show that the theorem is true when B is an elementary matrix. If B is an arbitrary nonsingular matrix, then B can be written as a product of elementary matrices by Exercise A.45, and then the result follows by induction on the number of elementary matrices in such a product. □

Corollary A.40. *If A is a nonsingular $n \times n$ matrix, then* $\det(A^{-1}) = (\det A)^{-1}$.

Proof. Just note that $1 = \det I_n = \det(AA^{-1}) = (\det A)(\det A^{-1})$. □

Corollary A.41. *A square matrix is singular if and only if its determinant is zero.*

Proof. One direction follows from Proposition A.38(h); the other from Corollary A.40. □

Corollary A.42. *If A and B are $n \times n$ matrices with B nonsingular, then* $\det(BAB^{-1}) = \det A$.

Proof. This is just a computation using Theorem A.39 and Corollary A.40:

$$\begin{aligned} \det(BAB^{-1}) &= (\det B)(\det A)(\det B^{-1}) \\ &= (\det B)(\det A)(\det B)^{-1} \\ &= \det A. \end{aligned}$$ □

The last corollary allows us to extend the definition of the determinant to linear maps on arbitrary finite-dimensional vector spaces. Suppose V is an n-dimensional vector space and $T\colon V \to V$ is a linear map. With respect to any choice of basis for V, T is represented by an $n \times n$ matrix. As we observed above, the matrices A and $\widetilde{A}$ representing T with respect to two different bases are related by $\widetilde{A} = BAB^{-1}$ for some nonsingular matrix B (see (A.3)). It follows from Corollary A.42, therefore, that $\det \widetilde{A} = \det A$. Thus we can make the following definition: For any linear map $T\colon V \to V$ from a finite-dimensional vector space to itself, we define the *determinant of T* to be the determinant of any matrix representation of T (using the same basis for the domain and range).

For actual computations of determinants, the formula in the following proposition is usually more useful than the definition.

Proposition A.43 (Expansion by Minors). *Let A be an $n \times n$ matrix, and for each i, j let M_i^j denote the $(n-1) \times (n-1)$ minor obtained by deleting the ith column and jth row of A. For any fixed i between 1 and n inclusive,*

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} A_i^j \det M_i^j. \tag{A.6}$$

Proof. It is useful to consider first a special case: Suppose A is an $n \times n$ matrix that has the block form

$$A = \begin{pmatrix} B & 0 \\ C & 1 \end{pmatrix}, \tag{A.7}$$

where B is an $(n-1) \times (n-1)$ matrix and C is a $1 \times n$ row matrix. Then in the defining formula (A.4) for $\det A$, the factor $A_n^{\sigma(n)}$ is equal to 1 when $\sigma(n) = n$ and zero otherwise, so in fact the only terms that are nonzero are those in which $\sigma \in S_{n-1}$, thought of as the subgroup of S_n consisting of elements that permute $\{1, \ldots, n-1\}$ and leave n fixed. Thus the determinant of A simplifies to

$$\det A = \sum_{\sigma \in S_{n-1}} (\operatorname{sgn} \sigma) A_1^{\sigma(1)} \cdots A_{n-1}^{\sigma(n-1)} = \det B.$$

Now let A be arbitrary, and fix $i \in \{1, \dots, n\}$. For each $j = 1, \dots, n$, let X_i^j denote the matrix obtained by replacing the ith column of A by the basis vector e_j. Since the determinant is a multilinear function of its columns,

$$\begin{aligned} \det A &= \det \left(A_1, \dots, A_{i-1}, \sum_{j=1}^n A_i^j e_j, A_{i+1}, \dots, A_n \right) \\ &= \sum_{j=1}^n A_i^j \det(A_1, \dots, A_{i-1}, e_j, A_{i+1}, \dots, A_n) \qquad \text{(A.8)} \\ &= \sum_{j=1}^n A_i^j \det X_i^j. \end{aligned}$$

On the other hand, by interchanging columns $n - i$ times and then interchanging rows $n - j$ times, we can transform X_i^j to a matrix of the form (A.7) with $B = M_i^j$. Therefore, by the observation in the preceding paragraph,

$$\det X_i^j = (-1)^{n-i+n-j} \det M_i^j = (-1)^{i+j} \det M_i^j.$$

Inserting this into (A.8) completes the proof. □

Formula (A.6) is called the *expansion of* $\det A$ *by minors along the* i*th column*. Since $\det A = \det A^T$, there is an analogous expansion along any row. The factor $(-1)^{i+j} \det M_i^j$ multiplying A_i^j in (A.6) is called the *cofactor* of A_i^j, and is denoted by cof_i^j.

Proposition A.44 (Cramer's Rule). *If A is a nonsingular $n \times n$ matrix, then A^{-1} is equal to $1/(\det A)$ times the transposed cofactor matrix of A. Thus its matrix entries are given by*

$$(A^{-1})_j^i = \frac{1}{\det A} \operatorname{cof}_i^j = \frac{1}{\det A} (-1)^{i+j} \det M_i^j. \qquad \text{(A.9)}$$

Proof. Let B_j^i denote the expression on the right-hand side of (A.9). Then

$$\sum_{j=1}^n B_j^i A_k^j = \frac{1}{\det A} \sum_{j=1}^n (-1)^{i+j} A_k^j \det M_i^j. \qquad \text{(A.10)}$$

When $k = i$, the summation on the right-hand side is precisely the expansion of $\det A$ by minors along the ith column, so the right-hand side of (A.10) is equal to 1. On the other hand, if $k \neq i$, the summation is equal to the determinant of the matrix obtained by replacing the ith column of A

by the kth column. Since this matrix has two identical columns, its determinant is zero. Thus (A.10) is equivalent to the matrix equation $BA = I_n$, where B is the matrix (B^i_j). By Exercise A.40(c), therefore, $B = A^{-1}$. □

A square matrix $A = (A^i_j)$ is said to be *upper triangular* if $A^i_j = 0$ for $i > j$ (i.e., the only nonzero entries are on and above the main diagonal). Determinants of upper triangular matrices are particularly easy to compute.

Proposition A.45. *If A is an upper triangular $n \times n$ matrix, then the determinant of A is the product of its diagonal entries:*

$$\det A = A^1_1 \cdots A^n_n.$$

Proof. When $n = 1$, this is trivial. So assume the result is true for $(n-1) \times (n-1)$ matrices, and let A be an upper triangular $n \times n$ matrix. In the expansion of $\det A$ by minors along the first column, there is only one nonzero entry, namely $A^1_1 \det M^1_1$. By induction, $\det M^1_1 = A^2_2 \cdots A^n_n$, which proves the proposition. □

Suppose X is an $(m+k) \times (m+k)$ matrix. We say X is *block upper triangular* if X has the form

$$X = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} \tag{A.11}$$

for some matrices A, B, C of sizes $m \times m$, $m \times k$, and $k \times k$, respectively.

Proposition A.46. *If X is the block upper triangular matrix given by* (A.11)*, then* $\det X = (\det A)(\det C)$.

Proof. If A is singular, then clearly the columns of X are linearly dependent, which implies that $\det X = 0 = (\det A)(\det C)$. So let us assume that A is nonsingular.

Consider first the following special case:

$$X = \begin{pmatrix} I_m & 0 \\ 0 & C \end{pmatrix}.$$

Expanding by minors along the first column, an easy induction shows that $\det X = \det C$ in this case. A similar argument shows that

$$\det \begin{pmatrix} A & 0 \\ 0 & I_k \end{pmatrix} = \det A.$$

The general case follows from these two observations together with the factorization

$$\begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & I_k \end{pmatrix} \begin{pmatrix} I_m & 0 \\ 0 & C \end{pmatrix} \begin{pmatrix} I_m & A^{-1}B \\ 0 & I_k \end{pmatrix}.$$ □

Inner Products and Norms

If V is a real vector space, an *inner product* on V is a map $V \times V \to \mathbb{R}$, usually written $(X, Y) \mapsto \langle X, Y \rangle$, that is

(i) Symmetric:

$$\langle X, Y \rangle = \langle Y, X \rangle;$$

(ii) Bilinear:

$$\langle aX + a'X', Y \rangle = a\langle X, Y \rangle + a'\langle X', Y \rangle,$$
$$\langle X, bY + b'Y' \rangle = b\langle X, Y \rangle + b'\langle X, Y' \rangle;$$

(iii) Positive definite:

$$\langle X, X \rangle > 0 \text{ unless } X = 0.$$

A vector space endowed with a specific inner product is called an *inner product space*. The standard example is, of course, $\mathbb{R}^n$ with its *dot product* or *Euclidean inner product*:

$$\langle x, y \rangle = x \cdot y = \sum_{i=1}^{n} x^i y^i.$$

Suppose V is an inner product space. For any $X \in V$, the *length* of X is the number $|X| = \sqrt{\langle X, X \rangle}$. A *unit vector* is a vector of length 1. The *angle* between two nonzero vectors $X, Y \in V$ is defined to be the unique $\theta \in [0, \pi]$ satisfying

$$\cos\theta = \frac{\langle X, Y \rangle}{|X|\,|Y|}.$$

Two vectors $X, Y \in V$ are said to be *orthogonal* $\langle X, Y \rangle = 0$. If both are nonzero, this is equivalent to the angle between them being $\pi/2$.

◇ **Exercise A.46.** Let V be an inner product space. Show that the length function associated with the inner product satisfies

$$\begin{aligned} |X| &> 0, && X \in V,\ X \neq 0,\\ |cX| &= |c|\,|X|, && c \in \mathbb{R},\ X \in V,\\ |X + Y| &\le |X| + |Y|, && X, Y \in V, \end{aligned}$$

and the *Schwartz inequality*:

$$|\langle X, Y \rangle| \le |X|\ |Y| \qquad X, Y \in V.$$

Suppose V is a finite-dimensional inner product space. A basis $(E_1, \dots, E_n)$ for V is said to be *orthonormal* if each E_i is a unit vector and E_i is orthogonal to E_j when $i \neq j$.

Proposition A.47 (The Gram-Schmidt Algorithm). *Every finite-dimensional inner product space V has an orthonormal basis. In fact, if*

$(E_1, \dots, E_n)$ is any basis for V, there is an orthonormal basis $(\widetilde{E}_1, \dots, \widetilde{E}_n)$ with the property that

$$\operatorname{span}(E_1, \dots, E_k) = \operatorname{span}(\widetilde{E}_1, \dots, \widetilde{E}_k) \text{ for } k = 1, \dots, n. \tag{A.12}$$

Proof. The proof is by induction on $n = \dim V$. If $n = 1$, there is only one basis element E_1, and then $\widetilde{E}_1 = E_1/|E_1|$ is an orthonormal basis.

Suppose the result is true for inner product spaces of dimension $n-1$, and let V have dimension n. Then $W = \operatorname{span}(E_1, \dots, E_{n-1})$ is an $(n-1)$-dimensional inner product space with the inner product restricted from V, so there is an orthonormal basis $(\widetilde{E}_1, \dots, \widetilde{E}_{n-1})$ satisfying (A.12) for $k = 1, \dots, n-1$. Define $\widetilde{E}_n$ by

$$\widetilde{E}_n = \frac{E_n - \sum_{i=1}^{n-1} \langle E_n, \widetilde{E}_i \rangle \widetilde{E}_i}{\left| E_n - \sum_{i=1}^{n-1} \langle E_n, \widetilde{E}_i \rangle \widetilde{E}_i \right|}.$$

A computation shows that $(\widetilde{E}_1, \dots, \widetilde{E}_n)$ is the desired orthonormal basis for V. □

◇ **Exercise A.47.** For $z, w \in \mathbb{C}^n$, define the *Hermitian dot product* by $z \cdot w = \sum_{i=1}^n z^i \overline{w^i}$. A basis $(E_1, \dots, E_n)$ for $\mathbb{C}^n$ (over $\mathbb{C}$) is said to be orthonormal if $E_i \cdot E_i = 1$ and $E_i \cdot E_j = 0$ for $i \neq j$. Show that the statement and proof of Proposition A.47 hold for the Hermitian dot product.

An isomorphism $T: V \to W$ between inner product spaces is called an *isometry* if it takes the inner product of V to that of W:

$$\langle TX, TY \rangle = \langle X, Y \rangle.$$

◇ **Exercise A.48.** Show that any isometry is a homeomorphism that preserves lengths, angles, and orthogonality, and takes orthonormal bases to orthonormal bases.

◇ **Exercise A.49.** If (E_i) is any basis for a finite-dimensional vector space V, show that there is a unique inner product on V for which (E_i) is orthonormal.

◇ **Exercise A.50.** Suppose V is a finite-dimensional inner product space and $E: \mathbb{R}^n \to V$ is the basis map determined by any orthonormal basis. If $\mathbb{R}^n$ is endowed with the Euclidean inner product, show that E is an isometry.

The preceding exercise shows that finite-dimensional inner product spaces are geometrically indistinguishable from the Euclidean space of the same dimension.

If V is a finite-dimensional inner product space and $S \subset V$ is a subspace, the *orthogonal complement* of S in V is the set

$$S^\perp = \{X \in V : \langle X, Y \rangle = 0 \text{ for all } Y \in S\}.$$

◇ **Exercise A.51.** Let V be a finite-dimensional inner product space and let $S \subset V$ be any subspace. Show that $S^\perp$ is a subspace and $V = S \oplus S^\perp$.

Thanks to the result of the preceding exercise, for any subspace S of an inner product space V, there is a natural projection $\pi\colon V \to S$ with kernel $S^\perp$. This is called the *orthogonal projection* of V onto S.

A *norm* on a vector space V is a function from V to $\mathbb{R}$, written $X \mapsto |X|$, satisfying

(i) POSITIVITY: $|X| \geq 0$ for every $X \in V$, and $|X| = 0$ if and only if $X = 0$.

(ii) HOMOGENEITY: $|cX| = |c|\,|X|$ for every $c \in \mathbb{R}$ and $X \in V$.

(iii) TRIANGLE INEQUALITY: $|X + Y| \leq |X| + |Y|$ for all $X, Y \in V$.

A vector space together with a specific choice of norm is called a *normed linear space*. Exercise A.46 shows that the length function associated with any inner product is a norm; thus, in particular, every finite-dimensional vector space possesses many norms.

Given a norm on V, the distance function $d(X, Y) = |X - Y|$ turns V into a metric space, yielding a topology on V called the *norm topology*. The set of all open balls in V is easily seen to be a basis for this topology. Two norms $|\cdot|_1$ and $|\cdot|_2$ are said to be *equivalent* if there are positive constants c, C such that

$$c|X|_1 \leq |X|_2 \leq C|X|_1 \text{ for all } X \in V.$$

◇ **Exercise A.52.** Show that equivalent norms on a vector space V determine the same topology.

◇ **Exercise A.53.** Show that any two norms on a finite-dimensional vector space are equivalent. [Hint: Choose an inner product on V, and show that the closed unit ball in any norm is compact with respect to the topology determined by the inner product.]

The preceding exercise shows that finite-dimensional normed linear spaces of the same dimension are topologically indistinguishable from each other. Thus any such space automatically inherits all the usual topological properties of Euclidean space, such as compactness of closed and bounded subsets.

If V and W are normed linear spaces, a linear map $T\colon V \to W$ is said to be *bounded* if there exists a positive constant C such that

$$|TX| \leq C|X| \text{ for all } X \in V.$$

◇ **Exercise A.54.** Show that a linear map between normed linear spaces is bounded if and only if it is continuous.

◇ **Exercise A.55.** Show that every linear map between finite-dimensional normed linear spaces is bounded and therefore continuous.

The vector space $\mathrm{M}(m \times n, \mathbb{R})$ of $m \times n$ real matrices has a natural Euclidean inner product, obtained by identifying a matrix with a point in $\mathbb{R}^{mn}$:

$$A \cdot B = \sum_{i,j} A^i_j B^i_j.$$

This yields a Euclidean norm on matrices:

$$|A| = \sqrt{\textstyle\sum_{i,j} (A^i_j)^2}. \tag{A.13}$$

Whenever we use a norm on a space of matrices, it will always be assumed to be this Euclidean norm.

◇ **Exercise A.56.** For any matrices $A \in \mathrm{M}(m \times n, \mathbb{R})$ and $B \in \mathrm{M}(n \times k, \mathbb{R})$, show that

$$|AB| \le |A|\ |B|\,.$$

Direct Products and Direct Sums

If $V_1, \dots, V_k$ are real vector spaces, their *direct product* is the vector space whose underlying set is the Cartesian product $V_1 \times \dots \times V_k$, with addition and scalar multiplication defined componentwise:

$$\begin{aligned}(X_1, \dots, X_k) + (X'_1, \dots, X'_k) &= (X_1 + X'_1, \dots, X_k + X'_k),\\ c(X_1, \dots, X_k) &= (cX_1, \dots, cX_k).\end{aligned}$$

This space is also sometimes called the *direct sum* of $V_1, \dots, V_k$, and denoted by $V_1 \oplus \dots \oplus V_k$, for reasons that will become clearer below. The basic example, of course, is the Euclidean space $\mathbb{R}^n = \mathbb{R} \times \dots \times \mathbb{R}$.

For some applications (chiefly in our treatment of de Rham cohomology in Chapters 15 and 16), it is important to generalize this to an infinite number of vector spaces. For this discussion, we return to the general setting of modules over a commutative ring $\mathcal{R}$. Linear maps between $\mathcal{R}$-modules are defined exactly as for vector spaces: If V and W are $\mathcal{R}$-modules, a map $F \colon V \to W$ is said to be $\mathcal{R}$-*linear* if $F(aX + bY) = aF(X) + bF(Y)$ for all $a, b \in \mathcal{R}$ and $X, Y \in V$. Throughout the rest of this section, we will assume that $\mathcal{R}$ is a fixed commutative ring. In all of our applications, $\mathcal{R}$ will be either the field $\mathbb{R}$ of real numbers, in which case the modules are real vector spaces and the linear maps are the usual ones, or the ring of integers $\mathbb{Z}$, in which case the modules are abelian groups and the linear maps are group homomorphisms.

If $\{V_\alpha\}_{\alpha \in A}$ is an arbitrary indexed family of sets, their Cartesian product, denoted by $\prod_{\alpha \in A} V_\alpha$, is defined as the set of functions $X \colon A \to \bigcup_{\alpha \in A} V_\alpha$

with the property that $X(\alpha) \in V_\alpha$ for each α. Thanks to the axiom of choice, the Cartesian product of a nonempty indexed family of nonempty sets is nonempty. If X is an element of the Cartesian product, we usually denote the value of X at $\alpha \in A$ by X_α instead of $X(\alpha)$; the element X itself is usually denoted by $(X_\alpha)_{\alpha\in A}$, or just (X_α) if the index set is understood. This can be thought of as an indexed family of elements of the sets V_α, or an "A-tuple." For each $\beta \in A$, we have a canonical *projection map* $\pi_\beta \colon \prod_{\alpha\in A} V_\alpha \to V_\beta$, defined by

$$\pi_\beta\left((X_\alpha)_{\alpha\in A}\right) = X_\beta.$$

Now suppose that $\{V_\alpha\}_{\alpha\in A}$ is an indexed family of $\mathcal{R}$-modules. The *direct product* of the family is the set $\prod_{\alpha\in A} V_\alpha$, made into an $\mathcal{R}$-module by defining addition and scalar multiplication as follows:

$$\begin{aligned}(X_\alpha) + (X'_\alpha) &= (X_\alpha + X'_\alpha),\\ c(X_\alpha) &= (cX_\alpha).\end{aligned}$$

The zero element of this module is the A-tuple with $X_\alpha = 0$ for every α. It is easy to check that each projection map π_β is $\mathcal{R}$-linear.

Lemma A.48 (Characteristic Property of the Direct Product). *Let $\{V_\alpha\}_{\alpha\in A}$ be an indexed family of $\mathcal{R}$-modules. Given an $\mathcal{R}$-module W and a family of $\mathcal{R}$-linear maps $G_\alpha \colon W \to V_\alpha$, there exists a unique $\mathcal{R}$-linear map $G \colon W \to \prod_{\alpha\in A} V_\alpha$ such that $\pi_\alpha \circ G = G_\alpha$ for each $\alpha \in A$.*

◇ **Exercise A.57.** Prove the preceding lemma.

Dual to direct products is the notion of direct sums. Given an indexed family $\{V_\alpha\}_{\alpha\in A}$ as above, we define the *direct sum* of the family to be the subset of their direct product consisting of A-tuples $(X_\alpha)_{\alpha\in A}$ with the property that $X_\alpha = 0$ for all but finitely many α. This is easily seen to be an $\mathcal{R}$-module with the operations of addition and scalar multiplication inherited from the direct product. The direct sum is denoted by $\bigoplus_{\alpha\in A} V_\alpha$, or in the case of a finite family by $V_1 \oplus \cdots \oplus V_k$. For finite families, the direct product and the direct sum are identical.

For each $\beta \in A$, there is a canonical $\mathcal{R}$-linear injection $\iota_\beta \colon V_\beta \to \bigoplus_{\alpha\in A} V_\alpha$, defined by letting $\iota_\beta(X)$ be the A-tuple $(X_\alpha)_{\alpha\in A}$ with $X_\beta = X$ and $X_\alpha = 0$ for $\alpha \neq \beta$. In the case of a finite direct sum, this just means $\iota_\beta(X) = (0, \dots, 0, X, 0, \dots, 0)$, with the only nonzero entry in position β. Typically, we identify V_β with its image under ι_β, and thus consider each V_β as a submodule of the direct sum. In particular, in a direct sum $V = V_1 \oplus V_2$ of two vector spaces, V_1 and V_2 are subspaces that span V and whose intersection is trivial, so our earlier definition of direct sums can be viewed as a special case of this one.

Lemma A.49 (Characteristic Property of the Direct Sum). *Let $\{V_\alpha\}_{\alpha\in A}$ be an indexed family of $\mathcal{R}$-modules. Given an $\mathcal{R}$-module W and a* *family of $\mathcal{R}$-linear maps $G_\alpha\colon V_\alpha \to W$, there exists a unique $\mathcal{R}$-linear map $G\colon \bigoplus_{\alpha\in A} V_\alpha \to W$ such that $G \circ \iota_\alpha = G_\alpha$ for each $\alpha \in A$.*

◇ **Exercise A.58.** Prove the preceding lemma.

If V and W are $\mathcal{R}$-modules, the set $\operatorname{Hom}(V, W)$ of all $\mathcal{R}$-linear maps from V to W is an $\mathcal{R}$-module under pointwise addition and scalar multiplication:

$$\begin{aligned}(F+G)(X) &= F(X) + G(X),\\ (aF)(X) &= aF(X).\end{aligned}$$

Lemma A.50. *Let $\{V_\alpha\}_{\alpha\in A}$ be an indexed family of $\mathcal{R}$-modules. For any $\mathcal{R}$-module W, there is a canonical isomorphism*

$$\operatorname{Hom}\left(\bigoplus_{\alpha\in A} V_\alpha, W\right) \cong \prod_{\alpha\in A} \operatorname{Hom}(V_\alpha, W).$$

Proof. Define a map $\Phi\colon \operatorname{Hom}\left(\bigoplus_{\alpha\in A} V_\alpha, W\right) \to \prod_{\alpha\in A}\operatorname{Hom}(V_\alpha, W)$ by setting $\Phi(F) = (F_\alpha)_{\alpha\in A}$, where $F_\alpha = F \circ \iota_\alpha$.

To prove that Φ is surjective, let $(F_\alpha)_{\alpha\in A}$ be an arbitrary element of $\prod_{\alpha\in A}\operatorname{Hom}(V_\alpha, W)$. This just means that for each α, F_α is an $\mathcal{R}$-linear map from V_α to W. The characteristic property of the direct sum then guarantees the existence of an $\mathcal{R}$-linear map $F\colon \bigoplus_{\alpha\in A} V_\alpha \to W$ satisfying $F \circ \iota_\alpha = F_\alpha$ for each α, which is equivalent to $\Phi(F) = (F_\alpha)_{\alpha\in A}$.

To prove that Φ is injective, suppose that $\Phi(F) = (F_\alpha)_{\alpha\in A} = 0$. By definition of the zero element of the direct product, this means that $F_\alpha = F \circ \iota_\alpha$ is the zero homomorphism for each α. By the uniqueness assertion in Lemma A.49, this implies that F itself is the zero homomorphism. □

Calculus

In this section, we summarize the main results from multivariable calculus and real analysis that are needed in this book. For details on most of the ideas touched on here, you can consult [Rud76] or [Apo74].

The Total Derivative

For maps between (open subsets of) finite-dimensional vector spaces, the most general notion of derivative is the total derivative.

Let V, W be finite-dimensional vector spaces, which we may assume to be endowed with norms. If $U \subset V$ is an open set, a map $F\colon U \to W$ is said to be *differentiable* at $a \in U$ if there exists a linear map $L\colon V \to W$ such that

$$\lim_{v\to 0} \frac{|F(a+v) - F(a) - Lv|}{|v|} = 0. \tag{A.14}$$

The norm in the numerator of this expression is that of W, while the norm in the denominator is that of V. However, because all norms on a finite-dimensional vector space are equivalent, this definition is independent of both choices of norms.

◇ **Exercise A.59.** Suppose $F\colon U \to W$ is differentiable at $a \in U$. Show that the linear map L satisfying (A.14) is unique.

If F is differentiable at a, the linear map L satisfying (A.14) is denoted by $DF(a)$ and is called the *total derivative* of F at a. Condition (A.14) can also be written

$$F(a+v) = F(a) + DF(a)v + R(v), \tag{A.15}$$

where the remainder term $R(v)$ satisfies $|R(v)|/|v| \to 0$ as $v \to 0$. Thus the total derivative represents the "best linear approximation" to $F(a+v) - F(a)$ near a. One thing that makes the total derivative so powerful is that it makes sense for arbitrary finite-dimensional vector spaces, without the need to choose a basis or even a norm.

◇ **Exercise A.60.** Suppose V, W are finite-dimensional vector spaces, $U \subset V$ is an open set, a is a point in U, and $F, G\colon U \to W$ and $f, g\colon U \to \mathbb{R}$ are maps.

(a) If F is differentiable at a, show that F is continuous at a.
(b) If F and G are differentiable at a, show that $F+G$ is also, and

$$D(F+G)(a) = DF(a) + DG(a).$$

(c) If f and g are differentiable at $a \in U$, show that fg is also, and

$$D(fg)(a) = f(a)Dg(a) + g(a)Df(a).$$

(d) If f is differentiable at a and $f(a) \neq 0$, show that $1/f$ is differentiable at a, and

$$D(1/f)(a) = -(1/f(a)^2)Df(a).$$

(e) If $F\colon V \to W$ is a linear map, show that F is differentiable at every point $a \in V$, with total derivative equal to F itself: $DF(a) = F$.

Proposition A.51 (The Chain Rule for Total Derivatives). *Suppose V, W, X are finite-dimensional vector spaces, $U \subset V$ and $\widetilde{U} \subset W$ are open sets, and $F\colon U \to \widetilde{U}$ and $G\colon \widetilde{U} \to X$ are maps. If F is differentiable at $a \in U$ and G is differentiable at $F(a) \in \widetilde{U}$, then $G \circ F$ is differentiable at a, and*

$$D(G \circ F)(a) = DG(F(a)) \circ DF(a).$$

Proof. Let $A = DF(a)$ and $B = DG(F(a))$. We need to show that

$$\lim_{v \to 0} \frac{|G(F(a+v)) - G(F(a)) - BAv|}{|v|} = 0. \tag{A.16}$$

Let us write $b = F(a)$ and $w = F(a+v) - F(a)$. With these substitutions, we can rewrite the quotient in (A.16) as

$$\begin{aligned} \frac{|G(b+w) - G(b) - BAv|}{|v|} &= \frac{|G(b+w) - G(b) - Bw + Bw - BAv|}{|v|} \\ &\leq \frac{|G(b+w) - G(b) - Bw|}{|v|} + \frac{|B(w - Av)|}{|v|}. \end{aligned} \tag{A.17}$$

Since A and B are linear, Exercise A.55 shows that there are constants C, C' such that $|Ax| \leq C|x|$ for all $x \in V$, and $|By| \leq C'|y|$ for all $y \in W$. The differentiability of F at a means that for any $\varepsilon > 0$, we can ensure that

$$|w - Av| = |F(a+v) - F(a) - Av| \leq \varepsilon |v|$$

as long as v lies in a small enough neighborhood of 0. Moreover, as $v \to 0$, $|w| = |F(a+v) - F(a)| \to 0$ by continuity of F. Therefore, the differentiability of G at b means that by making $|v|$ even smaller if necessary, we can also achieve

$$|G(b+w) - G(b) - Bw| \leq \varepsilon |w|.$$

Putting all of these estimates together, we see that for $|v|$ small, (A.17) is bounded by

$$\begin{aligned} \varepsilon \frac{|w|}{|v|} + C' \frac{|w - Av|}{|v|} &= \varepsilon \frac{|w - Av + Av|}{|v|} + C' \frac{|w - Av|}{|v|} \\ &\leq \varepsilon \frac{|w - Av|}{|v|} + \varepsilon \frac{|Av|}{|v|} + C' \frac{|w - Av|}{|v|} \\ &\leq \varepsilon^2 + \varepsilon C + C' \varepsilon, \end{aligned}$$

which can be made as small as desired. □

Partial Derivatives

Now we specialize to maps between Euclidean spaces. Suppose $U \subset \mathbb{R}^n$ is open and $f \colon U \to \mathbb{R}$ is a real-valued function. For any $a = (a^1, \ldots, a^n) \in U$ and any $j = 1, \ldots, n$, the jth *partial derivative* of f at a is defined to be the ordinary derivative of f with respect to x^j while holding the other variables fixed:

$$\begin{aligned} \frac{\partial f}{\partial x^j}(a) &= \lim_{h \to 0} \frac{f(a^1, \ldots, a^j + h, \ldots, a^n) - f(a^1, \ldots, a^j, \ldots, a^n)}{h} \\ &= \lim_{h \to 0} \frac{f(a + he_j) - f(a)}{h}, \end{aligned}$$

if the limit exists.

More generally, for a vector-valued function $F \colon U \to \mathbb{R}^m$, we write the coordinates of $F(x)$ as $F(x) = (F^1(x), \ldots, F^m(x))$; this defines m functions $F^1, \ldots, F^m \colon U \to \mathbb{R}$ called the *component functions* of F. The partial derivatives of F are defined simply to be the partial derivatives $\partial F^i/\partial x^j$ of its component functions. The matrix $(\partial F^i/\partial x^j)$ of partial derivatives is called the *Jacobian matrix* of F.

If $F\colon U \to \mathbb{R}^m$ is a map for which each partial derivative exists at each point in U and the functions $\partial F^i/\partial x^j : U \to \mathbb{R}$ so defined are all continuous, then F is said to be of *class* C^1 or *continuously differentiable*. If this is the case, we can differentiate the functions $\partial F^i/\partial x^j$ to obtain *second-order partial derivatives*

$$\frac{\partial^2 F^i}{\partial x^k \partial x^j} = \frac{\partial}{\partial x^k}\left(\frac{\partial F^i}{\partial x^j}\right),$$

if they exist. Continuing this way leads to higher-order partial derivatives—the partial derivatives of F of order k are the (first) partial derivatives of those of order $k-1$, when they exist.

In general, for $k \geq 0$, a function $F\colon U \to \mathbb{R}^m$ is said to be of *class* C^k or k *times continuously differentiable* if all the partial derivatives of F of order less than or equal to k exist and are continuous functions on U. (Thus a function of class C^0 is just a continuous function.) A function that is of class C^k for every $k \geq 0$ is said to be of *class* C^∞, *smooth*, or *infinitely differentiable*. Because existence and continuity of derivatives are local properties, clearly F is C^k or smooth if and only if it has that property in a neighborhood of each point in U.

We will be especially concerned with real-valued functions, that is, functions whose range is $\mathbb{R}$. If $U \subset \mathbb{R}^n$ is open, the set of all real-valued functions of class C^k on U is denoted by $C^k(U)$, and the set of all smooth real-valued functions by $C^\infty(U)$. Sums, constant multiples, and products of functions are defined pointwise: For $f, g\colon U \to \mathbb{R}$ and $c \in \mathbb{R}$,

$$\begin{aligned}(f+g)(x) &= f(x) + g(x),\\ (cf)(x) &= c(f(x)),\\ (fg)(x) &= f(x)g(x).\end{aligned}$$

◇ **Exercise A.61.** Let $U \subset \mathbb{R}^n$ be an open set, and suppose $f, g \in C^\infty(U)$ and $c \in \mathbb{R}$.

(a) Show that $f+g$, cf, and fg are smooth.
(b) Show that these operations turn $C^\infty(U)$ into a commutative ring and a commutative and associative algebra over $\mathbb{R}$ (see page 564).
(c) If g never vanishes on U, show that f/g is smooth.

The following important result shows that, for most interesting functions, the order in which we take partial derivatives is irrelevant. For a proof, see [Rud76].

Proposition A.52 (Equality of Mixed Partial Derivatives). *If U is an open subset of $\mathbb{R}^n$ and $F\colon U \to \mathbb{R}^m$ is a function of class C^2, then the mixed second-order partial derivatives of F do not depend on the order of* *differentiation:*

$$\frac{\partial^2 F^i}{\partial x^j \partial x^k} = \frac{\partial^2 F^i}{\partial x^k \partial x^j}.$$

Corollary A.53. *If $F\colon U \to \mathbb{R}^m$ is smooth, then the mixed partial derivatives of f of any order are independent of the order of differentiation.*

Next we study the relationship between total and partial derivatives. Suppose $U \subset \mathbb{R}^n$ is open and $F\colon U \to \mathbb{R}^m$ is differentiable at $a \in U$. As a linear map between Euclidean spaces $\mathbb{R}^n$ and $\mathbb{R}^m$, $DF(a)$ can be identified with an $m \times n$ matrix. The next lemma identifies that matrix as the Jacobian of F.

Lemma A.54. *Let $U \subset \mathbb{R}^n$ be open, and suppose $F\colon U \to \mathbb{R}^m$ is differentiable at $a \in U$. Then all of the partial derivatives of F at a exist, and $DF(a)$ is the linear map whose matrix is the Jacobian of F at a:*

$$DF(a) = \left(\frac{\partial F^j}{\partial x^i}(a)\right).$$

Proof. Let $B = DF(a)$. The fact that F is differentiable at a implies that each component of the vector-valued function $(F(a+v) - F(a) - Bv)/|v|$ goes to zero as $v \to 0$. Applying this to the jth component with $v = te_i$, we obtain

$$0 = \lim_{t \to 0} \frac{F^j(a+te_i) - F^j(a) - tB_i^j}{|t|}.$$

Considering $t > 0$ and $t < 0$ separately, we find

$$\begin{aligned}
0 &= \lim_{t \searrow 0} \frac{F^j(a+te_i) - F^j(a) - tB_i^j}{t} \\
&= \lim_{t \searrow 0} \frac{F^j(a+te_i) - F^j(a)}{t} - B_i^j. \\
0 &= -\lim_{t \nearrow 0} \frac{F^j(a+te_i) - F^j(a) - tB_i^j}{t} \\
&= -\left(\lim_{t \nearrow 0} \frac{F^j(a+te_i) - F^j(a)}{t} - B_i^j\right).
\end{aligned}$$

Combining these results, we obtain $\partial F^j/\partial x^i(a) = B_i^j$ as claimed. $\square$

◇ **Exercise A.62.** Suppose $U \subset \mathbb{R}^n$ is open. Show that a map $F\colon U \to \mathbb{R}^m$ is differentiable at $a \in U$ if and only if each of its component functions $F^1, \dots, F^m$ is differentiable at a, and

$$DF(a) = \begin{pmatrix} DF^1(a) \\ \vdots \\ DF^m(a) \end{pmatrix}.$$

In particular, a map $\gamma\colon (c,d) \to \mathbb{R}^m$ is differentiable if and only if its component functions are differentiable in the sense of one-variable calculus.

The next proposition gives the most important sufficient condition for differentiability; in particular, it shows that all of the usual functions of elementary calculus are differentiable. For a proof, see [Rud76].

Proposition A.55. *Let $U \subset \mathbb{R}^n$ be open. If $F\colon U \to \mathbb{R}^m$ is of class C^1, then it is differentiable at each point of U.*

For maps between Euclidean spaces, the chain rule can be rephrased in terms of partial derivatives.

Corollary A.56 (The Chain Rule for Partial Derivatives). *Let $U \subset \mathbb{R}^n$ and $\widetilde{U} \subset \mathbb{R}^m$ be open sets, and let $x = (x^1, \dots, x^n)$ denote the coordinates on U and $y = (y^1, \dots, y^m)$ those on $\widetilde{U}$.*

(*a*) *Any composition of C^1 functions $F\colon U \to \widetilde{U}$ and $G\colon \widetilde{U} \to \mathbb{R}^p$ is again of class C^1, with partial derivatives given by*

$$\frac{\partial (G^i \circ F)}{\partial x^j}(x) = \sum_{k=1}^{m} \frac{\partial G^i}{\partial y^k}(F(x)) \frac{\partial F^k}{\partial x^j}(x).$$

(*b*) *If F and G are smooth, then $G \circ F$ is smooth.*

◇ **Exercise A.63.** Prove Corollary A.56.

From the chain rule and induction one can derive formulas for the higher partial derivatives of a composite map as needed, provided the maps in question are sufficiently differentiable.

Now suppose $f\colon U \to \mathbb{R}$ is a smooth real-valued function on an open set $U \subset \mathbb{R}^n$, and $a \in U$. For any vector $v \in \mathbb{R}^n$, we define the *directional derivative* of f in the direction v at a to be the number

$$D_v f(a) = \left.\frac{d}{dt}\right|_{t=0} f(a + tv). \tag{A.18}$$

(This definition makes sense for any vector v; we do not require v to be a unit vector as one sometimes does in elementary calculus.)

Since $D_v f(a)$ is the ordinary derivative of the composite map $t \mapsto a + tv \mapsto f(a + tv)$, by the chain rule the directional derivative can be written more concretely as

$$D_v f(a) = \sum_{i=1}^{n} v^i \frac{\partial f}{\partial x^i}(a) = Df(a)v.$$

The fundamental theorem of calculus expresses one well-known relationship between integrals and derivatives. Another is that integrals of smooth functions can be differentiated under the integral sign. A precise statement is given in the next theorem; this is not the best that can be proved, but it is more than sufficient for our purposes. For a proof, see [Rud76].

Theorem A.57 (Differentiation Under an Integral Sign). *Let $U \subset \mathbb{R}^n$ be an open set, $a, b \in \mathbb{R}$, and let $f \colon U \times [a,b] \to \mathbb{R}$ be a continuous function such that the partial derivatives $\partial f/\partial x^i \colon U \times [a,b] \to \mathbb{R}$ are also continuous for $i = 1, \dots, n$. Define $F \colon U \to \mathbb{R}$ by*

$$F(x) = \int_a^b f(x,t)\,dt.$$

Then F is of class C^1, and its partial derivatives can be computed by differentiating under the integral sign:

$$\frac{\partial F}{\partial x^i}(x) = \int_a^b \frac{\partial f}{\partial x^i}(x,t)\,dt.$$

Theorem A.58 (First-Order Taylor's Formula with Remainder). *Let $U \subset \mathbb{R}^n$ be a convex open set, and let $a \in U$ be fixed. Suppose $f \in C^{k+1}(U)$ for some $1 \le k \le \infty$. Then*

$$f(x) = f(a) + \sum_{i=1}^n \frac{\partial f}{\partial x^i}(a)(x^i - a^i) + \sum_{i=1}^n g_i(x)(x^i - a^i), \tag{A.19}$$

for some functions $g_1, \dots, g_n \in C^k(U)$ satisfying $g_i(a) = 0$.

Proof. For any $a \in U$ and any $w \in \mathbb{R}^n$ small enough that $a + w \in U$, the fundamental theorem of calculus and the chain rule give

$$\begin{aligned} f(a+w) - f(a) &= \int_0^1 \frac{d}{dt} f(a+tw)\,dt \\ &= \int_0^1 \sum_{i=1}^n \frac{\partial f}{\partial x^i}(a+tw) w^i\,dt \\ &= \sum_{i=1}^n \frac{\partial f}{\partial x^i}(a) w^i + \sum_{i=1}^n w^i \int_0^1 \left(\frac{\partial f}{\partial x^i}(a+tw) - \frac{\partial f}{\partial x^i}(a) \right) dt. \end{aligned}$$

Substituting $w = x - a$ and

$$g_i(x) = \int_0^1 \left(\frac{\partial f}{\partial x^i}(a + t(x-a)) - \frac{\partial f}{\partial x^i}(a) \right) dt, \tag{A.20}$$

we obtain (A.19). It is obvious from the definition that $g_i(a) = 0$. Because the integrand is of class C^k in all variables, we can differentiate under the integral up to k times, and thus $g_i \in C^k(U)$. □

We will sometimes need to consider smoothness of maps whose domains are subsets of $\mathbb{R}^n$ that are not open. If $A \subset \mathbb{R}^n$ is *any* subset, a map $F \colon A \to \mathbb{R}^m$ is said to be smooth if it admits a smooth extension to an open neighborhood of each point; more precisely, if for every $x \in A$, there exist an open set $U_x \subset \mathbb{R}^n$ containing x and a smooth map $\widetilde{F}\colon U_x \to \mathbb{R}^m$ that agrees with F on $U_x \cap A$.

Multiple Integrals

In this section, we give a brief review of some basic facts regarding multiple integrals in $\mathbb{R}^n$. For our purposes, the Riemann integral will be more than sufficient. Readers who are familiar with the theory of Lebesgue integration are free to interpret all of our integrals in the Lebesgue sense, because the two integrals are equal for the types of functions we will consider. For more details on the aspects of integration theory described here, you can consult nearly any text that treats multivariable calculus rigorously, such as [Apo74, Fle77, Mun91, Rud76, Spi65].

A *rectangle* in $\mathbb{R}^n$ (also called a *closed rectangle*) is a product set of the form $[a^1, b^1] \times \cdots \times [a^n, b^n]$, for real numbers $a^i < b^i$. Analogously, an *open rectangle* is the interior of a closed rectangle, a set of the form $(a^1, b^1) \times \cdots \times (a^n, b^n)$. The *volume* of a rectangle A of either type, denoted by $\operatorname{Vol}(A)$, is defined to be the product of the lengths of its component intervals:

$$\operatorname{Vol}(A) = (b^1 - a^1) \cdots (b^n - a^n).$$

A rectangle is called a *cube* if all of its side lengths $|b_i - a_i|$ are equal.

A *partition* of a closed interval $[a, b]$ is a finite set $P = \{a_0, \dots, a_k\}$ of real numbers such that $a = a_0 < a_1 < \cdots < a_k = b$. Each of the intervals $[a_{i-1}, a_i]$ for $i = 1, \dots, k$ is called a *subinterval* of the partition. Similarly, a partition P of a rectangle $A = [a^1, b^1] \times \cdots \times [a^n, b^n]$ is an n-tuple $(P_1, \dots, P_n)$, where each P_i is a partition of $[a^i, b^i]$. Each rectangle of the form $I_1 \times \cdots \times I_n$, where I_j is a subinterval of P_j, is called a *subrectangle* of P. Clearly A is the union of all the subrectangles in any partition, and distinct subrectangles intersect only on their boundaries.

Suppose $A \subset \mathbb{R}^n$ is a closed rectangle and $f\colon A \to \mathbb{R}$ is a bounded function. For any partition P of A, we define the *lower sum* of f with respect to P by

$$L(f, P) = \sum_j \bigl(\inf_{R_j} f\bigr) \operatorname{Vol}(R_j),$$

where the sum is over all the subrectangles R_j of P. Similarly, the *upper sum* is

$$U(f, P) = \sum_j \bigl(\sup_{R_j} f\bigr) \operatorname{Vol}(R_j).$$

The lower sum with respect to P is obviously less than or equal to the upper sum with respect to the same partition. In fact, more is true.

Lemma A.59. *Let $A \subset \mathbb{R}^n$ be a rectangle, and let $f\colon A \to \mathbb{R}$ be a bounded function. For any pair of partitions P and P' of A,*

$$L(f,P) \le U(f,P').$$

Proof. Write $P = (P_1, \dots, P_n)$ and $P' = (P'_1, \dots, P'_n)$, and let Q be the partition $Q = (P_1 \cup P'_1, \dots, P_n \cup P'_n)$. Each subrectangle of P or P' is a union of finitely many subrectangles of Q. An easy computation shows

$$L(f,P) \le L(f,Q) \le U(f,Q) \le U(f,P'),$$

from which the result follows. □

The *lower integral* of f over A is

$$\underline{\int_A} f\,dV = \sup\{L(f,P) : P \text{ is a partition of } A\},$$

and the *upper integral* is

$$\overline{\int_A} f\,dV = \inf\{U(f,P) : P \text{ is a partition of } A\}.$$

Clearly both numbers exist because f is bounded, and Lemma A.59 implies that the lower integral is less than or equal to the upper integral.

If the upper and lower integrals of f are equal, we say that f is (*Riemann*) *integrable*, and their common value, denoted by

$$\int_A f\,dV,$$

is called the *integral* of f over A. The "dV" in this notation, like the "dx" in the notation for single integrals, has no meaning on its own; it is just a "closing bracket" for the integral sign. Other common notations are

$$\int_A f \quad \text{or} \quad \int_A f\,dx^1 \cdots dx^n \quad \text{or} \quad \int_A f(x^1, \dots, x^n)\,dx^1 \cdots dx^n.$$

In $\mathbb{R}^2$, the symbol dV is often replaced by dA.

There is a simple criterion for a bounded function to be Riemann integrable. It is based on the following notion. A subset $A \subset \mathbb{R}^n$ is said to have *measure zero* if for any $\delta > 0$, there exists a countable cover of A by open cubes $\{C_i\}$ such that $\sum_i \operatorname{Vol}(C_i) < \delta$. (For those who are familiar with the theory of Lebesgue measure, this is equivalent to the condition that the Lebesgue measure of A is equal to zero.)

Lemma A.60 (Properties of Sets of Measure Zero).

(*a*) *A countable union of sets of measure zero in $\mathbb{R}^n$ has measure zero.*

(*b*) *Any subset of a set of measure zero in $\mathbb{R}^n$ has measure zero.*

(*c*) *A set of measure zero in $\mathbb{R}^n$ can contain no nonempty open set.*

(*d*) *Any proper affine subspace of* $\mathbb{R}^n$ *has measure zero in* $\mathbb{R}^n$.

◇ **Exercise A.64.** Prove Lemma A.60.

Part (d) of this lemma illustrates that having measure zero is a property of a set in relation to a particular Euclidean space containing it, not of a set in and of itself—for example, an open interval in the x-axis has measure zero as a subset of $\mathbb{R}^2$, but not when considered as a subset of $\mathbb{R}^1$. For this reason, we sometimes say a subset of $\mathbb{R}^n$ has *n-dimensional measure zero* if we wish to emphasize that it has measure zero as a subset of $\mathbb{R}^n$.

The following proposition gives a sufficient condition for a function to be integrable. It shows, in particular, that every bounded continuous function is integrable.

Proposition A.61 (Lebesgue's Integrability Criterion). *Let* $A \subset \mathbb{R}^n$ *be a rectangle, and let* $f\colon A \to \mathbb{R}$ *be a bounded function. If the set*

$$S = \{x \in A : f \text{ is not continuous at } x\}$$

has measure zero, then f *is integrable.*

Proof. Let $\varepsilon > 0$ be given. By definition of measure zero sets, S can be covered by a countable collection of open cubes $\{C_i\}$ with total volume less than ε.

For each point $q \in A \smallsetminus S$, since f is continuous at q, there is a cube D_q centered at q such that $|f(x) - f(q)| < \varepsilon$ for all $x \in D_q \cap A$. This implies $\sup_{D_q} f - \inf_{D_q} f \le 2\varepsilon$.

The collection of all open cubes of the form $\operatorname{Int} C_i$ or $\operatorname{Int} D_q$ is an open cover of A. By compactness, finitely many of them cover A. Let us relabel these cubes as $\{C_1, \dots, C_k, D_1, \dots, D_l\}$. Replacing each C_i or D_j by its intersection with A, we may assume that each C_i and each D_j is a rectangle contained in A.

Since there are only finitely many rectangles $\{C_i, D_j\}$, there is a partition P with the property that each C_i or D_j is equal to a union of subrectangles of P. (Just use the union of all the endpoints of the component intervals of the rectangles C_i and D_j to define the partition.) We can divide the subrectangles of P into two disjoint sets $\mathcal{C}$ and $\mathcal{D}$ such that every subrectangle in $\mathcal{C}$ is contained in C_i for some i, and every subrectangle in $\mathcal{D}$ is contained in D_j for some j. Then

$$\begin{aligned}
&U(f,P) - L(f,P)\\
&\qquad = \sum_i (\sup_{R_i} f)\operatorname{Vol}(R_i) - \sum_i (\inf_{R_i} f)\operatorname{Vol}(R_i)\\
&\qquad = \sum_{R_i \in \mathcal{C}} (\sup_{R_i} f - \inf_{R_i} f)\operatorname{Vol}(R_i) + \sum_{R_i \in \mathcal{D}} (\sup_{R_i} f - \inf_{R_i} f)\operatorname{Vol}(R_i)\\
&\qquad \le (\sup_A f - \inf_A f) \sum_{R_i \in \mathcal{C}} \operatorname{Vol}(R_i) + 2\varepsilon \sum_{R_i \in \mathcal{D}} \operatorname{Vol}(R_i)
\end{aligned}$$

$$\leq (\sup_A f - \inf_A f)\varepsilon + 2\varepsilon \operatorname{Vol}(A).$$

It follows that

$$\overline{\int_A} f\,dV - \underline{\int_A} f\,dV \leq (\sup_A f - \inf_A f)\varepsilon + 2\varepsilon \operatorname{Vol}(A),$$

which can be made as small as desired by taking ε sufficiently small. This implies that the upper and lower integrals of f must be equal, so f is integrable. □

Remark. In fact, Lebesgue's criterion is both necessary and sufficient for Riemann integrability, but we will not need that.

Now suppose $D \subset \mathbb{R}^n$ is any bounded set, and $f\colon D \to \mathbb{R}$ is a bounded function. Let A be any rectangle containing D, and define $f_D\colon A \to \mathbb{R}$ by

$$f_D(x) = \begin{cases} f(x) & x \in D, \\ 0 & x \in A \smallsetminus D. \end{cases} \tag{A.21}$$

If the integral

$$\int_A f_D\,dV \tag{A.22}$$

exists, f is said to be *integrable over* D, and the integral (A.22) is denoted by $\int_D f\,dV$ and called the *integral of* f *over* D. It is easy to check that the value of the integral does not depend on the rectangle chosen.

In practice, we will be interested only in integrals of bounded continuous functions. However, since we will sometimes need to integrate them over domains other than rectangles, it is necessary to consider also integrals of discontinuous functions such as the function f_D defined by (A.21). The main reason for proving Proposition A.61 is that it allows us to give a simple description of domains on which all bounded continuous functions are integrable.

A subset $D \subset \mathbb{R}^n$ will be called a *domain of integration* if D is bounded and ∂D has n-dimensional measure zero. It is easy to check (using Lemma A.60) that any set whose boundary is contained in a finite union of proper affine subspaces is a domain of integration, and finite unions and intersections of domains of integration are again domains of integration. Thus, for example, any finite union of open or closed rectangles is a domain of integration.

Proposition A.62. *If $D \subset \mathbb{R}^n$ is a domain of integration, then every bounded continuous function on D is integrable over D.*

Proof. Let $f\colon D \to \mathbb{R}$ be bounded and continuous, and let A be a rectangle containing D. To prove the theorem, we need only show that the function $f_D\colon A \to \mathbb{R}$ defined by (A.21) is continuous except on a set of measure zero.

If $x \in \operatorname{Int} D$, then $f_D = f$ on a neighborhood of x, so f_D is continuous at x. Similarly, if $x \in A \smallsetminus \overline{D}$, then $f_D \equiv 0$ on a neighborhood of x, so again f is continuous at x. Thus the set of points where f_D is discontinuous is contained in ∂D, and therefore has measure zero. $\square$

Of course, if D is compact, then the assumption that f is bounded in the preceding proposition is superfluous.

If D is a domain of integration, the *volume* of D is defined to be

$$\operatorname{Vol}(D) = \int_D 1\, dV.$$

The integral on the right-hand side is often abbreviated $\int_D dV$.

The next two propositions collect some basic facts about volume and integrals of continuous functions.

Proposition A.63 (Properties of Volume). *Let $D \subset \mathbb{R}^n$ be a domain of integration.*

(*a*) $\operatorname{Vol}(D) \geq 0$*, with equality if and only if D has measure zero.*

(*b*) *If $D_1, \dots, D_k$ are domains of integration whose union is D, then*

$$\operatorname{Vol}(D) \leq \operatorname{Vol}(D_1) + \cdots + \operatorname{Vol}(D_k),$$

with equality if and only if $D_i \cap D_j$ has measure zero for each i, j.

(*c*) *If D_1 is a domain of integration contained in D, then* $\operatorname{Vol}(D_1) \leq \operatorname{Vol}(D)$*, with equality if and only if $D \smallsetminus D_1$ has measure zero.*

Proposition A.64 (Properties of Integrals). *Let $D \subset \mathbb{R}^n$ be a domain of integration, and let $f, g\colon D \to \mathbb{R}$ be continuous and bounded.*

(*a*) *For any $a, b \in \mathbb{R}$,*

$$\int_D (af + bg)\, dV = a \int_D f\, dV + b \int_D g\, dV.$$

(*b*) *If D has measure zero, then $\int_D f\, dV = 0$.*

(*c*) *If $D_1, \dots, D_k$ are domains of integration whose union is D and whose pairwise intersections have measure zero, then*

$$\int_D f\, dV = \int_{D_1} f\, dV + \cdots + \int_{D_k} f\, dV.$$

(*d*) *If $f \geq 0$ on D, then $\int_D f\, dV \geq 0$, with equality if and only if $f \equiv 0$ on* $\operatorname{Int} D$.

(*e*) $(\inf_D f) \operatorname{Vol}(D) \leq \int_D f\, dV \leq (\sup_D f) \operatorname{Vol}(D)$.

(*f*) $\left| \int_D f\, dV \right| \leq \int_D |f|\, dV$.

◇ **Exercise A.65.** Prove Propositions A.63 and A.64.

There are two more fundamental properties of multiple integrals that we will need. The proofs are too involved to be included in this summary, but you can look them up in the references listed at the beginning of this section if you are interested. Each of these theorems can be stated in various ways, some stronger than others. The versions we give here will be quite sufficient for our applications.

Theorem A.65 (Change of Variables). *Suppose D and E are compact domains of integration in $\mathbb{R}^n$, and $G\colon D \to E$ is a smooth map such that $G|_{\operatorname{Int} D}\colon \operatorname{Int} D \to \operatorname{Int} E$ is a bijective smooth map with smooth inverse. For any continuous function $f\colon E \to \mathbb{R}$,*

$$\int_E f\,dV = \int_D (f \circ G)\,|\det DG|\;dV.$$

Theorem A.66 (Evaluation by Iterated Integration). *Suppose $E \subset \mathbb{R}^n$ is a compact domain of integration and $g_0, g_1\colon E \to \mathbb{R}$ are continuous functions such that $g_0 \le g_1$ everywhere on E. Let $D \subset \mathbb{R}^{n+1}$ be the subset*

$$D = \{(x^1, \dots, x^n, y) \in \mathbb{R}^{n+1} : x \in E \text{ and } g_0(x) \le y \le g_1(x)\}.$$

Then D is a domain of integration, and

$$\int_D f\,dV = \int_E \left(\int_{g_0(x)}^{g_1(x)} f(x,y)\,dy \right) dV.$$

Of course, there is nothing special about the last variable in this formula; an analogous result holds for any domain D that can be expressed as the set on which one variable is bounded between two continuous functions of the remaining variables.

If the domain E in the preceding theorem is also a region between two graphs, the same theorem can be applied again to E. In particular, the following formula for an integral over a rectangle follows easily by induction.

Corollary A.67. *Let $A = [a^1, b^1] \times \cdots \times [a^n, b^n]$ be a closed rectangle in $\mathbb{R}^n$, and let $f\colon A \to \mathbb{R}$ be continuous. Then*

$$\int_A f\,dV = \int_{a^n}^{b^n} \left(\cdots \left(\int_{a^1}^{b^1} f(x^1, \dots, x^n)\,dx^1 \right) \cdots \right) dx^n,$$

and the same is true if the variables in the iterated integral on the right-hand side are reordered in any way.

Integrals of Vector-Valued Functions

If $D \subset \mathbb{R}^n$ is a domain of integration and $F\colon D \to \mathbb{R}^k$ is a bounded continuous vector-valued function, we define the integral of F over D to be the vector in $\mathbb{R}^k$ obtained by integrating F component by component:

$$\int_D F\,dV = \left(\int_D F^1\,dV, \dots, \int_D F^k\,dV\right).$$

The analogues of parts (a)–(c) of Proposition A.64 obviously hold for vector-valued integrals, just by applying them to each component. Part (f) holds as well, but requires a bit more work to prove.

Lemma A.68. *Suppose $D \subset \mathbb{R}^n$ is a domain of integration and $F\colon D \to \mathbb{R}^k$ is a bounded continuous vector-valued function. Then*

$$\left|\int_D F\,dV\right| \le \int_D |F|\,dV. \tag{A.23}$$

Proof. Let G denote the vector $\int_D F\,dV \in \mathbb{R}^k$. Then

$$\begin{aligned} |G|^2 = \sum_{i=1}^k (G^i)^2 &= \sum_{i=1}^k G^i \int_D F^i\,dV \\ &= \sum_{i=1}^k \int_D G^i F^i\,dV \\ &= \int_D (G \cdot F)\,dV. \end{aligned}$$

Applying Proposition A.64(f) to the scalar integral $\int_D (G \cdot F)\,dV$, we obtain

$$|G|^2 \le \int_D |G \cdot F|\;dV \le \int_D |G|\;|F|\;dV = |G| \int_D |F|\;dV.$$

If $G = 0$, the result is trivial; otherwise, dividing both sides by $|G|$ yields (A.23). □

As an application of this inequality, we prove an important estimate for the local behavior of a C^1 function in terms of its total derivative. If $U \subset \mathbb{R}^n$ is any subset, a function $F\colon U \to \mathbb{R}^m$ is said to be *Lipschitz continuous* on U if there is a constant C such that

$$|F(x) - F(y)| \le C|x - y| \qquad \text{for all } x, y \in U. \tag{A.24}$$

Any such C is called a *Lipschitz constant* for F.

Proposition A.69 (Lipschitz Estimate for C^1 Functions). *Let $U \subset \mathbb{R}^n$ be an open set, and let $F\colon U \to \mathbb{R}^m$ be of class C^1. Then F is Lipschitz continuous on any compact convex subset $B \subset U$. The Lipschitz constant can be taken to be $\sup_{x \in B} |DF(x)|$.*

Proof. Since $|DF(x)|$ is a continuous function of x, it is bounded on the compact set B. (The norm here is the Euclidean norm on matrices defined in (A.13).) Let $M = \sup_{x \in B} |DF(x)|$. For arbitrary $a, b \in B$, $a + t(b-a) \in B$

for all $t \in [0,1]$ because B is convex. By the fundamental theorem of calculus applied to each component of F, together with the chain rule,

$$\begin{aligned} F(b) - F(a) &= \int_0^1 \frac{d}{dt} F(a + t(b-a))\, dt \\ &= \int_0^1 DF(a + t(b-a))(b-a)\, dt. \end{aligned}$$

Therefore, by (A.23) and Exercise A.56,

$$\begin{aligned} |F(b) - F(a)| &\le \int_0^1 |DF(a + t(b-a))|\; |b-a|\; dt \\ &\le \int_0^1 M\, |b-a|\; dt \\ &= M|b-a|. \end{aligned}$$

□

Sequences and Series of Functions

We conclude with a summary of the most important facts about sequences and series of functions on Euclidean spaces.

Let $S \subset \mathbb{R}^n$ be any subset, and for each integer $i \ge 1$ suppose that $f_i \colon S \to \mathbb{R}^m$ is a function on S. The sequence $\{f_i\}$ is said to *converge pointwise* to $f \colon S \to \mathbb{R}^m$ if for each $a \in S$ and each $\varepsilon > 0$, there exists an integer N such that $i \ge N$ implies $|f_i(a) - f(a)| < \varepsilon$. The sequence is said to *converge uniformly* to f if N can be chosen independently of the point a: For each $\varepsilon > 0$ there exists N such that $i \ge N$ implies $|f_i(a) - f(a)| < \varepsilon$ for every $a \in S$. The sequence is *uniformly Cauchy* if for any $\varepsilon > 0$ there exists N such that $i, j \ge N$ implies $|f_i(a) - f_j(a)| < \varepsilon$ for all $a \in S$.

Theorem A.70 (Properties of Uniform Convergence). *Let $S \subset \mathbb{R}^n$, and suppose $f_i \colon S \to \mathbb{R}^m$ is continuous for each integer $i \ge 1$.*

(*a*) *If $f_i \to f$ uniformly, then f is continuous.*

(*b*) *If the sequence $\{f_i\}$ is uniformly Cauchy, then it converges uniformly to a continuous function.*

(*c*) *If $f_i \to f$ uniformly and S is a compact domain of integration, then*

$$\lim_{i \to \infty} \int_S f_i\, dV = \int_S f\, dV.$$

(*d*) *If S is open, each f_i is of class C^1, $f_i \to f$ pointwise, and $\{\partial f_i / \partial x^j\}$ converges uniformly on S as $i \to \infty$, then $\partial f / \partial x^j$ exists on S and*

$$\frac{\partial f}{\partial x^j} = \lim_{i \to \infty} \frac{\partial f_i}{\partial x^j}.$$

For a proof, see [Rud76].

An infinite series of functions $\sum_{i=0}^{\infty} f_i$ on $S \subset \mathbb{R}^n$ is said to converge pointwise to a function g if the corresponding sequence of partial sums converges pointwise:

$$g(x) = \lim_{M \to \infty} \sum_{i=0}^{M} f_i(x) \qquad \text{for all } x \in S.$$

The series is said to converge uniformly if its partial sums do so.

Proposition A.71 (Weierstrass *M*-test). *Suppose $S \subset \mathbb{R}^n$ is any subset, and $f_i \colon S \to \mathbb{R}^k$ are functions. If there exist positive real numbers M_i such that $\sup_S |f_i| \le M_i$ and $\sum_i M_i$ converges, then $\sum_i f_i$ converges uniformly on S.*

◇ **Exercise A.66.** Prove Proposition A.71.

References

[AM78] Ralph Abraham and Jerrold E. Marsden. *Foundations of Mechanics*. Benjamin/Cummings, Reading, MA, second edition, 1978.

[Apo74] Tom M. Apostol. *Mathematical Analysis*. Addison-Wesley, Reading, MA, second edition, 1974.

[BD65] William E. Boyce and Richard C. DiPrima. *Elementary Differential Equations and Boundary Value Problems*. Wiley, New York, 1965.

[BM58] Raoul Bott and John Milnor. On the parallelizability of the spheres. *Bull. Amer. Math. Soc.*, 64:87–89, 1958.

[Bre93] Glen E. Bredon. *Topology and Geometry*. Springer-Verlag, New York, 1993.

[Cha93] Isaac Chavel. *Riemannian geometry—a modern introduction*. Cambridge University Press, Cambridge, 1993.

[DK90] S. K. Donaldson and P. B. Kronheimer. *The Geometry of Four-Manifolds*. Clarendon Press, New York, 1990.

[FIS97] Stephen H. Friedberg, Arnold J. Insel, and Lawrence E. Spence. *Linear Algebra*. Prentice Hall, Upper Saddle River, NJ, third edition, 1997.

[Fle77] Wendell Fleming. *Functions of Several Variables*. Springer-Verlag, New York, second edition, 1977.

[FQ90] Michael Freedman and Frank Quinn. *Topology of 4-Manifolds*. Princeton University Press, Princeton, 1990.

[Her75] Israel N. Herstein. *Topics in Algebra*. Wiley, New York, second edition, 1975.

[HL82] Reese Harvey and H. Blaine Lawson, Jr. Calibrated geometries. *Acta Math.*, 148:47–157, 1982.

[Hun90] Thomas W. Hungerford. *Abstract Algebra: An Introduction*. Saunders College Publishing, Philadelphia, 1990.

[Ker60] Michel A. Kervaire. A manifold which does not admit any differentiable structure. *Comment. Math. Helv.*, 34:257–270, 1960.

[KM63] Michel A. Kervaire and John W. Milnor. Groups of homotopy spheres: I. *Annals of Math.*, 77:504–537, 1963.

[Lee97] John M. Lee. *Riemannian Manifolds: An Introduction to Curvature*. Springer-Verlag, New York, 1997.

[Lee00] John M. Lee. *Introduction to Topological Manifolds*. Springer-Verlag, New York, 2000.

[Mas89] William S. Massey. *Algebraic Topology: An Introduction*. Springer-Verlag, New York, 1989.

[Mil65] John W. Milnor. *Topology from the Differentiable Viewpoint*. Princeton University Press, Princeton, 1965.

[Moi77] Edwin E. Moise. *Geometric Topology in Dimensions 2 and 3*. Springer-Verlag, New York, 1977.

[Mos65] Jürgen Moser. On the volume elements on a manifold. *Trans. Amer. Math. Soc.*, 120:286–294, 1965.

[Mun60] James R. Munkres. Obstructions to the smoothing of piecewise differentiable homeomorphisms. *Annals of Math.*, 72:521–554, 1960.

[Mun84] James R. Munkres. *Elements of Algebraic Topology*. Addison-Wesley, Menlo Park, CA, 1984.

[Mun91] James R. Munkres. *Analysis on Manifolds*. Addison-Wesley, Redwood City, CA, 1991.

[Mun00] James R. Munkres. *Topology*. Prentice Hall, Upper Saddle River, NJ, 2000.

[Noe71] Emmy Noether. Invariant variation problems. *Transport Theory Statist. Phys.*, 1(3):186–207, 1971. Translated from the German (Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II 1918, 235–257).

[Pet98] Peter Petersen. *Riemannian geometry*. Springer-Verlag, New York, 1998.

[Rud76] Walter Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, New York, third edition, 1976.

[Sha97] R. W. Sharpe. *Differential geometry*. Springer-Verlag, New York, 1997. Cartan's generalization of Klein's Erlangen program.

[Sie92] Allan J. Sieradski. *An Introduction to Topology and Homotopy*. PWS-Kent, Boston, 1992.

[Spa89] Edwin H. Spanier. *Algebraic Topology*. Springer-Verlag, New York, 1989.

[Spi65] Michael Spivak. *Calculus on Manifolds*. W. A. Benjamin, New York, 1965.

[Spi79] Michael Spivak. *A Comprehensive Introduction to Differential Geometry*, volume I–V. Publish or Perish, Berkeley, second edition, 1979.

[Ste64] Shlomo Sternberg. *Lectures on Differential Geometry*. Prentice Hall, Englewood Cliffs, NJ, 1964.

[Var84] V. S. Varadarajan. *Lie Groups, Lie Algebras, and Their Representations*. Springer-Verlag, New York, 1984.

[War83] Frank W. Warner. *Foundations of Differentiable Manifolds and Lie Groups*. Springer-Verlag, New York, 1983.

[Whi36] Hassler Whitney. Differentiable manifolds. *Ann. of Math. (2)*, 37:645–680, 1936.

[Whi44a] Hassler Whitney. The self-intersections of a smooth n-manifold in $2n$-space. *Ann. of Math. (2)*, 45:220–246, 1944.

[Whi44b] Hassler Whitney. The singularities of a smooth n-manifold in $(2n-1)$-space. *Ann. of Math. (2)*, 45:247–293, 1944.

[Wol71] Joseph A. Wolf. On the geometry and classification of absolute parallelisms I. *J. Differential Geometry*, 6:317–342, 1971.

[Wol72] Joseph A. Wolf. On the geometry and classification of absolute parallelisms II. *J. Differential Geometry*, 7:19–44, 1972.

Index