I. Topological background

This is a set of notes for the final course in a three quarter graduate level sequence in topology and geometry.

Basic references for the course

We shall begin the official text for the course:

> **L. Conlon.** Differentiable Manifolds. (Second Edition), *Birkhäuser-Boston, Boston MA*, 2001, ISBN 0–8176–4134–3.

These notes will complement the text is several different ways that are described in Section I.A. Several files of exercises for the course also appear (or will appear) in the course directory.

At many points we shall assume material covered in the preceding two courses. Both of the latter are largely based upon the following standard textbook:

> **J. R. Munkres.** Topology. (Second Edition), *Prentice-Hall, Saddle River NJ*, 2000. ISBN: 0–13–181629–2.

We shall refer to this book in the notes as [MUNKRES1]).

In addition, a fairly complete set of lecture notes for the first course in the sequence along with many other documents are available for downloading in an online directory I have set up:

`http://math.ucr.edu/~res/math205A`

As indicated, this directory contains lecture notes for Mathematics 205A in the file(s)

`prelimtext.*`

where $*$ is `ps` for PostScript files and `pdf` for Acrobat PDF files. The preliminary course handouts contain information on how to download these and other files from the online directories for various courses including this one and Mathematics 205A. We shall refer to the lecture notes for 205A as the

ONLINE 205A NOTES

in this document. We should also note that the online directory for 205A contains numerous other items besides the ONLINE 205A NOTES, including a title page and table of contents (files of the form `prelimcontents.*`), homework exercises with solutions, and various files with supplementary pictures and written material. A summary of material in the directory appears at the end of the ONLINE 205A NOTES.

Occasionally it will be necessary or helpful to give references to some other standard references for point set topology. Here is a list of these books:

> **J. Dugundji.** Topology. (Reprint of the 1966 Edition. Allyn and Bacon Series in Advanced Mathematics.) *Allyn and Bacon, Boston MA-London (U.K.)-Sydney (Austr.)*, 1978. ISBN: 0-205-00271-4.

K. Jänich. Topology. (With a chapter by Theodor Bröcker. Translated from the German by Silvio Levy. Undergraduate Texts in Mathematics.) *Springer-Verlag, Berlin-Heidelberg-New York*, 1984. ISBN: 0-387-90892-7.

J. L. Kelley. General Topology. (Reprint of the 1955 Edition. Graduate Texts in Mathematics, No. 27.) *Springer-Verlag, Berlin-Heidelberg-New York*, 1975. ISBN: 0-387-90125-6.

L. A. Steen; J. A. Seebach, Jr. Counterexamples in Topology. (Reprint of the Second (1978) Edition.) *Dover, Mineola NY*, 1995. ISBN: 0-486-68735-X.

At a few points we shall also need to quote some basic results from algebraic topology beyond those in the last few chapters of [MUNKRES1]. Everything we need is contained in the following recent text, which is available online free of charge (for individual use only — see the Internet link for details):

A. Hatcher. Algebraic Topology. (Third Paperback Printing), *Cambridge University Press, New York NY*, 2002. ISBN: 0–521–79540–0.

Here is the Internet link to the online version of the book:

`www.math.cornell.edu/~hatcher/AT/ATpage.html`

Chapter 1 of Hatcher's book overlaps considerably with Chapters 9–14 of [MUNKRES1].

Finally, at a few other points we shall mention some basic results about integration theory on "nice" topological spaces. Our references for background material in this area will be three standard texts for the undergraduate and graduate real analysis sequences:

W. Rudin Principles of Mathematical Analysis. (Third Edition. International Series in Pure and Applied Mathematics.) *McGraw-Hill, New York-Auckland-Düsseldorf*, 1976. ISBN: 0-07-054235-X. [This book is sometimes known as "BABY RUDIN" or "LITTLE RUDIN."]

H. L. Royden. Real Analysis. (Third Edition.) *Macmillan, New York NY*, 1988. ISBN: 0-02-404151-3.

W. Rudin. Real and Complex Analysis. (Third Edition. Mc-Graw-Hill Series in Higher Mathematics.) *McGraw-Hill, Boston*-etc., 1987. ISBN: 0-07-054234-1. [This book is sometimes known as "BIG RUDIN."]

The graduate real analysis sequence is **not** a prerequisite or a corequisite for the geometry/topology sequence, and accordingly no detailed knowledge of definitions, theorems or proofs will be needed for the course (aside from some clearly marked exercises).

A few objectives

The main goal of this course is to develop the basic properties of certain mathematical objects known as

smooth manifolds

(also known as *differentiable* or *differential manifolds*). Roughly speaking, a smooth manifold is a topological space with extra structure that allows one to "do differential and integral calculus" on the given space. As indicated below, some work is needed in order to develop a logically sound and scientifically useful theory of smooth manifolds, but in the meantime some informal remarks might be worthwhile.

What should a smooth manifold be? Perhaps the simplest motivation appears in the ancient argument that the earth is flat because "it looks that way." The standard response is that it does indeed look flat if one only views a small portion of the earth, but it does not look flat if one looks at the larger portions of the earth's surface as a whole. This suggests that a manifold should be something that locally looks like Euclidean space of some fixed dimension. Furthermore, it also suggests that the usual 2-dimensional sphere, which is the standard mathematical model for the earth's surface, should be a fundamental example of a smooth manifold. From this viewpoint, one is led to ask which properties of the 2-sphere should figure in the definition of an abstract smooth manifold.

One way of defining the 2-dimensional sphere mathematically is to say that it is the set of all points $\mathbf{p} = (x,\, y,\, z) \in \mathbf{R}^3$ such that $x^2 + y^2 + z^2 = 1$. This is a topological space, and every point of this space has an open neighborhood that is homeomorphic to an open subset of $\mathbf{R}^2$ (For the sake of completeness here is a proof: Since some coordinate of $\mathbf{p}$ is nonzero, it follows that $\mathbf{p}$ does not lie on at least one of the circles defined by the intersection of the sphere with the $xy-$, $yz-$ or $xz-$planes; Each of these circles separates the sphere into two hemispheres, and each of these hemispheres is homeomorphic to the open disk of radius 1 about the origin — for example, if the separating circle lies on the $xy-$plane, then the homeomorphism sends a point $(x\, y)$ on the planar disk to $(x,\, y,\, \pm\sqrt{1 - x^2 - y^2}\,)$ in the upper or lower hemisphere, depending upon the sign of the third coordinate) .

A second important fact about the sphere comes from a basic theme in undergraduate multivariable calculus: *A great deal of differential and integral calculus for functions of two variables can be carried over to the 2-dimensional sphere.*

Despite our emphasis on the 2-sphere, it is definitely not the only significant example of a smooth manifold, and in fact many other curves and surfaces in the Euclidean plane and 3-space yield additional examples of smooth manifolds. A large amount of further discussion appears in Section I.1 of these course notes.

Some motivation

What are the reasons for studying smooth manifolds? We shall split the discussion into two parts, one of which involves looking forwards to further courses in mathematics and the other of which involves looking backwards to issues arising in some undergraduate courses.

LOOKING FORWARDS. Perhaps the simplest and most compelling answers to the lead question are that

(i) smooth manifolds arise in an extremely broad range of mathematical contexts,

(ii) they also play an important role in several aspects of classical and modern physics.

While smooth manifolds are not quite as ubiquitous mathematically as topological spaces, they arise in many different contexts. Furthermore, already in the nineteenth century mathematicians and physicists realized that one should consider manifolds of arbitrary finite dimension and not just curves and surfaces.

Smooth manifolds underlie virtually everything in differential geometry, and they are also a fundamentally important special class of objects in topology. They are also important in many aspects of analysis, including ordinary and partial differential equations. Within algebra, they arise naturally in algebraic geometry, which starts with the study of solutions of systems of polynomial equations in several variables, and they are also fundamental to the theory of nonassociative algebras.

Currently the uses of manifolds in physics are most visible in relativity theory, particularly in connection with questions about the structure of the universe. However, theoretical physicists have actually been interested in manifolds ever since they were introduced during the nineteenth century. In particular, manifolds provided the mathematical framework for discussing the states of various physical systems. One early example involved the possible configurations of mechanical systems like rigid systems of particles; in fact, such ideas play a significant role in studying applications to present day engineering problems like the motion of robot arms. The dimensions of the manifolds generally depended upon the complexity of the system. In another direction, the Hamiltonian approach to mechanics depends upon the notion of a *phase space*, which is essentially a special type of 6-dimensional manifold. Intuitively speaking, points of a phase space represent motion states of a single particle, with three coordinates describing the position and three describing velocity. Similarly, if one wishes to describe the motion states of a ball, one obtains an 8-dimensional object, with three coordinates for position, three for linear velocity, and two for angular velocity. Of course, one of the distinguishing features of relativity theory is its description of the universe as a 4-dimensional object, with three space dimensions and one time variable. However, this idea also arose during the nineteenth century in various classical and pre-modern contexts.

Although smooth manifolds are the key new concept introduced in this course, we shall also encounter a few other objects that appear frequently in mathematics and its applications to physics:

Tensors. Standard reference works in mathematics and physics often describe tensors as generalizations of vectors. The classical definitions are extremely messy: An n-dimensional tensor of rank r is defined classically in terms of n^r local coordinates that satisfy some basic compatibility relations. If $r = 0$ tensors reduce to scalars, and if $r = 1$ the references claim that one somehow obtains n-dimensional vectors. However, the "somehow" part is frequently not explained with much clarity. The material of this course includes a relatively simple and coordinate-free description of tensors, including a precise formulation of the appropriate concept of vector.

Lie algebras. Formally, a Lie algebra can be defined pretty directly as a vector space with a multiplication operation that satisfies the usual distributive laws and certain other identities that do **not** include the associative law. On the other hand, these objects originally arose in connection with certain questions in geometry and analysis, and currently they also play important roles in theoretical physics and combinatorics (the mathematical theory of counting). In this course we shall explain how Lie algebras arise in connection with systems of ordinary differential equations.

LOOKING BACKWARDS. Frequently the study of a more advanced area in mathematics provides new insights into certain topics from earlier courses, and in fact various observations from this course yield valuable new perspectives of this sort. We shall give several important examples from undergraduate multivariable calculus and related topics, using the following text as a background reference.

[MT] **J. E. Marsden and A. J. Tromba.** Vector Calculus. (Fifth Edition), *W. H. Freeman & Co., New York NY*, 2003. ISBN: 0–7147–4992–0.

We shall begin with an important algebraic operation in vector analysis.

Cross products for 3-dimensional vectors. Undergraduate linear algebra courses generalize most orf elementary 2- and 3-dimensional vector algebra to n-dimensions, where n is an arbitrary positive integer. The most notable exception is the theory of cross products. In fact, one can show that vector products from $\mathbf{R}^n \times \mathbf{R}^n$ to $\mathbf{R}^n$ satisfying several basic properties of the cross product can only exist when $n = 3$ or 7; further information is available online at the following sites:

`http://www.math.niu.edu/~rusin/known-math/95/prods`

`http://www.math.niu.edu/~rusin/known-math/96/octonionic`

On the other hand, in this course we shall construct *TWO* abstract n-dimensional algebraic settings into which the 3-dimensional cross product fits naturally. Each turns out to be surprisingly intricate. Although neither captures all the good algebraic properties of the ordinary cross product, the two settings play distinct, fundamentally important roles in modern mathematics. Details may be found in Sections IV.3 and VI.1 of these course notes.

Jacobians and related objects. We have already mentioned the standard result in multivariable calculus that a reasonable real valued function f of several variables has a good linear approximation of the form

$$f(x+h) \quad\approx\quad f(x) \;+\; \nabla f(x)\cdot h$$

if $|h|$ is sufficiently small. The analogous result for vector valued functions (compare pp. 134–135 of [MT]) is fundamental to this course. Specifically, the generalization replaces $\nabla f(x)\cdot h$ with $Df(x)[h]$, where $Df(x)$ is the matrix of partial derivatives of the coordinate functions at x. In the special case of n-dimensional vector valued functions of n variables, the invertibility of this matrix is equivalent to the nonvanishing of the Jacobian. Numerous constructions and results from this course will provide fresh perspectives on these concepts and certain related facts that sometimes appear in multivariable calculus courses but are usually not emphasized even if they are mentioned. These include the Inverse and Implicit Function Theorems (see Section 3.5 of [MT]) and Taylor's Formula for functions of several variables (see Section 3.2 of [MT]).

Solutions to systems of differential equations. For several reasons, basic undergraduate courses on differential equations focus heavily on writing down solutions to specific types of differential equations. While this emphasis on examples is necessary and important, it does not shed much light on general questions about solutions to ordinary differential equations. And these questions are important for many reasons. Perhaps the most compelling is that many basic examples of differential equations do not have solutions that can be expressed using simple formulas from calculus. In particular, one cannot do this for the classical THREE BODY PROBLEM in celestial mechanics, which grow systems comprised of just three objects. In this course it will be necessary to work with some important general properties of solutions to differential equations, and we shall develop everything that we shall need. These concepts and results play a fundamentally important role in many branches of geometry, topology, analysis and mathematical physics.

Vector fields. In classical physics, vector fields often arise from the flow curves of some mechanical or electromagnetic system, and as such they are closely related to differential equations. We have already discussed the latter, so we now turn to another issue. Ordinary vectors in undergraduate mathematics courses have magnitudes and directions, but vectors in undergraduate physics courses generally also have points of application. This course provides a mathematical framework in which one can discuss vectors that have points of application. Such objects play a fundamentally important role in many branches of modern mathematics and mathematical physics.

Integral theorems of vector analysis. Although the basic results of vector analysis known as Green's Theorem, Stokes' Theorem and the Divergence Theorem all arose from physical problems (*cf.* page 518 of [MT]), one can and should also view them formally as higher dimensional versions of the Fundamental Theorem of Calculus. It is natural to wonder how these might extend to even higher dimensions. This course establishes a general formula that includes all the previously mentioned results as special cases. In the process of carrying this out, we shall also describe higher dimensional analogs of line, surface and volume integrals. We shall also describe a general theory of **exterior differentiaion** that yields the classical divergence and curl operations on ordinary 2- and 3-dimensional vector fields.

Path independence of line integrals. Both physical and mathematical considerations lead one naturally to ask when a line integral of the form

$$\int_{\Gamma} \mathbf{F} \cdot d\,\mathbf{x}$$

depends only upon the integrand and the endpoints of the curve Γ. One assumes that the integrand is defined on some fixed connected open set containing the endpoints. For planar curves and regions this is treated in multivariable calculus courses (*cf.* Section 8.3 of [MT]), and on pages 551–553 of [MT]) there is a result for curves in 3-dimensional Euclidean space with finitely many points removed. If time permits, we shall establish a higher dimensional version of the result from [MT]. In any case, it will be covered in these course notes or another document in the course directory.

Starting point for the course

In mathematics, and especially in geometry, it is frequently necessary to make intuitive concepts logically rigorous. Often this requires more time and effort than one might first expect. The central objects of this course — smooth manifolds — are typical examples of this sort. We shall attempt to describe them in a fairly direct and uncomplicated manner, but as in many other parts of geometry there is no "royal road" to the mathematical theory. One major reason for this is the relatively broad range of mathematical background we need.

Although the theory of smooth manifolds is mainly topological in nature, it also requires nontrivial input from both algebra and analysis. Fortunately, the algebraic and analytic inputs are manageable because

(*i*) most of the algebraic and analytic material is relatively straightforward and may be viewed as an elaboration on familiar ideas from undergraduate courses,

(*ii*) there is no loss of logical continuity if one postpones the proofs of the deepest and most difficult background results and proceeds to derive their implications for the theory of smooth manifolds.

The preceding discussion suggests a two step approach to defining and studying smooth manifolds and related concepts; namely, summarizing a few important properties of their topological structure and reformulating some concepts and results from linear algebra and multivariable calculus in a fashion that is convenient for our purposes. Specifically, we shall need an approach that relies as little as possible upon standard Cartesian coordinates. The reasons will become apparent in Unit III when we state the central definitions of the course. In this first unit we shall discuss a class of topological spaces that include the underlying spaces of smooth manifolds, and in Unit II we shall describe the coordinate-free formulation of multivariable calculus. Issues from linear algebra will arise naturally in connection with the latter.

Hints for studying course materials

Because of the amount of input from algebra and analysis and the time constraints of the course, it is unlikely that every detail can be covered completely in the lectures, but of course the class notes will be complete. Some of the more elementary portions of the algebraic/analytic input will be treated as reading assignments, and the proofs of the "hard" theorems from algebra and analysis will be included for reference, with coverage in the lectures postponed and treated as time permits.

There will also be some new but fairly elementary material from point set topology that will be covered briefly, with some details left to the reader to complete outside of class. Much of this material will look pretty formal and probably unmotivated at first, but nevertheless it is important to go through the details conscientiously in order to be well prepared at certain crucial points of the course. One might view this as building the mathematical strength needed to master the central concepts and results of the course.

Illustrations are often extremely helpful aids to understanding the subject matter of this course. Unfortunately, it was not possible to include pictures in the class notes, but throughout the notes we shall give online references to selected sites on the World Wide Web containing illustrations that might be helpful. Here are a few general references to online lecture notes, including several that are particularly well illustrated.

`http://www.maths.ox.ac.uk/current-students/undergraduates/` [*continue*]
`lecture-material/c/diffman/`

`http://www.maths.tcd.ie/~zaitsev/ln.pdf`

`http://www.math.ucsd.edu/~lindblad/250a/250a.html`

`http://people.hofstra.edu/faculty/Stefan_Waner/ diff__geom/tc.html`

`http://www.math.snu.ac.kr/~hongjong/DiffMfd/0.pdf`

`www-math.mit.edu/~mrowka/math965lectnote.pdf`

`http://www/math.ntnu.no/ dundas/SIF5034/revII/dtII2.ps`

(*Note:* For ease of viewing, a `pdf` version formatted to American letter size paper has been inserted into the course directory as `dundasnotes.pdf`; the latter file is only intended for use in connection with these notes.)

In a few cases we shall also include references to separate graphics files in the course directory.

I.A : Guidelines for using these notes

> Some books are to be tasted, others to be swallowed, and some few to be chewed and digested; that is, some books are to be read only in parts; others to be read, but not curiously; and some few to be read wholly, and with diligence and attention.

Francis Bacon (1561-1626), *Of Studies*

(See the files `bacon.*` for further information)

These course notes have several goals, some of which are different enough that they might seem to conflict with each other. Certainly the most basic is to provide a script matching the content of the course lectures, particularly where the latter deviate from the text. At times the approaches in the lectures will be simplified versions of the text, at other times there will be explanations for crucial assertions that the text characterizes as "easy exercises" or something similar, and at still other times the approach in the notes might differ very significantly from that of the text. Another goal of the notes is to include extra background material. Some wil be reviews of topics from standard undergraduate courses, and other portions will cover topics that are a bit more advanced but should be viewed as supplementing familiar undergraduate course material. Yet another goal is to include further examples to illustrate the numerous concepts and results that appear in this course. As in Mathematics 205A, the examples will include some that satisfy a given set of conditions and others that satisfy some but not all of the given conditions. Finally, one additional objective is to discuss some fundamental questions arising directly from the concepts and results of this course, briefly describing what is know and providing further references for further information. The topics covered range from intermediate graduate course material to research discoveries during the past quarter century.

Clearly some of the goals for these notes have higher priority than others. From Bacon's perspective, the top priority material (which turns out to be quite substantial here!) must be understood thoroughly, at least a reasonable familiarity with medium priority material is necessary, and some understanding of lower priority material may be helpful although it is not essential. Since the relative priorities of topics in the notes might not be immediately clear in some places, we shall assign a priority index to every subsection according to the following basic pattern:

> [Not coded] Central material of the course. It is important to understand this material well, and to be able to work with it effectively. This includes knowing the main definitions, results and proofs well enough to reproduce them (aside from proofs designated as unusually complicated or difficult; in such cases an understanding of the logic of the proofs is sufficient). This also includes the ability to work all exercises not marked as particularly difficult.
>
> ($\star$) Topics that are logically or conceptually necessary but of subsidiary importance. In these cases it will generally suffice to have a passive understanding as opposed to the active mastery described above; on the other hand, it is strongly recommended to strive for a level understanding comparable to the that of the highest priority material.

(2⋆) Peripheral material that is potentially helpful in obtaining a broader or deeper perspective on the subject matter but not needed to understand the required material in the course.

In contrast to Bacon's priority classes, most sections in these notes have no code, with fewer items in the lower priorities.

From time to time there might be subsections with intermediate priority ratings like ($\frac{1}{2}\star$) or ($1\frac{1}{2}\star$). In the latter case the material is also optional (just like the case of (2⋆)) but there may be a few references to explicit items from such subsections in later non-optional sections of the notes.

We shall also use stars to identify proofs that are relatively complicated or difficult. If a proof is marked with one star (⋆), then the reader is expected to understand the argument but not expected to know it well enough to reproduce it. Proofs marked with two stars (⋆⋆) are still more challenging, and in these cases a reader might wish to skip the proof (at least at first reading) and to focus on understanding the statement and implications of the result. *In all these cases it is nevertheless recommended to understand as much of the proofs as possible.*

Remarks, notes, etc. of a peripheral nature will be marked with a double dagger (‡).

I.B : Prerequisites

(Conlon, *Preface to the Second Edition*)

We shall begin by summarizing material from the first two courses in the sequence that we shall need.

I.B.1 : *Background material from Mathematics* 205A

There is a one page summary of the key topics in the online document(s)

`www.math.ucr.edu/~res/math205C/ptsettop.*`

where $*$ = `ps` in PostScript format or `pdf` in Acrobat format.

These topics translate into the following portions of the ONLINE 205A NOTES in the previously cited Mathematics 205A directory: All of Units I–III, Section V.1 and Sections VI.1–VI.4. The material in Unit IV and Section V.2 of the outline will be covered in our course this term, and the final section of the outline (Section VI.6) will not be needed except as supplementary background material. Appendices A through D of the ONLINE 205A NOTES also contain background material that will be useful at various points of this course.

From the second course in the sequence, we shall need the following material:

The fundamental group [Munkres1, §§ 51–54, 58–59] (*Homotopy of paths, the fundamental group, path lifting and the fundamental group of the circle, deformation retracts and homotopy type, the fundamental group of the n-sphere*)

Covering spaces [Munkres1, §§ 53–54, 79–82] (*Covering spaces, path lifing and covering homotopy properties, equivalence of covering spaces, the universal covering space, covering transformations, existence of covering spaces*)

I.B.2 : *Additional input from multivariable calculus*

Since smooth manifolds are supposed to be topological spaces in which one can carry out many aspects of differential and integral calculus, it should not be surprising that the course will use methods and results from the calculus of functions of several variables. The amount of such material will increase as the course progresses, so we shall limit our comments now to material that will be needed right away. As before, we shall use the multivariable text by Marsden and Tromba (reference [MT] above) as a reference for background material.

For our purposes, two particularly important results from multivariable differential calculus are

(1) the linear approximation theorem for functions with continuous first partial derivatives,

(2) the multivariable version of the chain rule.

The first of these results for functions of two variables is treated on pages 133–139 of [MT], and the second is treated on pages 152–157 of [MT].

I.B.3 : *Additional input from linear algebra*

Given the usefulness of vector terminology in undergraduate courses on multivariable calculus, it is also predictable that linear algebra will pay a greater role in this course than in Mathematics 205A and 205B. Nearly everything we need is contained in the first linear algebra course, which covers the subject throughout the theory of determinants (Mathematics 131 at UC Riverside); the default outline for this course is available online from the course directory

`www.math.ucr.edu/~res/math205C`

in the files `math131.*`, where the $*$ is either `ps` in PostScript format or `pdf` in Acrobat PDF format.

However, in contrast to first undergraduate courses in linear algebra, the primary emphasis will be on *coordinate-free formulations* of the key concepts and results. One major reason is that linear algebra plays a crucial role in developing the coordinate-free approach to multivariable calculus upon which much of this course is based.

I.B.4 : *Links to classical differential geometry* $(2\star)$

Although the classical differential geometry of curves and surfaces is not a prerequisite for this course, we have already noted that it does provide important examples of smooth manifolds, and simple versions many basic concepts and constructions from this course appear in the classical differential geometry of curves and surfaces. The online directory

`http://math.ucr.edu/~res/math138A`

contains course notes and other materials for the first undergraduate course in differential geometry at UC Riverside (*i.e.*, Mathematics 138A). There might be some material in this directory that could be helpful to students taking this course.

I.1 : Topological manifolds

(Conlon, §§1.1–1.2, 1.7, Appendix A)

We have already mentioned that smooth manifolds should look like open subsets of some Euclidean space, at least locally. Therefore the first step is to formalize this concept.

I.1.1 : *The basic definitions*

Definition. A topological space X is said to be a **topological n-manifold** if it is Hausdorff and each point $x \in X$ has an open neighborhood that is homeomorphic to an open subset of $\mathbf{R}^n$.

The term "manifold" has evolved from G. F. B. Riemann's description of n-manifolds as n-fold extended quantities (roughly speaking, *manifold* = *many* + *fold*).

One would like to say that the integer n is the **dimension** of the topological manifold, but in order to do so one must dispose of the following question:

> If m and n are positive integers and M is both a topological m-manifold and n-manifold, does it follow that $m = n$?

In fact, the answer to this question is yes by a classical result from algebraic topology.

INVARIANCE OF DIMENSION THEOREM. (L. E. J. Brouwer, 1910) *Let U and V be subspaces of $\mathbf{R}^m$ and $\mathbf{R}^n$, and suppose that U is homeomorphic to V. Then $m = n$.*

Proof. $(2\star)$ This is a fairly direct consequence of another classical result in algebraic topology with a very similar name:

> BROUWER'S INVARIANCE OF DOMAIN THEOREM. *Let $U \subset \mathbf{R}^n$ be an open subset, and let $f : U \to \mathbf{R}^n$ be a continuous $1-1$ mapping. Then f is open.*

This result is Theorem 2B.3 on page 172 of Hatcher's online algebraic topology text.

Turning to the proof of Invariance of Dimension, we shall assume that $m \neq n$ and obtain a contradiction. We may as well assume that $m > n$; if we can establish this case then the other case will follow by systematically interchanging the roles of m and n and U and V in the argument. In mathematical writings, situations like this are often indicated by the phrase, "Without loss of generality, we may assume that $m > n$."

By hypothesis we know that there is a homeomorphism f from U to V. Let

$$j : V \longrightarrow \mathbf{R}^m$$

be the composite of the inclusions $V \subset \mathbf{R}^n$ and

$$\mathbf{R}^n \quad = \quad \mathbf{R}^n \times \{0\} \quad \subset \quad \mathbf{R}^n \times \mathbf{R}^{m-n} \quad \cong \quad \mathbf{R}^m \ .$$

Then the composite $j \circ f$ is a continuous 1–1 map from U to $\mathbf{R}^m$, and its image is $V \times \{0\} \subset \mathbf{R}^n \times \{0\}$. We claim this yields a contradiction. On one hand, Invariance of Domain implies that the image is open. On the other hand, the explicit description of the image implies that the latter is not open in $\mathbf{R}^m$; in fact, the image is nowhere dense because it is contained in $\mathbf{R}^n \times \{0\}$. The contradiction arises from the assumption that $m > n$, and therefore this cannot hold. As noted in the first paragraph of the argument, it follows that m must be equal to n.■

The following result is an immediate consequence of the definitions:

PROPOSITION. *Let X be a Hausdorff topological space. Then X is a topological n-manifold if and only if every point has a neighborhood base of open neighborhoods that are homeomorphic to (open balls/disks in) $\mathbf{R}^n$.*

Proof. The $(\Longleftarrow)$ implication follows immediately from the definition because $\mathbf{R}^n$ is open in itself, so we now turn to the $(\Longrightarrow)$ direction.

More generally, we have the following elementary observation: *If $x \in X$ has an open neighborhood homeomorphic to U such that x corresponds to $y \in U$ and $\{W_\alpha\}$ is a neighborhood base at y, then x has a neighborhood base consisting of sets homeomorphic to the sets in $\{W_\alpha\}$.* Specializing to the case where X is a topological n-manifold, we know that an arbitrary point x has an open neighborhood homeomorphic to an open subset $U \subset \mathbf{R}^n$, so it is only necessary to show that every point in U has a neighborhood base of the type described. But this follows immediately; an arbitrary point $y \in U$ has an open neighborhood base of sets $N_{1/k}(y)$ where k is a sufficiently large positive integer, so everything reduces to show that each of these neighborhoods is homeomorphic to $\mathbf{R}^n$. Note first that $N_\varepsilon(y)$ is homeomorphic to $N_1(0)$ by the map

$$h(u) \quad = \quad \left(\frac{1}{\varepsilon}\right) \ \cdot \ (u - y) \ .$$

Finally, note that the $N_1(0)$ is homeomorphic to $\mathbf{R}^n$ by the map

$$k(v) \quad = \quad \left(\frac{1}{1 - |v|}\right) \ \cdot v \ .$$

This completes the proof.■

Example. Given that we have added the Hausdorff condition in the definition of a topological manifold, one might expect that there are spaces that satisfy the main condition in the definition (locally Euclidean) but are not Hausdorff. The **Forked Line**, or something homeomorphic to it, is the standard example in the 1-dimensional case. Similar examples exist for all dimensions ≥ 1.

Here is the basic construction. Let X be the quotient space of $\mathbf{R} \times \{0,\ 1\}$ modulo the equivalence relation whose equivalence classes are given by the two point sets

$$\{\ (y, 0),\ (y, 1)\ \}$$

for $y \neq 0$ and the one point sets given by $(0, 0)$ and $(0, 1)$, By construction, X is locally Euclidean of dimension 1. However, we claim that *the images of $(0, 0)$ and $(0, 1)$ do not have disjoint open neighborhoods*, or equivalently if we are given open neighborhoods U_0 and U_1 of these respective points then $U_0 \cap U_1 \neq \emptyset$.

Let U_0 and U_1 be open neighborhoods in X for the equivalence classes determined by $(0, 0)$ and $(0, 1)$ respectively, and let

$$q : \mathbf{R} \times \{0,\ 1\} \longrightarrow X$$

be the quotient space projection. Then $q^{-1}(U_0)$ and $q^{-1}(U_1)$ are open subsets of $\mathbf{R} \times \{0,\,1\}$ that are unions of equivalence classes and contain $(0,0)$ and $(0,1)$ respectively. Since $q^{-1}(U_0)$ is an open subset containing $(0,0)$ it must contain an open interval of the form $(-a,a) \times \{0\}$, and since it is also a union of equivalence classes it must also contain the interval

$$\Big(\ (-a,0)\ \cup\ (0,a)\ \Big)\ \times\ \{1\}\ .$$

Similarly, we must have

$$q^{-1}(U_1)\quad \supset\quad \{\ (0,1)\ \}\ \ \cup\ \ \Big(\ (-b,0)\ \cup\ (0,b)\ \Big)\ \times\ \{0,\,1\}$$

for some $b > 0$. Therefore, if c denotes the smaller of a and b, then we know that $U_0 \cap U_1 \ \supset\ q(J_c)$, where

$$J_c \quad = \quad \Big(\ (-c,0)\ \cup\ (0,c)\ \Big)\ \times\ \{0,\,1\}\ .$$

In particular, U_0 and U_1 cannot be disjoint, and therefore X cannot be a Hausdorff space.■

The online MathWorld encyclopedia entry

`http://mathworld.wolfram.com/TopologicalManifold.html`

gives a reference to Hawking and Ellis for uses of non-Hausdorff locally Euclidean spaces in theoretical physics; however, we shall not need such objects subsequent in this course aside from a few exercises. For the sake of completeness, here is a detailed bibliographic description of the book mentioned above:

> **S. W. Hawking and G. F. R. Ellis.** The Large Scale Structure of Space-Time. (Cambridge Monographs on Mathaemtical Physics.) *New York: Cambridge University Press, New York, NY*, 1975. ISBN: 0-521-09906-4.

(*Note.* The trade and reader reviews of this book on `www.amazon.com` are definitely worth reading.)

Before proceeding further, we introduce a fundamental concept.

Definition. If X is a topological n-manifold and $x \in X$, then a **topological coordinate chart** at x is a pair (U,h) consisting of an open set U in $\mathbf{R}^n$ and a 1–1 continuous open map $h : U \to X$ such that $x \in h(U)$.

The following is an almost trivial consequence of the definitions:

PROPOSITION. *If X is a topological n-manifold and $x \in X$, then for each $x \in X$ there exists a topological coordinate chart (U_x, h_x) at x.*■

I.1.2 : *Separation and metrization properties*

The preceding discussion of the Hausdorff spaces leads naturally to questions about other separation properties of topological manifolds. We begin with some immediate consequences of the definitions and standard results in point set topology.

PROPOSITION. *If X is a topological n-manifold, then the following hold:*

(*i*) *The space X is locally compact and Hausdorff, and hence X is also $\mathbf{T}_3$; in fact, it is also completely regular.*

(ii) The space X is locally arcwise connected.

(iii) Every point $x \in X$ has a simply connected open neighborhood.

Verification of this result is left to the reader.■

COROLLARY. *If X is a topological n-manifold, then the connected components are the same as the path components, and these are open sets ($\Longrightarrow$ topological n-manifolds themselves).*

Everything except the statement in parentheses follows from local arcwise connectedness, and the statement in parentheses follows by combining the prior portion of the conclusion with the first part of the previous proposition.■

Metrizability. A standard topological counterexample called the **Long Line** shows that a topological manifold in the sense of these notes is not necessarily metrizable (in fact, not necessarily $\mathbf{T}_4$). This example requires a considerable amount of background about well-ordered sets that we shall not otherwise need, and therefore the construction and proofs have been placed into the course directory file(s) `longline.`$*$ where $*$ is `ps` or `pdf` as usual. For our purposes the Long Line's significance is that it leads directly to the following question: *Under what conditions on the topology of a topological manifold X is the latter metrizable?*

The following results contain the answers to this question:

THEOREM 1. *If X is a connected topological n-manifold that is metrizable, then X is second countable.*$\diamondsuit$

THEOREM 2. *If X is a topological n-manifold, then X is metrizable if and only if each component of X is second countable.*$\diamondsuit$

THEOREM 3. *If X is a topological n-manifold, then X is metrizable if and only if X is paracompact.*$\diamondsuit$

We shall prove these results in the Appendix to Section I.2. The third result implies that all compact topological n-manifolds are metrizable. A more direct proof of this result appears as Theorem 36.2 on pages 326–327 of [MUNKRES1].

I.I.3 : *Further examples and nonexamples*

It is now time to give examples of some spaces that are topological manifolds and others that are not.

EXAMPLES OF TOPOLOGICAL MANIFOLDS.

Example 0. Every open subset U of $\mathbf{R}^n$ is a topological n-manifold. — For each $x \in X$ one can take the "nice" neighborhood to be U itself.■

Example 1. More generally, if U is an open subset of X and X is a topological n-manifold, then U is a topological n-manifold. — The proof of this is left as an exercise.■

Example 2. We already mentioned that the standard 2-dimensional sphere is a topological 2-manifold. More generally, the following argument shows that the standard n-dimensional sphere S^n is a topological n-manifold:

By definition the standard n-dimensional unit sphere S^n is the set of all points x in $\mathbf{R}^n$ such that $|x|^2 = 1$, where $|v|$ denotes the length of v as a vector in $\mathbf{R}^n$. If $n = 1$ or 2 these definitions yield the standard circle in sphere in $\mathbf{R}^2$ and $\mathbf{R}^3$ respectively. — By construction the space S^n is Hausdorff, so we need to prove it is locally Euclidean. For $\sigma = \pm$, let U_j^σ be the set of points on S^n such that the j^{th} coordinate is positive for $\sigma = +$ and negative for $\sigma = -$. Now every point on S^n must have at least one nonzero coordinate, and since this coordinate is either positive or negative it follows that every point lies in (at least) one of the sets U_j^σ. Furthermore, each of the sets U_j^σ is open because it is the intersection of S^n with the open set in $\mathbf{R}^n$ consisting of all points whose j^{th} coordinates lie in either $(-\infty, 0)$ or $(0, +\infty)$ depending upon the choice of σ. To complete the verification that S^n is a topological manifold, it will suffice to prove that each set U_j^σ is homeomorphic to $N_1(0) \subset \mathbf{R}^n$. Let $Q_j : \mathbf{R}^{n+1} \to \mathbf{R}^n$ be the linear transformation whose value on the standard unit vector $\mathbf{e}_i$ (with a one in the i^{th} coordinate and zeros elsewhere) is equal to $\mathbf{e}_i$ if $i < j$, zero if $i = j$, and $\mathbf{e}_{i-1}$ if $i > j$, and let k_j^σ be the restriction of Q_j to U_j^σ. We claim these maps define homeomorphisms from the sets U_j^σ onto $N_1(0)$. In fact, we shall construct explicit inverse mappings h_j^σ as follows: Let $S_j : \mathbf{R}^n \to \mathbf{R}^{n+1}$ be the linear transformation whose value on the standard unit vector $\mathbf{e}_i$ is $\mathbf{e}_i$ if $i < j$ and $\mathbf{e}_{i+1}$ if $i \geq j$. Then elementary calculations show that the continuous map

$$h_j^\sigma(x) \quad = \quad S_j(x) \ + \ \sigma \sqrt{1 - |x|^2}\, \mathbf{e}_j$$

is an inverse to k_j^σ.

Since the formulas for k_j^σ and h_j^σ are given in relatively concise form, the following descriptions of the functions when $n = j = 2$ might be helpful:

$$k_j^\sigma(x, y, z) \quad = \quad (x, z)$$

$$h_j^\sigma(u, v) \quad = \quad (u, \sigma \sqrt{1 - u^2 - v^2}, v) \ \blacksquare$$

Remark. One important feature of the n-sphere is that it is a **compact** topological manifold, in contrast to nonempty open subsets of $\mathbf{R}^n$ which are always noncompact [PROOF: If U is a compact open subset of $\mathbf{R}^n$ then it is closed, so by the connectedness of $\mathbf{R}^n$ we either have $U = \emptyset$ or $U = \mathbf{R}^n$. Since $\mathbf{R}^n$ is not compact, it follows that U must be empty.]

MORE EXAMPLES OF TOPOLOGICAL MANIFOLDS.

Example 3. Another example of a compact topological 2-manifold is the 2-**torus**, which by definition is equal to $S^1 \times S^1$. More generally, if X is a topological n-manifold and Y is a topological m-manifold, then $X \times Y$ is a topological $(m+n)$-manifold. — The proof of this is also left as an exercise.

Example 4. Still more generally, if for each j such that $1 \leq j \leq m$ the space X_j is a topological n_j-manifold, then the product $\prod_j X_j$ is a topological d-manifold, where $d = \sum_j n_j$. in the special case where $X_j = S^1$ for each j, this product is known as the n-**torus** and denoted by T^n.$\blacksquare$

Remark. (‡) Various considerations in topology and geometry lead to a converse question: *If X and Y are spaces such that $X \times Y$ is a topological manifold for some n, are X and Y topological manifolds?* — There are many examples showing that the answer to this question is no. Here are two classical references:

[1] R. H. Bing, *The cartesian product of a certain nonmanifold and a line is E^4*. Ann. of Math. (2) **70** (1959), 399–412.

[2] R. M. Fox, *On a problem of S. Ulam concerning Cartesian products*, Fund. Math. **34** (1947), 278–287.

[3] J. Glimm, *Two Cartesian products which are Euclidean spaces*, Bull. Soc. Math. France **88** (1960), 131–135.

[4] K. W. Kwun, *Products of Euclidean spaces modulo an arc*, Ann. of Math. **79** (1964), 104–108.

Example 5. A Hausdorff space is a topological 0-manifold if and only if it is discrete. This is also left as an exercise (in fact the proof is almost trivial).■

Example 6. If E and X are connected Hausdorff spaces and $p : E \to X$ is a covering space projection, then E is a topological n-manifold if and only if X is. — Once again the proof is left as an exercise.■

Before proceeding to give examples of spaces that are not topological manifolds, we shall note one simple but important consequence of the preceding observation:

PROPOSITION. *If X is a connected topological n-manifold, then X has a simply connected covering space that is also a topological n-manifold.*

The existence of a simply connected covering space follows from earlier observations that a topological manifold is locally arcwise connected and every point has a neighborhood base of simply connected open subsets; by the observation above this simply connected covering space must also be a topological n-manifold.■

SOME SPACES THAT ARE **NOT** TOPOLOGICAL MANIFOLDS.

Example 7. A figure 8 curve is an example of a Hausdorff space that is not a topological manifold of any dimension. One specific example of such a curve is given by the parametric equations

$$\gamma(t) \quad = \quad (\sin 2t, \, \sin t)$$

where t lies in some open interval containing $[0, 2\pi]$. Detailed discussions of an equivalent curve (x- and y-coordinates switched, one axis compressed by a factor of $\frac{1}{2}$) may be found at the following online sites:

`http://www-gap.dcs.st-and.ac.uk/~history/Curves/Eight.html`

`http://www.xahlee.org/SpecialPlaneCurves_dir/LemniscateofGerono_dir/` [*continue*]
`lemniscateofGerono.html`

If one simply looks at the character

8

it seems likely that the crossing point in the center cannot have a neighborhood that is homeomorphic to an open n-disk in any Euclidean space. Formally, we CLAIM that *this space is not locally Euclidean*

Proof. ($\star$) The argument is based upon the following two observations.

FACT 1. *If E denotes the image of γ as above and U is an open subset of E contained inside $N_1(0)$, then $U - \{0\}$ has at least 4 connected components.*

FACT 2. *If X is a topological n-manifold and $x \in X$, then the following hold:*

[0] *If $n = 0$ and U is an open subset containing x, then there is an open subset V such that $x \in V \subset U$ and $V - \{x\} = \emptyset$.*

[1] *If $n = 1$ and U is an open subset containing x, then there is an open subset V such that $x \in V \subset U$, the deleted neighborhood $V - \{x\}$ has exactly 2 components, and every open subset W satisfying $x \in W \subset V$ contains points in both connected components of $V - \{x\}$.*

[2] *If $n \geq 2$ and U is an open subset containing x, then there is an open subset V such that $x \in V \subset U$ and $V - \{x\}$ is connected.*

Since the space E does not satisfy any of the conditions [0]–[2], it follows that E cannot be a topological manifold of any dimension.

Verification of Fact 1. The intersection of E with the coordinate axes is given by the points $(\pm 1, 0)$ and $(0, 0)$. If W denotes the complement of the coordinate axes, then W has exactly four components; namely, the standard four quadrants defined by the following systems of simultaneous inequalities:

FIRST QUADRANT:	$x > 0,$	$y > 0$
SECOND QUADRANT.	$x < 0,$	$y > 0$
THIRD QUADRANT.	$x < 0,$	$y < 0$
FOURTH QUADRANT.	$x > 0,$	$y < 0$

Given an open neighborhood U of the origin in E, one can find some $\varepsilon < 1$ such that $N_\varepsilon(0) \cap E \subset U$. We claim that $N_\varepsilon(0) \cap E$ contains points in each quadrant. Visually this is obvious because E looks like the mathematical symbol ∞, and writing everything down analytically is relatively straightforward. By continuity there is some $\delta > 0$ such that γ maps the open intervals with radii δ and centers 0 and π into $N_\varepsilon(0)$, and in fact one can say even more:

The open interval $(0,\, \delta)$ is mapped into the first quadrant.

The open interval $(\pi - \delta,\, \pi)$ is mapped into the fourth quadrant.

The open interval $(\pi,\, \pi + \delta)$ is mapped into the second quadrant.

The open interval $(2\pi - \delta,\, 2\pi)$ is mapped into the third quadrant.

We now know that $N_\varepsilon(0) \cap E \subset U \cap W$ and contains points in each component of W. Suppose that a, b, c, d are points of $N_\varepsilon(0) \cap E$ such that each lies in a different component of W. Then it is impossible for a pair of these points to lie in the same connected component of $N_\varepsilon(0) \cap E$, for if they did then they would lie int the same connected component of W. In particular, it follows that $N_\varepsilon(0) \cap E$ must have at least four connected components.■

Verification of Fact 2. There are three cases to consider. If $n = 0$ the definition implies that each $x \in X$ has an open neighborhood that is homeomorphic to an open subset of $\mathbf{R}^0 = \{0\}$. Such a neighborhood must consist only of x itself, and hence one may simply take $V = \{x\}$. If $n = 1$, then there is an open neighborhood $V \subset U$ and a homeomorphism $h : (a - \delta,\, a + \delta) \to V$ such that $h(a) = x$. It follows that $V - \{x\}$ is the disconnected set $h(a - \delta,\, 0) \cup h : (0,\, a + \delta)$. Similarly, if $n \geq 2$ then there is an open neighborhood $V \subset U$ and a homeomorphism $h : N_\delta(a) \to V$ such that $h(a) = x$, and it follows that $V - \{x\}$ is the connected set $N_\delta(a) - \{a\}$.■

Example 8. We claim that the set $[0, +\infty)$ of nonnegative real numbers is also not a topological n-manifold for any value of n. We shall show this using Fact 2 and some refinements of the latter. Since every point of $[0, +\infty)$ is a limit point of this set, it follows that this space cannot be a topological 0-manifold. To see that $[0, +\infty)$ is not a topological 1-manifold, observe that every open neighborhood of 0 in $[0, +\infty)$ contains an open subneighborhood of the form $V = [0, \delta)$. For each such V we know that $V - \{0\}$ is connected; but this is inconsistent with [1], and therefore it follows that $[0, +\infty)$ cannot be a topological 1-manifold. Finally, if $0 < x \in [0, +\infty)$, then for every open neighborhood U of x we know that $U - x$ is disconnected, and therefore by [2] we know that $[0, +\infty)$ cannot be a topological n-manifold for any $n \geq 2$.■

Example 9. ($\star$) Using Invariance of Domain, one can extend the preceding (non)example: *If X_k^n is the set of all points in $\mathbf{R}^n$ whose first k coordinates are nonnegative (where $1 \leq k \leq n$), then X_k^n is not a topological manifold.* The first step is to notice that X_k^n contains $(0, \infty)^n$ as an open subset, and since the latter is a topological n-manifold, it follows that X_k^n cannot be a topological k-manifold unless $k = n$. Suppose now that it is a topological n-manifold; let $0 \in X_k^n$ be the zero point, and let W be an open neighborhood of 0 in X_k^n. Since W contains no points for which at least one of the first k coordinates is negative, it follows that W cannot be open in $\mathbf{R}^n$. Invariance of Domain then implies that W cannot be homeomorphic to an open subset of $\mathbf{R}^n$. Therefore X_k^n cannot be a topological n-manifold.

Example 10. ($2\star$) The final nonexample requires some input from topological dimension theory; everything we need is contained in Section 50 of [MUNKRES1], either in the text itself or in the accompanying exercises. Let $\mathbf{HQ}$ (the *Hilbert cube*) be a cartesian product of $\aleph_0$ copies of the unit interval $[0, 1]$, and for each positive integer n let $\mathbf{HQ}_n$ be the subspace of all points whose i^{th} coordinate vanishes for each $i \geq n$ (hence $\mathbf{HQ}_n$ is homeomorphic to $[0, 1]^n$). — Suppose that $\mathbf{HQ}$ is a topological n-manifold for some $n \geq 0$. By Corollary 50.8 on page 314 of [MUNKRES1], it then follows that $\mathbf{HQ}$ must be homeomorphic to a subset of $\mathbf{R}^{2n+1}$. Let $f : \mathbf{HQ} \to \mathbf{R}^{2n+1}$ be a homeomorphism onto a compact subset. Also, let $h : [0, 1]^{2n+2} \to \mathbf{HQ}_{2n+2}$ be the homeomorphism mentioned earlier, let W denote the image of $(0, 1)^{2n+2}$ under h, and let W_0 denote the image of $(0, 1)^{2n+1} \times \{\frac{1}{2}\}$. Then Invariance of Domain implies that $f(W_0)$ is open in $\mathbf{R}^{2n+1}$. On the other hand, W_0 is also nowhere dense in $\mathbf{HQ}$, and hence $f(W_0)$ must also be nowhere dense in $\mathbf{R}^{2n+1}$. In particular, the latter implies that $f(W_0)$ cannot be an open subset of $\mathbf{R}^{2n+1}$, yielding a contradiction. The latter means that the assumption that $\mathbf{HQ}$ is a topological n-manifold must be false. It follows that $\mathbf{HQ}$ is not a topological n-manifold for any choice of n.

I.1.4 : *Another way of constructing manifolds* ($\star$)

The remaining material in this section is only needed to work one of the exercises for this section, and except for this it may be skipped until it is needed in Unit III.

The following result turns out to be very useful for finding examples of topological manifolds. This result uses the concept of a continuous group action on a topological space; this notion is defined on page 199 of [MUNKRES1] at the beginning of Exercise 8.

WEAK QUOTIENT SPACE PRINCIPLE. *If X is a topological n-manifold and G is a finite group that acts freely on X (in other words, if $1 \neq g \in G$ then $g \cdot x \neq x$ for all $x \in X$), then the quotient space X/G is also a topological n-manifold.*

The quotient projection from X to X/G will be denoted by p.

Proof. Let $x \in X$, and let $1 \neq g \in G$. Since X is Hausdorff there are disjoint open sets U_g and V_g such that $x \in U_g$ and $g \cdot x \in V_g$. Let

$$W = \bigcap_{g \neq 1} \left(U_g \cap g^{-1} \cdot V_g \right) .$$

We claim that $a \cdot W \cap b \cdot W = \emptyset$ if $a \neq b$ in G. Since $a \cdot W \cap b \cdot W = \emptyset$ if and only if $W \cap a^{-1} b \cdot W = \emptyset$, it suffices to consider the case where $a = 1$. But in this case we have $W \subset U_b$ and $b \cdot W \subset V_b$, so that $W \cap b \cdot W \subset U_b \cap V_b = \emptyset$.

Since W is open, the set

$$p^{-1}(\,p(W)\,) \quad = \quad \bigcap_{g\in G} \; g\cdot W$$

is also open, and therefore $p(W)$ is open in X/G by the definition of the quotient topology. Let U be an open neighborhood of x such that $U \subset W$ and U is homeomorphic to an open subset of $\mathbf{R}^n$. We claim that U is open in X/G; this follows from the same argument showing that $p(W)$ is open in X/G, and in fact it also follows that for each open subset $V \subset U$ the set $p(V)$ is also open in X/G.

We have now shown that $p|U$ is a continuous open mapping from X to X/G. To show that X/G is locally Euclidean of dimension n, it will suffice to verify that $p|U$ is 1–1. But if $p(y) = p(z)$ then $z = g\cdot y$ for some $g \in G$; the latter implies that $z \in g\cdot U \subset g\cdot W$; since $y \in U \subset W$ and $W \cap g\cdot W = \emptyset$ if $g \neq 1$, this means that g must be equal to 1 and therefore z must be equal to y.

Finally, we need to show that X/G is Hausdorff. Let s and t be distinct points of X/G, and let x and y be points in X which map to s and t respectively. Since X is Hausdorff, the points x and y have disjoint open neighborhoods U and V. In fact, we claim that x and y have open neighborhoods $U_0 \subset U$ and $V_0 \subset V$ such that $U_0 \cap g\cdot V_0 = \emptyset$ for all $g \in G$. To wee this, let U_g and V_g be disjoint open neighborhoods containing x and y respectively; if $U_0 = \cap_g \; U_g$ and $V_0 = \cap_g \; g^{-1}\cdot V_g$, then it follows that $U_0 \cap g\cdot V_0 \subset U_o \cap V_g = \emptyset$.

As in the preceding discussion from the previous paragraphs, one can find open subneighborhoods $U_0' \subset U_0$ and $V_0' \subset V_0$ such that $U_0' \cap g\cdot U_0' = V_0' \cap g\cdot V_0' = \emptyset$ if $g \neq 1$. This discussion also shows that $p(U_0')$ and $p(V_0')$ are open neighborhoods of $s = p(x)$ and $t = p(y)$ respectively.

To complete the proof, we need to show that $p(U_0') \cap p(V_0')$ is empty. Suppose to the contrary that we have some point r in the intersection. Then $r = p(z)$ for some $z \in U_0'$ and $r = p(w)$ for some $w \in V_0'$. Since w and z have the same image under p, it follows that $w = g\cdot z$ for some $g \in G$. Therefore we have $w \in g\cdot U_0' \cap V_0'$, so that

$$U_0 \;\cap g^{-1}\cdot V_0 \quad\supset\quad U_0' \;\cap g^{-1}\cdot V_0' \neq \emptyset\ .$$

But we have constructed U_0 and V_0 so that $U_0 \;\cap g^{-1}\cdot V_0 \quad \supset U_0' = \emptyset$, so this is a contradiction, and consequently it follows that $p(U_0') \cap p(V_0') = \emptyset$.■

Example. The condition that $g \neq 1 \implies g\cdot x \neq x$ for all $x \in X$ (*i.e.*, the group acts *freely*) is not needed to prove that X/G is Hausdorff, but it is needed to prove that X/G is locally Euclidean. Consider the action of the multiplicative group $g = \{\pm 1\}$ on $\mathbf{R}$ by ordinary multiplication; in this case $0 = (-1)\cdot 0$, so the group does not act freely. We claim that there is an standard homeomorphism from $\mathbf{R}/G$ to $[0,+\infty)$ sending the equivalence class of $\pm x$ to $|x|$.

(PROOF: There is an obvious well defined set-theoretic map $h : \mathbf{R}/G \to [0,+\infty)$ sending the equivalence class of x to $|x|$, and this map is continuous by the fundamental properties of the quotient topology. This map is also 1–1 and onto, so all that remains is to check that it is a homeomorphism. The latter will hold if the absolute value map from $\mathbf{R}$ to $[0,+\infty)$ is either an open mapping or a closed mapping. Verification that the mapping is open is an elementary exercise that will be left to the reader; one important step is to notice that the image of an open interval of the form $(-h,h)$ is the half-open interval $[0,h)$, which is an open subset of $[0,+\infty)$.)

Since we have already verified that $[0,+\infty)$ is not a topological manifold of any dimension, this yields an example of a finite group action on a manifold such that the quotient space is not a manifold.■

I.1.5 : *Manifolds with boundary* ($2\star$)

For many purposes it is convenient to view spaces like $[0, +\infty)$ or the closed unit disk $D^n \subset \mathbf{R}^n$ ($=$ all x such that $|x| \leq 1$) as mild generalizations of topological manifolds as previously defined. The interiors of these sets in $\mathbf{R}$ and $\mathbf{R}^n$ (respectively) are open ($\Longrightarrow$ topological manifolds of the same dimension), and their point-set-theoretic frontiers are manifolds of one dimension lower. Intuitively, one might view these as examples of **manifolds with** [a nice] **boundary**. Such objects play an extremely important role in mathematics, and basic information on them is presented in Sections 1.6 and 3.6 of Conlon. However, we shall not define or work with such objects in these notes in order to avoid introducing an additional layer of complexity into the course.

I.1.6 : *The topological classification problem* ($1\frac{1}{2}\star$)

In many parts of theoretical mathematics, it is interesting and important to study *classification problems*. For example, in the theory of finite groups, one natural question is to describe all groups of a fixed order n up to isomorphism, and every undergraduate course in abstract algebra answers this question if n is prime (all finite groups of prime order are cyclic). The corresponding problem for many other values of n arises frequently in graduate abstract algebra courses, and it is answered completely if the prime factorization of n is not too complicated (for example, if n is a square of a prime or twice and odd prime). For topological manifolds, or subclasses satisfying suitable restrictions, the corresponding question involves classification up to homeomorphism:

Classification Problem for Manifolds. *Let $\mathcal{A}$ be a class of topological manifolds. Find an explicitly describable subclass $\mathcal{A}_0 \subset \mathcal{A}$ such that every space in $\mathcal{A}$ is homeomorphic to a unique space in $\mathcal{A}_0$.*

Both [MUNKRES1] and Conlon include discussions of the important special case where $\mathcal{A}$ is the class of all compact topological 2-manifolds (also known as *compact* or *closed surfaces*). Specific references are Chapter 11 (Sections 74–78) in [MUNKRES1] and Section 1.3 in Conlon. Of course, one can also pose similar questions about classifying topological n-manifolds for other values of n, and we shall conclude this section by describing known results in these cases.

If $n = 0$ the classification is completely trivial because a topological 0-manifold is a discrete space; therefore, if $\mathcal{A}$ is the class of all second countable topological 0-manifolds, then one can take $\mathcal{A}_0$ to consist of one discrete space of each cardinality up to and including $\aleph_0$. If $n = 1$ and $\mathcal{A}$ is the class of all connected topological 1-manifolds, then one can use the methods of point set topology to prove that every such manifold is homeomorphic to the real line or the circle (and of course the latter are not homeomorphic because the second is compact and the first is not). The main ideas behind the proof of this result appear in the texts listed below; specifically, the reference in Hocking and Young is Section 2–5 on pages 52–55 with background material in the preceding section, and the reference in Christensen and Voxman is Section 9.A on pages 227–232, with accompanying exercises on page 251, and closely related material in Section 5.A on pages 127–128.

Hocking, John G.; Young, Gail S. Topology. (Second edition.) *Dover, New York NY*, 1988. ISBN: 0-486-65676-4.

Christenson, Charles O.; Voxman, William L. Aspects of Topology. [FIRST EDITION.] (Pure and applied Mathematics, Vol. 39.) *Marcel Dekker, New York-Basel*, 1977.

If $n = 3$ and $\mathcal{A}$ is the class of all compact topological 3-manifolds, then the answer to the classification question is unknown, and in fact it is closely tied to one of the seven Millenium

Problems on the list published by the Clay Mathematical Institute in Cambridge, Massachusetts (*i.e.*, the POINCARÉ CONJECTURE); see the online site `http://www.claymath.org/millennium/` for additional information). Recently G. Perelman (a Russian mathematician from St. Petersburg) announced results that would yield an answer to this classification question. Other mathematicians believe that Perelman has made some extremely important contributions, and currently serious efforts are underway to check the logical accuracy and completeness of his work.

If $n = 4$ and $\mathcal{A}$ is some reasonable class of all compact topological 4-manifolds, then the answer to the classification question is beyond the unknown: It turns out to be mathematically unsolvable. The reasons are essentially algebraic and depend upon the recursive unsolvability of certain group-theoretic questions; one example is the impossibility of finding uniform criteria to decide whether two "reasonable" groups are isomorphic. One uses the fundamental group to reduce the topological questions to group-theoretic ones. A detailed discussion may be found in the following standard reference for the subject:

C. F. Miller, III. On group-theoretic decision problems and their classification. Annals of Mathematics Studies, No. 68. *Princeton University Press, Princeton NJ*, 1971. ISBN: 0-691-08091-7.

I.2 : Partitions of unity

(Conlon, §§1.4–1.5)

In our study of manifolds we shall need some methods for constructing global versions of structures that are known to exist locally. The purpose of this section is to develop a particularly useful technique for carrying out such constructions.

I.2.1 : *A global existence problem*

It seems worthwhile to begin with a question that motivates the constructions of this section.

INTERPOLATION PROBLEM. *Suppose that X is a topological space and we are given an open neighborhood U of $X \times \{0\}$ in $X \times [0, \infty)$. Is there a continuous real valued function $g : X \to (0, \infty)$ such that the set*

$$\{ \ (x, t) \in X \times [0, \infty) \mid t < f(x) \ \}$$

is contained in U?

If X is compact then results from Mathematics 205A yield an affirmative answer, and in fact one can take g to be a constant function. On the other hand, if we take X to be the real line then it is easy to construct an open neighborhood of $X \times \{0\}$ in $X \times [0, \infty)$ that contains no subsets of the form $X \times [0.\varepsilon)$ regardless of how small one takes $\varepsilon > 0$; in particular, the set of points (x, t) such that $t < \min\{1, \ 1/x\}$ is a simple counterexample. An affirmative answer to the Interpolation Problem would yield an attractive and convenient "neighborhood base" for $X \times \{0\}$ in $X \times [0, \infty)$, and later in the course we shall encounter similar problems of this sort.

Let us examine the Interpolation Problem a bit further. For each $x \in X$ we can find an open neighborhood V of x and an $\varepsilon(x) > 0$ such that $V \times [0, \varepsilon(x)) \subset U$, and therefore we can define a function that solves the Interpolation Problem over V. We would like to have some means of assembling these locally defined functions into a globally defined function with the desired property. There are also many other situations where one can construct something locally and would like to piece the local constructions together and obtain a corresponding global object. Partitions of unity provide a solid mathematical framework for doing so in a reasonably broad class of contexts. In an Appendix to this section, we shall use partitions of unity to help prove the metrization theorems for topological manifolds stated in Section I.1.

Definition. Let $\mathcal{U} = \{ U_\alpha \}$ be an open covering of the topological space X with indexing set A. A *partition of unity subordinate to* $\mathcal{U}$ is an indexed family of continuous functions $\varphi_\alpha : X \to [0,1]$ such that

(i) For each α the **support** of φ_α — namely the *closure* of the set of points where $\varphi_\alpha > 0$ — is contained in U_α,

(ii) for each $x \in X$ there is an open neighborhood V_x of x such that only finitely functions φ_α are nonzero on V_x,

(iii) the sum of the functions φ_α — which is defined and continuous by (ii) – is identically equal to 1 on X.

The motivation and potential usefulness of this concept are probably unclear the first time one encounters the definition. Perhaps the best approach is to defer these issues until the discussion of applications.

I.2.2 : *Paracompactness and manifolds*

Paracompactness is an extremely useful condition on topological spaces that formally generalizes compactness. The class of paracompact Hausdorff spaces includes all compact Hausdorff spaces, all metrizable spaces, and nearly all the spaces that arise in algebraic and geometric topology; on the other hand, since one can prove that paracompact Hausdorff spaces are normal (see Theorem 41.1 on page 253 of [MUNKRES1]), it follows that there are many significant examples of spaces from point set topology that are not paracompact.

Sections 39, 41 and 42 of [MUNKRES1] summarize many of the main facts about paracompact spaces (and still further information may be found in the standard reference works by Dugundji, Kelley, and Steen–Seebach). For our purposes, it will be useful to focus on paracompactness as a condition on topological manifolds. One advantage of this approach is that many of the proofs simplify dramatically if one restricts to topological manifolds, and another is that the simplified arguments will provide valuable insights when we move on to study smooth manifolds.

The following definitions come directly from the ONLINE 205A NOTES.

Definition. A family of subsets $\mathcal{A} = \{ A_\alpha \}$ in a topological space X is said to be *locally finite* if for each $x \in X$ there is an open neighborhood U such that $U \cap A_\alpha \neq \emptyset$ for only finitely many A_α.

Examples. Aside from finite families, perhaps the most basic examples of locally finite families are given by the following families of subsets of $\mathbf{R}^n$.

(1) For each positive integer n let A_n be the closed annulus consisting of all points x such that $n - 1 \leq |x| \leq n$. Then for each $y \in \mathbf{R}^n$ the set $N_{1/2}(y)$ only contains points from at most two closed sets in the family (verify this!).

(2) For each positive integer n let V_n be the open annulus consisting of all points x such that $n-2 < |x| < n+1$. The details for this example are left to the reader as an exercise.

Local finiteness is just one concept that is needed to define paracompactness, so we shall continue to develop the concepts we need.

Definition. Let X be a topological space and let $\mathcal{U} = \{ U_\alpha \}$ be an open covering of X. Then $\mathcal{U}$ is said to be *locally finite* if for each $x \in X$ there is an open neighborhood V_x such that $V_x \cap U_\alpha \neq \emptyset$ for only finitely many α.

Of course, finite coverings are locally finite. If $X = \mathbf{R}^n$, then the open covering consisting of the ring shaped regions U_k defined by $k-1 < |x| < k+2$ is locally finite, for if we choose a positive integer m such that $|x| < m$ then the open set $N_m(x)$ intersects U_k nontrivially if and only if $k \leq m$. On the other hand, the open covering of $\mathbf{R}^n$ by the open disks $N_k(0)$ is not locally finite; in fact, if $x \in X$ such that $x < m$ for some nonnegative integer m, then $x \in N_k(0)$ for all $k \geq m$ and hence every open neighborhood of x intersects each of the sets $N_k(0)$ nontrivially.

Definition. Let X be a topological space and let $\mathcal{U} = \{ U_\alpha \}$ and $\mathcal{V} = \{ V_\beta \}$ be open coverings of X. Then $\mathcal{V}$ is said to be a *refinement* of $\mathcal{U}$ if for each V_β there is some U_α such that $V_\beta \subset U_\alpha$.

If $\mathcal{U}$ is an open covering of X and $\mathcal{V}$ is a subcovering, then $\mathcal{V}$ is a refinement of $\mathcal{U}$ because we can simply take U_α to be V_β itself. Here are some less trivial examples: Suppose that X is a compact metric space and $\mathcal{U}$ is an open covering of X. Then the Lebesgue covering lemma (see Section III.1 in the ONLINE 205A NOTES) guarantees the existence of a number $\eta > 0$ such that every set of diameter less than η is contained in some element of U. Therefore any open covering of X consisting of open neighborhoods of the form $N_{\eta/2}(x_\alpha)$ will be a refinement of $\mathcal{U}$.

The following results provide further insight into the notion of refinement; the proofs are elementary and left to the reader as exercises:

PROPOSITION. *If X is a topological n-manifold and $\mathcal{U}$ is an open covering of X, then there is a refinement $\mathcal{V}$ of $\mathcal{U}$ such that every subset of $\mathcal{V}$ is homeomorphic to $\mathbf{R}^n$.*■

TRANSITIVITY PROPERTY. *If the open covering $\mathcal{V}$ is a refinement of $\mathcal{U}$ and the open covering $\mathcal{W}$ is a refinement of $\mathcal{V}$, then $\mathcal{W}$ is a refinement of $\mathcal{U}$.*■

We are finally ready to define paracompactness.

DEFINITION OF PARACOMPACTNESS. A topological space X is said to be *paracompact* if every open covering has a locally finite refinement.

Example and consequences. Since subcoverings are refinements, compact spaces are automatically paracompact. If X is an infinite discrete space, then X is not compact, but it is paracompact. To see the latter, given any open covering $\mathcal{U}$ of X, it follows immediately that the covering $\mathcal{V}$ consisting of one point sets is an open refinement. It is also locally finite because if $x \in X$ then $\{x\}$ is an open neighborhood of X that is disjoint from all but one element of $\mathcal{V}$ (namely, the set $\{x\}$ itself!).

The next result provides a standard criterion for concluding that a topological manifold is paracompact.

PARACOMPACTNESS THEOREM. *If X is a locally compact Hausdorff space that is second countable, then X is paracompact.*

Proof. Suppose that X is locally compact Hausdorff, let $\mathcal{B}$ be a base for the topology on X, and let $\mathcal{B}' \subset \mathcal{B}$ be the set of all elements with compact closures. We claim that $\mathcal{B}'$ is also a base for X.

To see this, let $U \in \mathcal{B}$ and let $x \in U$. Then we can find an open set V such that $x \in V \subset \overline{V} \subset U$ such that $\overline{V}$ is compact. Since $\mathcal{B}$ is a base for the topology we can find another open subset $W \in \mathcal{B}$ such that $x \in W \subset V$. Since $\overline{W} \subset \overline{V}$ and the latter is compact, it follows that $\overline{W}$ is also compact. Therefore $\mathcal{B}_0$ is a base for the topology on X.■

If we now assume that X is second countable and let $\mathcal{B}$ be a countable base for X, then clearly $\mathcal{B}_0$ is also a countable base. Therefore we may list the elements of $\mathcal{B}_0$ in a sequence $\{ W_k \}$. We shall recursively define a sequence of compact subsets $A_k \subset X$ (where k is a positive integer) with the following properties:

(1) For each k the set A_k lies in the interior of A_{k+1}.

(2) For each k the set A_k contains $\overline{W_k}$.

Since $X = \cup_k W_k$ it will follow that $X = \cup_k A_k$. If $X = \mathbf{R}^n$, a sequence of this type is given by taking the closed disks of radius k about the origin, where k runs over all positive integers.

The construction begins by taking $A_1 = \overline{W_1}$. Given A_k, construct A_{k+1} as follows: The family $\mathcal{B}_0$ determines an open covering of $A_k \cup \overline{W_{k+1}}$, and therefore there is a positive integer m such that this subset is contained in $\cup_{j=1}^m W_j$. If we take A_{k+1} to be the compact set $\cup_{j=1}^m \overline{W_j}$, then A_{k+1} will have all the required properties.

The next step is to write X as a locally finite union of open subsets obtained from the compact sets A_k. In our previous example, where $X = \mathbf{R}^n$ and A_k is the closed disk of radius k centered at the origin, these open sets will be the neighborhoods of the ring shaped sets (or annuli)

$$\{ x \in \mathbf{R}^n \mid k - 1 \leq |x| \leq k \}$$

defined by the inequalities

$$k - 2 < |x| < k + 1 .$$

Note that if Q_k is the open set given by the preceding inequality, then $Q_k \cap Q_j = \emptyset$ unless $|j-k| \leq 2$; in fact, if $k \leq |x| < k+1$ then $N_1(x) \cap Q_j \neq \emptyset$ only if $|j-k| \leq 2$. Therefore the family of open subsets $\{ Q_k \}$ is indeed locally finite.

To imitate this construction in the GENERAL CASE, let

$$Q_k \quad = \quad \mathrm{Int}(A_{k+1}) \ - \ A_{k-2} \ .$$

Once again we have that $Q_k \cap Q_j = \emptyset$ unless $|j-k| \leq 2$. Given $x \in X$, choose $k > 0$ to be the least positive integer such that $x \in \mathrm{Int}(A_{k+1})$. Then $x \notin \mathrm{Int}(A_k)$ implies and $x \notin A_{k-1}$, so that x must belong to Q_k. Since Q_k has a nonempty intersection with only finitely many of the sets Q_j it follows that the family of open subsets $\{ Q_k \}$ is also locally finite in this case.

Using the preceding decompositions of X we may construct a locally finite refinement of our original open covering as follows: Let $\mathcal{U}$ be an open covering of X. Since each set of the form $C_k = A_k - \mathrm{Int}(A_{k-1})$ is compact, there is a finite subset $\mathcal{U}_k$ of $\mathcal{U}$ such that C_k is contained in the union of the open sets in $\mathcal{U}_k$. Let $\mathcal{V}_k$ denote the intersections of the open sets in $\mathcal{U}_k$ with the open set Q_k. Then $\mathcal{V}_k$ is a finite family of open sets covering C_k such that each subset is contained in Q_k. Now every point in X belongs to some subset C_k, and therefore the union $\mathcal{V}$ of the families $\mathcal{V}_k$ defines an open covering of X. Since each open set in $\mathcal{V}_k$ is a subset of a set in $\mathcal{U}$ by construction, it follows that $\mathcal{V}$ is a refinement of $\mathcal{U}$. Finally, we need to show that $\mathcal{V}$ is locally finite. Suppose that $x \in C_k \subset X$, then Q_k is an open neighborhood of x, and its intersection with an open subset from $\mathcal{V}_j$ is nonempty if and only if $|j-k| \leq 2$ because $V \in \mathcal{V}_j$ implies $V \subset Q_j$. Since each collection $\mathcal{V}_j$

is finite, it follows that Q_k has a nonempty intersection with only finitely many members of $\mathcal{V}$, and therefore the open covering $\mathcal{V}$ is locally finite.■

Examples of the sets A_k, Q_k and C_k are depicted in the picture file `sigmacompact.*` (where $*$ is `jpg`, `ps` or `pdf`) from the course directory.

For topological manifolds one can prove a somewhat stronger conclusion that turns out to be extremely important.

COMPLEMENT. *If X is a topological manifold, then it is possible to find a locally finite refinement $\mathcal{V}$ such that the following hold:*

(i) Each V_α in $\mathcal{V}$ is homeomorphic to the open disk of radius 2 centered at the origin.

(ii) If $h_\alpha : N_2(0) \to V_\alpha$ is a homeomorphism as in (i), then the sets $W_\alpha = h_\alpha(\,N_1(0)\,)$ also form an open covering of X.

Note that $\overline{W_\alpha}$ is compact and that $\overline{W_\alpha} \subset V_\alpha$ by construction.

Proof of Complement. The argument is the same except for the construction of a locally finite refinement, which proceeds as follows. Given a positive integer k and a point $x \in C_k$, one can find an open neighborhood M_x of x such that M_x is homeomorphic to $N_2(0)$ and M_x is contained in $W_k \cap U_\alpha$ for some U_α in $\mathcal{U}$. Let $Q_x \subset M_x$ denote the image of $N_1(0)$. Then the sets Q_x form an open covering of C_k such that each subset M_x lies in $W_k \cap U_\alpha$ for some U_α, and therefore one can extract a finite subcovering, say $\mathcal{N}_k = \{\ Q_i\ \}$. Let $\mathcal{M}_k$ be the corresponding family of open subsets M_i.

As in the proof of the theorem we take $\mathcal{M}$ and $\mathcal{N}$ to be the unions of the families $\mathcal{M}_k$ and $\mathcal{N}_k$ respectively; the reasoning in the theorem then implies that $\mathcal{M}$ and $\mathcal{N}$ are locally finite refinements of $\mathcal{U}$. Note also that we have $\overline{Q_i} \subset M_i$ for these sets.■

Note. The open refinements constructed in the theorem and complement are all countable.

I.2.3 : *Constructing partitions of unity*

The preceding results yield the following important tool for piecing together local constructions to form a global object.

EXISTENCE OF PARTITIONS OF UNITY. *Let X be a topological manifold, and let $\mathcal{M}$ and $\mathcal{N}$ be countable open coverings as in the Complement above. Then there is a family of continuous functions $\varphi_j : X \to [0,1]$ such that*

(i) the **support** *of φ_j — that is, the closure of the set on which $\varphi \neq 0$ – is a compact subset of M_j,*

(ii) we have $\sum_j \varphi_j = 1$.

One says that such a family of continuous functions forms a *partition of unity* subordinate to the open covering $\mathcal{M}$. Note that *there is no convergence problem with the sum even if there are infinitely many sets in the open covering $\mathcal{M}$*; each point has a neighborhood that meets only finitely many M_j's nontrivially, and on this neighborhood the sum reduces to a finite sum.

Proof. For each j let $h_j : N_2(0) \to M_j$ be a homeomorphism. Let ω be the continuous real valued function on the interval $[0,2]$ such that $\omega = 1$ on $[0,1]$, ω decreases linearly from 1 to 0 on $[1.\frac{3}{2}]$, and $\omega = 0$ on $[\frac{3}{2},2]$. Then the function $f_j : M_j \to [0,1]$ defined by $f_j(x) = \omega(\,|h_j^{-1}(x)|\,)$

extends to X by taking $f_j = 0$ on the complement of M_j [PROOF: Let E_j denote the image of the closed disk of radius $\frac{3}{2}$; then f_j on M_j and the zero function on $X - E_j$ combine to form a continuous function on X because their restrictions to $W - E_j$, which is the image under h_j of the set of points satisfying $|x| \in (\frac{3}{2}, 2)$, are both equal to zero.]

Since each point $x \in X$ lies in some Q_j it follows that $f_j(x) > 0$ for some j. As noted in the paragraph before the beginning of the proof, the local finiteness of $\mathcal{M}$ ensures that one can add the functions f_j to obtain a well defined continuous function (in fact, near each point the sum reduces to a finite sum of continuous functions). The sum $f = \sum_j f_j$ is always positive by the first sentences of this paragraph, and therefore we may define

$$\varphi_j \quad = \quad \frac{f_j}{f} \ .$$

It follows immediately that the functions φ_j have all the required properties.■

I.2.4 : *Two simple applications*

We shall begin by answering the question at the beginning of this section.

PROPOSITION. *Suppose that X is a topological manifold and U is an open neighborhood of $X \times \{0\}$ in $X \times \mathbf{R}$. Then there is a continuous real valued function $f : X \to (0, \infty)$ such that the set*

$$\{ \ (x, t) \in X \times [0, \infty) \mid t < f(x) \ \}$$

is contained in U.

The additional exercises for this section contain some basic geometric implications of this result.

Proof. For each $x \in X$ there is an open neighborhood U_x and an $\varepsilon_x > 0$ such that $(x, 0) \in U_x \times (-\varepsilon_x, \varepsilon_x) \subset U$. The family $\mathcal{U} = \{ U_x \}$ forms an open covering of X, and consequently one has a locally finite refinement $\mathcal{M}$ of the type described above. As in the previous discussion, let $\{ \varphi_j \}$ be a partition of unity subordinate to $\mathcal{M}$. For each $M_j \in \mathcal{M}$ choose some x such that $M_k \subset U_x$, and define ε_j to be equal to ε_x. By construction it follows that $M_j \times (-\varepsilon_j, \varepsilon_j) \subset U$ for all j.

Let $f = \frac{1}{2} \sum_j \ \varepsilon_j \, \varphi_j$; the local finiteness of $\mathcal{M}$ implies that this reduces to a finite sum on a neighborhood of each point and therefore is continuous. We need to show that $(x, f(x)) \in U$ for all $x \in X$. Let W be an open neighborhood of x such that $W \cap M_j \neq \emptyset$ for only finitely many j, and choose $\ell > 0$ so that $W \cap M_j = \emptyset$ for $j > \ell$. If ε^* denotes the largest of the numbers $\varepsilon_1, \ \cdots, \varepsilon_\ell$, then it follows that $\{x\} \times (-\varepsilon^*, \varepsilon^*) \subset U$. Since

$$f(x) \quad = \quad \tfrac{1}{2} \sum_{j=1}^{\infty} \ \varepsilon_j \, \varphi_j(x) \quad = \quad \tfrac{1}{2} \sum_{j=1}^{\ell} \ \varepsilon_j \, \varphi_j(x) \quad \leq \quad \tfrac{1}{2} \sum_{j=1}^{\ell} \ \varepsilon^* \, \varphi_j(x) \quad \leq \quad \frac{\varepsilon}{2} \ .$$

Therefore $f(x) \in (0, \varepsilon/2] \subset (0, \varepsilon)$, and by the last sentence of the previous paragraph this means that $(x, f(x)) \in U$ as required.■

Here is another example of a construction using a partition of unity.

PROPOSITION. *Let X be a topological manifold. Then there is a continuous function $f : X \to [0, +\infty)$ such that for each $K > 0$ the inverse image $f^{-1}([0, K])$ is compact (in other words, f is a **proper** map).*

Proof. Let $\mathcal{M}$ be a locally finite covering as before, let $\mathcal{N}$ be the associated refinement, and let $\{ \varphi_j \}$ be a partition of unity subordinate to $\mathcal{M}$. By construction φ_j is positive on $\overline{N_j}$ for each j, and since the latter set is compact the function has a positive minimum m_j there. Define

$$f(x) \quad = \quad \sum_{j=1}^{\infty} \frac{j}{m_j} \cdot \varphi_j(x) \ .$$

Then local finiteness again implies that f is continuous. Since $f^{-1}(\,[0,K]\,)$ is closed, it suffices to show that it is contained in a compact subset of X.

By construction, if $x \in N_j$ then $f(x) \geq j$ (in fact, the j^{th} summand of f is $\geq j$ on $\overline{N_j}$). Therefore, if K_0 is the smallest integer $\geq K$ then we know that

$$f^{-1}\big(\,[0,K_0]\,\big) \quad \subset \quad \bigcup_{j=1}^{K_0} \overline{N_j}$$

and since each summand on the right hand side is compact it follows that the left hand side is contained in a compact set.∎

I.2.5 : *Standing assumption*

Nonmetrizable topological manifolds are rarely studied in mathematics or physics except to acknowledge their existence. Furthermore, by the metrization results we know that every metrizable topological manifold is a union of pairwise disjoint open subsets that are second countable. Therefore we shall make the following assumption for the rest of these notes:

UNLESS SPECIFICALLY STATED OTHERWISE, ALL TOPOLOGICAL MANIFOLDS CONSIDERED HENCEFORTH ARE ASSUMED TO BE SECOND COUNTABLE.

I.2. Appendix : Metrization proofs

The starting point for studying the metrizability of a topological n-manifold is that such a space is *locally metrizable*: Every point has a neighborhood base of metrizable open subsets. The Long Line is an example of a space that is Hausdorff and locally metrizable but not metrizable, so it is clear that some additional conditions are needed to ensure that a locally metrizable space is metrizable. We begin with a couple of results about this problem.

I.2.A.1 : *Local versus global metrizability* ($\star$)

Our first result reduces the global metrizability question for topological manifolds to the connected case.

PROPOSITION. *Suppose that X is a Hausdorff topological space whose connected components are all open. Then X is metrizable if and only if each component is metrizable.*

Proof. (2$\star$) The ($\Longrightarrow$) implication follows because subspaces of metrizable spaces are metrizable, so we focus on the other direction for the rest of the proof. Write $X = \cup_\alpha\ X_\alpha$ where each X_α is an open connected metrizable subset. It follows immediately that V is open in X if and only if $V = V_\alpha$ where V_α is open in X_α for each α.

Given a metric space, one can always find another metric with diameter ≤ 1 that defines the same topology; therefore for each α we can find a metric d_α for X_α with diameter ≤ 1. Using these metrics, we shall define a candidate for a metric on X that will define the same topology on X as the original one. Specifically, for $p.q \in X$ define $d(p,q)$ to be $d_\alpha(p,q)$ if there is a (necessarily unique) α such that both p and$q$ belong to X_α and set $d(p,q) = 2$ otherwise. It is a routine exercise to verify that d defines a metric on the set X; verification of the Triangle Inequality is the least trivial part, and this is done on a case by case basis depending upon whether points lie in the same or different components of X.

We must now show that the d-metric topology is equal to the original topology, which we shall call $\mathbf{U}$. First of all we claim that every ε disk for d belongs to $\mathbf{U}$. If $\varepsilon \leq 2$ and $p \in X_\alpha$, then the d-disk about p in X and the $_\alpha$-disks about p in X_α are identical; since the later is open in X_α and hence X, this proves the result when $\varepsilon \leq 2$. On the other hand, if $varepsilon > 2$ then the d-disk about p of radius ε in X is equal to X. This implies that the d-topology is contained in $\mathbf{U}$. Conversely, if $V \in \mathbf{U}$ then $V = V_\alpha$ where V_α is open in X_α for each α. By choice of the metrics d_α we know that V_α is also d_α-open in X_α for each α. Furthermore, by construction the d-open and d_α-open subsets of X_α are equal. Therefore each X_α is d-open in X_α, and by the openness of the latter in X it also follows that each V_α is d-open in X. Thus we have shown that every $\mathbf{U}$-open subset is d-open. Since we have already shown the converse, it follows that the d-topology is equal to $\mathbf{U}$.∎

COROLLARY. *Let X be a topological manifold. Then X is metrizable if and only if each connected component is.*∎

The preceding results are essentially a special case of the next one, but we have included a proof of the former because it can be established by a fairly elementary argument.

LOCAL-GLOBAL THEOREM. *Let X be a paracompact Hausdorff space. Then X is metrizable if and only if X is locally metrizable.*

The ($\Longrightarrow$) implication is trivial, and a proof of the ($\Longleftarrow$) implication using partitions of unity is given in the online document `smirnov.`$*$ where as usual $* =$ `ps` or `pdf`.∎

We are now able to prove the third of the metrization results from Section 1: *A topological manifold is metrizable if and only if it is paracompact.*

The ($\Longrightarrow$) implication follows from A. H. Stone's result that metric spaces are paracompact (see [MUNKRES1], Theorem 41.4, p. 257), and the ($\Longleftarrow$) implication follows from the Local-Global Theorem stated above.∎

I.2.A.2 : *Second countability and metrizability* ($\star$)

In order to complete our discussion of metrizability, we need to show that metrizability and second countability are equivalent for topological manifolds.

Second countability $\Longrightarrow$ metrizability. Previously in this section we showed that a locally compact, Hausdorff, second countable space is paracompact. Since topological manifolds are locally metrizable, the preceding results show that a topological manifold is metrizable if it is paracompact. Combining these, we obtain the desired implication.■

The proof of the reverse implication depends upon the following result:

BASIC σ-COMPACTNESS THEOREM. *Let X be a space that is paracompact $\mathbf{T}_2$, locally compact and connected. Then there is a countable family of compact subsets $K_n \subset X$ such that $X = \cup_n \, K_n$.*

The proof of this theorem requires the following auxiliary result.

LEMMA. *Let X be a topological space, let K be a compact subset of X, and let $\{U_\alpha\}$ be a locally finite open covering of X. Then there are only finitely many open sets U_β in the open covering such that $K \cap U_\beta = \emptyset$.*

Proof of the Lemma. For each $x \in K$ there is an open neighborhood V_x whose intersection with all but finitely many of the sets U_α is empty. By compactness K is contained in a finite union of the form

$$V_{x_1} \cup \cdots \cup V_{k_m}$$

and the intersection of this finite union with U_α is empty for all but finitely many α. Therefore the intersection of K with U_α is also empty for all but finitely many α.■

Proof of the Basic σ-Compactness Theorem. $(2\star)$ Let $\{U_\alpha\}$ be an open covering of X by subsets whose closures are compact. Such a covering exists because X is locally compact. Since every open covering has a locally finite refinement, we may as well assume that $\{U_\alpha\}$ itself is locally finite (note that the condition about compact closures is true for refinements of an open covering if it is true for the covering itself).

Choose W_0 to be an arbitrary nonempty set U_β from the open covering. Define a sequence of subspaces $\{W_n\}$ recursively by

$$W_n = \cup_\alpha \, \{U_\alpha \mid U_\alpha \cap W_{n-1} \neq \emptyset\}.$$

By construction this is an increasing sequence of open subsets. We claim that $X = \cup_k \, W_n$. Since the right hand side is nonempty and open, it suffices to show that $\cup_k \, W_n$ is closed. Suppose that x lies in the closure of $\cup_k W_n$. Then $x \in U_\alpha$ for some α, and since the closure of a set is the union of that set and its limit points it follows that

$$U_\alpha \bigcap (\cup_k \, W_n) \neq \emptyset.$$

The latter in turn implies that $U_\alpha \cap W_{n_0} \neq \emptyset$ for some n_0. But this implies that $x \in W_{n_0+1}$. Therefore all points in the closure $\overline{\cup_k \, W_n}$ in fact lie in $\cup_n W_n$, and hence the latter is closed.

We shall now show that the sets $\overline{W_n}$ is compact by induction on n; if $k = 0$ this holds because $\overline{U_\beta}$ is compact. If $\overline{W_n}$ compact, then by the lemma there are only finitely many U_α such that $U_\alpha \cap \overline{W_n} \neq \emptyset$; call these $U_{\alpha_1}, \cdots, U_{\alpha_p}$. It then follows that

$$\overline{W_{n+1}} \;=\; \overline{U_{\alpha_1}} \cup \cdots \cup \overline{U_{\alpha_p}} \;.$$

Since each of the closures on the right hand side is compact, it follows that the left hand side is a finite union of compact subsets and therefore is compact. Therefore if we set $K_n = \overline{W_n}$ then we know that K_n is compact, $K_n \subset K_{n+1}$ for all n, and $X = \cup_n \, K_n$.■

Proof that metrizable and connected $\Longrightarrow$ second countable. Assume that X is a topological manifold with these properties. As noted above, a space X is paracompact if it is metrizable. Therefore by the Basic σ-Compactness Theorem we know that $X = \cup_n K_n$ where each K_n is compact. Furthermore, since X is metrizable each K_n is also metrizable. The latter implies that each K_n has a countable dense subset D_n, and therefore the countable subset $\cup D_n$ is dense in X. Since X is metric, this means that it is also second countable.■

I.3 : The Contraction Lemma

(Conlon, Appendix B)

The so-called (Banach) Contraction Lemma is a basic result on metric spaces. Although it is elementary to state, it has far-reaching consequences for solving various sorts of equations in many different ways. In this course we shall use it in two highly distinct ways. One is the proof of the classical Inverse Function Theorem from multivariable calculus, and the other is the Picard Method of Successive Approximations which yields existence and uniqueness theorems for solutions to systems of ordinary differential equations. These and further applications appear in Sections II.2 and II.4 below.

I.3.1 : *Statement and proof*

It is convenient to formulate the complete metric space. Background material on the latter may be found in Section III.2 of the ONLINE 205A NOTES.

CONTRACTION LEMMA. *Let X be a complete metric space, and let $T : X \to X$ be a map such that $\mathbf{d}(T(x), T(y)) \leq \alpha \cdot \mathbf{d}(x, y)$ for some fixed $\alpha \in (0, 1)$ and all $x, y \in X$ (in particular, T is uniformly continuous). Then there is a unique $z \in X$ such that $T(z) = z$ (in other words, a* **unique fixed point** *for T).*

Remarks on the hypotheses.

(A) To see that the result does not hold if $\alpha = 1$, consider the antipodal map on the circle taking (x, y) to $(-x, -y)$.

(B) To see the need for completeness, consider the open interval $(0, 1)$ and let T be multiplication by $\frac{1}{2}$.

Proof. The idea is beautifully simple. One starts with an arbitrary point $x \in X$ and considers the sequence of points $x,\ T(x),\ T^2(x),\ \cdots$. This sequence is shown to be a Cauchy sequence, and the limit z of this sequence turns out to be the unique fixed point.

More formally, we begin by noting that T has at most one fixed point. If $z, w \in X$ satisfy $T(z) = z$ and $T(w) = w$, then we have

$$0 \leq \mathbf{d}(z, w) = \mathbf{d}(T(z), T(w)) \leq \alpha\, \mathbf{d}(z, w)$$

and since $0 < \alpha < 1$ this can only happen if $\mathbf{d}(z, w) = 0$; *i.e.*, if $z = w$.

We now follow the idea described in the first paragraph of the proof. By induction on n we have

$$\mathbf{d}\left(T^n(x), T^{n+1}(x)\right) \quad \leq \quad \alpha^n \mathbf{d}\left(x, T(x)\right)$$

and therefore by the triangle inequality for $m > n$ we also have

$$\mathbf{d}\left(T^n(x), T^m(x)\right) \quad \leq \quad \sum_{i=n+1}^{m} \alpha^i \, \mathbf{d}\left(x, T(x)\right) \quad =$$

$$\frac{\alpha^{n+1}(1-\alpha^{m-n})}{1-\alpha} \cdot \mathbf{d}\left(x, T(x)\right) \quad \leq \quad \frac{\alpha^{n+1}}{1-\alpha} \cdot \mathbf{d}\left(x, T(x)\right)$$

which implies that the sequence $\{\, T^n(x) \,\}$ is a Cauchy sequence. By the completeness of X there is a point z such that $z = \lim_{n\to\infty} T^n(x)$.

By Theorem 23.1 on page 130 of [MUNKRES1] we have

$$T(z) \quad = \quad \lim_{n\to\infty} T\left(T^n(x)\right) \quad = \quad \lim_{n\to\infty} T^{n+1}(x)$$

and by a change of variable (specifically, take $k = n+1$) the right hand side is equal to $\lim_{k\to\infty} T^k(x)$, which by construction is z. Therefore we have $T(z) = z$.■

I.3.2 : *Examples* ($\star$)

A few simple applications of the Contraction Lemma to solving polynomial equations in one variable appear in the file(s) `cubicroots.*` in the course directory. In this subsection we shall give a somewhat different example.

If we plot the graphs of the functions $y = \frac{1}{2}(1 + \cos x)$ and $y = x$ on the same coordinate grid, we quickly see that they meet at one point somewhere between $x = 0$ and $x = \frac{\pi}{3}$; *i.e.*, there is a unique solution to the equation $z = \frac{1}{2}(1 + \cos z)$ such that z lies between the two given values of x. In fact, since the first function minus the second is positive for $x = 0$ and negative for $x = \frac{\pi}{3}$, one can prove this rigorously using the Intermediate Value Theorem for real valued functions of one real variable. We shall prove that this also follows from the Contraction Lemma and note one important advantage of the alternative approach: Using the latter, one can determine the numerical value of the solution fairly efficiently using a relatively simple scientific calculator.

If $T(x) = \frac{1}{2}(1 + \cos x)$ then clearly $T(x)$ maps the whole real line into $[0, 1]$ and hence must map $[0, 1]$ to itself. In order to prove that the hypothesis of the Contraction Lemma holds, by the Mean Value Theorem from ordinary single variable calculus it suffices to show that $|T'(x))| < \alpha < 1$ for some $\alpha \in (0, 1)$. However, this follows easily because $|T'(x)| < \frac{1}{2}$ everywhere. Therefore the Contraction Lemma implies the existence of a unique solution to the equation we are considering.

According to the proof of the Contraction Lemma, one can find the solution by forming the sequence $\{T^n(x)\}$ where $x \in [0, 1]$ is arbitrary and then taking the limit. Here is a partial list of the numerical values one obtains starting with $x = 0$

Iteration 01:	0.7701512
Iteration 02:	0.8589027
Iteration 03:	0.8266343
Iteration 04:	0.8386778
Iteration 05:	0.8342234
Iteration 06:	0.8358766
Iteration 07:	0.8352638
Iteration 08:	0.8354910
Iteration 09:	0.8354068
Iteration 10:	0.8354380
Iteration 11:	0.8354265
Iteration 12:	0.8354307
Iteration 13:	0.8354292
Iteration 14:	0.8354297
Iteration 15:	0.8354295
Iteration 16:	0.8354296
Iteration 17:	0.8354296

Since we obtained the same value twice in a row for the sixteenth and seventeenth iterations, it is safe to conclude that this gives the numerical value of the solution to seven decimal places.■

I.4 : Basic topological constructions revisited

(*cf.* [MUNKRES1], §§ 15, 16, 19, 22)

In point set topology the notions of subspace, quotient space and Cartesian product of two spaces can be defined simply and unambiguously. However, the analogous constructions for smooth manifolds are unavoidably more complicated, and in fact *these are better understood in terms of their fundamental properties rather than their formal definitions*. Similarly, if one wishes to define the Cartesian product of three or (finitely many!) more spaces, there are several ways of doing this; each has its own advantages, but the important points are that

(a) the various constructions are canonically homeomorphic to each other,

(b) once again the fundamental properties are more important than the formal definitions.

In all these cases the crucial properties are expressible in terms of continuous functions, and the purpose of this section is to describe the relevant properties of subspaces, quotient spaces and finite Cartesian products. This is essentially a review of elementary material from a more sophisticated viewpoint; nothing is really new and the results probably do not seem interesting for their own right, but they serve as a simple model for subsequent material that is more complicated. The Wright brothers' development of the airplane provides a good nonmathematical analogy: Their initial flight was over dry land along the beach and not straight out over the ocean.

IMPORTANT NOTE. The material in this section may — and perhaps should — be omitted in a first reading of these notes. It provides some basic topological background for Section III.2 below, and at that point it will be necessary to know something about the content of this section. There will be references in Section III.2 to indicate when we are explicitly using something from the material below.

I.4.1 : *Previous examples* ($1\frac{1}{2}\star$)

In the first two courses of this sequence there were instances where fundamentally important objects were awkward or even difficult to construct. Sometimes the specific method of construction remains important for further work. However, in other cases the methods are not particularly helpful in such further studies and for these purposes one only needs to understand some basic "axiomatic" properties. Here are some key examples.

The real number system. Formally, the real numbers are generally constructed from the rationals by one of two methods. However, as indicated in the online document

`math.ucr.edu/~res/math205A/realnumbers.*`

(where as usual $*$ = ps or pdf) the mathematical study of the real number system is based upon the existence of a complete ordered field and the result in

`math.ucr.edu/~res/math205A/uniqreals.*`

(with $*$ as before) showing that all complete ordered fields are isomorphic. Once we have verified the existence of a complete ordered field using either of the standard constructions, the latter are not needed for any further purposes.

Completions and compactifications of metric and topological spaces. These are similar to the preceding example. In both cases the objective is to construct a larger space containing a "suitable copy" of the given one as a dense subset. For completions the phrase "suitable copy" means an isometric copy, and for compactifications the phrase "suitable copy" means a homeomorphic copy. In all these cases the constructions are a bit awkward, but once we have the desired completions or compactifications and their basic properties the details of the constructions are not needed for subsequent work.

Counterexamples in point set topology. Throughout point set topology one is interested in the logical relationships between various properties that a given space might or might not possess. For example, it is important to understand that second countability, separability and the Lindelöf Property are not logically interchangeable for arbitrary topological spaces even though this is the case for metrizable spaces. In these cases it is relatively simple to find spaces that are Lindelöf or separable but not second countable, but in other situations it can be moderately or extremely challenging to find examples showing that one property does not imply another. However, once one knows that such examples exist it is often enough to recognize their existence when writing out logical proofs (*e.g.*, not saying, "Since the space is separable it is second countable, and therefore there is a countable neighborhood base at each point.").

Simply connected universal covering spaces. The entire theory of Hausdorff covering spaces for locally path connected spaces is elegant and conceptual with one important exception; namely, the proof that a simply connected covering space actually exists, at least if all points have sufficiently many simply connected neighborhoods. This construction is awkward, but once it yields the existence of simply connected covering spaces the details of the construction do not play a role in subsequent work.

I.4.2 : *Defining threefold Cartesian products* ($\star$)

Given three sets A, B and C, we clearly want their threefold Cartesian product to be the set of all suitably defined ordered triples (a, b, c) such that $a \in A$, $b \in B$ and $c \in C$, and (finite)

Cartesian products of four or more sets should also have similar description. Here are two reasonable approaches:

COMPREHENSIVE APPROACH. Let A be a set with at least three elements, and let $\{X_\alpha\}$ be an indexed family of sets with indexing set A. Denote the union $\cup_\alpha X_\alpha$ by $\mathbf{X}$. The *Cartesian product* $\prod_\alpha X_\alpha$ is defined to be all subsets $\mathbf{y}$ of $\mathbf{X} \times A$ such that the following hold:

(i) For each $\alpha \in A$ there is a unique $y_\alpha \in \mathbf{X}$ such that $(y_\alpha, \alpha) \in \mathbf{y}$ (*i.e.*. $\mathbf{y}$ is the graph of a function A from $\mathbf{X}$).

(ii) For each $\alpha \in A$ and $(y, \alpha) \in Y$ we have $y \in X_\alpha$.

Note. This differs slightly from the definition in the ONLINE 205A NOTES, which describe the elements of the products as functions rather than graphs of functions. The definition here is essentially the one given in [MUNKRES1]. One advantage of the definition here is that it has the following reassuring property:

Compatibility with inclusions. *If for each α we have a subset $W_\alpha \subset X_\alpha$, then the product $\prod_\alpha W_\alpha$ is a subset of $\prod_\alpha X_\alpha$.*■

We now describe another approach to defining finite products.

RECURSIVE APPROACH. Let $n \geq 2$, and suppose that k-fold Cartesian products have been defined such that a $1-$fold product is just the unique set in the indexed family and $2-$fold products are defined in the standard fashion (either by postulates or by some formal set-theoretic construction). Given a sequence of spaces $\{X_1, \cdots, X_{n+1}\}$, define the $(n+1)$-fold Cartesian product recursively by the following formula:

$$X_1 \times \cdots \times X_{n+1} \quad = \quad (X_1 \times \cdots \times X_n) \times X_{n+1}$$

It should be intuitively clear that both constructions yield equivalent objects that can be interpreted as ordered $m-$tuples for some appropriate value of m. The comprehensive approach works for arbitrary indexed families but the formal definition is awkward, while the recursive approach has a simpler definition but only works for finite products. One important advantage of the recursive approach is that it reduces many questions about finite products to questions about twofold products using finite induction (for example, this is helpful in proving that if each X_j is finite then the cardinal number of the product is the product of the cardinal numbers).

Certainly one could discuss the pros and cons of both approaches at great length, but our purpose here is to formulate a concept of n-fold cartesian product that is not stated in terms of **either** approach but views each approach as a valid manifestation of a common basic concept. More precisely, we shall

(i) give axioms for the Cartesian product and the projection maps from such a product onto its factors,

(ii) prove that all systems satisfying the product axioms are mathematically equivalent in a strong sense, and

(iii) verify that both constructions satisfy the relevant axioms.

It will follow that both constructions are equally valid ways of defining a Cartesian product even though they might seem rather different when viewed at close range. We now proceed to the first step.

Definition. Let $\{X_\alpha\}$ be an indexed family of sets with indexing set A. A (categorical) **direct product** of the indexed family is pair $(P,\ \{p_\alpha\})$, where P is a set and p_α is a function from P to X_α for each α, such that the following **Universal Mapping Property** holds:

Given an arbitrary set Y and functions $f : Y \to X_\alpha$ for each α, there is a unique function $f : Y \to X$ such that $p_\alpha \circ f = f_\alpha$ for each α.

The adjective "categorical" is added in parentheses because this definition of a direct product comes from category theory. Although we shall not discuss the latter formally at this point, we shall frequently use the category-theoretic characterization of abstract mathematical constructions as structures with suitable "universality properties." The latter turn out to characterize the construction uniquely up to a suitably defined notion of equivalence.

As indicated previously, we would like categorical direct products to be essentially unique. The following results describes the strong uniqueness property that we shall need.

UNIQUENESS THEOREM. *Let $\{X_\alpha\}$ be an indexed family of sets with indexing set A, and let suppose that $(P,\ \{p_\alpha\})$ and $(Q,\ \{q_\alpha\})$ be direct products of the indexed family $\{X_\alpha\}$. Then there is a unique $1-1$ correspondence $h : Q \to P$ such that $p_\alpha \circ h = q_\alpha$ for all α.*

Proof. First of all, we claim that a function $\varphi : P \to P$ is the identity if and only if $p_\alpha \circ \varphi = p_\alpha$ for all α, and likewise $\psi : Q \to Q$ is the the identity if and only if $q_\alpha \circ \psi = q_\alpha$ for all α. These are immediate consequences of the Universal Mapping Property.

Since $(P,\ \{p_\alpha\})$ is a direct product, the Universal Mapping Property implies there is a unique function $h : Q \to P$ such that $p_\alpha \circ h = q_\alpha$ for all α, and likewise since $(Q,\ \{q_\alpha\})$ is a direct product, the Universal Mapping Property implies there is a unique function $f : P \to Q$ such that $q_\alpha \circ h = p_\alpha$ for all α. We claim that h and f are inverse to each other; this is equivalent to the identities $h \circ f = \mathrm{id}_Q$ and $f \circ h = \mathrm{id}_P$.

To verify these identities, first note that for all α we have

$$p_\alpha \quad = \quad q_\alpha \circ h \quad = \quad p_\alpha \circ f \circ h$$

for all α and similarly

$$q_\alpha \quad = \quad p_\alpha \circ f \quad = \quad q_\alpha \circ h \circ f$$

for all α. By the observations in the first paragraph of the proof, it follows that $f \circ h = \mathrm{id}_P$ and $h \circ f = \mathrm{id}_Q$.■

We now need to show that the axioms for categorical direct products hold for the constructions described above (and also for the usual cartesian product of two sets!).

PROPOSITION. *The following are examples of direct products:*

(*i*) *The usual Cartesian product of two sets together with the usual coordinate projection maps into the factors.*

(*ii*) *The Comprehensive Definition of products for indexed families of sets with the projections p_α defined so that $p_\alpha(\mathbf{y})$ is the unique ordered pair (x, β) in $\mathbf{y}$ whose second coordinate is equal to α.*

(*iii*) *The Recursive Definition of products for indexed families of finite sets*

$$X_1 \times \ \cdots \ \times X_{n+1} = (X_1 \times \ \cdots \ \times X_n) \times X_{n+1}$$

with the projections p_j defined as follows: If π and ρ denote the projections of the right hand side onto the factors $X_1 \times \cdots \times X_n$ and X_{n+1}, and q_j is the projection of $X_1 \times \cdots \times X_n$ onto X_j for $1 \leq j \leq n$, then $p_j = q_j \circ \pi$ for $j \leq n$ and $p_{n+1} = \rho$.

If we combine this proposition with the Uniqueness Theorem, we see that the Comprehensive and Recursive Definitions of direct product yield essentially the same object.

Note on terminology. In part (iii) of the proposition, the projections π and ρ come from the definition of a twofold Cartesian product, and the projections q_j are just the coordinate projections associated to the product structure on $X_1 \times \cdots \times X_n$.

Proof of Proposition. We shall prove the three parts in order. All of the arguments are elementary, but we include all the details because of their importance.

PROOF OF (i) : Let A and B be two sets, and let $p_A : A \times B \to A$ and $p_B : A \times B \to B$ be projections onto the first and second factors. To prove the Universal Mapping Property, suppose that X is a set and $f_A : X \to A$ and $f_B : X \to B$ be functions. Then the map $f : X \to A \times B$ defined by

$$f(x) \quad = \quad \big(f_A(x),\, f_B(x) \big)$$

satisfies $p_A \circ f = f_A$ and $p_B \circ f = f_B$. Furthermore, if $g : X \to A \times B$ is an arbitrary function such that $p_A \circ g = f_A$ and $p_B \circ g = f_B$, then we may verify that $g = f$ as follows: Given $x \in X$ write $g(x) = (u, v)$. We then have

$$u \quad = \quad p_A \circ g(x) \quad = \quad f_A(x) \qquad , \qquad v \quad = \quad p_B \circ g(x) \quad = \quad f_B(x)$$

which shows that

$$g(x) \quad = \quad (u, v) \quad = \quad \big(f_A(x),\, f_B(x) \big) \quad = \quad f$$

for all $x \in X$. But this implies $g = f$ and verifies the uniqueness statement.∎

PROOF OF (ii) : Given the indexed family of sets $\{X_\alpha\}$, let p_α be the projection from $P = \prod_\alpha X_\alpha$ to X_α. Suppose that we are given functions $f_\alpha : Y \to X_\alpha$. If we define $f : Y \to P$ by taking $f(y)$ to be the set of all ordered pairs of the form $(f_\alpha(y),\, \alpha)$, then by construction we know that $p_\alpha \circ f = f_\alpha$. Suppose now that $g : Y \to P$ is an arbitrary function such that $p_\alpha \circ g = f_\alpha$ for all α. Given an arbitrary $y \in Y$ its image $g(y)$ consists of pairs of the form (u_β, β) where $\beta \in A$, and by definition $p_\alpha \circ g(y) = u_\alpha$ for all α. On the other hand, since $p_\alpha \circ f_\alpha$ we must have $u_\alpha = f_\alpha(y)$. This implies that g must be equal to f and hence f must be unique.∎

PROOF OF (iii) : Suppose that we are given functions $f_j : Y \to X_j$ for $1 \leq j \leq n+1$. Then there is a unique function $F : Y \to X_1 \times \cdots \times X_n$ such that $q_j \circ F = f_j$ for $1 \leq j \leq n$. By the Universal Mapping Property for twofold cartesian products, it follows that there is a unique function

$$f : Y \to (X_1 \times \cdots \times X_n) \times X_{n+1}$$

such that $\pi \circ f = F$ and $\rho \circ f = f_{n+1}$.

It follows immediately that

$$p_j \circ f \quad = \quad q_j \circ \pi \circ f \quad = \quad q_j \circ f_j$$

for $j \leq n$ and $p_{n+1} \circ f = \rho \circ F = f_{n+1}$. To conclude the proof we need to show that any map g satisfying $p_j \circ g = f_j$ for all j must be equal to f. Let $G : Y \to X_1 \times \cdots \times X_n$ be the composite $\pi \circ g$. We claim that $G = F$; this is true because $q_j \circ G = q_j \circ \pi \circ g = p_j \circ g = f_j$ if $1 \leq j \leq n$.

By the Universal Mapping Property for twofold products, we know that $f = g$ if $\pi \circ f = \pi \circ g$ and $\rho \circ f = \rho \circ g$. We have just verified the first of these conditions, and the second holds because $\rho = p_{n+1}$ and we know that both $\rho \circ f$ and $\rho \circ g$ are equal to f_{n+1}.■

ANALOGOUS RESULTS FOR TOPOLOGICAL PRODUCTS. For our purposes the following observation is important:

The preceding discussion goes through for products of topological spaces with only minor changes.

In order to verify this, one must replace sets by topological spaces throughout the discussion and stipulate that all functions be continuous. With these modifications the proof of the Uniqueness Theorem goes through unchanged. To prove analogs of the previous proposition and show that the usual constructions on topological spaces yield topological direct products, one uses the fact that a map from a space into the product is continuous if and only if its projections onto each of the factors is continuous. Specifically, in (i) and (ii) this implies that the "universal" map f is continuous, while in (iii) this principle shows in turn that F and f are continuous. Writing out the complete proof is left to left to the reader as an exercise.

In the discussion for sets we mentioned an advantage of the recursive definition for set-theoretic products, and we shall close by noting one advantage of the recursive definition for topological products: Namely, it reduces the proof that a finite product of connected spaces is connected to the case of a product with two factors.

I.4.3 : *Quotient spaces and quotient maps* $(2\star)$

Both [MUNKRES1] and the ONLINE 205A NOTES treat quotient topologies, but there is a difference in emphasis: In the ONLINE 205A NOTES the central concept is the quotient space while in [MUNKRES1] the central concept is the quotient map. These are related as follows:

> *If $\mathcal{R}$ is an equivalence relation on a topological space and $X/\mathcal{R}$ is the set of $\mathcal{R}$-equivalence classes with the quotient topology, then the equivalence class projection $\pi(\mathcal{R}) : X \to X/\mathcal{R}$ is a quotient map. Conversely, if $f : X \to Y$ is a (surjective) quotient map then there is a unique equivalence relation $\mathcal{R}_f$ on X and a unique homeomorphism $h : X/\mathcal{R}_f \to Y$ such that $f = h \circ \pi(\mathcal{R}_f)$.*

This is essentially the definition on page 83 in the ONLINE 205A NOTES.

As in the preceding discussion of products, the approach to quotients in [MUNKRES1] emphasizes the role of morphisms rather than the underlying spaces. We would like to carry this further and give another description of quotients that characterizes their behavior with respect to continuous mappings with even less reliance on the definition of the quotient topology.

PROPOSITION. *Let X and Y be topological spaces, and let $f : X \to Y$ be a continuous onto map. Then f is a quotient map if and only if the it has the following* Universal Mapping Property*: If $g : X \to Z$ is a continuous mapping such that $f(x) = f(x') \Longrightarrow g(x) = g(x')$, then there is a unique continuous mapping $h : Y \to Z$ such that $g = h \circ f$.*

Proof. The $(\Longrightarrow)$ implication is an immediate consequence of the proposition and definition on page 83 of the ONLINE 205A NOTES. Conversely, suppose that the Universal Mapping Property holds. Let $\mathcal{R}$ be the equivalence relation $\mathcal{R}_f$, and let $p : X \to X/\mathcal{R}$ be the continuous projection onto the set of equivalence classes. As noted in Section V.1 of the ONLINE 205A NOTES. we then have unique continuous map $h : X/\mathcal{R} \to Y$ such that $f = h \circ p$, and this map is 1–1 onto.

On the other hand, since f has the Universal Mapping Property we also know that there is a unique continuous map $k : Y \to X/\mathcal{R}$ such that $p = k \circ f$. We then have

$$f \quad = \quad h \circ p \quad = \quad h \circ k \circ f$$

and therefore by the Universal Mapping Property it follows that $h \circ k = \mathrm{id}_Y$. Since h is 1–1 onto, it follows that k must be a set-theoretic inverse to h, and since we already know k is continuous it also follows that h is a homeomorphism. Therefore f is a quotient map if it has the Universal Mapping Property.■

I.4.4 : *Subspaces and topological embeddings* $(2\star)$

There is a similar characterization of subspace topologies in terms of morphisms. This is less important to point set topology than the morphism characterization of quotients, but the analog for smooth manifolds will be useful later in this course.

The crucial idea is a concept of *topological embedding* that fits into an analogy

TOPOLOGICAL EMBEDDING *is to* SUBSPACE INCLUSION *as*

QUOTIENT MAP *is to* QUOTIENT SPACE PROJECTION.

Formally, we do this as follows:

Definition. Let X and Y be topological spaces, and let $f : X \to Y$ be a continuous 1–1 map. Then f is a *topological embedding* if the only open subsets of X are sets of the form $f^{-1}(U)$ where U is open in Y.

BASIC EXAMPLES. If A is a subset of Y, then the inclusion $i : A \subset Y$ is a topological embedding by definition of the subspace topology.

We then have the following result relating topological embeddings to topological subspaces that parallels the previously stated result relating quotient maps to topological quotient spaces.

PROPOSITION. *If $f : X \to Y$ is a topological embedding then there is a unique subset $A \subset Y$ and a unique homeomorphism $h : A \to X$ such that $f \circ h$ is the inclusion of A in Y.*

Proof. The argument is extremely straightforward. Let $A = f(X)$ and take h to be the 1–1 correspondence whose inverse $k : X \to A$ is defined by the formula $k(x) = f(x)$; the only difference between k and f is that the codomain of k is A while the codomain of f is Y. It is a routine exercise to verify that A is unique and that h is the unique 1–1 correspondence of sets from A to X such that $f \circ h$ is the inclusion of A in Y.

It remains to verify that h is a homeomorphism. By hypothesis, a set U is open in X if and only if $U = f^{-1}(W_0)$ where W_0 is open in Y. By definition of the subspace topology, $V \subset A$ is open if and only if V has the form $W_1 \cap A$ where W_1 is open in A. We use this first to show that h is continuous: If U is open in X and $U = f^{-1}(W_0)$ as above, then

$$h^{-1}(U) \quad = \quad h^{-1}\left(f^{-1}(W_0)\right) \quad = \quad (f \circ h)^{-1}(W_0) \quad = \quad A \cap W_0$$

shows that h is continuous. We next show that h is open. If U is open in A, then by definition $U = W_1 \cap A$ where W_1 is open in Y. If $V = h(U)$ then $U = h^{-1}(V)$, and to show that h is open we need to show that V is the inverse image of an open subset of Y. But now we have

$$V \quad = \quad h(U) \quad = \quad h\,(W_1 \cap A) \quad = \quad h\,\left((f \circ h)^{-1}(W_1)\right) \quad =$$

$$h\left(h^{-1}\left(f^{-1}(W_1)\right)\right) \quad = \quad f^{-1}(W_1)$$

which shows that V is open in X.■

We also have a characterization of topological embeddings that is similar to the characterization of quotient maps.

PROPOSITION. *Let X and Y be topological spaces, and let $f : X \to Y$ be a continuous $1-1$ map. Then f is a topological embedding if and only if the it has the following* Universal Mapping Property*: If $g : Z \to Y$ is a continuous mapping such that $g(Z) \subset f(X)$, then there is a unique continuous mapping $h : Z \to X$ such that $g = f \circ h$.*

Proof. Suppose that f is a topological embedding, and suppose also that $g : Z \to Y$ is a continuous mapping such that $g(Z) \subset f(X)$. Define a set-theoretic map $h : Z \to X$ by setting $h(z)$ equal to the unique $x \in X$ such that $f(x) = g(z)$; the existence of such an x is ensured by the condition $g(Z) \subset f(X)$, and it is unique because f is 1–1. CLAIM: h is continuous. Suppose that U is open in X. By definition we know that $U = f^{-1}(V)$ where V is open in Y. We then have

$$h^{-1}(U) \quad = \quad h^{-1}\left(f^{-1}(U)\right) \quad = \quad g^{-1}(U)$$

$$(\text{because } g \quad = \quad f \circ h)$$

and since g is continuous the set in the displayed equations is open. This proves that h is continuous.

Conversely, suppose that the Universal Mapping Property holds for f. Let $A = f(X)$ and let $j; A \to Y$ be the inclusion map, where A has the subspace topology. We know that j is a topological embedding and hence by the first part of the proof it also has the Universal Mapping Property. Therefore we have unique continuous mappings $h : A \to X$ and $k : X \to A$ such that $f \circ h = j$ and $j \circ k = f$.

Taking composites, we find that $f = f \circ h \circ k$ and $j = j \circ k \circ h$. If we apply the Universal Mapping Property for $f : X \to Y$ in the special case $g = f$, we see that the only continuous map $\varphi : X \to X$ satisfying $f \circ \varphi$ is the identity, and therefore $h \circ k$ must be the identity on X. Similarly, the only continuous map $\psi : A \to A$ satisfying $j \circ \psi$ is the identity, and therefore $k \circ h$ must be the identity on A. It follows that h and k must be inverses to each other, and since both are continuous they are homeomorphisms.

It follows that U is open in X if and only if $k(U)$ is open in A, and the latter is true if and only if $k(U) = W \cap A$ for some open subset in Y. The preceding condition is in turn equivalent to saying that $U = k^{-1}(W \cap A)$ for some open subset $W \subset Y$; since $W \cap A = j^{-1}(W)$, the condition on U is true if and only if $U = k^{-1}\left(j^{-1}(W)\right)$ for some W open in Y. Since $f = j \circ k$, it follows that the right hand side of the preceding equation is equal to $f^{-1}(W)$, and from this we see that U is open in X if and only if $U = f^{-1}(W)$, where W is open in Y. This is exactly the defining condition for a topological embedding, and consequently we have shown that f is indeed a topological embedding if it has the Universal Mapping Property.■

II. Local theory of smooth functions

This unit contains the basic results from multivariable calculus that we shall need. Some of this material already appears in second year advanced calculus courses, other topics may be mentioned in such courses but not in much detail, and still others rarely if ever appear in undergraduate courses. In any case, the approach differs significantly from the usual one in advanced calculus courses; we shall generally try to do everything in a coordinate free manner whenever this is possible and not unreasonably opaque or awkward.

To motivate the emphasis on a coordinate-free approach, it is worthwhile to look back at undergraduate linear algebra. Coordinates are absolutely necessary for computational purposes, but for many other purposes it is often necessary or useful to find a different approach that avoids the massive complications that coordinates often generate. Perhaps the most basic examples involve the use of linear transformations to analyze many fundamental questions about matrices. Since linear approximations to smooth functions are a fundamental concept for this course, the value of a coordinate-free approach to multivariable calculus should be apparent.

II.1 : Differentiability

(Conlon, §§ 2.1, 2.3–2.4)

Much of this section may be a review of familiar concepts, but we are starting at a very basic level for the sake of completeness.

II.1.1 : *Linear approximations*

For functions of one real variable, a function is continuous at a point if it has a derivative at that point, but there are standard examples of functions of two variables that have partial derivatives defined near a point but are not continuous there. On the other hand, a basic result in multivariable calculus shows that functions that have continuous partial derivatives near a point are necessarily continuous at the point in question.

We shall begin by establishing a version of this result that holds for functions of an arbitrary (finite) number of real variables. It will be convenient to adopt some notation first. The unit vector in $\mathbf{R}^n$ whose i^{th} coordinate is 1 and whose other coordinates are 0 will be denoted by $\mathbf{e}_i$. If U is an open set in $\mathbf{R}^n$ and $f : U \to \mathbf{R}^n$ is a (not necessarily continuous) function such that all first partial derivatives exist at some point $p \in U$, then the *gradient* $\nabla f(p)$ will denote the vector whose i^{th} coordinate is the partial derivative of f with respect to the i^{th} variable at p. We shall use $\langle u, v\rangle$ to denote the usual dot product of two vectors $u, v \in \mathbf{R}^n$.

PROPOSITION. *Let U be an open subset in $\mathbf{R}^n$, let $x \in \mathbf{R}^n$, and let $f : U \to \mathbf{R}^n$ be a (not necessarily continuous) function such that f has continuous partial derivatives on some open subset of U containing x. Then for all sufficiently small $h \neq 0$ in $\mathbf{R}^n$ one can define a function $\theta(h)$ such that*

$$f(x+h) - f(x) \quad = \quad \langle \nabla f(x), h\rangle \; + \; |h|\,\theta(h)$$

where $\lim_{h\to 0}\theta(h) = 0$.

Proof. Write $h = \sum_i t_i h_i$ for suitable real numbers t_i, take $\delta > 0$ so that f has continuous partial derivatives on $N_\delta(x)$, and assume that $0 < |h| < \delta$. Define h_i for $0 \leq i \leq n$ recursively by $h_0 = 0$ and $h_{i+1} = h_i + t_{i+1}\mathbf{e}_{i+1}$. Then $h_n = h$ and we have

$$f(x+h) - f(x) = \sum_i f(x+h_i) - f(x_{i-1})$$

and if we apply the ordinary Mean Value Theorem to each summand we see that the right hand side is equal to

$$\sum_i \left(\frac{\partial}{\partial x_i} f\left(x + h_{i-1} + K_i(x)t_i\right)\right) \cdot t_i$$

for some numbers $K_i(x) \in (0,1)$. The expression above may be further rewritten in the form

$$\langle \nabla f(x), h\rangle \quad + \quad \sum_i \left(\frac{\partial}{\partial x_i} f\left(x + h_{i-1} + K_i(x)t_i\right) - \frac{\partial}{\partial x_i} f(x)\right) \cdot t_i$$

and therefore an upper estimate for $|f(x+h) - f(x) - \langle \nabla f(x), h\rangle\ |$ is given by

$$\sum_i \left|\frac{\partial}{\partial x_i} f\left(x + h_{i-1} + K_i(x)t_i\right) - \frac{\partial}{\partial x_i} f(x)\right| \ \cdot\ |t_i|.$$

Since the partial derivatives of f are all continuous at x, for every $\varepsilon > 0$ there is a $\delta_1 > 0$ such that $\delta_1 < \delta$ and $|h| < \delta_1$ implies that each of the differences of partial derivatives has absolute value less than ε/n. Since $h \neq 0$, if we define $\theta(h)$ as in the statement of the conclusion, the preceding considerations imply that

$$|\theta(h)| \quad < \quad \sum_i \frac{\varepsilon \cdot |t_i|}{n \cdot |h|}$$

and since

$$|t_i| \quad \leq \quad \left(\sum_i t_i^2\right)^{1/2}$$

it follows that $|\theta(h)| < \varepsilon$ when $0 < |h| < \delta_1$.■

COROLLARY. *If f satisfies the conditions of the proposition, then f is continuous at x.*■

The conclusion of the theorem indicates the right generalization of differentiability to functions of more than one variable:

Definition. Let U be an open subset in $\mathbf{R}^n$ and let $f : U \to \mathbf{R}^n$ be a function (with no further assumptions at this point). The function f is said to be *differentiable* at the point $x \in U$ if there is a linear transformation $L : \mathbf{R}^n \to \mathbf{R}^m$ such that for all sufficiently small vectors h we have

$$f(x+h) - f(x) \quad = \quad L(h) \ + \ |h|\,\theta(h)$$

where $\lim_{h\to 0}\theta(h) = 0$.

Immediate Consequence. *If $m = 1$ in the definition above and we restrict h so that $h = t\mathbf{e}_i$ for $|t|$ sufficiently small then if f is differentiable at x it follows that all first partial derivatives of f are defined at x and*

$$\frac{\partial f(x)}{\partial x_i} \quad = \quad L(\mathbf{e}_i)\ .$$

In particular, the preceding shows that there is at most one choice of L for which the differentiability criterion is true, at least if $m = 1$.■

Of course the proposition above implies that a function is differentiable if it has continuous partial derivatives.

The differentiability of a function turns out to be determined completely by the differentiability of its coordinate functions:

PROPOSITION. *let U be open in $\mathbf{R}^n$, let $f : U \to \mathbf{R}^m$ be a function, and express f in coordinates as $f(x) = \sum_j y_j(x)\mathbf{e}_j$. Then f is differentiable at x if and only if each y_j is differentiable at x, and in this case the linear transformation L is given by*

$$L(u) \quad = \quad \sum_i \langle \nabla y_i(x), u\rangle \, \mathbf{e}_i \ .$$

It follows that there is at most one choice of L for an arbitrary value of m. If we write $u = \sum_j u_j \mathbf{e}_j$ then this yields the fundamental identity

$$L(u) \quad = \quad \sum_i \sum_j \frac{\partial y_i(x)}{\partial x_j} \, u_j \mathbf{e_i}$$

for the derivative linear transformation. In words, the (i,j) entry of the matrix representing L is the j^{th} partial derivative of the i^{th} coordinate function.

Proof of proposition. Suppose that f is differentiable at x. For i between 1 and m, let $\mathbf{P}_i$ be projection onto the i^{th} coordinate. If we apply $\mathbf{P_i}$ to the formula for $f(x+h) - f(x)$ we obtain the following relationship:

$$y_i(x+h) - y_i(x) \quad = \quad \mathbf{P_i}\big(\, L(h)\,\big) \ + \ |h| \cdot \mathbf{P_i}\big(\,\theta(h)\,\big)$$

The composite $\mathbf{P}_i L$ is linear because both factors are linear, and the relation

$$\big|\mathbf{P_i}\theta(h)\big| \quad \leq \quad \big|\theta(h)\big|$$

shows that $\lim_{h\to 0} \mathbf{P_i}\theta(h) = 0$, so that $y_i = \mathbf{P}_i f$ is differentiable at X and

$$\nabla y_i(x) \quad = \quad \mathbf{P}_i \circ L$$

as required.

Now suppose that each y_i is differentiable at x, and write

$$f(x+h) - f(x) \quad = \quad \sum_i \big(y_i(x+h) - y_i(x)\big)\mathbf{e_i} \quad =$$

$$\sum_i \, \langle \nabla y_i(x), h\rangle \mathbf{e_i} \ + \ \sum_i \, |h|\theta_i(h)\mathbf{e_i}$$

where $h = \sum_j \, t_j \mathbf{e}_j$ and $\lim_{h\to 0} \theta_i(h) = 0$ for all i. Choose $\delta > 0$ so that $N_\delta(x) \subset U$ and $0 < |h| < \delta$ implies $|\theta_i(h)| < \varepsilon/n$ for all i. Then $0 < |h| < \delta$ implies $|\theta(h)| < \varepsilon$, whosing that f is differentiable at x and the linear transformation L has the form described in the proposition.■

II.1.2 : *Smoothness classes of functions*

If U and V are open sets in Euclidean spaces and $f : U \to V$ is a function, then we say that f is *(smooth of class)* $\mathcal{C}^1$ if Df exists everywhere and is continuous. For $r \geq 2$ we inductively define f to be *(smooth of class)* $\mathcal{C}^r$ if Df is (smooth of class) $\mathcal{C}^{r-1}$, and we say that f is *(smooth of class)* $\mathcal{C}^\infty$ if it is smooth of class $\mathcal{C}^r$ for all positive integers r. For the sake of notational uniformity we often say that every continuous function is of class $\mathcal{C}^0$.

It is elementary to check that a function is of class $\mathcal{C}^r$ if and only if all its coordinate functions are and that

$$\mathcal{C}^\infty \implies \mathcal{C}^r \implies \mathcal{C}^{r-1} \implies \mathcal{C}^0$$

for all r. Every polynomial function is obviously of class $\mathcal{C}^\infty$, and for each r there are many examples of functions that are $\mathcal{C}^r$ but not $\mathcal{C}^{r+1}$. If $r = 0$ the absolute value function $f_0(x) = |x|$ is an obvious example, and inductively one can construct an example f_r which is $\mathcal{C}^r$ but not $\mathcal{C}^{r+1}$ by taking an antiderivative of f_{r-1}.

One important point in ordinary and multivariable courses is that standard algebraic operations on differentiable (or smooth) functions yield differentiable (or smooth) functions. In particular, this applies to addition, subtraction, multiplication, and division (provided the denominator is nonzero in this case). We shall use these facts without much further comment. The smoothness properties of composites of differentiable and smooth functions will be discussed fairly soon.

II.1.3 : *Matrix operations*

Plenty of examples of smooth functions can be found in multivariable calculus books, so we concentrate here on some basic examples that will be needed shortly.

PROPOSITION. *Addition addition of $m \times n$ matrices is a $\mathcal{C}^\infty$ map from*

$$\mathbf{R}^{2mn} \cong (\mathbf{M}(m,n))^2$$

to $\mathbf{R}^{mn} \cong \mathbf{M}(m,n)$, scalar multiplication of $m \times n$ matrices is a $\mathcal{C}^\infty$ map from

$$\mathbf{R}^{mn+1} \cong \mathbf{R} \times \mathbf{M}(m,n)$$

to $\mathbf{R}^{mn} \cong \mathbf{M}(m,n)$, transposition of $m \times n$ matrices defines a $\mathcal{C}^\infty$ map from $\mathbf{M}(m,n)$ to $\mathbf{M}(n,m)$, and matrix multiplication from $\mathbf{M}(m,n) \times \mathbf{M}(n,p)$ to $\mathbf{M}(m,p)$ is also a $\mathcal{C}^\infty$ map.

This simply reflects the fact that that the entries of a matrix sum or product are given by addition and multiplication operations on the entries of the original matrices (or matrix and scalar).■

The next result is slightly less trivial but still not difficult.

PROPOSITION. *The set of invertible $n \times n$ matrices $GL(n,\mathbf{R})$ is an open subset of $\mathbf{R}^{n^2} \cong \mathbf{M}(n,n)$, and the map from $GL(n,\mathbf{R})$ to itself sending a matrix to its inverse is a $\mathcal{C}^\infty$ map.*

Proof. We shall prove this using coordinates; it is possible to prove the result without using coordinates, but the proof using coordinates is shorter and simpler.

Recall that the determinant of a square matrix is a polynomial function in the entries of the matrix and that a matrix is invertible if and only if its determinant is nonzero. The former implies that the determinant function is continuous (and in fact $\mathcal{C}^\infty$), while the second observation and the continuity of the determinant imply that the set of invertible matrices, which is equal to the set $\det^{-1}(\mathbf{R} - \{0\})$, is open. But Cramer's Rule implies that the entries of the inverse to a matrix are rational expression in the entries of the original matrix, and thus the entries of an inverse matrix are $\mathcal{C}^\infty$ functions of the entries of the original matrix. Therefore the matrix inverse is a $\mathcal{C}^\infty$ function from $GL(n, \mathbf{R})$ to itself.■

II.1.4 : *Matrix norms* ($\star$)

Before discussing the metric and topological properties of $\mathcal{C}^r$ functions it is necessary to know a little about the corresponding properties of their linear approximations. The most basic property is a strong form of uniform continuity.

PROPOSITION. *If $L : \mathbf{R}^n \to \mathbf{R}^m$ is a linear transformation, then there is a constant $b > 0$ such that $|L(u)| \leq b\,|u|$ for all $U \in \mathbf{R}^n$.*

Proof. The easiest way to see this is to note that L is continuous, and therefore the restriction of the function $h(u) = |L(u)|$ to the (compact!) unit sphere in $\mathbf{R}^n$ assumes a maximum value, say c, so that $|u| = c$. Every vector u may be written as a product cv where c is nonnegative and $|v| = 1$. It then follows that

$$|L(u)| \quad = \quad |cL(v)| \quad = \quad c \cdot |L(v)| \quad = \quad |u| \cdot |L(v)| \leq b \cdot |u|$$

as required.■

Notation. The maximum value b is called the *norm* of L and written $\parallel L \parallel$.

PROPOSITION. *The norm of a matrix (or linear transformation) makes the space of $m \times n$ matrices into a normed vector space.*

Proof. ($\star$) By definition the norm is nonnegative, and if it is zero then $L(v) = 0$ for all $v \in \mathbf{R}^n$ satisfying $|v| = 1$; it follows that $L(v) = 0$ for all v (why?). If a is a scalar and L is a linear transformation, then the maximum value $\parallel L \parallel$ of $|aL(v)|$ for $|v| = 1$ is simply $|a| \cdot \parallel L \parallel$. Finally, if L_1 and L_2 are linear transformations and v is a unit vector such that $|[L_1 + L_2](v)| = \parallel L_1 + L_2 \parallel$, then we have

$$\parallel L_1 + L_2 \parallel \quad = \quad \big|\,[L_1 + L_2](v)\,\big| \quad \leq \quad |L_1(v)| + |L_2(v)| \quad \leq \quad \parallel L_1 \parallel + \parallel L_2 \parallel \quad .$$

Thus the norm as defined above satisfies the conditions for a normed vector space.■

The matrix norm has the following additional useful property:

PROPOSITION. *If A is an $m \times n$ matrix and B is an $n \times p$ matrix, then we have*

$$\parallel AB \parallel \quad \leq \quad \parallel A \parallel \cdot \parallel B \parallel \quad .$$

Proof. ($\star$) Let v be a unit vector in $\mathbf{R}^p$ at which the function $f(x) = ABx$ takes a maximum value. Then we have

$$\parallel AB \parallel \quad = \quad |ABx| \quad \leq \quad \parallel A \parallel \cdot |Bx| \quad \leq \quad \parallel A \parallel \cdot \parallel B \parallel$$

as required.■

II.1.5 : *Comparisons of norms* $(\star)$

Although there are many different norms that can be defined on $\mathbf{R}^n$, the following result shows that they all yield the same open sets.

THEOREM. *Let $|...|$ denote the standard Euclidean norm on $\mathbf{R}^n$, and let $\| \ldots \|$ denotes some other norm. Then there are positive constants A and B such that*

$$\| x \| \leq A\,|x| \ , \quad |x| \leq B \,\| x \|$$

for all $x \in \mathbf{R}^n$. In particular, the identity maps of normed vector spaces from Euclidean space $(\mathbf{R}^n, |...|)$ to $(\mathbf{R}^n, \| \ldots \|)$ and vice versa are uniformly continuous.

Proof. $(\star)$ Given a typical vector x, write it as $\sum_i x_i \mathbf{e}_i$.

Choose M such that $\| \mathbf{e}_i \| \leq M$ for all i. Then we have

$$\| x \| \leq \sum_i |x_i| \cdot \| \mathbf{e}_i \| \leq nM \sum_i |x_i|$$

and by the Cauchy-Schwarz-Buniakovsky Inequality the summation is less than or equal to $n^{1/2}|x|$. Therefore $\| x \| \leq n^{3/2} M |x|$.

The preceding paragraph implies that the function $f(x) = \| x \|$ is a continuous function on $\mathbf{R}^n$ with respect to the usual Euclidean metric. Let $c > 0$ be the minimum value of f on the unit sphere defined by $|x| = 1$. It then follows that $\| x \| \geq c \cdot |x|$ for all x and hence that

$$|x| \leq \frac{1}{c} \| x \|$$

for all $x \in \mathbf{R}^n$.∎

COROLLARY. *The same conclusion holds if the Euclidean norm is replaced by a second arbitrary norm.*

Proof. $(\star)$ Let α and β denote arbitrary norms on $\mathbf{R}^n$ and let $|...|$ denote the Euclidean norm. Then there are positive constants A_α, A_β, B_α, B_β such that the following hold for all $x \in \mathbf{R}^n$:

$$\alpha(x) \leq A_\alpha |x| \ , \quad |x| \leq B_\alpha \,\alpha(x)$$

$$\beta(x) \leq A_\beta |x| \ , \quad |x| \leq B_\beta \,\beta(x)$$

These immediately imply $\alpha(x) \leq A_\alpha B_\beta \beta(x)$ and $\beta(x) \leq A_\beta B_\alpha \alpha(x)$.∎

The following observation will be useful later.

PROPOSITION. *If A is an $n \times n$ matrix such that $\| A \| < 1$, then $I - A$ is invertible.*

Proof. $(\star)$ Suppose that the conclusion is false, so that $I - A$ is not invertible. Then there is a nonzero vector $v \in \mathbf{R}^n$ such that $(I - A)v = 0$. The latter implies that $Ax = x$ for some x such that $|x| = 1$, which in turn implies that $\| A \| \geq 1$.∎

Note. If $\| A \| < 1$, then the the previously stated inequalities for the matrix norm show that

$$\| A^k \| \leq \| A \|^k$$

and the latter implies that the inverse to $I - A$ may be computed using the geometric series:

$$(I - A)^{-1} \quad = \quad \sum_k A^k$$

II.1.6 : *The Chain Rule*

We shall now generalize two important principles from elementary calculus. One is a version of the Chain Rule, and the other is a general form of a basic consequence of the Mean Value Theorem. As an application of these results we shall show that the restriction of a smooth map to a compact set satisfies a metric inequality called a *Lipschitz condition* that generalizes the strong form of uniform continuity which holds for linear maps of (finite-dimensional) Euclidean spaces.

We begin with a multivariable version of the Chain Rule. Undergraduate multivariable calculus courses generally state one or more extensions of the single variable formula

$$[g \circ f]'(x) \quad = \quad g'(f(x)) \cdot f'(x)$$

to functions of several real variables. Linear transformations provide a conceptually simple way of summarizing the various generalizations of this basic fact from single variable calculus:

CHAIN RULE. *Let U and V be open in $\mathbf{R}^n$ and $\mathbf{R}^m$ respectively, let $f : U \to V$ be a map that is differentiable at x, and let $g : V \to \mathbf{R}^p$ be differentiable at $f(x)$. Then $g \circ f$ is differentiable at x and*

$$D[g \circ f](x) \quad = \quad \big(Dg(f(x)) \big) \circ Df(x) \ .$$

Proof. By the definition of differentiability for g at $y = f(x)$, for $|k|$ sufficiently small we have

$$g(y + k) - g(y) \quad = \quad [Dg(y)](k) \ + \ |k| \cdot \alpha(k)$$

where $\lim_{k \to 0} \alpha(k) = 0$. If we take k so that $y + k = f(x + h)$ for $|h|$ sufficiently small, then we have $k = f(x + h) - f(x)$, and by the differentiability of f at x we have

$$k \quad = \quad f(x + h) - f(x) \quad = \quad [Df(x)](h) \ + \ |h| \cdot \beta(h)$$

where $\lim_{h \to 0} \beta(h) = 0$. If we make this substitution into the first equation in the proof we obtain the relation

$$[g \circ f](x + h) \ - \ [g \circ f](x) \quad = \quad [Dg(f(x))]([Df(x)](h) \ + \ |h| \cdot \beta(h)) \ + \ |k| \cdot \alpha(f(x + h) - f(x))$$

and for $h \neq 0$ the right hand side may be rewritten as follows:

$$[Dg(f(x))]([Df(x)](h)) \ + \ [Dg(f(x))](|h| \cdot \beta(h)) \ + \ |h| \cdot \frac{|k|}{|h|} \, \alpha\big(f(x + h) - f(x)\big)$$

Let $\varepsilon > 0$ be given. We need to show there is a $\delta > 0$ such that $|h| < \delta$ implies the following inequalities:

$$\| \, Dg\,(f(x)) \, \| \cdot |\beta(h)| \quad < \quad \frac{\varepsilon}{2}$$

$$\frac{|k|}{|h|} \quad < \quad \| \, Df(x) \, \| \; + 1$$

$$\bigl|f(x+h) - f(x)\bigr| \quad < \quad \frac{\varepsilon}{2(\| \, Df(x) \, \| \; + 1)}$$

The first of these can be realized by the limit condition on β, and the third can be realized by the limit condition on α and the continuity of f at x. To deal with the second condition, note that

$$|k| \quad = \quad \bigl|f(x+h) - f(x)\bigr| \quad = \quad \bigl|\,[Df(x)](h) + |h|\beta(h)\,\bigr| \quad \leq$$

$$\| \, Df(x) \, \| \cdot |h| \quad + \quad |h|\bigl|\beta(h)\bigr|$$

so that

$$\frac{|k|}{|h|} \quad \leq \quad \| \, Df(x) \, \| \quad + \quad \bigl|\beta(h)\bigr|$$

and thus we may realize the second inequality if we choose δ so that $0 < |h| < \delta$ implies

$$\bigl|\beta(h)\bigr| \quad < \quad \frac{\varepsilon}{2(\| \, Df(x) \, \| + 1)} \; .$$

If we combine all these we see that

$$[g \circ f](x+h) \;-\; [g \circ f](x) \quad = \quad \bigl[Dg\,(f(x))\bigr]\bigl([Df(x)](h)\bigr) \;+\; |h|\;\gamma(h)$$

where $\lim_{h \to 0} \gamma(h) = 0$.■

COROLLARY. *In the notation of the preceding result, if f is $\mathcal{C}^r$ on U and g is $\mathcal{C}^r$ on V, then $g \circ f$ is $\mathcal{C}^r$.*

Proof. Suppose that $r = 1$. Then the Chain Rule formula, the continuity of the derivatives of f and g and the continuity of f show that $D(g \circ f)$ is continuous.

Suppose now that $r \geq 2$ is an integer and we have shown the Corollary inductively for $\mathcal{C}^s$ functions for $1 \leq s \leq r-1$. Then the functions $Dg \circ f$ and Df are $\mathcal{C}^{r-1}$ by the induction hypothesis and the fact that f is $\mathcal{C}^r$, and the matrix product of these functions is also $\mathcal{C}^r$ because matrix multiplication is $\mathcal{C}^\infty$.■

Since the result is true for all finite r, it follows immediately that the conclusion is also true if $r = \infty$.

Example. Suppose that U is open in $\mathbf{R}^n$ and $f : U \to \mathbf{R}^m$ is $\mathcal{C}^r$ for some $r \geq 1$; let $a, x \in U$ be such that $N_{2|x-a|}(a) \subset U$. Then the function

$$g(t) \quad = \quad f\,(a + (x-a))$$

is a $\mathcal{C}^r$ function on some interval $(-\delta, 1+\delta)$ and

$$g'(t) \quad = \quad \bigl[Df\,(a + t(x-a))\,\bigr](x-a) \; .$$

II.1.7 : *Mean Value Estimate*

The Mean Value Theorem for real-valued differentiable functions of one real variable does not generalize directly to other situations, but some of its important consequences involving derivatives and definite integrals can be extended. The following example is important in many contexts:

PROPOSITION. *Let U be open in $\mathbf{R}^n$, let $f : U \to \mathbf{R}^m$ be a $\mathcal{C}^1$ function, and suppose that $a \in U$ and $\delta > 0$ are such that $|x - a|, |y - a| \leq \delta$ implies $x, y \in U$. Then for all such x we have the following inequality:*

$$|f(x) - f(y)| \quad \leq \quad \max_{|z-a|\leq\delta} \| \, Df(z) \, \| \cdot |x - y|$$

Proof. ($\star$) Let $h = x - y$, and set $g(t) = f(y + th)$; since open disks are convex we know that $y + th$ also satisfies $|y + th - a| \leq \delta$. Then we have

$$f(x) - f(y) \quad = \quad \int_0^1 g'(t) \, dt$$

and therefore

$$\bigl|f(x) - f(y)\bigr| \quad \leq \quad \int_0^1 \bigl|g'(t)\bigr| \, dt \ .$$

As indicated before, by the Chain Rule we know that

$$g'(t) \quad = \quad \bigl[Df\,(y + t(x - y))\,\bigr]\,(x - y)$$

and therefore we have the estimate

$$\int_0^1 \bigl|g'(t)\bigr| \, dt \quad \leq \quad \max{}_{0\leq t\leq 1} \| \, Df\,(y + t(x - y)) \, \| \cdot \bigl|x - y\bigr|$$

which immediately yields the inequality in the proposition.■

II.1.8 : *Lipschitz conditions* ($\star$)

The restriction of a smooth function (say of class $\mathcal{C}^r$) to a compact set satisfies a strong form of uniform continuity that generalizes the matrix inequality $|Ax| \leq \| A \| \cdot |x|$.

THEOREM. *Let U be open in $\mathbf{R}^n$, let $f : U \to \mathbf{R}^m$ be a $\mathcal{C}^1$ function, and let $K \subset U$ be compact. Then there is a constant $B > 0$ such that*

$$|f(u) - f(v)| \ \leq \ B \cdot |u - v|$$

for all $u, v \in K$ such that $u \neq v$.

The displayed inequality is called a *Lipschitz condition* for f. This strong form of uniform continuity associates to each $\varepsilon > 0$ a corresponding δ equal to ε/B. An example of a uniformly continuous function not satisfying any Lipschitz condition is given by $h(x) = \sqrt{x}$ on the closed unit interval $[0, 1]$ (use the Mean Value Theorem and $\lim_{t\to 0^+} h'(t) = +\infty$). Incidentally, the inverse of this map is a homeomorphism that does satisfy a Lipschitz condition (*e.g.*, we can take $B = 2$); verifying this is left as an exercise.

The inequality

$$\Big|u - v\Big| \quad \geq \quad \Big||u| - |v|\Big|$$

for $u, v \in \mathbf{R}$ shows that $f(x) = |x|$ is a function that satisfies a Lipschitz condition but is not $\mathcal{C}^1$.

A *Lipschitz constant* for f on a set K (not necessarily compact) is a number $B > 0$ such that $|f(u) - f(v)| \leq B\,|u - v|$ for all $u, v \in K$ such that $u \neq v$. Note that Lipschitz constants are definitely nonunique; if B is a Lipschitz constant for f on a set K and $C > B$, then C is also a Lipschitz constant for f on a set K.

To summarize the preceding discussion, we have implications

$$\text{smooth } \mathcal{C}^1 \Longrightarrow \text{Lipschitz on compact sets}$$

$$\text{Lipschitz} \Longrightarrow \text{uniformly continuous}$$

but the reverse implications do not hold.

Proof of theorem. $(\star)$ For each $x \in K$ there is a $\delta(x) > 0$ such that $N_{2\delta(x)}(x) \subset U$. By compactness there are finitely many points $x_1, \cdots, x_q$ such that the sets $N_{\delta(x_i)}(x_i)$ cover K. Let B_i be the maximum of $\| Df \|$ for $|y - x_i| \leq \delta(x_i)$. If $B_i = 0$ for all i then $Df = 0$ on an open set containing K and therefore f is constant on K, so that the conclusion of the theorem is trivial. Therefore we shall assume some $B_i > 0$ for the rest of the proof.

By the Mean Value Estimate we know that $y, z \in N_{\delta(x_i)}(x_i)$ implies that $|f(y) - f(z)| \leq B_i|y - z|$.

Let $\eta > 0$ be a Lebesgue number [MUNKRES1, p. 175] for the open covering of K determined by the sets $N_{\delta(x_i)}(x_i)$, and let $M \subset K \times K$ be the set of all points $(u, v) \in K \times K$ such that $|u - v| \geq \eta/2$. The function $\Delta(u, v) = |u - v|$ is continuous on $K \times K$, and consequently it follows that M is a closed and thus compact subset of $K \times K$. Consider the continuous real-valued function on M defined by

$$h(u, v) \quad = \quad \frac{|f(u) - f(v)|}{|u - v|} \ .$$

Since the denominator is positive on M, this is a continuous function and therefore attains a maximum value A.

Let B be the maximum of the numbers $A, B_1, \cdots, B_k$, and suppose that $(u, v) \in K \times K$. If $(u, v) \in M$, then by the preceding paragraph we have $|f(u) - f(v)| \leq A \cdot |u - v|$. On the other hand, if $(u, v) \notin M$, then $|u - v| < \eta/2$ and thus there some i such that $u, v \in N_{\delta(x_i)}(x_i)$. By the Mean Value Estimate we know that $|f(u) - f(v)| \leq B_i \cdot |u - v|$ in this case. Therefore $|f(u) - f(v)| \leq B \cdot |u - v|$ for all u and v.■

II.1.9 : *Higher order derivatives* $(\star)$

We shall conclude this section by giving a coordinate free version of the symmetric identity for mixed second partial derivatives:

$$\frac{\partial^2 f}{\partial x\, \partial y} \quad = \quad \frac{\partial^2 f}{\partial y\, \partial x}$$

(For a discussion and proof of this see pp. 183–184 of [MT], and especially Theorem 1 on page 183.)

Some preliminary discussion of second derivatives from our viewpoint will be helpful. If U is an open subset in $\mathbf{R}^n$ and $f : U \to \mathbf{R}^m$ is a smooth $\mathcal{C}^2$ function, then $D^2 f = D(Df)$ is a continuous map from U to the vector space

$$\mathcal{L}(\mathbf{R}^n,\ \mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)\ )\ .$$

Given an element Φ of this vector space and two vectors u and v in $\mathbf{R}^n$, one obtains a vector $\Phi^{\$}(u,v) \in \mathbf{R}^m$ by the formula

$$\Phi^{\#}(u,v) \quad = \quad [\Phi(u)]\,(v)$$

and *this expression is a* bilinear *function of u and v (linear in each variable) with values in $\mathbf{R}^m$.*

From our perspective, equality of mixed partials is expressible as follows:

PROPOSITION. *In the notation of the preceding paragraph we have*

$$\left[D^2 f(x)\right]^{\#}(u,v) \quad = \quad \left[D^2 f(x)\right]^{\#}(v,u)$$

for all $x \in U$ and $u, v \in \mathbf{R}^m$; in other words $\left[D^2 f(x)\right]^{\#}(u,v)$ is a SYMMETRIC *bilinear function.*

Proof. We begin with some standard linear algebra. Suppose that $\mathcal{E} = \{\mathbf{e}_1, \ \cdots\ , \mathbf{e}_n\}$ is the standard basis of unit vectors for $\mathbf{R}^n$. Express the vectors u and v as linear combinations $u = \sum_i u_i\, \mathbf{e}_i$ and $v = \sum_i v_i\, \mathbf{e}_i$ respectively. If φ is an arbitrary bilinear form on $\mathbf{R}^n$ with values in some other vector space, then bilinearity yields the following identity

$$\varphi(u,v) \quad = \quad \sum_{j,k} u_j\, v_k\, \varphi(\mathbf{e}_j,\, \mathbf{e}_k)$$

which shows that φ is completely specified by its values at pairs of standard unit vectors. CLAIM: The bilinear function φ is symmetric if and only if

$$\varphi(\mathbf{e}_j,\, \mathbf{e}_k) \quad = \quad \varphi(\mathbf{e}_k,\, \mathbf{e}_j)$$

for all j and k. — This follows by applying the preceding observations to the given bilinear form φ and the reverse order form $\varphi^{\mathbf{op}}$ defined by the rule

$$\varphi^{\mathbf{op}}(u,v) \quad = \quad \varphi(v,u)\ .$$

In the special case where $\varphi = D^2 f(x)^{\#}$, the value at $(\mathbf{e}_j,\, \mathbf{e}_k)$ has coordinates given by the second partial derivatives of the respective coordinate functions of f, where on takes partial derivatives first with respect to x_k and then with respect to x_j. By equality of mixed partials for $\mathcal{C}^2$, functions, this is exactly the same as the result obtained by taking partial derivatives in the reverse order. Combining all these observations, we see that $D^2 f(x)^{\#}$ is symmetric as claimed in the statement of the proposition.■

Standard generalization. One has similar symmetry properties for the r^{th} derivative of a $\mathcal{C}^r$ function if $r \geq 3$ (compare the comments at the top of page 185 in [MT]). In this case $D^r f(x)$ may be viewed as an r-multilinear function on $\mathbf{R}^n$ with values in $\mathbf{R}^m$, and the conclusion is that this function is symmetric with respect to the r variables.■

II.1.10 : *Diffeomorphisms*

For continuous maps on topological spaces, one has the fundamental notion of homeomorphism; the analogous concept for smooth maps of class $\mathcal{C}^r$ (where $1 \le r \le \infty$) is denoted by $\mathcal{C}^r$**-diffeomorphism**.

Definition. Let U and V be open subsets of Euclidean spaces, and let $f : U \to V$ be a small $\mathcal{C}^r$ mapping, where as usual $1 \le r \le \infty$. The map f is said to be a *(smooth) diffeomorphism of class* $\mathcal{C}^r$ if f is a homeomorphism and f^{-1} is also smooth of class $\mathcal{C}^r$.

Note. By the Chain Rule have the identity

$$D\left[f^{-1}\right](y) \quad = \quad \left(Df\left[f^{-1}(y)\right]\right)^{-1}$$

and since a linear transformation $L : \mathbf{R}^n \to \mathbf{R}^m$ is invertible if and only if $m = n$, it follows that if f is a diffeomorphism then $m = n$; in other words, this yields a simple alternate proof of invariance of dimension for smooth functions.■

The following observations are formal and elementary, but they will also be useful for our purposes.

PROPOSITION. *Let* $1 \le r \le \infty$.

(i) If U *is open in* $\mathbf{R}^n$*, then the identity map* id_U *is a* $\mathcal{C}^\infty$ *diffeomorphism.*

(ii) If U *and* V *are open in* $\mathbf{R}^n$ *and* $f : U \to V$ *is a* $\mathcal{C}^r$*-diffeomorphism, then so is* f^{-1}*.*

(iii) If U*,* V *and* W *are open in* $\mathbf{R}^n$*, and* $f : U \to V$ *and* $g : V \to W$ *are* $\mathcal{C}^r$ *diffeomorphisms, then so is* $g \circ f$*.*

Verifications of these statements are elementary and left to the reader as exercises.■

Example. If $f : \mathbf{R} \to \mathbf{R}$ is the function $f(x) = x^3$, then f is a homeomorphism that is smooth of class $\mathcal{C}^\infty$ but f^{-1} (the cube root mapping) is not $\mathcal{C}^1$ because it is not differentiable at $x = 0$.

One can take this even further.

PROPOSITION. *For each* r *such that* $1 \le r < \infty$*, then there is a smooth* $\mathcal{C}^r$ *mapping* $f : \mathbf{R} \to \mathbf{R}$ *that is a smooth diffeomorphism of class* $\mathcal{C}^r$ *but is not smooth of class* $\mathcal{C}^{r+1}$.■

Sketch of proof. $(\star)$ We begin with some elementary inequalities that follow immediately from the infinite series expansion for the hyperbolic cosine function $\cosh x$: For all $x > 0$ and even integers s we have $\cosh x > x^s/s!$, and for all odd integers s we have

$$\sinh x \quad > \quad \frac{1 + x^{s+1}}{(s+1)!} \quad > \quad \frac{x^s}{(s+1)!}$$

(the right hand inequality follows because $y^s < 1 + y^{s+1}$ for all $y > 0$).

In an entirely different direction, if $g(x) = x^k\,|x|$, then by induction and the Leibniz rule it follows that $g'(x) = kx^{k-1}\,|x|$. If we combine this with the previous paragraph we obtain the following conclusions:

$$\cosh x \quad > \quad \frac{g(x)}{(k+2)!}, \qquad (\text{where} \quad x,\ k \quad > \quad 0)$$

$$\lim_{x \to \pm\infty} \sinh x \;+\; \frac{g(x)}{(k+2)!} \;=\; \pm\infty$$

Therefore, if we let

$$f(x) \;=\; \sinh x \;+\; \frac{g(x)}{(k+2)!}$$

then it follows that f is a strictly increasing continuously differentiable function from $\mathbf{R}$ to itself that has a positive derivative everywhere and a continuous inverse. Since $g(x)$ is smooth of case $\mathcal{C}^k$ but not $\mathcal{C}^{k+1}$, the same is true of f. One can prove that the function f has a $\mathcal{C}^k$ inverse for $k \leq r$ by induction and the standard formulas from single variable calculus for differentiation of inverse functions:

$$g \;=\; f^{-1} \;\Longrightarrow\; g'(y) \;=\; \frac{1}{f'(g(y))} \quad \blacksquare$$

The formula for differentiating inverses leads directly to the topics covered in the next section.

II.2 : Implicit and inverse function theorems

(Conlon, Appendix B, §§ 2.4–2.5)

One standard problem in point set topology is to recognize when a 1–1 onto continuous function from one space to another has a continuous inverse. There are also many situations where it is useful to know simply whether a **local inverse** exists. For real valued functions on an interval, the Intermediate Value Property from elementary calculus implies that local inverses exist for functions that are strictly increasing or strictly decreasing (we have not actually proved this yet, however). Since the latter happens if the function has a derivative that is everywhere positive or negative close to a given point, one can use the derivative to recognize very quickly whether local inverses exist in many cases, and in these cases one can even compute the derivative of the inverse function using the standard formula quoted at the end of the previous section:

$$g \;=\; f^{-1} \;\Longrightarrow\; g'(y) \;=\; \frac{1}{f'(g(y))}$$

Of course this formula requires that the derivative of f is not zero at the points under consideration.

II.2.1 : *The multivariable Inverse Function Theorem*

There is a far-reaching generalization of the single variable inverse function theorem for functions of several real variables. It is covered in many but not all courses on multivariable calculus or undergraduate courses on the theory of functions of a real variable, but even when it is covered the treatment is sometimes incomplete (for example, only worked out for functions of two or at most three variables).

INVERSE FUNCTION THEOREM. *Let U be open in $\mathbf{R}^n$, let $a \in U$, and let $f : U \to \mathbf{R}^n$ be a $\mathcal{C}^r$ map (where $1 \leq r \leq \infty$) such that $Df(a)$ is invertible. Then there is an open set W containing a such that the following hold:*

(i) The restriction of f to W is $1-1$ and its image is an open subset V.

(ii) There is a $\mathcal{C}^r$ inverse map $g : V \to U_0$ such that $g(f(x)) = x$ on U_0.

In the previous section we noted that $Df(x)$ is invertible if the conclusions of the Inverse Function Theorem hold (see the discussion after the definition of diffeomorphism), and the Inverse Function Theorem is in some sense a local converse to this observation.

Proof. ($\star$) The proof given here uses the Contraction Lemma.

It is convenient to reduce the proof to the special case where $a = f(a) = 0$ and $Df(a) = I$. Suppose we know the result in that case. Let $A = Df(a)$, and define f_1 by the formula

$$f_1(x) \quad = \quad A^{-1}\left(f(x+a) - f(a)\right) \ .$$

Then $f_1(0) = 0$ and by the Chain Rule we have $Df_1(0) = I$. Then assuming the conclusion of the theorem is known for f_1, we take W_1, V_1, g_1 as in that conclusion. If we take $W = a + W_1$, $V = AV_1 + f(a)$, and

$$g(z) \quad = \quad g_1\left(A^{-1}\left(y - f(a)\right)\right) + a$$

then it follows immediately that the function $f(y) = Af_1(y-a) + f(a)$ satisfies the conclusions of the theorem.

Since f has a continuous derivative, there is a $\delta > 0$ such that $|x| \leq \delta$ implies $\| \, Df(x) - I \, \| < \frac{1}{2}$. For each y on the closed disk D of radius $\delta/2$ about the origin, define a map T on D by the formula $T(x) = x + y - f(x)$, and observe that $T(x) = x$ if and only if $y = f(x)$.

We want to apply the Contraction Lemma to T. The first step is to show that T maps D to itself. Let φ be the function $\varphi(x) = x - f(x)$; then $\varphi(0) = 0$ and $D\varphi = I - Df$, and consequently by the Mean Value Estimate we have that $|\varphi(x)| \leq |x|/2$ if $|x| \leq \delta$. Since $T(x) = y + \varphi(x)$ and $|y| < \delta/2$ it follows that $|T(x)| \leq \delta$ and hence $T(D) \subset D$.

We now need to estimate $|T(x_1) - T(x_0)|$ in terms of $|x_1 - x_0|$. Since T and φ differ by a constant it follows from the Mean Value Estimate that

$$\big|T(x_1) - T(x_0)\big| \quad = \quad \big|\varphi(x_1) - \varphi(x_0)\big| \quad \leq \quad \tfrac{1}{2}|x_1 - x_0|$$

and therefore the Contraction Lemma implies the existence of a unique point $x \in D$ such that $T(x) = x$, which is equivalent to $f(x) = y$.

Let $g : N_{\delta/2}(0) \to D$ be the inverse map sending a point y to the unique x such that $f(x) = y$. We claim that g is continuous. The first step is to show that $|x_0|$, $|x_1| \leq \delta/2$ implies $|f(x_1) - f(x_0)| \geq \frac{1}{2}|x_1 - x_0|$. To see this, use the identity $f(x) = x - \varphi(x)$ and use the equation and inequalities

$$|f(x_1) - f(x_0)| \quad \geq \quad |x_1 - x_0| - \big|\varphi(x_1) - \varphi(x_0)\big| \quad \geq \quad |x_1 - x_0| - \tfrac{1}{2}|x_1 - x_0| \quad = \quad \tfrac{1}{2}|f(x_1) - f(x_0)| \ .$$

If we set $y_i = f(x_i)$ so that $x_i = g(y_i)$ then we have

$$|x_1 - x_0| \quad = \quad |g(y_1) - g(y_0)| \quad \leq \quad 2 \cdot |y_1 - y_0|$$

and thus g is uniformly continuous.

Let U_0 be the image of g; we claim that U_0 is open. Suppose that $x \in U_0$, so that $f(x) = y$ where $|y| < \delta/2$. Then one can find some $\eta > 0$ so that $|z - x| < \eta$ implies $|f(z)| < \delta/2$ (why?)

and the identity $g(f(z)) = z$ then implies that $z \in \mathrm{Image}(g)$. Thus we may take $V = N_{\delta/2}(0)$ and $U_0 = g(V)$.

Finally, we need to show that g is a $\mathcal{C}^r$ function if f is a $\mathcal{C}^r$ function. Given $y \in V$ and k such that $y + k \in V$, write $y = f(x)$ and $y + k = f(x+h)$. Since $\| Df(x) - I \| < 1$ it follows that $Df(x)$ is invertible. Let L be its inverse. Then we have

$$g(y+k) - g(y) - L(k) \quad = \quad h - L(k) \quad = \quad -L\left(f(x+h) - f(x) - Df(x)h\right)$$

and the right hand side is equal to

$$L\left(|h| \cdot \theta(h)\right)$$

where $\lim_{|h|\to 0} \theta(h) = 0$. Since $h = g(y+k) - g(y)$ we know that $|h| \leq 2|k|$ and therefore we also have

$$\lim_{|k|\to 0} \frac{1}{|k|} \cdot L\left(|h| \cdot \theta(h)\right) \quad = \quad 0$$

(where $h = g(y+k) - g(y)$ as above), which shows that g is differentiable at y and satisfies a familiar looking formula:

$$Dg(y) \quad = \quad \left(Df\left(g(y)\right)\right)^{-1}$$

Since the entries of an inverse matrix are rational expressions in the inverse of the original matrix, the continuity of g and the $\mathcal{C}^1$ property of f imply that g is also $\mathcal{C}^1$.

If f is a $\mathcal{C}^r$ function, one can now prove that g is a $\mathcal{C}^s$ function for all $s \leq r$ inductively as follows: Suppose we know that f is $\mathcal{C}^r$ and g is $\mathcal{C}^s$ for $1 \leq s < r$. By the formula for the derivative of g we know that Dg is formed by the composite of g, Df and matrix inversion. We know that g is $\mathcal{C}^s$, that Df is too because f is $\mathcal{C}^{s+1}$ (recall that $s+1 \leq r$), and that inversion is $\mathcal{C}^\infty$ because its entries are given by rational functions, and therefore it follows that $D[g \circ f]$ is also $\mathcal{C}^s$, which means that $g \circ f$ is $\mathcal{C}^{s+1}$.∎

COROLLARY. *Let U and V be open in $\mathbf{R}^n$, and let $f : U \to V$ be $1-1$ onto and $\mathcal{C}^r$ where $1 \leq r \leq \infty$. Then f^{-1} is also $\mathcal{C}^r$.*∎

As indicated earlier, a similar result holds when $r = 0$, but the proof requires entirely different methods which come from algebraic topology.

The Inverse Function Theorem also has the following purely topological consequence for $\mathcal{C}^1$ mappings:

COROLLARY. *Let $f : U \to \mathbf{R}^n$ be a $\mathcal{C}^1$ function $(r \geq 1)$, where U is open in $\mathbf{R}^n$, and assume that $Df(x)$ is invertible for all $x \in U$. Then f is open.*∎

We have already noted that Brouwer's Invariance of Domain Theorem generalizes this result to the $\mathcal{C}^0$ case provided f is locally 1–1.

Proof. Let W be open in U and let $x \in W$. Then the Inverse Function Theorem implies that there is an open subset $W_0(x) \subset W$ containing x such that f maps $W_0(x)$ onto an open subset $V(x)$ in $\mathbf{R}^n$. Therefore it follows that

$$f(W) \quad = \quad \bigcup_x f\big(W_0(x)\big) \quad = \quad \bigcup_x V_x$$

which is open in $\mathbf{R}^n$.

Examples. Consider the complex exponential mapping f from $\mathbf{R}^2$ to itself sending (x,y) to $(e^x \cos y, e^x \sin y)$. The derivative of this map is invertible at every point but the map is not 1–1 because every nonzero point in $\mathbf{R}^2$ is the image of infinitely many points; specifically, for every (x,y) and integer k we have $f(x, y+2k\pi) = f(x,y)$.

Another example of this type is the complex square mapping $f : \mathbf{R}^2 - \{0\} \to \mathbf{R}^2$ sending (x,y) to $(x^2 - y^2, 2xy)$, which has the property that $f(-x,-y) = f(x,y)$ for all (x,y); note that if we write $z = u + iv$, then $f(z) = z^2$.

Final Remark. Given a $\mathcal{C}^1$ function $f : U \to \mathbf{R}^n$ with U open in $\mathbf{R}^n$, if we write the coordinate functions of f as $y_1, \cdots y_n$ then $\det Df(p)$ is just the classical *Jacobian function*

$$\frac{\partial(y_1, \cdots, y_n)}{\partial(x_1, \cdots, x_n)}(p)$$

and with this terminology the condition on $Df(p)$ in the Inverse Function Theorem may be rephrased to state that the Jacobian at p is nonzero.

II.2.2 : *The Implicit Function Theorem* ($\star$)

There is a close relation between the Inverse Function Theorem and the standard Implicit Function Theorem from ordinary and multivariable calculus. In its simplest form the Implicit Function Theorem states that **locally** one can solve an equation $F(x,y) = 0$ uniquely for y in terms of x; more precisely, if $F(a,b) = 0$ and the second partial derivative of F is nonzero at (a,b), then on some open interval $(a-\delta, a+\delta)$ there is a unique function $f(x)$ such that $y = f(x) \iff F(x,y) = 0$ (hence $f(a) = b$), and its derivative is given by

$$\frac{df}{dx} \quad = \quad - \frac{\left(\frac{\partial F}{\partial x}\right)}{\left(\frac{\partial F}{\partial y}\right)} \quad .$$

Here is a general version of this result:

IMPLICIT FUNCTION THEOREM. *Let U and V be open in $\mathbf{R}^n$ and $\mathbf{R}^m$ respecitvely, and let $f : U \times V \to \mathbf{R}^m$ be a smooth function such that for some $(x,y) \in U \times V$ we have $f(x,y) = 0$ and the partial derivative of f with respect to the last m coordinates is invertible. Then there is an open neighborhood U_0 of x and a smooth function $g : U_0 \to V$ such that $g(x) = y$ and for all $u \in U_0$ we have $f(u,v) = 0$ if and only if $v = g(u)$.*

For the sake of completeness we note that the partial derivative of f with respect to the last m coordinates is the derivative of the function $f^*(v) = f(x,v)$, and that smooth means smooth of class $\mathcal{C}^r$ for some r such that $1 \leq r \leq \infty$.

Proof. Define $h : U \times V \to \mathbf{R}^m \times \mathbf{R}^n$ by $h(u,v) = (f(u,v), u)$. Then the hypotheses imply that $Dh(x,y)$ is invertible, and therefore by the Inverse Function Theorem there is a local inverse

$$k : \text{Int } (\varepsilon D^m) \times U_0 \longrightarrow U \times V \ .$$

Since the second coordinate of $h(u,v)$ is u, it follows that the first coordinate of the inverse $k(z,w)$ is w so that we may write $k(z,w) = (w, \, Q(z,w)\,)$ for some smooth function Q.

On one hand we have $g\,(\,k(z,w)\,) = (z,w)$, but on the other hand we also have

$$g\,(k(z,w)\,) \quad = \quad g\,(w, Q(z,w)) \quad = \quad \big(f\,(w, Q(z,w))\,, w\big).$$

In particular, this means that

$$z \quad = \quad f\big(u, Q(z,u)\big)$$

for all z and u. If we take $g(u) = Q(0,u)$ it follows that $y = g(x)$ and $f(u,v) = 0$ if and only if $v = g(u)$.■

One can use the Chain Rule to calculate $Dg(u)$ as follows: If $\varphi(u) = f(u, g(u))$ and $\mathbf{p} \in \mathbf{R}^n$, then the Chain Rule yields the formula

$$[D\varphi(u)](\mathbf{p}) \quad = \quad [D_1 f\,(u, g(u)\,)]\,(\mathbf{p}) \;+\; [D_2 f\,(u, g(u)\,)]\,(\,[Dg(u)](\mathbf{p})\,)$$

where D_1 and D_2 refer to partial derivatives with respect to the first and last sets of variables. Since $\varphi = 0$ it also follows that $[D\varphi(u)](\mathbf{p}) = 0$. Furthermore, if u and v are sufficiently close to x and y then the second partial derivative is invertible. Therefore one obtains the formula

$$Dg \quad = \quad -\,(D_2 f)^{-1} \circ D_1 f$$

which generalizes the formula from elementary multivariable calculus.■

II.2.3 : *Smooth maps of maximum rank*

Suppose that U is open in $\mathbf{R}^n$ and $f : U \to \mathbf{R}^m$ is smooth, and in addition assume that $m \neq n$. Then the Inverse Function Theorem cannot apply to f because the derivative of f is never invertible, but one can still ask what happens when $Df(x)$ has maximum rank for some $x \in U$. We shall use the Inverse Function Theorem to draw strong conclusions in such cases.

If $n \leq m$ the maximum rank condition means that the derivative is always 1–1, while if $n \geq m$ this means that the derivative is always onto. In these cases we say that f is an *immersion* or *submersion* at x respectively.

The Inverse Function Theorem yields important information on the local behavior of immersions and submersions.

IMMERSION STRAIGHTENING PROPOSITION. *Let $f : U \to V$ be a smooth map on open subsets in Euclidean spaces. Then f is an immersion at $x \in U$ if and only if one can find open neighborhoods U_0 and V_0 of x and $f(x)$ and a diffeomorphism $h : V_0 \to U_0 \times \mathrm{Int}\, D^{m-n}$ so that*

$$h(f(y)) \quad = \quad (y, 0)$$

for all $y \in U_0$.

IMPORTANT NOTATIONAL CONVENTION THROUGHOUT THIS COURSE. Given a function G we shall repeatedly use "G" to denote a function defined by the same rules as G but possibly defined an a subset of the domain of G with a codomain that is possibly a subset of the codomain of G.

Proof. We begin with the $(\Longrightarrow)$ implication. Let $T : \mathbf{R}^{m-n} \to \mathrm{Image}\, Df(x)^{\perp}$ be a linear isomorphism and define $g : U \times \mathbf{R}^{m-n} \to \mathbf{R}^m$ by $g(y,z) = f(y) + T(z)$. Then $Dg(x) = Df(x) + T$, which is an isomorphism, and hence it is a diffeomorphism on some open set of the form $U_0 \times \mathrm{Int}\, \varepsilon D^{m-n}$. By construction the image of f corresponds to points with vanishing second coordinate.

We now consider the $(\Longleftarrow)$ implication: If h exists then D"$h \circ f$"(x) is 1–1 and $Df(y) = [Dh(f(x))]^{-1} \cdot [D$"$h \circ f$"$(x)]$. Therefore $Df(x)$ is also 1–1.■

SUBMERSION STRAIGHTENING PROPOSITION. *Let $f : U \to V$ be a smooth map on open subsets in Euclidean spaces. Then f is a submersion at $x \in U$ if and only if ons can find open neighborhoods U_0 and V_0 of x and $f(x)$ and a diffeomorphism $k : V_0 \times \mathrm{Int}\, D^{n-m} \to U_0$ so that*

$$f(k(y,z)) \quad = \quad y$$

for all $(y,z) \in V_0 \times \mathrm{Int}\, D^{n-m} \to U - 0$.

Proof. The argument is similar to the preceding one.

$(\Longleftarrow)$ If k exists then D"$f \circ k$"(y) is onto and $Df = [D$"$f \circ k$"$] \circ Dk^{-1}$. This implie that $Df(x)$ is onto.

$(\Longrightarrow)$ By hypothesis the kernel of $Df(x)$ is $(n-m)$-dimensional. Let $S_1 : \mathrm{Kernel}\, Df(x) \to \mathbf{R}^{n-m}$ be a linear isomorphism, let

$$S_2 : \mathbf{R}^m \longrightarrow \mathrm{Kernel}\, Df(x)$$

be perpendicular projection, and define

$$g : U \longrightarrow V \times \mathbf{R}^{n-m}$$

by the formula

$$g(u) \quad = \quad \big(f(u), S_1 S_2(u)\big) \ .$$

By construction $Dg(x)$ is invertible, and therefore by the Inverse Function Theorem there is a local inverse $k : V_0 \times \mathrm{Int}\, \varepsilon D^{n-m} \to U_0$. If P denotes projection onto V then $f = P \circ g$, so that $f(k(y,z)) = P(g(k(y,z)))$. Since g is inverse to k, the latter reduces to $P(y,z) = y$ as required.■

II.3 : Bump functions

(Conlon, § 2,6)

In this course it will be very important to have smooth versions of the partitions of unity described in Section I.2; of course this requires a notion of smoothness, so at this point we shall only attempt to do this for open subsets of Euclidean spaces (but we shall also do this for smooth manifolds in subsequent units). As in Section I.1, we need these in order to take differentiable functions defined on pieces of a space and to construct something out of the pieces that is globally differentiable. Such constructions are needed in a wide range of geometric and analytical contexts.

II.3.1 : *Input from single variable calculus*

Our construction of smooth partitions of unity begins with the following example from single variable calculus:

PROPOSITION. *There is an infinitely differentiable function $f : \mathbf{R} \to \mathbf{R}$ such that $f(x) > 0$ for $x > 0$ and $f(x) = 0$ for $x < 0$.*

Since this function is infinitely differentiable it follows that the higher order derivatives satisfy $f^{(n)}(0) = 0$ for all n even though the function is not constant in any open neighborhood of zero. In particular, this means that there cannot be an infinite series expansion for f as a convergent power series.

Sketch of Proof. Consider the function

$$f(t) \quad = \quad \exp\left(-\frac{1}{t^2}\right)$$

which is defined and infinitely differentiable for $t > 0$. If we can show that $f^{(n)}(0) = 0$ for all n, then we can extend f to an infinitely differentiable function on the whole real line by taking it to be zero for $t \leq 0$.

Since the iterated derivatives of f have the form $g \cdot f$ where g is a rational function of t (use the Leibniz rule repeatedly), the result will follow if we can show $\lim_{t\to 0+} g(t) \cdot f(t) = 0$. This is a straightforward (but perhaps somewhat messy) consequence of L'Hospital's Rule.■

The previous result allows us to construct a large number of infinitely differentiable functions that are constant on entire intervals.

PROPOSITION. *There is an infinitely differentiable function $B : \mathbf{R} \to \mathbf{R}$ such that $B(t) > 0$ for $t \in (0, 1)$ and $B(t) = 0$ elsewhere.*

Proof. Take $B(t) = f(t) \cdot f(1 - t)$.■

PROPOSITION. *There is an infinitely differentiable function $C : \mathbf{R} \to \mathbf{R}$ such that $C(t) = 0$ for $t \leq 0$, C is strictly increasing on $[0, 1]$, and $C(t) = 1$ for $t \geq 1$.*

Proof. Let $L = \int_0^1 B(t)\, dt$ and set

$$C(t) \quad = \quad \frac{1}{L} \int_0^t B(u)\, du \ .■$$

COROLLARY. *A similar result holds if $[0, 1]$ is replaced by an arbitrary closed interval $[a, b]$.*■

The following result is a simple but important consequence of the existence of bump functions.

GERM EXTENSION THEOREM. *Let $U \subset \mathbf{R}^n$ be open, let $x \in U$, and let $f : U \to \mathbf{R}^m$ be a smooth C^r function. Then there is an open neighborhood $V \subset U$ of x such that the restriction $f|V$ extends to a smooth C^r function from $\mathbf{R}^n$ to $\mathbf{R}^m$.*

Proof. Choose $\delta > 0$ such that $N_\delta(x) \subset U$, and let V be the subdisk of radius $\delta/3$ centered at x. By the previous results there is a smooth function $\varphi : [0, \delta] \to \mathbf{R}$ such that $\varphi = 1$ on $[0, \frac{\delta}{3}]$ and $\varphi = 0$ on $[\frac{2\delta}{3}, \delta]$. If we define g by $g(y) = \varphi(y) \cdot f(y)$ for $|y - x| < \delta$ and $g(y) = 0$ for all other points on $\mathbf{R}^n$, then g is a well defined smooth C^r function on all of $\mathbf{R}^n$ and its restriction to V is equal to $f|V$.■

Here is an explanation of the terminology in the statement of the result. Given a topological space X and a point $p \in X$, we say that two $\mathbf{R}^m$ valued continuous functions f, g defined on (possibly different) open neighborhoods of p *have the same germ at the point* p if there is a smaller neighborhood W such that $f|W = g|W$. In this terminology, the theorem states that every smooth $\mathcal{C}^r$ function defined on a neighborhood of x in $\mathbf{R}^n$ has the same germ as a smooth $\mathcal{C}^r$ function defined on all of $\mathbf{R}^n$. — It is elementary to check that a similar result holds for continuous functions with $\mathbf{R}^n$ replaced by an arbitrary $\mathbf{T}_{3\frac{1}{2}}$ space (this is left as an exercise for the reader).

II.3.2 : *Smooth partitions of unity*

The results on bump functions lead to the existence of smooth partitions of unity for open subsets of Euclidean spaces. To formulate the result, we use notation similar to that of Section I.2. Let U be open in $\mathbf{R}^n$, and let $\mathcal{U}$ be an open covering of U. Then the methods of Section I.2 yield a locally finite refinement $\mathcal{V}$ such that the following hold:

(*i*) Each V_α in $\mathcal{V}$ is an open disk.

(*ii*) The sets W_α, defined as open disks whose centers are the same as those of the corresponding V_α and whose radii are half those of the corresponding V_α, also form an open covering of U.

As in the general case, $\overline{W_\alpha}$ is compact and $\overline{W_\alpha} \subset V_\alpha$ by construction.

EXISTENCE OF SMOOTH PARTITIONS OF UNITY, VERSION 1. *Let U be an open subset of $\mathbf{R}^n$, and let $\mathcal{M}$ and $\mathcal{N}$ be countable open coverings satisfying the properties of $\mathcal{V}$ and $\mathcal{W}$ as above. Then there is a family of smooth $\mathcal{C}^\infty$ functions $\varphi_j : U \to \mathbf{R}$ with values in $[0,1]$ such that*

(*i*) *the* **support** *of φ_j — that is, the closure of the set on which $\varphi \neq 0$ — is a compact subset of M_j,*

(*ii*) *we have $\sum_j \varphi_j = 1$.*

Such a family of continuous functions is called a *smooth* ($\mathcal{C}^\infty$) *partition of unity* subordinate to the open covering $\mathcal{M}$. As in the continuous case, *there is no convergence problem with the sum even if there are infinitely many sets in the open covering* $\mathcal{M}$; each point has a neighborhood that meets only finitely many M_j's nontrivially, and on this neighborhood the sum reduces to a finite sum.

Proof. The argument is analogous to the proof of the topological result in Section I.2, so we shall concentrate on the changes that are needed to make that proof work in the present situation.

For each j, let x_j denote the center of M_j and let r_j be its radius. For each j let $h_j : N_2(0) \to M_j$ be the standard "linear" homeomorphism defined by

$$h_j(y) \quad = \quad x_j \; + \; \left(\frac{r_j}{2}\right) \cdot y \ .$$

Let ω be the smooth $\mathcal{C}^\infty$ bump function on the interval $[0,2]$ such that $\omega = 1$ on $[0,1]$, ω decreases linearly from 1 to 0 on $[1.\frac{3}{2}]$, and $\omega = 0$ on $[\frac{3}{2}, 2]$. Then the function $f_j : M_j \to [0,1]$ defined by $f_j(x) = \omega(\,|h_j^{-1}(x)|\,)$ extends to a smooth function on U by taking $f_j = 0$ on the complement of M_j by the same type of argument used in the continuous case.

As in the topological case, for each point $x \in U$ we know that $f_j(x) > 0$ for some j; furthermore, the local finiteness of $\mathcal{M}$ ensures that one can add the functions f_j to obtain a well defined smooth function. This sum $f = \sum_j f_j$ is always positive by the first sentences of this paragraph, and therefore as before we may define

$$\varphi_j \quad = \quad \frac{f_j}{f} \ .$$

It follows immediately that the functions φ_j have all the required properties.■

We now have smooth analogs of the two applications of continuous partitions of unity that appeared in Section I.2. section.

FIRST PROPOSITION, SMOOTH VERSION 1. *Suppose that U is open in $\mathbf{R}^n$ and Ω is an open neighborhood of $U \times \{0\}$ in $X \times \mathbf{R}$. Then there is a smooth $\mathcal{C}^\infty$ real valued function $f : U \to (0, \infty)$ such that the set*

$$\{\ (x,t) \in U \times [0,\infty) \mid t < f(x)\ \}$$

is contained in Ω.■

SECOND PROPOSITION, SMOOTH VERSION 1. *Let U be open in $\mathbf{R}^n$. Then there is a smooth $\mathcal{C}^\infty$ function $f : U \to \mathbf{R}$ with values in $[0, +\infty)$ such that for each $K > 0$ the inverse image $f^{-1}(\,[0, K]\,)$ is compact (in other words, f is a* **proper** *smooth map).*■

II.3.3 : *Regular curves in Euclidean regions* $(1\frac{1}{2}\star)$

Suppose that U is an open subset of $\mathbf{R}^n$. In Section III.5 of the ONLINE 205A NOTES we noted that U is connected if and only if it is arcwise connected, and in fact we proved that if it is connected then every pair of points can be joined by a special type of piecewise smooth curve known as a *broken line*; to save time and space we refer the reader to the section in question for details. Towards the end of Section III.5 in the ONLINE 205A NOTES we also posed the following question:

> *Suppose that U is a connected open subset of $\mathbf{R}^n$ and $x, y \in U$. Is there a continuous curve $\gamma : [a, b] \to U$ such that $\gamma(a) = x$, $\gamma(b) = y$, the curve γ has (continuous) derivatives of all orders, the tangent vector $\gamma'(t)$ is never zero?*

A little physical experimentation — either with pencil and paper or with wires or strings — strongly indicates that one can modify the broken line curve into a smooth curve with the desired properties. In fact, **it is always possible to do construct such curves**, and a mathematically complete argument appears in the file(s) `goodcurves.*` in the course directory. The main idea behind this is fairly intuitive: One would like to smooth out a broken line curve near the finite sets of points where there is an abrupt change of direction. Some work is needed to verify that everything works as expected, and the bump functions defined in this section play an important role in the formal proof.

II.4 : Integral flows

(Conlon, §§ 2.7–2.8, Appendix C.1–C.3)

Many important geometric properties and laws of physics are best expressed mathematically using ordinary differential equations. In order to understand the significance of such properties and laws, it is necessary to know something about the solutions of these differential equations. Therefore we shall begin by establishing the standard existence and uniqueness theorems for such solutions. For several reasons we also need information on the extent to which the solution curves depend upon the initial condition.

Similar examples lead one to expect the dependence upon initial conditions is extremely regular. For example, if we consider the differential equation $y' = y$ and let b be the initial condition at $t = 0$, then the general solution has the form $y(t) = b\,e^t$ and hence the solution depends smoothly (in fact *linearly*) ion the initial condition. Similar conclusions follow if one considers other elementary differential equations from introductory undergraduate courses on the subject.

Much of this section is devoted to proving completely general regularity results on the dependence of solutions on initial conditions.

II.4.1 : *Existence and uniqueness of solutions*

Of all the consequences of the Contraction Lemma in Section I.3, by far the most important and far-reaching is Picard's basic existence and uniqueness theorem for solutions of first order ordinary differential equations.

Before stating the main result, it wil be useful to state a variant of Lipschitz conditions as discussed in Section III.1. Suppose that J is an open interval in $\mathbf{R}$ and that U is open in $\mathbf{R}^n$. A continuous function

$$\mathbf{F} : J \times U \longrightarrow \mathbf{R}^m$$

is said to be *Lipschitz on U uniformly with respect to J* if there is a constant $K > 0$ such that

$$|\mathbf{F}(t,x) \;-\; \mathbf{F}(t,y)| \;\;\leq\;\; A \cdot |x \;-\; y|$$

for all $t \in J$ and $x, y \in U$.

FUNDAMENTAL EXAMPLE. *Let F be a smooth function of class $\mathcal{C}^1$ on $J \times U$, and suppose we are given a point $a \in J$ and a compact subset $K \subset U$. Then there is a subinterval J_0 containing a and an open set U_0 such that $K \subset U_0 \subset U$ such that $\mathbf{F}|J_0 \times U_0$ is uniformly Lipschitz on U_0 with respect to J_0.*

Sketch of proof. This follows by the same sort of argument involving the Mean Value Estimate that was used in Section II.1 to establish ordinary Lipschitz conditions for smooth functions.∎

The preceding observation implies that the main existence and uniqueness result below applies to $\mathcal{C}^1$ functions, the only change being the need to restrict the function $\mathbf{F}$ to a smaller open subset $J_0 \times U_0$ at the beginning of the proof.

PICARD SUCCESSIVE APPROXIMATION METHOD FOR THE SOLUTIONS OF DIFFERENTIAL EQUATIONS. *Let J be an interval in $\mathbf{R}$ containing a given point a, let U*

be open in $\mathbf{R}^n$, *and let* $F : J \times U \to \mathbf{R}^n$ *be a continuous function that is uniformly Lipschitz on* U *with respect to* J. *Then for each* $(a,b) \in J \times U$ *there is a positive real number* $\delta > 0$ *such that there is a unique solution of the differential equation*

$$\frac{dy}{dx} \quad = \quad \mathbf{F}(x,y)$$

on the interval $(a-\delta, a+\delta)$ *satisfying the initial condition* $y(a) = b$.

Proof. To motivate the proof, note first that a function f is a solution of the differential equation with the given initial value condition if and only if

$$f(x) \quad = \quad b \; + \; \int_a^x \mathbf{F}\,(t, f(t))\; dt$$

where as usual the integral is zero if $x = a$, while if $x < a$ the integral from a to x is defined to be the negative of the integral from x to a.

The idea of the proof is to use the right hand side to define a map of bounded continuous functions and then to apply the Contraction Lemma. However, one needs to be a bit careful in order to specify exactly which sorts of functions form the space upon which the mapping is defined and in order to ensure that the map has the contraction property.

Choose $h, k > 0$ so that

$$S \quad = \quad [a-h, a+h] \times [b-k, b+k] \quad \subset \quad U$$

so that $\mathbf{F}$ and its (first) partial derivatives are bounded on S. Let L be an upper bound for $\mathbf{F}$. By the assumption that $\mathbf{F}$ is uniformly Lipschitz on U with respect to J have that

$$\left|\mathbf{F}(x,y_1) - \mathbf{F}(x,y_2)\right| \quad \leq \quad A \cdot |y_1 - y_2|$$

for some $A > 0$, all $x \in J$, and all $y_1, y_2 \in U$.

Next choose $\delta > 0$ so that $\delta \leq h$, $L\,\delta < k$ and $A\,\delta < 1$. Define M to be the metric space of all bounded continuous functions g on $(a-\delta, a+\delta)$ for which $|g - b| \leq L\,\delta$, where as usual we identify a real number with the constant function whose value is that number.

For every metric space Z, every $z \in Z$ and every positive real number B, the set of points w with $\mathbf{d}(z,w) \leq B$ is closed (why?), and therefore M is a complete metric space. We need to show that the map

$$[T(g)](x) \quad = \quad b \; + \; \int_a^x \mathbf{F}\,(t, g(t))\; dt$$

is defined for all $g \in X$, it maps

$$X \quad \subset \quad \mathbf{BC}\big(\ (a-\delta, a+\delta)\ \big)$$

into itself, and it satisfies the hypothesis of the Contraction Lemma on M.

First of all, it follows immediately that $T(g)$ is continuous whenever g is continuous (fill in the details here). Next, by the boundedness of $\mathbf{F}$ on the closed solid rectangle S we have

$$|T(g) - b| \quad = \quad \left|\int_a^x \mathbf{F}\,(t, g(t))\; dt\right| \quad \leq \quad L \cdot \left|\int_a^x \; dt\right| \quad \leq \quad L\,\delta$$

so that $g \in M$ implies $T(g) \in M$.

Finally, let $g_1, g_2 \in X$ and consider $|T(g_1) - T(g_2)|$. By definition the latter is equal to the least upper bound of the numbers

$$\left| \int_a^x \left(\mathbf{F}\left(t, g_1(t)\right) - \left(t, g_2(t)\right)\right) \, dt \right| \quad \leq$$

$$\int_a^x \left| \mathbf{F}\left(t, g_1(t)\right) - \left(t, g_2(t)\right) \right| \, dt \quad \leq \quad A \, \delta \cdot |g_1 - g_2| \ .$$

Since $A \, \delta < 1$, all the hypotheses of the Contraction Lemma apply so that there is a unique fixed point, and as noted above this unique fixed point must be the (necessarily unique) solution of the original differential equation with the prescribed boundary condition.■

Note. One can prove an existence theorem with the weaker hypothesis that $\mathbf{F}$ is continuous (compare Exercise 25 on pages 170–171 of "LITTLE RUDIN"), but uniqueness does not follow. For example, if $\mathbf{F}(x, y) = y^{1/2}$, then the zero function and $x^2/4$ are both solutions to the differential equation with initial condition $y(0) = 0$.■

II.4.2 : *Higher-order differential equations* ($\star$)

For a multitude of purposes in mathematics, the other sciences (including the social sciences) and engineering, it is necessary to work with second order differential equations of the form

$$y'' \quad = \quad F(x, y, y')$$

for some reasonable continuous function F, and there are also numerous situations where one encounters differential equations of even higher orders. Formally, if k is a positive integer, then a k^{th} order differential equation may be written as

$$y^{(k)} \quad = \quad F(x, y, y', \ \cdots \ , y^{(k-1)})$$

where F again is some reasonable function. As in the case of first order differential equation one can view a system of $n \geq 2$ differential equation in n "unknown" scalar functions as a single n-dimensional vector differential equation.

For both practical and theoretical purposes, the study of higher order differential equations relies heavily on the following basic **Reduction of Order Principle:** *A system of n differential equations of order k is equivalent to a system of nk first order differential equations.*

DESCRIPTION OF THE EQUIVALENCE. We discuss second order differential equations first because they arise so often in mathematics and the sciences, it is particularly easy to describe this case, and simple versions of this trick already appear in undergraduate differential equations courses. Specifically, given the second order system of differential equations

$$y'' \quad = \quad F(x, y, y')$$

if we set $p = y'$ then we may rewrite the original system of n second order equations as the following system of $2n$ first order equations with "unknowns" y and p. This idea frequently appears in the first chapter of undergraduate differential equations texts as a tool for reducing the solving of certain second order equations to first order equations, and it also arises explicitly in the study of *phase planes*.

The corresponding trick for k^{th} order equations is similar but the notation is more complicated. Specifically, if we set $p_0 = y$ and define p_i recursively by $p_i = p'_{i-1}$ for $1 \leq i \leq k-1$, then a given system of n differential equations of order k is equivalent to the following system of nk first order equations:

$$\begin{aligned} p'_j &= p_{j-1} \qquad (1 \leq j \leq k-1) \\ p'_{k-1} &= F\big(x, p_0, p_1, \cdots, p_{k-1)}\big) \end{aligned}$$

In all of these cases, equivalence means that a solution of the k^{th} order system leads very directly to a solution of the larger first order system and vice versa.■

The preceding discussion yields the following existence and uniqueness result for higher order (systems of ordinary) differential equations.

THEOREM. *Let J be an open interval in the real line, let U_i be open in $\mathbf{R}^n$ for $1 \leq i \leq k$, let F be a smooth function from $J \times \prod_i U_i$ to $(\mathbf{R}^n)^k$. Given $a \in J$ and $b_i \in U_i$ for $1 \leq i \leq k$, there is a locally unique solution to the system*

$$y^{(k)} = F(x, y, y', \cdots, y^{(k-1)})$$

with initial conditions $y^{(i)}(a) = b_i$ for $1 \leq i \leq k$.■

II.4.3 : *Joint continuity of solutions*

As noted above, for the purposes of this course it is important to know that the solution curves for a (system of ordinary) differential equations

$$\frac{dy}{dx} = F(x, y)$$

depend continuously on the initial value and in fact depend smoothly on the initial value if F is smooth. The next goal of this section is to prove these results. Although the result itself is indispensable, an understanding of the proof is not needed in the rest of the course, As in the proof for the existence of solutions, we assume that F satisfies a Lipschitz condition.

Suppose now that $\varphi : J_0 \to U$ is a smooth function of class $\mathcal{C}^1$ and $\varepsilon \geq 0$. We shall say that φ is an ε-approximate solution of F on J_0 if for all $t \in J_0$ we have

$$|\varphi'(t) - F(t, \varphi(t))| \leq \varepsilon .$$

PROPOSITION. *Let φ_1 and φ_2 be ε_1- and ε_2-approximate solutions of the differential equation $y' = F(t, y)$ on J_0, and let $\varepsilon = \varepsilon_1 + \varepsilon_2$. Assume that F is Lipschitz on U and uniformly bounded with respect to J_0 with constant K, or else that the partial derivative D_2F exists and is bounded by K on $J_0 \times U$. Let $t_0 \in J_0$. The for all $t \in J_0$ we have*

$$|\varphi_1(t) - \varphi_2(t)| \leq \exp(K\,|t - t_0|) \cdot \Big(|\varphi_1(t_0) - \varphi_2(t_0)| + \frac{\varepsilon}{K}\Big) .$$

Proof. (2⋆) By the Triangle inequality we have

$$|\varphi_1'(t) - \varphi_2'(t) + F(t,\,\varphi_2(t)) - F(t,\,\varphi_1(t))| \quad \leq \quad \varepsilon\ .$$

Let $t = t_0 + |t - t_0|$, and let $\psi(t)$ and $\omega(t)$ be equal to $|\varphi_1'(t) - \varphi_2'(t)$ and $F(t,\,\varphi_2(t)) - F(t,\,\varphi_1(t))|$ respectively. Using the inequality $|u - v| \geq |u| - |v|$ and some standard integral estimates and identities we obtain the following inequalities:

$$|\psi(t) - \psi(t_0)| \;\leq\; \varepsilon\,(t - t_0) + \int_{t_0}^{t} \omega(s)\,ds \;\leq\; \varepsilon\,(t - t_0) + K\int_{t_0}^{t} \psi(s)\,ds \;\leq\; K\int_{t_0}^{t} \left[\psi(s) + \frac{\varepsilon}{K}\right] ds$$

The latter immediately implies that

$$\psi(t) \quad \leq \quad \psi(t_0) \;+\; K\int_{t_0}^{t} \left[\psi(s) + \frac{\varepsilon}{K}\right] ds$$

and the final step of the proof is the following inequality:

LEMMA. *Let g be a positive real valued continuous function on an interval that is bounded by a constant L. Let $t_0 \leq t$ be in the interval and assume that there are nonnegative numbers A and K such that*

$$g(t) \quad \leq \quad A \;+\; K\int_{t_0}^{t} g(s)\,ds\ .$$

Then for all integers $n \geq 1$ we have

$$g(t) \quad \leq \quad A\left[1 \;+\; \frac{K(t - t_0)}{1!} \;+\; \cdots \;+\; \frac{K^{n-1}(t - t_0)^{n-1}}{(n-1)1!}\right] \;+\; \frac{L\,K^n(t - t_0)^n}{(n!}\ .$$

This may be established by induction on n.■

As noted above, this completes the proof of the proposition.■

COROLLARY. *Let $F : J \times U \to \mathbf{R}^n$ be continuous and satisfy a Lipschitz condition on U uniformly with respect to J. Let $x_0 \in U$. Then there exists and open subinterval $J_0 \subset J$ containing 0, and an open subset $U_0 \subset U$ containing x_0, such that F has a unique flow $\alpha : J_0 \times U_0 \to U$. We can select J_0 and U_0 so that α is continuous and satisfies a Lipschitz condition on $J_0 \times U_0$.*

Proof. (2⋆) Given $x,\,y \in U_0$ let $\varphi_1(t) = \alpha(t,x)$ and $\varphi_2(t) = \alpha(t,y)$. In this case $\varepsilon_1 = \varepsilon_2 = 0$. If $s,\,t \in J_0$ then

$$|\alpha(t,x) - \alpha(s,y)| \quad \leq \quad |\alpha(t,x) - \alpha(t,y)| \;+\; |\alpha(t,y) - \alpha(s,y)| \quad \leq \quad |x - y|\,e^K \;+\; |t - s|\,L$$

provided the diameter of J_0 is less than 1 and L is a Lipschitz constant for F (filling in the details is left to the reader).■

COROLLARY. *Let J be an open interval in the real line containing 0 and let U be open in $\mathbf{R}^n$. Let $F : J \times U \to \mathbf{R}^n$ be a continuous map which is Lipschitz on U uniformly for every compact subinterval of J. Let $t_0 \in J$ and let φ_1 and φ_2 be two $\mathcal{C}^1$ maps with the same values at t_0 and satisfying $\varphi_i'(t) = F(t,\,\varphi_i(t)\,)$ for $t = 1,\,2$. Then $\varphi_1 = \varphi_2$.*

Proof. Substitute $\varepsilon = 0$ in the previous proposition.■

Consider now the restriction of Φ to a slice of the form $\{t\} \times W$. By construction the map $\Phi|\{a\} \times W$ is just the inclusion map, and for other values of t the map $\Phi|\{t\} \times W$ takes a point $w \in W$ to its image under the integral curve at parameter value t. Joint continuity implies that this is a continuous map from W to $\mathbf{R}^n$. Physically one can visualize the maps $\Phi|\{t\} \times W$ as describing a flow of motion associated to the differential equation. For example, it the original differential equation defines the motion within a fluid, then $\Phi| \ (\{t\} \times W)$ shows how the particles in W move around from parameter value a to parameter value t. The notion of **integral flow** arising in this fashion will play a fundamental role throughout the course.

Example. Suppose that $\mathbf{F}(u, v) = (-v, u)$. Then the solution curves are given by the formula

$$\Phi(t; u, v) \quad = \quad (u \cos t - v \sin t,\, u \sin t + v \cos t)$$

(the reader is invited to verify this using the methods of undergraduate differential equations courses). If we hold u and v constant then the formula implies that the unique solution with initial condition (u, v) is a counterclockwise circle. If we hold t constant then we see that the flow transformation $\Phi_t = \Phi|\{t\} \times \mathbf{R}$ is just counterclockwise rotation through t radians.■

Further examples as well as interactive pictures may be found at the University of British Columbia's "Living Mathematics" online site, which is given after the discussion of vector fileds below.

II.4.4 : *Vector fields and autonomous differential equations*

A differential equation (or system of n differential equations)

$$x' \quad = \quad F(t, x)$$

is said to be *autonomous* if the right hand side is expressible as a function of x alone. In this case the right hand side may also be viewed as a **vector field** on some open subset U of $\mathbf{R}^n$; *i.e.*, a continuous or smooth function $F : U \to \mathbf{R}^n$. This vector field is closely related to the solutions of differential equations. If $\gamma(t)$ is a solution curve for the equation $x' = F(x)$, then by definition $F(\gamma(t))$ is the tangent vector $\gamma'(t)$. Most differential equations in this course will be autonomous, and some basic properties of their solution curves will be discussed in Section IV.2. Further discussion and some interactive graphics packages appear in the following online references:

`http://en.wikipedia.org/wiki/Vector_field`

`http://www.sunsite.ubc.ca/LivingMathematics/V001N01/UBCExamples/Flow/Flow.html`

II.4.5 : *Smooth dependence on initial values* $(\star)$

Having constructed the continuous map Φ, it is natural to ask whether this map is smooth. The rest of this section is devoted to establishing the smoothness properties that will be needed later in the course.

We shall being with a result on linear differential equations that we shall need:

PROPOSITION. *Let J be an open interval in the real line containing* 0, *let U be an open subset of $\mathbf{R}^n$, let $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^n)$ be the space of linear transformations from $\mathbf{R}^n$ to itself (identified as usual with n^2-dimensional Euclidean space), and let $g : J \times U \to \mathcal{L}(\mathbf{R}^n, \mathbf{R}^n)$ be continuous. Define*

$$G : J \times U \times \mathbf{R}^n \longrightarrow \mathbf{R}^n$$

by $G(t,x,z) = [g(t,x)](z)$. Then given $\mathbf{x}_0 \in U$ there exist $d > 0$ and $b > c$ such that $J_0 = (-b, b) \subset J$, $U_0 = N_d(x_0) \subset U$ and a unique mapping

$$\lambda : J_0 \times U_0 \longrightarrow \mathcal{L}(\mathbf{R}^n, \mathbf{R}^n)$$

such that for each $x_0 \in U$ the map $J_0 \times \mathbf{R}^n \to \mathbf{R}^n$ sending (t,z) to $\lambda(t,x)z$ is an integral curve of

$$(t,z) \longrightarrow G(t,x,z)$$

with initial condition z.

Proof. $(2\star)$ Shrink J to J_1 and U to a neighborhood U_0 of x_0 such that the matrix norm $\|g\|$ is bounded on $J_1 \times U_0$ by 1. Consider the open disk $V = N_1(0) \subset \mathbf{R}^n$. let $G_x : J_1 \times V \to \mathbf{R}^n$ be defined by $G_x(t,y) = G(t,x,y) = [g(t,x)](y)$ for $x \in U_0$. Then

$$\|G_x(t,y)\| \;=\; \|[g(t,x](y)\| \;\leq\; \|g(t,x)\|\,|y| \;<\; 1$$

and similarly

$$\|G_x(t,y) \;-\; G_x(t,z)\| \;=\; \|[g(t,x](y-z)\| \;\leq\; \|g(t,x)\|\,|y-z| \;\leq\; |y-z| \;.$$

Thus G_x satisfies a Lipschitz condition such that the Lipschitz constant is independent of x. By the continuity of solutions with respect to the initial conditions, if we choose $1 > a > b > 0$ such that $b < \frac{1}{2}$, then we know that there exists a unique flow $\alpha_x : J_b \times N_a(0) \to \mathbf{R}^n$ for G_x that is a continuous map.

Let $V_0 = N_a(0)$ and $J_0 = J_b$; consider the map α defined on $J_0 \times U_0 \times V_0$ by $\alpha(t,x,y) = \alpha_x(t,y)$. We claim that for fixed t and x the map $\alpha(t,x,...)$ is locally linear on V_0; *i.e.*, it satisfies the conditions for a linear transformation on V_0. First of all, if $y,\, y' \in V_0$ such that $y + y' \in V_0$, then the curve $\gamma(t) = \alpha(t,x,y) + \alpha(t,x,y')$ satisfies the conditions to be an integral curve for G_x with initial condition $y + y'$. This and local uniqueness of solutions prove additivity. A similar argument (left to the reader) verifies homogeneity (the identity associated to multiplication by a scalar).

Now if we are given a map h from an open neighborhood of 0 in $\mathbf{R}^m$ to $\mathbf{R}^n$ that is linear on this neighborhood, there is a unique extension to a linear transformation from $\mathbf{R}^m$ to $\mathbf{R}^n$; to see this, choose a basis for $\mathbf{R}^m$ that lies in the neighborhood, define a linear transformation whose value on the basis are given by h, and note that the restriction of this linear transformation to the original neighborhood is just the original map. Let $\lambda(t,x)$ be the extension of $\alpha(t,x,...)$ obtained in this fashion.

We claim that λ is continuous. Let x_1 be a fixed point of U_0, let x be another point of U_0, and let $z \in V_0$. If $\varphi(t)$ and $\varphi_1(t)$ denote $[\lambda(t,x)](z)$ and $[\lambda(t,x_1)](z)$ respectively, then $\varphi(0) = \varphi_1(0) = z$ while $\varphi'(t)$ and $\varphi_1'(t)$ are equal to $[g(t,x)\lambda(t,x)](z)$ and $[g(t,x)\lambda(t,x_1)](z)$ respectively. If I is a compact subinterval of J_0, then there exists a $Q > 0$ such that $|\,[\lambda(t,x_1)](z)\,| < Q$ for all $t \in I$ and $z \in V_0$. Suppose an $\varepsilon > 0$ is given. Then there exists a $\delta > 0$ such that $|x - x_1| < \delta$ implies

$$\|\, g(t,x) - g(t,x_1)\| \;<\; \frac{\varepsilon}{Q}$$

(for each $t_0 \in I$ choose a $\delta(t_0)$ such that this inequality holds for all t in an open neighborhood of t_0 — then compactness yields a finite cover of I by these sets and we can take δ to be the minimum of the $\delta(t_i)$).

We now have

$$\left|\,\varphi_1' - [g(t,x)\lambda(t,x_1)](z)\,\right| \quad \leq$$

$$\left|\,(|\varphi_1'(t) - [g(t,x_1)\lambda(t,x_1)](z)\,) \;+\; ([g(t,x_1)\lambda(t,x_1)z|(z) - [g(t,x)\lambda(t,x_1)](z)\,)\,\right| \quad \leq$$

$$\|\,g(t,x_1) - g(t,x)\| \cdot |\lambda(t,x_1)z| \quad \leq \quad \varepsilon$$

if $|x - x_1| < \delta$, $t \in I$ and $z \in V_0$. Therefore $\alpha(t,x_1,z)$ is an approximate solution of the differential equation defined by $\alpha(t,x,z)$, and consequently by the continuity estimate we know that

$$\left|\,[\lambda(t,x)](z) - [\lambda(t,x_1)](z)\,\right| \quad \leq \quad \varepsilon\, Q_1$$

for some constant Q_1. Let $(t_1,x_1,z_1) \in I \times U_0 \times V_0$. Then

$$\left|[\lambda(t,x)](z) - [\lambda(t_1,x_1)](z_1)\right| \quad \leq$$

$$\left|[\lambda(t,x)](z) - [\lambda(t,x_1)](z)\right| \;+\; \left|[\lambda(t,x_1)](z) - [\lambda(t_1,x_1)](z)\right| \;+\; \left|[\lambda(t_1,x_1)](z) - [\lambda(t_1,x_1)](z_1)\right|\,.$$

For $\varepsilon > 0$ there is a $\delta > 0$ such that $|x - x_1| < \delta$ implies that the first term on the right hand side is less than $\varepsilon\, Q_1$. If t is close to t_1, then the second term will be small because $[\lambda(t,x_1)](z)$ is continuous (even differentiable) in t. Finally, if $|z - z_1|$ is small then the third term will be small. Therefore α is continuous on $I \times U_0 \times V_0$ for every compact subinterval of J_0 and hence α is continuous on $J_0 \times U_0 \times V_0$ (why?).

Since α is continuous, it follows that the map λ from $J_0 \times U_0$ to $\mathcal{L}(\mathbf{R}^n,\,\mathbf{R}^n)$ is also continuous (look at the functions corresponding to the matrix entries).∎

We proceed to examine the derivative of the flow for F in the second variable.

LEMMA. *Let J be an open interval in the real line containing 0, let U be open in $\mathbf{R}^n$, and let $F : J \times U \to \mathbf{R}^n$ be a $\mathcal{C}^1$ mapping. Given $x_0 \in U$, there exist open disks J_0 and U_0 centered at 0 and x_0 such that F has a unique local flow $\alpha : J_0 \times U_0 \to U$. We can select J_0 and U_0 such that α is Lipschitz and satisfies the following condition: Given $\alpha \in U_0$ and $\theta(t,h) = \alpha(t,x+h) - \alpha(t,x)$, then*

$$F\,(t,\,\alpha(t,x+h)\,) \;-\; F\,(t,\,\alpha(t,x)\,) \;\;=\;\; D_2F\,(t,\,\alpha(t,x)\,)\,\theta(t,h) \;+\; |h|\,\psi(t,h)$$

where ψ has the following property: Given $\varepsilon > 0$ there is a $\delta > 0$ such that $|h| < \delta$ implies $|\psi(t,h)| < \varepsilon$ for all $t \in J_0$.

Proof. (2⋆) Find J_1 and U_0 such that F is bounded and has a unique Lipschitz flow on $J_1 \times U_0$. We then choose an open subinterval $J_0 \subset J_1$ containing 0 such that the closure of J_0 is compact and contained in J_1. Fix $u_0 \in U_0$. By the Mean Value Theorem we have

$$\left|F\,(t,\,\alpha(t,x+h)\,) \;-\; F\,(t,\,\alpha(t,x)\,) \;-\; D_2F\,(t,\,\alpha(t,x)\,)\,\theta(t,h)\right| \quad \leq$$

$$|h| \cdot \sup_{0 \leq s \leq 1} \;\|D_2F(t,y_s) \;-\; D_2F\,(t,\,\alpha(t,x))\,\|$$

where $y_s = s\,\theta(t,h) - \alpha(t,x)$.

Since α is Lipschitz for each $u \in J_0$ we can find an open neighborhood $W_u \subset J_1$ such that if $\varepsilon > 0$ is given there exist a $\delta(u, \varepsilon)$ such that $|h| < \delta(u, \varepsilon)$ and $t \in W_u$ imply

$$\|D_2F(t, y_s) \;-\; D_2F\,(t,\,\alpha(t,x))\,\| \;\;<\;\; \varepsilon\;.$$

This follows from the continuity of D_2F and the Lipschitz condition on α (uniformly on compact subintervals of J_1. Cover J_1 by a finite number of these intervals W_u and let δ b the minimum of the numbers $\delta(u, \varepsilon)$ for this cover.■

We can now prove the smoothness property for local flows.

THEOREM. *Let J be an open interval on the real line containing* 0*, let U be open in $\mathbf{R}^n$, and let $F : J \times U \to \mathbf{R}^n$ be a smooth $\mathcal{C}^p$ map for $p \geq 1$. Then for each $x_0 \in U$ there is a unique local flow for f at x_0, and one can select open neighborhoods $J_0 \subset J$ and $U_0 \subset U$ of* 0 *and x_0 respectively such that the unique local flow $\alpha : J_0 \times U_0 \to U$ is smooth of class $\mathcal{C}^p$.*

Proof. $(2\star)$ Let $G : J \times U \times \mathbf{R}^n \to \mathbf{R}^n$ be the $\mathcal{C}^{p-1}$ map given by

$$G(t,x,z) \;\;=\;\; [D_2F(t,\,\alpha(t,x)\,)](z)\;.$$

We claim that the partial derivative $D_2\alpha$ exists locally and satisfies the equation

$$[D_1D_2\alpha(t,x)]\,(z) \;\;=\;\; [D_2F\,(t,\,\alpha(t,x)\,)\circ D_2\alpha(t,x)]\,(z)\;.$$

Apply the preceding proposition to find J_0 and U_0 such that the differential equation defined by G has a unique continuous solution $\beta(t,x,z) = [\lambda(t,x)](z)$ with λ bounded on $J_0 \times U_0$.

Shrink J_0 and U_0 so that F has a unique flow on the product and is Lipschitz. Shrink J_0 still further to satisfy the conditions of the preceding lemma. CLAIM: If $t, x) \in J_0 \times U_0$ then $D_2\alpha(t,x)$ exists and is equal to $\lambda(t,x)$.

Once again, let $\theta(t,h) = \alpha(t, x+h) - \alpha(t,x)$. Then the first partial derivative of θ is equal to

$$D_2F\,(t,\,\alpha(t,x)\,)\,\theta(t,h) \;+\; |h|\,\psi(t,h)\;.$$

Since x is fixed, write $\lambda(t) = \lambda(t,x)$ and $\beta(t,z) = \beta(t,x,z)$. Then $\lambda(t)$ is bounded on J_0, and thus the linear differential equation on $J_0 \times \mathbf{R}^n$ defined by

$$H(t,z) = [D_2F\,(t,\,\alpha(t,x)\,)]\,(z)$$

is Lipschitz on $\mathbf{R}^n$ uniformly with in J_0. Apply the estimate concerning approximate solutions to $\theta(t,h)$ and $\beta(t,h)$. It follows that given $\varepsilon > 0$ there exists a δ such that for every h satisfying $|h| < \delta$ and all $t \in J_0$ we have

$$|\theta(t,h) - \beta(t,h)| \;\;\leq\;\; \varepsilon\,|h|\;.$$

Thus the "partial derivative" $D_2\,\alpha$ exists and is equal to λ, and hence it is continuous. Since α has continuous partials with respect to both variables, it follows that α is $\mathcal{C}^1$. Finally, since

$$\alpha(t,x) \;\;=\;\; x \;+\; \int_0^t \; F\,(u,\,\alpha(u,x)\,)\;du$$

induction shows that α is also smooth of class $\mathcal{C}^p$.■

III. Global theory of smooth manifolds and mappings

The objective of this unit is to construct a mathematical framework which combines the concepts of the previous two units, yielding topological manifolds with additional structure which allows one to generalize the coordinate-free multivariable calculus from Unit II.

Given the amount of background information presented thus far and the repeated need for lengthy discussions of technical points in this unit, the following translated quotation from the online site

`http://www.mathstat.uottawa.ca/Profs/Rossmann/Differential%20Geometry%20book.htm`

seems worth including:

> Let me quote a piece of advice by Hermann Weyl from his classic **Raum - Zeit - Materie** [Space - Time - Matter] of 1923 (my [the writer's] translation). *Many will be horrified by the flood of formulas and indices which here drown the main idea[s] ... (in spite of the author's honest effort for conceptual clarity). It is certainly regrettable that we have to enter into purely formal matters in such detail and give them so much space; but this cannot be avoided. Just as we have to spend laborious hours learning language and writing to freely express our thoughts, so the only way that we can lessen the burden of formulas here is to master the tool[s] ... to such a degree that we can turn to the real problems that concern us without being bothered by formal matters.*

Here is some bibliographic information for the standard English translation of Weyl's book:

H. Weyl. Space - Time - Matter. *Dover, New York NY*, 1952. ISBN: 0–486–80267–2.

III.1 : Basic definitions and examples

(Conlon, §§ 3.1–3.2, 3.5)

Aside from open subsets in Euclidean spaces, the most basic examples of topological manifolds include the unit sphere S^n in Euclidean $(n+1)$-space. Before giving the definition, we shall describe more general objects that should also be smooth manifolds and make some important observations regarding these sets. Our observations will yield precisely the extra structure needed for an abstract definition of smooth manifold.

III.1.1 : *Level Sets of Regular Values*

By definition, unit sphere S^n in Euclidean $(n+1)$-space is the set of all x such that $f(x) = 1$, where $f(x)$ is the sum of the squares of the coordinates of x. The equation $f(x) = 1$ has an important regularity property; namely, $Df(x) \neq 0$ whenever $f(x) = 1$. Our first objective here is to generalize this example, showing that one obtains a topological n-manifold whenever this regularity property holds. We shall base our proof on results from Section II.2, particularly the Submersion Straightening Proposition. This proof yields a distinguished class of topological coordinate charts at all points of the level set, and one property of this class will be the key concept in the definition of smooth manifolds.

More generally, if $n > m$ and $f : \mathbf{R}^n \to \mathbf{R}^m$ is a smooth $\mathcal{C}^r$ map such that $Df(x)$ has rank m whenever $f(x) = y$, then we say that y is a *regular value* of f. The inverse image $f^{-1}(\{y\})$ is said to be the *level set* for the value y.

Our definition of regular value implies that y is a regular value if the level set is empty; of course, we are primarily interested in situations where this is not the case.

THEOREM ON LEVEL SETS. *Let f be as above, and assume that y is a regular value of f. Then the level set $f^{-1}(\{y\})$ is a second countable topological $(n-m)$-manifold. Furthermore, one has a family of topological charts (U_α, k_α) whose image covers the level set such that the maps "$k_\beta^{-1} \circ k_\alpha$" from $U_\alpha \cap h_\beta^{-1}(U_\beta)$ to $U_\beta \cap h_\alpha^{-1}(U_\alpha)$ are diffeomorphisms of class $\mathcal{C}^r$.*

Proof. The first step is to prove that the level set is a topological manifold; for convenience of notation we shall call this set L.

By hypothesis $Df(x)$ has rank m for all $x \in L$; in fact, we claim this also holds for all points in some open neighborhood of L. One simple but inelegant way of seeing this is as follows: If we let $x \in L$ and consider the $n \times m$ matrix for $Df(x)$ formed by the partial derivatives of the coordinate functions then the assumption on its rank implies that one can form an $m \times m$ submatrix of columns $A(x)$ that is invertible. Consider the square matrices $A(z)$ formed by taking the corresponding columns of $Df(z)$ for other choices of $z \in \mathbf{R}^n$. Since the partial derivatives are continuous and invertibility corresponds to the nonvanishing of the Jacobian $\det A(z)$, it follows that $A(z)$ is invertible for all z sufficiently close to x and consequently that the rank of $A(z)$ is m. Since x is arbitrary this means there is an open neighborhood G of L on which Df always has rank m; *i.e.*, the map $f|G$ is a smooth submersion.

We can now apply the Submersion Straightening Proposition to analyze the behavior of f at points of L. Given a point $x \in L$, it follows that we can find open neighborhoods U of x and V of y, an open set $W \subset \mathbf{R}^{n-m}$, and a diffeomorphism $k : V \times W \to U$ so that $f(k(v, w)) = v$ for all $(v, w) \in V \times W$. It then follows that k maps $\{y\} \times W$ homeomorphically to $L \cap U$, which is an open neighborhood of x in L. Since $\{y\} \times W$ is homeomorphic to W, this means that the neighborhood $U \cap L$ is homeomorphic to an open subset $W \subset \mathbf{R}^{n-m}$. To complete the proof that L is a topological manifold, one must check that it is Hausdorff, but this follows quickly because it is a subset of a Euclidean space; the same sorts of considerations show that L is also second countable.

The preceding argument yields a specific family of topological coordinate charts for L arising from the diffeomorphisms constructed in the first part of the proof. We wish to study this family in greater detail. As in Section II.2, it we are given a function B we shall frequently use "B" to denote a function defined by the same rules as B but possibly defined on a subset of the domain of B with a codomain that is possibly a subset of the codomain of B.

Suppose that we are given two diffeomorphisms $k_i : V_i \times W_i \to U_i$ as in the third paragraph of this proof, where $1 = 1, 2$ and U_i is open in G, and suppose that the images of k_1 and k_2 have a nonempty intersection $U_1 \cap U_2 = k_1(V_1 \times W_1) \cap k_2(V_2 \times W_2)$.

It then follows that there is a diffeomorphism

$$\Psi : (V_1 \times W_1) \cap k_1^{-1}(k_2(V_2 \times W_2)) \to (V_2 \times W_2) \cap k_2^{-1}(k_1(V_1 \times W_1))$$

such that Ψ is given by "$k_2^{-1} \circ k_1$" in the sense of the previous paragraph.

Since $f(k_i(v,w)) = v$ for $i = 1, 2$ it follows that Ψ maps

$$(\{y\} \times W_1) \cap k_1^{-1}(k_2(\{y\} \times W_2))$$

to the corresponding set

$$(\{y\} \times W_2) \cap k_2^{-1}(k_1(\{y\} \times W_1)).$$

Therefore we may write $\Psi(y,w) = (y, \psi_0(w))$, and it will follow immediately that ψ_0 is smooth (the coordinate functions of ψ_0 are given by those of Ψ and the latter have smooth partials. Applying the same considerations to the inverse map "$k_1^{-1} \circ k_2$" $= \Psi^{-1}$ we see that ψ_0 also has a smooth inverse so that ψ_0 is a diffeomorphism. Therefore if we choose a family of smooth diffeomorphims k_α as above so that the image sets $k_\alpha(V_\alpha \times W_\alpha)$ form an open covering of (a neighborhood of) L, it will follow that the corresponding charts $h_\alpha : W_\alpha \to L$ defined by $h_\alpha(w) = k_\alpha(y,w)$ define a smooth atlas for L. This completes the proof.■

III.1.2 : *Atlases and their maximal enlargements*

In order to define a smooth manifold, we need an abstract formulation of the final conclusion in the Theorem on Level Sets. It will be convenient to start in a purely topological setting.

Given a topological n-manifold M, an *atlas* (more precisely, a *topological atlas*) for M is a collection $\mathcal{A}$ of ordered pairs (U_α, h_α) such that each U_α is homeomorphic to an open subset in $\mathbf{R}^n$, each h_α is a homeomorphism from U_α to an open subset of M, and the sets $h_\alpha(U_\alpha)$ form an open covering of M. The pairs (U_α, h_α) are called the *charts* of the atlas.

By construction every topological manifold has a topological atlas, and there is an obvious maximal topological atlas: Simply take all (U, h) so that U is open in $\mathbf{R}^n$ and h is a homeomorphism from U onto an open subset of M. Clearly this is also the **only** maximal topological atlas.

The related concept of *smooth atlas* (of class $\mathcal{C}^r$) is fundamental to this course. A topological atlas $\mathcal{A}$ is said to be *smooth of class* $\mathcal{C}^r$ (where $1 \le r \le \infty$) if for all pairs (U_α, h_α) and (U_β, h_β) in $\mathcal{A}$ the transition homeomorphisms

$$\psi_{\beta\alpha} : h_\alpha^{-1}(h_\beta(U_\beta)) \to h_\beta^{-1}(h_\alpha(U_\alpha))$$

defined by $\psi_{\beta\alpha}(x) = h_\beta^{-1}(h_\alpha(x))$ are smooth diffeomorphisms (of class $\mathcal{C}^r$). By the Theorem on Level Sets, it follows that the level set of $\mathcal{C}^r$ functions always has a smooth atlas of class $\mathcal{C}^r$.

A smooth atlas (of class $\mathcal{C}^r$) turns out to be the additional structure one needs to talk about smooth mappings on manifolds. However, if we simply say that a smooth structure is given by a topological manifold M and a smooth atlas for M (of class $\mathcal{C}^r$), we will end up with a lot of redundancy that is at best clumsy and at worst confusing. Perhaps the simplest examples involve open subsets in Euclidean spaces. If U is open in $\mathbf{R}^n$ then the simplest example of a smooth atlas (of class $\mathcal{C}^r$) is just (U, id_U). However, if U_α is an arbitrary open covering of U and $i_\alpha : U_\alpha \to U$ is the inclusion map, then each family (U_α, i_α) is also a smooth atlas (of class $\mathcal{C}^r$). Our definition of smooth structure should be formulated so that an open subset in $\mathbf{R}^n$ has a unique associated smooth structure rather than a multitude of smooth structures given by all possible open coverings as well as even larger atlases.

Similar considerations apply for the sphere S^n. One example of a smooth atlas of class $\mathcal{C}^\infty$ for S^n is given by the stereographic projection charts in Example 1.2.3 of Conlon (see p. 3 of that reference; also see the material on stereographic projections in the ONLINE 205A NOTES.). Another was given in Section I.1, with charts of the form

$$h_{i\pm} : \{|x| < 1\} \to S^n$$

such that $h_{i\pm}$ sends the first $(i-1)$ coordinates to themselves, shifts the remaining coordinates to the $(i+1)$ through $(n+1)$ coordinates on $S^n \subset \mathbf{R}^{n+1}$ and inserts $\pm\sqrt{1-|x|^2}$ in the i-th coordinate.

In analogy with the topological case, one would like to have a **universal** smooth atlas of class $\mathcal{C}^r$ associated to a smooth structure. This is given by the following result:

THEOREM ON MAXIMAL ATLASES. *Let r be a positive integer or ∞. If $\mathcal{A}$ is a smooth atlas of class $\mathcal{C}^r$ for the n-manifold M, then there is a unique* MAXIMAL *smooth atlas $\mathcal{A}'$ of class $\mathcal{C}^r$ containing $\mathcal{A}$. A chart (V, k) belongs to $\mathcal{A}'$ if and only if for each chart (U_α) in $\mathcal{A}$ the associated transition maps from $h_\alpha^{-1}(k(V)))$ to $k^{-1}(h_\alpha(U_\alpha))$ and vice versa are diffeomorphisms of class $\mathcal{C}^r$.*

Proof. By construction the set $\mathcal{A}'$ contains $\mathcal{A}$. There are three things to prove:

(1) If $\mathcal{A}'$ is defined as in the statement of the theorem, then it is a smooth atlas of class $\mathcal{C}^r$.

(2) $\mathcal{A}'$ is a maximal smooth atlas of class $\mathcal{C}^r$.

(3) $\mathcal{A}'$ is the only maximal smooth atlas of class $\mathcal{C}^r$ containing $\mathcal{A}$.

Verification of (1). We need to show that if (V, k) and (W, ℓ) are charts in $\mathcal{A}'$ then the transition map "$\ell^{-1} \circ k$" from $k^{-1}(\ell(W))$ to $\ell^{-1}(k(V))$ is smooth of class $\mathcal{C}^r$. Let x be an arbitrary point in the first subset.

By the definition of a smooth atlas there is a smooth chart (U_α, h_α) of class $\mathcal{C}^r$ in $\mathcal{A}$ such that $k(x) \in h_\alpha(U_\alpha)$; *i.e.*, $x \in k^{-1}\ell(W) \cap k^{-1}(h_\alpha(U_\alpha))$. Under the transition map from $k^{-1}(\ell(W))$ to $\ell^{-1}(k(V))$ the subset $k^{-1}(\ell(W)) \cap k^{-1}(h_\alpha(U_\alpha))$ is mapped to $\ell^{-1}(k(V)) \cap \ell^{-1}(h_\alpha(U_\alpha))$. On these subsets one can express "$\ell^{-1} \circ k$" as a composite ("$\ell^{-1} \circ h_\alpha$")$\circ$("$h_\alpha^{-1} \circ k$") and by the defining condition for $\mathcal{A}'$ it follows that both of these composites are diffeomorphisms. Therefore the transition map "$\ell^{-1} \circ k$" is smooth of class $\mathcal{C}^r$ on the open set $k^{-1}(\ell(W)) \cap k^{-1}(h_\alpha(U_\alpha))$ and hence is smooth of class $\mathcal{C}^r$ near x. Since x was arbitrary this implies that the transition map is smooth of class $\mathcal{C}^r$ everywhere.

Verification of (2). If we try to add another chart (N, φ) to $\mathcal{A}'$ then the defining condition for the latter implies that at least one of the transition maps "$\varphi^{-1} \circ h_\alpha$" is not a diffeomorphism of class $\mathcal{C}^r$. But this means that $\mathcal{A}'$ with the extra chart does not form a smooth atlas of class $\mathcal{C}^r$.

Verification of (3). Suppose that $\mathcal{H}$ is an arbitrary maximal smooth atlas of class $\mathcal{C}^r$ containing $\mathcal{A}$, and let (N, φ) be a chart in $\mathcal{H}$. Since the latter contains $\mathcal{A}$ it follows that all of the transition maps "$\varphi^{-1} \circ h_\alpha$" are diffeomorphisms. But this implies that (N, φ) is a chart in $\mathcal{A}'$, which in turn implies $\mathcal{H} \subset \mathcal{A}'$. Since both are maximal they must be equal.■

III.1.3 : *Definitions of smooth manifolds and maps*

We are **FINALLY** ready to define the main concepts of this course.

FUNDAMENTAL DEFINITION I. Suppose that r is a positive integer or ∞. A **smooth n-manifold** of class $\mathcal{C}^r$ is a pair $(M, \mathcal{A})$ where M is a topological n-manifold and $\mathcal{A}$ is a maximal smooth atlas on M of class $\mathcal{C}^r$. Frequently one says that $\mathcal{A}$ is a *smooth structure* (of class $\mathcal{C}^r$) on M. The charts in the maximal atlas are said to be *smooth charts* (of class $\mathcal{C}^r$) for the smooth manifold or smooth structure.

STANDARD NOTATIONAL CONVENTION. If the maximal smooth atlas $\mathcal{A}$ is implicit from the context, we shall often refer to $(M, \mathcal{A})$ simply as M.

In many books and papers the term **differential manifold** is used as a synonym for **smooth manifold**, and to a lesser extent the same is true for the term **differentiable manifold**. However, the latter is also sometimes used simply to denote a manifold that has a smooth structure.

The following result is an elementary but very useful consequence of the definitions:

CHART RESTRICTION LEMMA. *Let $\mathcal{A}$ be a maximal atlas for M, let (U, h) be a chart in $\mathcal{A}$, and let V be an open subset of U. Then $(V, h|V)$ is also a chart in $\mathcal{A}$.*■

Proof. Let $j : V \to U$ be the inclusion map, and let (W, k) be another chart in $\mathcal{A}$. We need to show that the map "$k^{-1} \circ (h \circ j)$" is a diffeomorphism. This map is a homeomorphism from $h^{-1}\,(\,k(W)\,) \cap V$ to $k^{-1}\,(\,h(V)\,)$. By construction, if

$$j_0 : h^{-1}\,(\,k(W)\,) \cap V \longrightarrow j_0 : h^{-1}\,(\,k(W)\,) \qquad \text{and} \qquad \ell : k^{-1}\,(\,h(V)\,) \longrightarrow k^{-1}\,(\,h(U)\,)$$

are the inclusion maps, then we have $\ell \circ$ "$k^{-1} \circ (h \circ j)$" $=$ "$k^{-1}h$" $\circ j_0$, and since the second is a composite of smooth maps it follows that the maps on both sides of the equation are smooth. But ℓ is also an inclusion of an open subset, and therefore it follows that "$k^{-1} \circ (h \circ j)$" is smooth. To show this map is a diffeomorphism, look at the derivatives; by the Inverse Function Theorem it suffices to show that the derivative of "$k^{-1} \circ (h \circ j)$" is invertible. Since ℓ is an inclusion, the map $D\ell(z)$ is the identity for all x, so it suffices in turn to show that the derivative of $\ell \circ$ "$k^{-1} \circ (h \circ j)$" is invertible. "$k^{-1}h$" $\circ j_0$; since Dj_0 equals the identity and the derivative of "$k^{-1}h$" is invertible whenever "$k^{-1}h$" is is defined, it follows that all derivatives in sight are invertible. Therefore we have shown that "$k^{-1} \circ (h \circ j)$" is a diffeomorphism, which is what we wanted to prove.■

A FEW EXAMPLES.

Example 1. As noted above, if U is an open subset of $\mathbf{R}^n$ then we have a smooth $\mathcal{C}^\infty$ atlas $\mathcal{E}_U$ whose sole chart is (U, id_U), and the **standard $\mathcal{C}^\infty$ smooth structure** on U is given by the maximal $\mathcal{C}^\infty$ atlas containing $\mathcal{E}_U$.

Example 2. If U is open in $\mathbf{R}^{m+n}$ and $f : U \to \mathbf{R}^m$ is a smooth $\mathcal{C}^r$ function such that $f(x) = 0 \Longrightarrow Df(x)$ is onto, then the Theorem on Level Sets yields a canonical smooth $\mathcal{C}^r$ atlas on the level set $L = f^{-1}(\{0\})$. The standard $\mathcal{C}^r$ structure on L associated to f is given by the maximal atlas containing the canonical smooth $\mathcal{C}^r$ atlas.

Example 3. The following example of a nonstandard smooth structure on $\mathbf{R}$ looks extremely formal, but it illustrates some extremely important points. In general, if U is an open subset of $\mathbf{R}^n$ and $h : U \to U$ is a homeomorphism, then one can form a smooth $\mathcal{C}^\infty$ atlas $\mathcal{E}_h$ whose sole chart is (U, h). This always yields a $\mathcal{C}^\infty$ atlas because the only transition map is the identity $(= h \circ h^{-1} = h^{-1} \circ h)$. We shall consider the special case where $U = \mathbf{R}$ and $h(x) = x^3$.

CLAIM: *The maximal $\mathcal{C}^\infty$ atlas containing this chart is not equal to the maximal $\mathcal{C}^\infty$ atlas associated to the standard smooth structure in the first example.* To see this, assume the contrary, so that (U, h) would be a smooth chart in the standard atlas and hence the transition map $h^{-1} =$

$h^{-1} \circ \mathrm{id}_U$ would be a $\mathcal{C}^\infty$ diffeomorphism. However, the map h^{-1} sends a real number x to its unique real cube root, and we know this map does not even have a derivative at $x = 0$. Note that everything would work just the same if we replaced ∞ by an arbitrary positive integer r.

Although the smooth structures determined by $\mathcal{E}_1$ and $\mathcal{E}_h$ are distinct for $h(x) = x^3$ on the real line, one exercise for this section will prove that the smooth manifolds $(\mathbf{R}, \mathcal{E})$ and $(\mathbf{R}, \mathcal{E}_h)$ are equivalent in an appropriate sense.

Example 4. This is actually a negative example to show that the Forked Line, which is not Hausdorff, has a smooth atlas as defined above. In fact, the existence of such an atlas follows immediately from the construction in Section I.1. Specifically, if we view the forked line **FL** as a quotient space of $\mathbf{R} \times \{0,1\}$ such that $(x,0)$ is identified with $(x,1)$ for all $x \neq 0$, consider the 1–1 continuous open mappings

$$h_i : \mathbf{R} \to \mathbf{R} \times \{0,1\} \to \mathbf{FL}$$

such that $h_i(x)$ is the image of $(x,0)$ in **FL**. These two maps define an atlas for the Forked Line, and direct inspection shows that each transition maps is merely the identity on $\mathbf{R} - \{0\}$. Therefore the atlas consisting of $(\mathbf{R}, h_0)$ and $(\mathbf{R}, h_1)$ is $\mathcal{C}^\infty$ according to the definition given above, and accordingly we may view **FL** as a non-Hausdorff smooth $\mathcal{C}^\infty$ manifold.

The next step for us is to **define smooth maps of smooth manifolds**. This will require some preparation. Suppose that $(M, \mathcal{A})$ and $(N, \mathcal{B})$ are smooth manifolds, let $\mathcal{B}_0$ be a smooth subatlas of $\mathcal{B}$, and let $f : M \to N$ be a continuous function. The preceding lemma and continuity imply the existence of a smooth atlas $\mathcal{A}_0 \subset \mathcal{A}$ such that for each smooth chart (U,h) in $\mathcal{A}_0$ there is a smooth chart $(V,k) \in \mathcal{B}_0$ such that $f(h(U)) \subset k(V)$ (explain in detail why this is true!).

FUNDAMENTAL DEFINITION II. Suppose that r is a positive integer or ∞ and let $(M, \mathcal{A})$ and $(N, \mathcal{B})$ be smooth manifolds of class $\mathcal{C}^r$. A **smooth map of class** $\mathcal{C}^r$ from $(M, \mathcal{A})$ to $(N, \mathcal{B})$ is a continuous map $f : M \to N$ of the underlying topological spaces with the following additional property:

($\mathbf{L}\mathcal{C}^r$) Given smooth atlases $\mathcal{B}_0$ and $\mathcal{A}_0$ as in the preceding paragraph and charts (U,h) in $\mathcal{A}_0$ and $(V,k) \in \mathcal{B}_0$ such that $f(h(U)) \subset k(V)$, the associated map "$k^{-1} \circ f \circ h$": $U \to V$ is smooth.

This definition is very useful for proving that a continuous map is smooth because it allows one to choose the smooth atlases $\mathcal{B}_0$ and $\mathcal{A}_0$ in a convenient manner. However, this freedom of choice is also a potential shortcoming because one can ask what might happen if another pair of atlases is chosen. The next result implies that if $(\star)$ holds for one choice of atlases then it holds for all such choices.

WEAK SMOOTHNESS CRITERION. *Let $(M, \mathcal{A})$ and $(N, \mathcal{B})$ be smooth manifolds of class $\mathcal{C}^r$, let $f : M \to N$ be continuous map, and assume that ($\mathbf{L}\mathcal{C}^r$) holds for suitably chosen subatlases of $\mathcal{A}$ and $\mathcal{B}$. Then ($\mathbf{L}\mathcal{C}^r$) also holds for all smooth charts (U_α, h_α) in $\mathcal{A}$ and $(V_\beta, k_\beta) \in \mathcal{B}$ such that $f(h_\alpha(U)) \subset k_\beta(V)$.*

Proof. Take arbitrary charts (U_α, h_α) in $\mathcal{A}$ and $(V_\beta, k_\beta) \in \mathcal{B}$ such that $f(h_\alpha(U)) \subset k_\beta(V)$, and let $x \in U_\alpha$. Choose charts (N,p) and (Ω, q) in $\mathcal{A}_0$ and $\mathcal{B}_0$ respectively such that $x \in p(N)$, $f(x) \in q(\Omega)$ and $f(p(N)) \subset q(\Omega)$. It will suffice to show that the restriction of "$k_\beta^{-1} \circ f \circ h_\alpha$" to $h_\alpha^{-1}(p(N))$ is smooth (because x is an arbitrary point in U_α and the set in question is an open neighborhood of x).

By hypothesis the map "$q^{-1} \circ f \circ p$" is smooth. Furthermore, as in the proof of the preceding theorem, on the open subset $h_\alpha^{-1}(p(N))$ the composite "$k_\beta^{-1} \circ f \circ h_\alpha$" may be written as a composite

$(\text{“}k_\beta^{-1} \circ q\text{”}) \circ (\text{ “}q^{-1} \circ f \circ p\text{”}) \circ (\text{ “}p^{-1} \circ h_\alpha\text{”})$; the middle factor is smooth by hypothesis, and the first and last factors are smooth because they are transition maps in the maximal atlases. Therefore the factors of the threefold composite are all smooth maps, and hence the composite itself is smooth.■

One can now define diffeomorphisms of class $\mathcal{C}^r$ for smooth manifolds of class $\mathcal{C}^r$ exactly as for open sets in Euclidean spaces: They are smooth maps of class $\mathcal{C}^r$ with smooth inverses of class $\mathcal{C}^r$. It is also possible to formulate definitions of submersions and immersions that work for smooth manifolds, but these will be easier to state in terms of tangent spaces, which will be constructed in Section III.5 below. The global definitions appear at the beginning of Section III.6.

III.1.4 : *Some elementary but important results*

At this point it might seem natural to consider additional examples of smooth manifolds and mappings, possibly along the lines of the examples given for topological manifolds in Section I.1. However, just as the definition of a smooth manifold is more delicate than one might first expect, the construction of some basic examples is also more complicated than for their topological counterparts. In this subsection we shall merely discuss one description of smooth manifolds that is often taken as the definition in undergraduate textbooks and then proceed to list some basic properties of smooth functions that will expedite the discussion of examples in Section III.2.

VARIANT OF THE THEOREM ON LEVEL SETS. Frequently textbooks, particularly at the undergraduate level (*cf.* the book by Edwards listed below), define a smooth n-manifold in a manner equivalent to the following: *A subset X of some Euclidean space $\mathbf{R}^K$ such that every point $x \in X$ has an open neighborhood U in $\mathbf{R}^K$ such that $X \cap U$ is the level set of a regular value for some smooth function $g_U : U \to \mathbf{R}^{K-n}$.*

The proof of the Theorem on Level Sets extends with only a few changes to show that such a set X is a second countable topological n-manifold and has a canonical smooth atlas.■

In fact, one can show that the definition described above is equivalent to the one adopted for this course. A proof of equivalence requires the Euclidean Embedding Property from Section III.6.

Here is the reference data for the book by Edwards cited above:

> **C. H. Edwards, Jr.** Advanced Calculus of Several Variables. (Corrected Reprint of 1973 Edition). *Dover, Mineola NY*, 1994. ISBN: 0–496–68336–2.

ELEMENTARY FORMAL PROPERTIES OF SMOOTH MAPPINGS. The following result shows that smooth $\mathcal{C}^r$ maps of smooth $\mathcal{C}^r$ manifolds have several of the same basic properties that hold for similar maps on open subsets of Euclidean spaces.

THEOREM. *Let $1 \leq r \leq \infty$.*

(i) If $(M, \mathcal{A})$ is a smooth $\mathcal{C}^r$ manifold, then the identity map on M is smooth of class $\mathcal{C}^r$.

(ii) If $(M, \mathcal{A})$ and $(N, \mathcal{B})$ are smooth $\mathcal{C}^r$ manifolds and $g : M \to N$ is a constant map, then $g : (M, \mathcal{A}) \to (N, \mathcal{B})$ is smooth of class $\mathcal{C}^r$.

(iii) If $f : (M, \mathcal{A}) \to (N, \mathcal{B})$ and $g : (N, \mathcal{B}) \to (L, \mathcal{E})$ are smooth $\mathcal{C}^r$ maps of smooth $\mathcal{C}^r$ manifolds, then the composite $g \circ f$ is smooth of class $\mathcal{C}^r$.

Proof. Throughout these arguments “smooth” will mean smooth of class $\mathcal{C}^r$.

Smoothness of the identity map. If (U,h) is an arbitrary smooth chart for $(M,\mathcal{A})$, then id_M maps $h(U)$ into itself and the associated map in local coordinates "$h^{-1}\circ \mathrm{id}_M \circ h$" is merely the identity map of U. Since the latter is smooth and we began with an arbitrary chart, it follows that id_M satisfies the local criterion for smoothness.■

Smoothness of the constant map. Let $p_0 \in N$ be the unique value for the constant map g, let (V_0, k_0) be a smooth chart in $\mathcal{B}$ at p_0, and choose q_0 to be the unique point in V_0 such that $p_0 = k_0(q_0)$. Then for each chart (U,h) in $\mathcal{A}$ we know that $h(U) \subset k_0(V_0)$, and in fact the map in local coordinates "$k_0^{-1}\circ g \circ h$" is merely the constant function whose value at every point is q_0. Since constant maps on open subsets of Euclidean spaces, it follows that the map "$k_0^{-1}\circ g\circ h$" is smooth, and therefore it follows that g satisfies the local criterion for smoothness.■

Smoothness of composites. By the continuity of g and the maximality of $\mathcal{B}$, there is a subatlas $\mathcal{B}_0$ of $\mathcal{B}$ such that for each chart (V_1, k_1) in $\mathcal{B}$ one can find a smooth chart (W_1, ℓ_1) in $\mathcal{E}$ such that $g(\,k_1(V_1)\,) \subset \ell_1(W_1)$. Likewise, by the continuity of f and the maximality of $\mathcal{A}$, there is a subatlas $\mathcal{A}_0$ of $\mathcal{A}$ such that for each chart (U_0, h_0) in $\mathcal{A}$ one can find a smooth chart (V_0, k_0) in $\mathcal{B}_0$ such that $f(\,h_0(U_0)\,) \subset k_0(V_0)$. Given this choice for (V_0, k_0), let (W_0, ℓ_0) in $\mathcal{E}$ be a smooth chart such that $g(\,k_0(V_0)\,) \subset \ell_0(W_0)$.

In order to prove that $g\circ f$ is smooth, it suffices to prove that all composites of the form "$\ell_0^{-1}\circ(g\circ f)\circ h_0$" are smooth. The key to doing this is to observe that such maps of open subsets of Euclidean spaces have factorizations of the form "$i\left(\ell_0^{-1}\circ g\circ k_0\right)$" $\circ$ "$\left(k_0^{-1}\circ f\circ h_0\right)$" for which each of the factors is also a map of open subsets of Euclidean spaces. However, by our smoothness assumptions on f and g we know that each of these factors is smooth as map of open subsets of Euclidean spaces, and therefore their composite must also be smooth in the same sense. This means that the original composite $g\circ f$ also satisfies the local criterion for smoothness and therefore that $g\circ f$ must also be smooth.■

III.1.5 : *Relations among smoothness classes* $(\star)$

Once again suppose that r is a positive integer or ∞, and assume the same for s with $r > s$. Then every $\mathcal{C}^r$-atlas for a manifold is automatically a $\mathcal{C}^s$-atlas (WARNING: A maximal $\mathcal{C}^r$ atlas will never be a maximal $\mathcal{C}^s$-atlas — the proof of this is left to the exercises). Furthermore, we also have the following elementary result:

PROPOSITION. *Let $(M,\mathcal{A})$ and $(N,\mathcal{B})$ be smooth manifolds of class $\mathcal{C}^r$, where r is above. Similarly, let $s < r$ as above, and let $(M,\mathcal{A}_s)$ to $(N,\mathcal{B}_s)$ be the associated smooth $\mathcal{C}^s$-manifolds (hence $\mathcal{F}_s$ is the maximal $\mathcal{C}^s$ atlas containing $\mathcal{F}$ for $\mathcal{F} = \mathcal{A},\mathcal{B}$). If $f : (M,\mathcal{A}) \to (N,\mathcal{B})$ is a smooth $\mathcal{C}^r$ map then $f : (M,\mathcal{A}_f) \to (N,\mathcal{B}_f)$ is a smooth $\mathcal{C}^s$ map.*

Proof. $(\star)$ Checking that f is $\mathcal{C}^s$ reduces to checking the local condition for appropriately defined subatlases, and we can use $\mathcal{B}$ and a judiciously chosen subatlas of $\mathcal{A}$ for this purpose. Since f is locally $\mathcal{C}^r$ with respect to the associated coordinate charts, it is also $\mathcal{C}^s$. Therefore $f : (M,\mathcal{A}_f) \to (N,\mathcal{B}_f)$ must be a $\mathcal{C}^s$ map as desired.■

In view of the preceding observations, it is natural to ask the following question: *If $1 \leq s < r$ as above and we are given a smooth manifold of class $\mathcal{C}^s$, under what conditions does it come from a smooth manifold of class $\mathcal{C}^r$?*

The answer to this question turns out to be affirmative in a very strong sense.

THEOREM ON RAISING DIFFERENTIABILITY CLASSES. *Suppose that* $1 = \leq s < r \leq \infty$ *and that* $(M, \mathcal{A})$ *is a smooth manifold of class* $\mathcal{C}^s$*. Then there is a smooth manifold* $(M, \mathcal{A}^*)$ *of class* $\mathcal{C}^r$ *such that* $(M, \mathcal{A})$ *is equal to* $(M, \mathcal{A}^*_s)$*. Furthermore,* $(M, \mathcal{A}^*)$ *is unique up to* $\mathcal{C}^r$ *diffeomorphism.*■

A proof of this result is given in Section 4 of the following classic book by Munkres:

> **J. R. Munkres**. Elementary differential topology. Lectures given at Massachusetts Institute of Technology, Fall, 1961. Revised edition. Annals of Mathematics Studies, No. 54 Princeton University Press, Princeton, N. J., 1966. ISBN: 0-691-09093-9.

There will be many places in these notes where we shall refer to this book, and henceforth we shall call it [MUNKRES2]).

Because of the Theorem on Raising Differentiability Classes, no essential examples of smooth manifolds are lost if we restrict attention to $\mathcal{C}^\infty$ manifolds. Since there will be many instances throughout the rest of this course where things simplify considerably if one only considers $\mathcal{C}^\infty$ manifolds and mappings, at this point we shall focus almost exclusively on the latter. Formally, this will be stated as follows:

DEFAULT CONVENTION. *Henceforth, unless stated otherwise explicitly,* ALL SMOOTH MANIFOLDS AND MAPPINGS WILL BE ASSUMED TO BE SMOOTH OF CLASS $\mathcal{C}^\infty$.

Reminder. Even in the case of open subsets in Euclidean spaces, for each finite value of r we know that there are smooth $\mathcal{C}^r$ maps that are not smooth of class $\mathcal{C}^{r+1}$. — Our Default Convention means that we shall generally avoid considering such maps. For many purposes this is not a problem because in many cases one can approximate a smooth $\mathcal{C}^r$ map ($r < \infty$) with certain good properties by a $\mathcal{C}^\infty$ map with the same good properties. A detailed account of such results is beyond the scope of this course, but some general results on $\mathcal{C}^\infty$ approximations appear in [MUNKRES2].

III.1.6 : *Manifolds with no smooth structures* ($2\star$)

It is also natural to ask whether an arbitrary topological manifold has a smooth structure. For manifolds of dimension ≤ 3, the answer is always yes. There is a proof of this in the book by R. C. Kirby and L. C. Siebenmann listed below; however, the result itself had been known decades earlier. The first example of a topological manifold with no smooth structure was discovered by M. Kervaire in the late nineteen fifties (see the reference below), and the book by Kirby and Siebenmann translates the existence question for smooth structures to a question in algebraic topology for manifolds of dimension $\neq 4$. In particular, results in this book showed the existence of nonsmoothable topological manifolds in all dimensions ≥ 5. Subsequent research on 4-manifolds proved that

(i) every **noncompact** and **connected** 4-dimensional manifold admits a smooth structure,

(ii) there exist compact 4-dimensional manifolds with no smooth structures, (*iii*) the higher dimensional result translating the existence of smooth structures to questions in algebraic topology does **not** extend to dimension 4.

Additional information on the 4-dimensional case appears in the book by M. H. Freedman and F. S. Quinn and the paper by R. Lashof and L. Taylor listed below.

Although there are many examples of topological manifolds that have no smooth structures, for manifolds of dimension $\neq 4$ a result of D. Sullivan shows that one can always find a topological atlas for which the transition maps are **bi-Lipschitz homeomorphisms**; *i.e.*, the homeomorphisms and their inverses satisfy Lipschitz conditions. Furthermore, Sullivan's results show that this structure is unique up to a suitably defined notion of equivalence. Subsequent work of Sullivan and S. Donaldson shows that both existence and uniqueness fail for 4-dimensional manifolds.

Here are references for some of the points mentioned in this subsection:

[1] S. K. Donaldson and D. P. Sullivan, *Quasiconformal 4-manifolds*, Acta Math. **163** (1989), 181–252.

[2] M. H. Freedman and F. S. Quinn III, Topology of 4-manifolds (Princeton Mathematical Series, Vol. 39). Princeton University Press, Princeton, N. J., 1990. ISBN: 0-691-08577-3 .

[3] M. A. Kervaire, *A manifold which does not admit any differentiable structure*, Comment. Math. Helv. **34** (1960), 257–270.

[4] R. C. Kirby and L. C. Siebenmann, Foundational essays on topological manifolds, smoothings, and triangulations (With notes by John Milnor and Michael Atiyah; Annals of Mathematics Studies, No. 88). Princeton University Press, Princeton, N.J., 1977. ISBN: 0-691-08191-3.

[5] R. Lashof and L. Taylor *Smoothing theory and Freedman's work on four-manifolds* Algebraic topology, Aarhus 1982 (Conference Proceedings, Aarhus, 1982), 271–292, Lecture Notes in Math., 1051, Springer, New York, 1984.

[6] D. P. Sullivan, *Hyperbolic geometry and homeomorphisms*, Geometric Topology (Proc. Georgia Topology Conf., Athens, Ga., 1977), pp. 543–555, Academic Press, New York-London, 1979.

III.1.7 : *Nondiffeomorphic smooth structures on a given manifold* ($2\star$)

In the previous subsection we have noted that a topological manifold does not necessarily have a smooth structure. Furthermore, at the beginning of this section we also gave an example to show that $\mathbf{R}$ has at least two smooth structures, but we also mentioned that the associated smooth manifolds are diffeomorphic. These two observations lead naturally to the following question:

CLASSIFICATION QUESTION FOR SMOOTH STRUCTURES. *Given a topological manifold M let $\mathbf{S}^{\#}(M)$ be the set of equivalence classes of smooth manifolds $(M, \mathcal{A})$ modulo the equivalence relation identifying $(M, \mathcal{B}_1)$ with $(M, \mathcal{B}_2)$ if and only if the latter are diffeomorphic. What can one say about $\mathbf{S}^{\#}(M)$ if we know it is nonempty? In particular, is it possible to find examples for which $\mathbf{S}^{\#}(M)$ contains more than one element?*

Results from the previously listed book by Kirby-Siebenmann implies that $\mathbf{S}^{\#}(M)$ contains exactly one element if $\dim M \leq 3$ (in fact, this result had already been known before the appearance of the Kirby-Siebenmann book, but a detailed historical discussion is beyond the scope of this course). During the nineteen fifties J. Milnor wrote a short, elegant and totally unanticipated paper which showed that $\mathbf{S}^{\#}(S^7)$ contained more than one element. Subsequent results, culminating with the book by Kirby and Siebenmann, led to a partial answer for the classification question in terms of algebraic topology, subject to some inherent redundancies and an assumption that the dimension is ≥ 5. Here are three qualitative consequences that are expressible in very simple terms:

EXISTENCE OF EXOTIC STRUCTURES. *(i) For all $k \geq 2$ the set $\mathbf{S}^{\#}(S^3 \times T^k)$ contains more than one element.*

(ii) For all odd integers $2m+1 \geq 7$ such that $2m+4$ is not a power of 2, the set $\mathbf{S}^{\#}(S^{2m+1})$ contains more than one element.■

NOTE: There dimensions in part *(ii)* are definitely not the only ones for which the set $\mathbf{S}^{\#}(S^k)$ contains more than one element. On the other hand, it is known that this set contains exactly one element if $k < 7$ with the possible exception of $k = 4$ (this unknown case is known as the *smooth 4-dimensional Generalized Poincaré Conjecture*).

FINITENESS PROPERTY. *If M is a compact topological manifold of dimension $\neq 4$, then $\mathbf{S}^{\#}(M)$ is finite.*■

NONCOMPACT ANALOG. *If M is a noncompact second countable topological manifold of dimension $\neq 4$, then $\mathbf{S}^{\#}(M)$ is countable.*■

We shall explain why several of the references below combine to imply this result in the file(s) `countsmoothstructures.*` (where as usual $*$ = `ps` or `pdf`) in the course directory.

Here is another consequence of the work of Kirby and Siebenmann (which had been known for a while before their breakthroughs in the structure theory of topological manifolds):

UNIQUENESS OF SMOOTHINGS FOR EUCLIDEAN SPACES. *If $n \neq 4$ then every smooth manifold that is homeomorphic to $\mathbf{R}^n$ is diffeomorphic to $\mathbf{R}^n$.*■

In contrast, if $n = 4$ all these results turn out to be systematically false. If M is a 4-manifold, then results of J. Cerf, Hirsch-Mazur and Lashof-Rothenberg combine with the second half of [MUNKRES2] to show that $\mathbf{S}^{\#}(M)$ is countable if M is compact and has cardinality $\leq 2^{\aleph_0}$ in the noncompact case (we shall also explain this in the file(s) `countsmoothstructures.*` mentioned above). In fact, one can find compact 4-manifolds for which $\mathbf{S}^{\#}(M)$ is infinite, and it is also known that the cardinality of $\mathbf{S}^{\#}(\mathbf{R}^4)$ is equal to $2^{\aleph_0}$. The result in the compact case is due to R. Friedman and J. Morgan, while the result for smooth structures on $\mathbf{R}^4$ is due to C. Taubes.

Here is a list of the references cited above:

[1] J. Cerf, Sur les difféomorphismes de la sphére de dimension trois ($\Gamma_4 = 0$). Lecture Notes in Mathematics, No. 53 Springer, New York, 1968.

[2] S. K. Donaldson, *Irrationality and the h-cobordism conjecture*, J. Diff. Geom. **26** (1987), 141–168.

[3] R. Friedman and J. W. Morgan, *On the diffeomorphism types of certain algebraic surfaces. I*, J. Diff. Geom. **27** (1988), 297–369.

[4] R. Friedman and J. W. Morgan, *On the diffeomorphism types of certain algebraic surfaces. II*, J. Diff. Geom. **27** (1988), 371–398.

[5] R. E. Gompf, *Three exotic $\mathbf{R}^4$'s and other anomalies*, J. Diff. Geom. **18** (1983), 317–328.

[6] M. W. Hirsch and B. Mazur, Smoothings of piecewise linear manifolds. Annals of Mathematics Studies, No. 80. Princeton University Press, Princeton, N. J., 1974. ISBN: 0-691-08145-X.

[7] M. A. Kervaire and J. W. Milnor, *Groups of homotopy spheres.I*, Ann. of Math. (2) **77** (1963), 504–537.

[8] R. Lashof, and M. Rothenberg, *Microbundles and smoothing*, Topology **3** (1965), 357–388.

[9] J. W. Milnor, *On manifolds homeomorphic to the* 7*-sphere*, Ann. of Math. (2) **64** (1956), 399–405.

[10] C. H. Taubes, *Gauge theory on asymptotically periodic* 4*-manifolds*, J. Diff. Geom. **25** (1987), 363–430.

III.2 : Constructions on smooth manifolds

(Conlon, §§ 1.7, 3.7)

In the theory of topological spaces it is often useful and important to construct new examples of topological spaces out of old ones. Not all of the constructions from point set topology have analogs for smooth manifolds, but there are several important cases where such analogs exist.

Although several basic constructions on topological spaces do have smooth analogs, in the latter case the formal definitions are considerably less straightforward. As indicated in Section I.4, it will be much easier to work with the properties of these objects involving smooth mappings rather than the details of the constructions. From this perspective, the constructions are uniquely characterized up to diffeomorphism by certain mapping properties. At some point one might need to do some messy work in order to show that constructions with the desired properties exist, but once existence has been established the explicit methods for showing existence are often of secondary importance because they are not needed for further work. In this respect the existence proofs function like a workman's ladders: They provide the means for someone to complete a job, it is necessary for a user to understand something about how and why they are structured the way they are, but once the job is done it is convenient to store them in a safe place that is out of the way.

Here is a summary of the constructions presented in this section:

(1) Restrictions of smooth structures to open subsets of the underlying manifolds.

(2) Direct (or cartesian) products of two smooth manifolds $(M, \mathcal{A})$ and $(N, \mathcal{B})$ such that the underlying topological manifold is $M \times N$. — As indicated in Section I.4, one can extend this to give a recursive construction for direct products of finitely many smooth manifolds.

(3) Construction of a canonical smooth structure associated to a smooth manifold $(M, \mathcal{A})$ and a Hausdorff covering space E of M.

(4) For a special class of equivalence relations on the underlying spaces of smooth manifolds, constructions of naturally associated smooth structures on the quotient spaces (of course, one basic necessary condition on the equivalence relations is that their quotient spaces should be topological manifolds, but there are also further conditions).

We shall consider these in the order listed.

III.2.1 : *Restrictions to open subsets*

We know that an open subset of a topological manifold is always a topological manifold. The corresponding result for smooth manifolds is also true, but the verification is not quite as simple (even though it is basically elementary).

RESTRICTIONS TO OPEN SUBSETS. *Let $(M, \mathcal{A})$ be a smooth manifold, and let Ω be an open subset of M. Then there is a smooth atlas $\mathcal{A}|\Omega$ on Ω such that*

(i) the inclusion map $j : (\Omega, \mathcal{A}|\Omega) \to (M, \mathcal{A})$ is smooth,

(ii) if $(L, \mathcal{E})$ is a smooth manifold and $g : L \to \Omega$ is continuous, then g is smooth if and only if $j \circ g$ is smooth.

Proof. If one simply takes $\mathcal{A}|\Omega$ to be the set of all charts (U, h) in $\mathcal{A}$ such that $h(U) \subset \Omega$, then it follows immediately that $(\Omega, \mathcal{A}|\Omega)$ is smooth. To prove that the inclusion is smooth, let $x \in \Omega$ and let (U, h) be a smooth chart at x in $\mathcal{A}|\Omega)$ and note that (U, h) is also a chart in $\mathcal{A}$. It then follows that the local map "$h^{-1} \circ j \circ h$" is just the identity map, which we know is smooth. By the weak criterion for smoothness, this implies that j is smooth.

To prove the second part, choose a smooth atlas $\mathcal{E}_0$ for L such that each chart (W, ℓ) in $\mathcal{E}_0$ gets mapped into the image of a chart in $\mathcal{A}$, the whose of which is contained in Ω; this is possible by continuity and the fact that the image of g is contained in Ω. If (U, h) is a chart of the prescribed type in $\mathcal{A}$, then it follows that "$h^{-1} \circ (j \circ g) \circ \ell$" is smooth because $j \circ g$ is smooth. However, by construction the chart (U, h) also belongs to $(\Omega, \mathcal{A}|\Omega)$ and the local map "$h^{-1} \circ (j \circ g) \circ \ell$" is identical to "$h^{-1} \circ g \circ \ell$" so that the weak smoothness criterion implies g must be smooth.∎

COROLLARY. *In the notation of the preceding result, if $f : (M, \mathcal{A}) \to (N, \mathcal{B})$ is smooth, then the restriction $f|\Omega$ defines a smooth map from Ω to N (with respect to the given smooth structures.*

This follows because the inclusion $\Omega \subset M$ is smooth and the composite of smooth mappings is smooth.∎

The next result shows that the inclusion $j : \Omega \to M$ satisfies an analog of the embedding property in Section I.4.

PROPOSITION. *Suppose that $(Q, \mathcal{P})$ is a smooth manifold and $g : Q \to M$ is a smooth map (with respect to the given smooth structures) such that $g(Q) \subset \Omega$, where Ω is open in M. Then there is a unique smooth map $g' : Q \to \Omega$ such that $g = g' \circ h$.*

Proof. The existence of a unique continuous map $g' : Q \to \Omega$ as above is verified in Section I.4. We need to show that this map is smooth. But this follows from (ii) in the result on restrictions to open subsets.∎

Another important property of open subset restrictions is the following smooth analog of a basic criterion for recognizing continuous functions:

THEOREM. *Suppose that $(M, \mathcal{A})$ is a smooth manifold and $\{\, \Omega_\lambda \,\}$ is an open covering of M with indexing set Λ. Let $(N, \mathcal{B})$ be a smooth manifold, and let $f : M \to N$ be continuous. Then f is smooth if and only if for each λ the restriction $f|\Omega_\lambda$ is smooth.*

Proof. We have already demonstrated the $(\Longrightarrow)$ implication, so we turn now to the $(\Longleftarrow)$ direction. Consider the subatlas $\mathcal{A}_0 \subset \mathcal{A}$ consisting of all smooth charts (U, h) such that $h(U) \subset \Omega_\lambda$ for some λ and there is some chart (V, k) in $\mathcal{B}$ such that $h(U) \subset k(V)$. By the preceding results it suffices to check that each map "$k^{-1} \circ f \circ h$" is smooth. But this follows from the hypothesis that the restriction of f to each open subset Ω_λ is smooth.■

We conclude this subsection with a basic observation that will be useful later.

CHART DIFFEOMORPHISM LEMMA. *Let $(M, \mathcal{A})$ be a smooth manifold, let (U, h) be a smooth chart in $\mathcal{A}$, let $j : h(U) \to M$ be inclusion, and assume that $h(U)$ has the smooth structure given by restriction of $\mathcal{A}$. If $h' : U \to h(U)$ is the unique map such that $h = j \circ h'$, then h' is a diffeomorphism.*

Proof. The statement of the lemma looks almost tautological, but it is still necessary to verify some things. Results from Section I.4 imply that h' is a homeomorphism; let k be its inverse. We need to check that the weak smoothness criterion holds for h' and k.

Both U and $h(U)$ have smooth atlases consisting of exactly one chart; the chart for U is the just (U, id_U) and the chart for $h(U)$ is just (U, h). By the weak smoothness criterion it suffices to check that both "$h^{-1} \circ h' \circ \mathrm{id}_U$" and "$(\mathrm{id}_U)^{-1} \circ k \circ h$" define smooth maps from U to itself. Direct evaluation of these maps at an arbitrary element of U shows that each is equal to the identity on U, and since the latter is smooth it follows that h' and its inverse k are also smooth.■

COROLLARY. *Let $(M, \mathcal{A})$ and $(N, \mathcal{B})$ be smooth manifolds and let $f : M \to N$ be continuous. Then f is smooth if for some subatlas $\mathcal{A}_0 \subset \mathcal{A}$ the composites $f \circ h_\alpha$ are smooth for all $(U_\alpha, h_\alpha) \in \mathcal{A}_0$.*

Proof. Let (U_α, h_α) be a smooth chart in $\mathcal{A}_0$. As in the preceding lemma, let $k_\alpha : h_\alpha(U_\alpha) \to U_\alpha$ be the inverse to the diffeomorphism in the other direction associated to h_α. Then we have $f|h_\alpha(U_\alpha) = f \circ h_\alpha \circ k_\alpha$. Since $f \circ h_\alpha$ and k_α are both smooth, it follows that their composite, which is the restriction, must also be smooth. Since we began with an arbitrary chart and the images of the charts in $\mathcal{A}_0$ form an open covering of M, by a previous theorem f must also be smooth.■

Here is another result along the same lines that is sometimes useful:

COROLLARY. *Let $(M, \mathcal{A})$ and $(N, \mathcal{B})$ be smooth manifolds, let $\mathcal{A}_0$ be a subatlas of $\mathcal{A}$, and assume that we are given smooth maps $f_\alpha : U_alpha \to N$ satisfying the consistency condition $f_\beta \circ$ "$h_\beta^{-1} h_\alpha$" $= f_\alpha$ on $h_\alpha^{-1}(h_\beta(U_\beta))$ for all α and β. Then there is a unique smooth function $f : M \to N$ such that $f \circ h_\alpha = f_\alpha$ for all α.*

Proof. (*Sketch*) For each α, once again let $k_\alpha : h_\alpha(U_\alpha) \to U_\alpha$ be the inverse to the diffeomorphism from U_α TO $h_\alpha(U_\alpha)$ determined by h_α, and let g_α be the composite $f_\alpha \circ k_\alpha$; then g_α defines a smooth map from $h_\alpha(U_\alpha)$ to N.

We first claim that the maps g_α fit together to yield a continuous map f from M to N. This follows because the consistency condition implies g_α and g_β have the same restriction to $h_\alpha(U_\alpha) \cap h_\beta(U_\beta)$ for all α and β. This map is smooth because its restriction to each set $h_\alpha(U_\alpha)$ is smooth and the latter form an open covering of M.■

III.2.2 : *Products of two smooth manifolds*

We have already seen that product constructions yield an important means for constructing new topological manifolds out of old ones. The corresponding statement is also true for smooth manifolds.

DIRECT PRODUCTS OF SMOOTH STRUCTURES. *Let $(M, \mathcal{A})$ and $(N, \mathcal{B})$ be smooth manifolds. Then there is a smooth atlas $\mathcal{P}$ on $M \times N$ such that the coordinate projection maps*

$\pi_M : M \times N \to M$ and $\pi_N : M \times N \to N$ are smooth, and more generally a continuous map $f : P \to M \times N$ is smooth if and only if the coordinate functions $\pi_M \circ f$ and $\pi_N \circ f$ are smooth.

The idea of the construction is simple: A smooth atlas for the product $M \times N$ can be constructed using charts of the form $(U_\alpha \times V_\beta, h_\alpha \times k_\beta)$ where (U_α, h_α) ranges over all charts in the maximal atlas for M and (V_β, k_β) ranges over all charts in the maximal atlas for N. ∎

The previous result yields the following **Universal Mapping Property** in analogy with the topological case (see Section I.4).

COROLLARY. *Let $(M, \mathcal{A})$ and $(N, \mathcal{B})$ be smooth manifolds, and let $(M \times N, \mathcal{R})$ be the direct product constructed above. Suppose that $(Q, \mathcal{E})$ is a smooth manifold and that $f : Q \to M$ and $g : Q \to N$ are smooth. Then there is a unique smooth map $\Phi : Q \to M \times N$ such that $\pi_M \circ \Phi = f$ and $\pi_N \circ \Phi = g$.*

Proof. Since all maps in sight are continuous, the existence of a unique continuous map Φ follows as in the topological case (see Section I.4 for the details). Thus it is only necessary to prove that Φ is smooth.

Let $\mathcal{A} \prod \mathcal{B}$ be the product atlas for $M \times N$ and let $\mathcal{E}_0$ be a subatlas of $\mathcal{E}$ consisting of smooth charts (U_γ, h_γ) such that each image $f \circ h_\gamma(U_\gamma)$ lies in a chart of $\mathcal{A}$ and each image $g \circ h_\gamma(U_\gamma)$ lies in a chart of $\mathcal{B}$. For a fixed value of γ choose specific charts (V_α, k_α) and (W_β, ℓ_β) such that $f \circ h_\gamma(U_\gamma) \subset k_\alpha(V_\alpha)$ and $g \circ h_\gamma(U_\gamma) \subset \ell_\beta(W_\beta)$. Then the coordinate functions for the local mapping "$(k_\alpha \times \ell_\beta)^{-1} \circ \Phi \circ h_\gamma$" are "$k_\alpha^{-1} \circ f \circ h_\gamma$" and "$\ell_\beta^{-1} \circ g \circ h_\gamma$" respectively. Since both of the latter are smooth, it follows that "$(k_\alpha \times \ell_\beta)^{-1} \circ \Phi \circ h_\gamma$" is smooth; since γ is an arbitrary chart in $\mathcal{E}_0$, it follows that Φ must be smooth.∎

For our purposes it is important to have the following converse to the Universal Mapping Property.

CHARACTERIZATION OF DIRECT PRODUCTS. *Let $(M, \mathcal{A})$ and $(N, \mathcal{B})$ be smooth manifolds, let $(P, \mathcal{R})$ be another smooth manifold, and let $p_M : (P, \mathcal{R}) \to (M, \mathcal{A})$ and $p_N : (P, \mathcal{R}) \to (N, \mathcal{B})$ be smooth mappings satisfying the following Universal Mapping Property:* Given a smooth manifold Q and smooth functions $f : Q \to M$ and $g : Q \to N$, THEN there is a unique smooth map $\Phi : Q \to P$ such that $p_M \circ \Phi = f$ and $p_N \circ \Phi = g$. *Then there is a unique diffeomorphism $H : P \to M \times N$ such that $\pi_M \circ H = p_M$ and $\pi_N \circ H = p_N$.*

Proof. The argument is entirely formal and parallel to the corresponding result for topological spaces. By the Universal Mapping Property there is a unique smooth function $H : P \to M \times N$ such that $\pi_M \circ H = p_M$ and $\pi_N \circ H = p_N$. We need to show that H is a diffeomorphism.

As in the topological case, we shall find an inverse to H using the universal mapping property for P with respect to p_M and p_N. In particular, the latter implies the existence of a unique map $K : M \times N \to P$ such that $p_M \circ K = \pi_M$ and $p_N \circ K = \pi_N$.

Taking composites, we find that

$$\pi_M \;=\; p_M \circ K \;=\; \pi_M \circ H \circ K \;=\; \pi_M \circ \mathrm{id}_{M \times N}$$

$$\pi_N \;=\; p_N \circ K \;=\; \pi_N \circ H \circ K \;=\; \pi_N \circ \mathrm{id}_{M \times N}$$

and by the uniqueness part of the Universal Mapping Property it follows that $H \circ K = \mathrm{id}_{M \times N}$. Likewise, we have that

$$p_M \;=\; \pi_M \circ H \;=\; \pi_M \circ K \circ H \;=\; p_M \circ \mathrm{id}_P$$

$$p_N \quad = \quad \pi_N \circ H \quad = \quad p_N \circ K \circ H \quad = \quad p_N \circ \mathrm{id}_P$$

and once again by uniqueness part of the Universal Mapping Property it follows that follows that $K \circ H = \mathrm{id}_P$. Therefore H and K must be inverse to each other and accordingly H is a diffeomorphism.■

Examples. If one applies this to known examples such as spheres, one obtains standard smooth structures on a products of spheres $S^p \times S^q$. Likewise, if M is a smooth manifold then $M \times \mathbf{R}^k$ also is a smooth manifold and has a distinguished smooth structure given by the original smooth structure on M and the usual smooth structure on $\mathbf{R}$.

REMARK. (‡) A fairly deep result in manifold theory yields a partial converse to the second class of examples: If M is a topological manifold of dimension $\neq 4$ and $M \times \mathbf{R}^k$ has a smooth structure for some $k > 0$, the M also has a smooth structure. The previously cited references by Kirby-Siebenmann and Freedman-Quinn provide further information about this. There are examples of compact 4-dimensional manifolds such that $M \times \mathbf{R}$ has a smooth structure but M does not.

III.2.3 : *Products of three or more smooth manifolds.*

Perhaps the simplest way to define products of three or more smooth manifolds is by recursion. For example, the product of three smooth manifolds $(M_1, \mathcal{A}_1)$, $(M_2, \mathcal{A}_2)$ and $(M_3, \mathcal{A}_3)$ could be set equal to

$$\Big((M_1, \mathcal{A}_1) \times (M_2, \mathcal{A}_2) \Big) \ \times \ (M_3, \mathcal{A}_3)$$

and similarly the product of n items would be the twofold product of (1) the previously defined product of the first $n-1$ manifold with (2) the final manifold on the list.

One important conceptual objection to this recursive definition is that there are other natural candidates that look equally valid. For example, one might try to define a threefold product instead by the expression

$$(M_1, \mathcal{A}_1) \ \times \ \Big((M_2, \mathcal{A}_2) \times (M_3, \mathcal{A}_3) \Big)$$

and similarly for products of more manifolds. One might view this as a left-handed definition as compared to the right-handed one given earlier. There is no obvious *a priori* advantage of one definition over the other. With products of four or more manifolds, the possibilities become even more extensive. For example, given A, B, C, D it is also reasonable to propose $(A \times B) \ \times \ (C \times D)$ as a definition for a fourfold product.

We shall avoid such questions by taking a neutral approach: Products of manifolds will be characterized axiomatically using a universal mapping property generalizing the one we gave for twofold products, and all the definitions proposed above (as well as many others) will be shown to satisfy the given axioms for a product.

One important advantage of this approach is that there will be no need to work directly with smooth atlases at any point in the arguments.

The following definition contains the desired axioms for a product of smooth manifolds.

Definition. Let $\{X_\alpha\}$ be an indexed family of smooth manifolds with finite indexing set A. A **direct product** of the indexed family is pair $(P, \{p_\alpha\})$, where P is a smooth manifold and each

p_α is a smooth function from P to X_α for each α, such that the following **Universal Mapping Property** holds:

> *Given an arbitrary smooth manifold Y and smooth functions $f : Y \to X_\alpha$ for each α, there is a unique smooth function $f : Y \to X$ such that $p_\alpha \circ f = f_\alpha$ for each $\alpha \in A$.*

As in Section I.4 and the preceding discussion, we would like direct products of smooth manifolds to be essentially unique. The following results describes the strong uniqueness property that we shall need.

UNIQUENESS THEOREM. *Let $\{X_\alpha\}$ be an indexed family of smooth manifolds with indexing set A, and suppose that $(P, \{p_\alpha\})$ and $(Q, \{q_\alpha\})$ are direct products of the indexed family $\{X_\alpha\}$. Then there is a unique diffeomorphism $h : Q \to P$ such that $p_\alpha \circ h = q_\alpha$ for all α.*

Proof. First of all, we claim that a smooth function $\varphi : P \to P$ is the identity if and only if $p_\alpha \circ \varphi = p_\alpha$ for all α, and likewise $\psi : Q \to Q$ is the the identity if and only if $q_\alpha \circ \psi = q_\alpha$ for all α. These are immediate consequences of the Universal Mapping Property.

Since $(P, \{p_\alpha\})$ is a direct product, the Universal Mapping Property implies there is a unique smooth function $h : Q \to P$ such that $p_\alpha \circ h = q_\alpha$ for all α, and likewise since $(Q, \{q_\alpha\})$ is a direct product, the Universal Mapping Property implies there is a unique smooth function $f : P \to Q$ such that $q_\alpha \circ h = p_\alpha$ for all α. We claim that h and f are inverse to each other; this is equivalent to the identities $h \circ f = \mathrm{id}_Q$ and $f \circ h = \mathrm{id}_P$.

To verify these identities, first note that for all α we have

$$p_\alpha \quad = \quad q_\alpha \circ h \quad = \quad p_\alpha \circ f \circ h$$

for all α and similarly

$$q_\alpha \quad = \quad p_\alpha \circ f \quad = \quad q_\alpha \circ h \circ f$$

for all α. By the observations in the first paragraph of the proof, it follows that $f \circ h = \mathrm{id}_P$ and $h \circ f = \mathrm{id}_Q$.■

We now need to show that the axioms for direct products hold for the constructions described above. This will be a consequence of the next result, which can be summarized informally as saying that "a product of products is a product."

THEOREM. *Let $\{X_\alpha\}$ and $\{Y_\beta\}$ be indexed families of smooth manifolds with indexing sets A and B, and suppose that $(P, \{p_\alpha\})$ and $(Q, \{q_\alpha\})$ are direct products of the indexed families $\{X_\alpha\}$ and $\{Y_\beta\}$ respectively. Let $P \times Q$ be the usual product with projection maps π_P and π_Q. Then $P \times Q$ and the associated maps $p_\alpha \circ \pi_P$ and $q_\beta \circ \pi_Q$ form a direct product of the indexed family formed by combining $\{X_\alpha\}$ and $\{Y_\beta\}$.*

Proof. It is only necessary to verify the Universal Mapping Property.

Suppose that we are given functions $f_\alpha : W \to X_\alpha$ and $g_\beta : W \to Y_\beta$. Then there are unique functions $F : W \to P$ and $G : W \to Q$ such that $p_\alpha \circ F = f_\alpha$ and $q_\beta \circ G = g_\beta$. By the Universal Mapping Property for twofold products, it follows that there is a unique function $\Phi : W \to P \times Q$ such that $\pi_P \circ \Phi = F$ and $\pi_Q \circ \Phi = G$. It follows immediately that

$$p_\alpha \circ \pi_P \circ \Phi \quad = \quad p_\alpha \circ F \quad = \quad f_\alpha \qquad \text{and} \qquad q_\beta \circ \pi_Q \circ \Phi \quad = \quad q_\alpha \circ G \quad = \quad g_\beta$$

for all α and β.

To conclude the proof we need to show that any map Ψ satisfying $p_\alpha \circ \pi_P \Psi = f_\alpha$ for all α and $q_\beta \circ \pi_Q \Psi = g_\beta$ for all β must be equal to Φ. Let F' and G' denote the composites $\pi_P \Psi$ and $\pi_Q \Psi$ respectively. Direct computation and the conditions on Ψ then show that $p_\alpha F' = f_\alpha$ for all α and $q_\beta G' = g_\beta$ for all β. Since P and Q (with the corresponding projections) are assumed to be products, the uniqueness part of the Universal Mapping Property then implies $F' = F$ and $G' = G$. Knowing this, we can apply the same property for the twofold cartesian product of P and Q to conclude that Ψ must be equal to Φ.■

REMARK. Product constructions are possible for smooth manifolds only if the number of factors is finite. One reason for this is that a cartesian product of infinitely many topological manifolds is almost never a topological manifold (the only exception being the trivial case in which all but finitely many factors are one point spaces). If there are infinitely many factors with positive dimensions, this can be shown by a variant of the argument proving that the Hilbert Cube **HQ** is not a topological manifold (see Example 10 in Section I.1). On the other hand, if there are infinitely many zero-dimensional factors, this follows from one of the exercises for Section I.1.

III.2.4 : *Covering manifolds* ($\star$)

For covering space projections, the basic principle is that if one is given a covering space projection and a smooth structure on the base, then the smooth structure can be lifted to the domain of the covering space map.

COVERINGS OF SMOOTH STRUCTURES. *Let $p : E \to M$ be a covering space projection where M is a topological manifold and E is Hausdorff, and let $\mathcal{A}$ be a maximal smooth atlas for M. Then there is a smooth atlas $\mathcal{E}$ on E such that if (U, h) is a coordinate chart for M such that $h(U)$ is evenly covered, then the restriction of p to each sheet of $p^{-1}(h(U))$ is a diffeomorphism.*

Proof. Here again the basic idea is fairly straightforward. Consider the subatlas $\mathcal{A}_0$ of $\mathcal{A}$ consisting of all charts (U, h) so that $h(U)$ is evenly covered. For each sheet W_γ of $p^{-1}(h(U))$ let s_γ be an inverse to $p|W_\gamma$. Then $\mathcal{E}_0$ is defined to be the set of all pairs $(U, s_\gamma \circ h)$ where (U, h) is a chart in $\mathcal{A}_0$. We claim that $\mathcal{E}_0$ is a smooth atlas.

Let $(U, s_\gamma \circ h)$ and $(V, s_\beta \circ k)$ be two charts in $\mathcal{E}_0$ where (U, h) and (V, k) are smooth charts in $\mathcal{A}_0$. Since s_β and s_γ are local inverses to p, the transition map "$(s_\beta \circ k)^{-1} \circ (s_\gamma \circ h)$" = "$k^{-1} \circ s_\beta^{-1} \circ s_\gamma \circ h$" is also expressible in the form "$k^{-1} \circ p \circ s_\gamma \circ h$" which further reduces to "$k^{-1} \circ h$". The latter is smooth by our assumption that both charts lie in the same smooth atlas, and therefore we have shown that the transition maps for $\mathcal{E}_0$ are also smooth.

By the preceding paragraph we know that $\mathcal{E}_0$ is a smooth atlas, and it follows that it is contained in maximal atlas. Therefore we may define $\mathcal{E}$ to be the unique maximal atlas containing $\mathcal{E}_0$.

Finally, we need to prove that if (U, h) is a coordinate chart for M such that $h(U)$ is evenly covered, then the restriction of p to each sheet of $p^{-1}(h(U))$ determines a diffeomorphism onto $h(U)$. Let s_γ be as above, and let p_0 be the homeomorphism from $s_\gamma \circ h(U)$ to $h(U)$ defined by p. By the lemma at the end of the last subsection we know that the maps $(s_\gamma \circ h)'$ and h' define diffeomorphisms from U to $s_\gamma \circ h(U)$ and $h(U)$ respectively, and we also know that $p_0 \circ (s_\gamma \circ h)' = h'$. Therefore we have

$$p_0 \quad = \quad h' \circ \big((s_\gamma \circ h)' \big)^{-1}$$

and since each factor on the right hand side is a diffeomorphism it follows that p_0 is also a diffeomorphism.■

Since covering space projections are open and open mappings are quotient maps, it follows that every covering space projection is a quotient map. The next result states that smooth covering space projections satisfy a smooth analog of a basic property for topological quotient maps.

PROPOSITION. *In the setting above, suppose that $(N, \mathcal{B})$ is a smooth manifold and that $f : N \to E$ is continuous. Then f is smooth if and only if $p \circ f$ is smooth.*

Proof. The ($\Longrightarrow$) implication follows because composites of smooth maps are smooth. To prove the opposite implication, let $\mathcal{A}_0$ and $\mathcal{E}_0$ be the atlases described above, and let $\mathcal{B}$ be a smooth atlas for N such that f maps the image of each chart in $\mathcal{B}$ to the image of a chart in $\mathcal{E}_0$. Suppose now that we have charts (U, h) in $\mathcal{A}_0$, $(U, s_\gamma \circ h)$ in $\mathcal{E}_0$ and $(W \ell)$ in $\mathcal{B}$ such that p maps $s_\gamma \circ (U)$ diffeomorphically onto $h(U)$ and $f \circ \ell(W) \subset s_\gamma \circ (U)$. We need to check that the local map "$(s_\gamma \circ h)^{-1} \circ f \circ \ell$" is smooth. As in the proof of the previous result, the local map may also be expressed in the form "$h^{-1} \circ (p \circ f) \circ \ell$" and the latter is smooth by the smoothness of $p \circ f$. Therefore the originally local map is indeed smooth, and it follows that f itself is also smooth.■

The preceding result has the following elementary but important consequence:

COROLLARY. *If we are given a smooth map into the base of a covering space projection and the map lifts continuously, then the lifting is always smooth.*■

The topological conditions for the existence of liftings are established in Lemma 79.1 on pages 478–480 of [MUNKRES1].

III.2.5 : *Quotients* ($\star$)

If $M$is a topological manifold and $\mathcal{R}$ is an equivalence relation on M, then the quotient $M/\mathcal{R}$ is not necessarily a topological manifold; one extremely simple example was given in Section I.1, and it is possible to construct an extremely broad range of examples with some highly unusual properties (compare the paper by Bing cited above and the related references). As noted in Section V.1 of the ONLINE 205A NOTES, it is even possible to construct quotients of manifolds that are finite but not Hausdorff. Therefore it should be clear that only some quotient spaces of smooth manifolds have a chance of supporting natural smooth structures, and it should be equally clear that any search for such smooth quotients should begin with families of quotients that are already known to be manifolds. For our purposes, the quotients by free actions of finite groups from Section I.1 provide an excellent starting point:

QUOTIENTS BY FREE SMOOTH ACTIONS OF FINITE GROUPS. *Let $(M, \mathcal{A})$ be a smooth manifold, let G be a finite group, and let G act freely on G by diffeomorphisms: More precisely there is a family of diffeomorphisms Φ_g indexed by G so that $\Phi_{gh} = \Phi_g \circ \Phi_h$, Φ_1 is the identity, and if $g \neq 1$ then $\Phi_g(x) \neq x$ for all $x \in M$. Let M/G be the quotient space of M by this action; then topological considerations imply that the quotient projection map $p : M \to M/G$ is a covering space projection and M/G is a topological manifold of the same dimension as M. Then there is a smooth atlas $\mathcal{A}'$ on M/G such that*

(i) $p : (M, \mathcal{A}) \to (M/G, \mathcal{A}')$ is smooth,

(ii) if $(N, \mathcal{B})$ is a smooth manifold and $f : M/G \to N$ is continuous, then f is smooth if and only if $f \circ p$ is smooth.

The following general fact will be needed to prove the result on quotient structures:

LEMMA. *Let $(M, \mathcal{A})$ be a smooth manifold, let (U, h) be a smooth chart in the maximal atlas $\mathcal{A}$, and let $f : M \to M$ be a diffeomorphism. Then the pair $(U, f \circ h)$ also belongs to $\mathcal{A}$.*

Proof of Lemma. By the hypotheses we know that $(U, f \circ h)$ is a topological chart for M. To prove that this topological chart belongs to the maximal atlas $\mathcal{A}$, it is necessary to show that for an arbitrary smooth chart (V, k) the transition map "$k^{-1} \circ (f \circ h)$" is a diffeomorphism; this requires a closer look at the given transition map.

Since f is a homeomorphism the map "$k^{-1} \circ (f \circ h)$" sends $U_0 = (f \circ h)^{-1}\,(\,k(V)\,) \subset U$ homeomorphically onto $V_0 = k^{-1}\,(\,f \circ h(U)\,) \subset V$. Let $h_0 = h|U_0$ and $k_0 = k|V_0$, so that f maps U_0 homeomorphically to V_0. Since f is a diffeomorphism, the definition of smoothness implies that both "$k_0^{-1} \circ f \circ h_0$)" and its inverse "$h_0^{-1} \circ f^{-1} \circ k_0$)" are both smooth. Now the local maps "$k^{-1} \circ f \circ h$" and "$k_0^{-1} \circ f \circ h_0$)" are the same by construction, and therefore we conclude that "$k^{-1} \circ f \circ h$" is a diffeomorphism as required.∎

Proof of the main result. In this case one chooses the atlas $\mathcal{A}'$ to consist of all charts $(U, \overline{h})$ such that $\overline{h}(U)$ is evenly covered and there is a smooth chart (U, h) in $\mathcal{A}$ such that $\overline{h} = p \circ h$. It follows immediately that the images of the charts in $\mathcal{A}'$ form an open covering of M/G.

We claim that $\mathcal{A}'$ is in fact a smooth atlas. Given charts $(U, \overline{h})$ and $(V, \overline{k})$ in $\mathcal{A}'$ with liftings (U, h) and (V, k), we need to show that the transition map "$\overline{k}^{-1} \circ \overline{h}$" is a diffeomorphism. Let $W \subset M/G$ be the open subset $\overline{h}(U) \cap \overline{k}(V) \subset M/G$, and let W^* denote the union of the pairwise disjoint open subsets $W_g^* = h(U) \cap \Phi_g \circ h(V)$ in M, where g runs through all the elements of G. It then follows that $p(W^*) = W$ and the open subsets $p(W*_g)$ decompose W into a union of pairwise disjoint open subsets. This implies in turn that the transition map "$\overline{k}^{-1} \circ \overline{h}$" admits a similar decomposition, and a direct examination shows that the transition maps on the corresponding pieces are given by "$(\Phi_g \circ k)^{-1} \circ h$" where once again g runs through all the elements of G. By the previous lemma we know that each pair $(V, \varphi_g \circ k)$ is a smooth chart in $\mathcal{A}$, and therefore the transition maps in the previous sentence must be diffeomorphisms. It follows that the union of these transition maps, which is just the map "$\overline{k}^{-1} \circ \overline{h}$" we started with, must also be a diffeomorphism. This completes the proof that $\mathcal{A}'$ is a smooth atlas.

The smoothness of the projection map follows because p maps $h(U)$ diffeomorphically to $\overline{h}(U)$ (this argument is similar to a step in the construction of smooth structures on covering spaces). Similarly the statement in (ii) follows from the analogous discussion for covering spaces.∎

Before discussing some important examples of quotient spaces, we need some background information about smooth maps into the level set examples considered at the beginning of Section III.1.

PROPOSITION. *Suppose that M is a smooth manifold, Ω is an open subset of $\mathbf{R}^n$, and $L \subset \Omega$ is the level set associated to some smooth map $g : \Omega \to \mathbf{R}^m$ where $m < n$ (hence $L = g^{-1}(\{0\})$ and $g(z) = 0 \Longrightarrow Dg(z)$ has rank m). If $f : M \to \Omega$ is a smooth map such that $f(M) \subset L$, then the associated map $f_0 : M \to L$, with $f_0(x) = f(x)$, is also smooth.*

Proof. By the Submersion Straightening Proposition, an atlas for L is given by first taking pairs $(V \times W, \ell)$ such that V is open in $\mathbf{R}^{n-m}$, W is open in $\mathbf{R}^m$, ℓ is a 1–1 onto open continuous map that is a diffeomorphism onto its image, and $g \circ \ell(v, w) = w$ on $V \times W$, and L is contained in the union of the images $\ell(V \times W)$ over all $(V \times W, \ell)$. The atlas for L then consists of associated charts (V, k) such that $k(v) = \ell(v, 0)$.

Let (U, h) be a coordinate chart in an atlas for M such that f maps $h(U)$ into some set of the form $\ell(V \times W)$. Since f is smooth we know that "$\ell^{-1} \circ f \circ h$" is smooth. The assumption $f(M) \subset L$ implies that the local map has the form $(\varphi(u), 0)$ for some function φ that must also be smooth by the smoothness of the local map. However, direct inspection also shows that the local map "$k^{-1} \circ f' \circ h$" is also equal to φ, and since φ is smooth it follows that the local map "$k^{-1} \circ f_0 \circ h$" must also be smooth. By the weak smoothness criterion it follows that f_0 must be smooth.∎

COROLLARY. *If $h : \mathbf{R}^n \to \mathbf{R}^n$ is a smooth map such that $h(S^{n-1}) \subset S^{n-1}$, then the associated map $h_0 : S^{n-1} \to S^{n-1}$, defined by $h_0(x) = h(x)$, is also smooth. If in addition h is a diffeomorphism and maps S^{n-1} onto itself, then h_0 is also a diffeomorphism.*

Proof. The first assertion is an immediate consequence of the proposition. To prove the second assertion, let k be the inverse of h and note that the additional hypothesis implies that k also maps the unit sphere to itself. Therefore the associated map k_0 is also smooth. Since h and k are inverses to each other, the same must be true of h_0 and k_0 ($h \circ k = \mathrm{id} \Longrightarrow h_0 \circ k_0 = \mathrm{id}$, and likewise for the reverse composites).∎

Important Special Cases of Free Quotients. The most basic example of this situation is given by the *real projective plane* $\mathbf{RP^2}$, which can be viewed as the quotient of S^2 by the action of the group $G = \{\pm 1\}$ by scalar multiplication. This defines a free action because a nonzero vector in a real vector space is never equal to its negative. To verify this is a smooth action, it suffices to note that multiplication by -1 defines an invertible linear transformation (hence diffeomorphism) from $\mathbf{R}^3$ to itself which sends S^2 onto itself; one can then apply the preceding corollary to prove smoothness. Similarly, for each positive integer n we can define *real projective n-space* to be the quotient of S^n by the action of $\{\pm 1\}$ via scalar multiplication.

In odd dimensions one has other important examples along the same line known as **lens spaces**. Given a finite cyclic group $\mathbf{Z}_k$ of order k, and a positive integer n, let $(m_1, \cdots, m_n)$ be an ordered n-tuple of positive integers less than k such that each m_j is prime to k. Then we claim that one can define a free action of $\mathbf{Z}_k$ on $S^{2n-1} \subset \mathbf{C}^n \cong \mathbf{R}^{2n}$ by the formula

$$g^j(z_1, \cdots, z_n \quad = \quad (\alpha^{m_1} z_1, \cdots, \alpha^{m_n} z_n)$$

where g denotes a standard generator of $\mathbf{Z}_k$ and $\alpha = \exp(2\pi i/k)$. To prove this claim, first observe that the maps g^j are invertible (complex) linear transformations from $\mathbf{C}^n$ to itself and they map the unit sphere onto itself; therefore by the preceding corollary the maps g^j define diffeomorphisms from the unit sphere onto itself. Verification that $g^j\mathbf{z} \neq \mathbf{z}$ if $j \not\equiv 0(k)$ and $\mathbf{z} \neq \mathbf{0}$ is elementary and left to the reader as an exercise. The associated quotient space of the sphere by the given free smooth action of $\mathbf{Z}_k$ is called the **lens space** of type $(k; m_1, \cdots, m_n)$ and is often denoted by notation such as $L^{2n-1}(k; m_1, \cdots, m_n)$.

The proof of the result on quotients actually yields a stronger conclusion that is extremely useful in many contexts.

GENERALIZATION OF THE QUOTIENT EXAMPLES. *Suppose we have a regular covering space projection $p : E \to B$ where both spaces are topological n-manifolds; the regularity condition implies that there is a group of covering transformations Γ acting on E such that M is homeomorphic to the quotient space E/Γ. Assume also that we have a smooth structure $\mathcal{B}$ on E such that the covering transformations are all diffeomorphisms from E to itself. If we let $\mathcal{A}$ be the set of all charts (U, k) on M such that $k(U)$ is evenly covered and there is a smooth chart (U, h) in $\mathcal{B}$ such that $k = p \circ h$, then $\mathcal{A}$ is a smooth atlas for M. Properties (i) and (ii) above remain true in this* *setting. Furthermore, if we construct the maximal smooth atlas $\mathcal{E}$ on E associated to $\mathcal{A}$ as above, then $\mathcal{E}$ is equal to $\mathcal{B}$.*■

An important special case of examples (which includes a well-known and important 2-manifold known as the **Klein bottle**) is discussed in the exercises for this section (specifically, see the exercise concerning the *mapping torus* construction).

REMARKS.

1. There are numerous examples of regular covering spaces $E \to B$ for which E has a smooth structure but B does not; in such cases it follows that **at least one** of the covering transformations in Γ is **not** a diffeomorphism. Note however that at least one — the identity — is automatically a diffeomorphism.

2. Although the orbit space M/G of a nonfree smooth action of a finite group G is not necessarily a manifold, there are several important cases where it is a topological manifold and in some of these cases it is even possible to construct reasonably well behaved smooth structures. Further information on some cases is contained in the following paper:

[1] R. Schultz, *Exotic spheres admitting circle actions with codimension four stationary sets*, "Proceedings of the Northwestern Homotopy Theory Conference (Evanston, Ill., 1982)", 339–368, Contemp. Math., 19, Amer. Math. Soc., Providence, RI, 1983.

OTHER IMPORTANT CONSTRUCTIONS. Sections III.4.B and III.5 of these notes contain some additional constructions on smooth manifold that will play a crucial role in the rest of this course.

III.2. Appendix : Alternate definition of smooth structures

One obvious problem with our definition of a smooth structure is that smooth atlases are often awkward to handle. For example, when constructing new smooth structures from old ones we usually cannot expect to get a new maximal atlas directly out of the old one(s), and when proving the basic properties of constructions one frequently needs to choose smooth atlases on an *ad hoc* basis. For this and other reasons, many books and papers define smooth structures in a different manner that does not require the introduction of atlases. In this appendix we shall give this alternate definition and prove it is equivalent to the one we have formulated in terms of atlases. Each formulation has specific advantages and disadvantages; we shall not attempt to discuss these at length.

NOTE. None of the discussion below will be needed subsequently in these notes.

The basic idea behind the alternate definition is to view a smooth structure in terms of its smooth real valued functions. This approach is frequently useful when one uses smooth manifolds in other branches of mathematics or physics (in the latter, smooth functions often reflect properties that can be measured experimentally). In particular, the alternate approach is essential if one wishes to compare and relate the theory of smooth manifolds to traditional mathematical subjects like algebraic geometry and modern developments like noncommutative topology and geometry.

Strictly speaking, the mathematical formulation of the alternate approach is presented most efficiently using **sheaf theory**, but we do not need a great deal of input from this subject so we can and shall describe what we need without mentioning the latter explicitly. For the sake of completeness, here are two online references for the definition of a sheaf. Further information can also be found in any of several standard books on sheaf theory.

`http://en.wikipedia.org/wiki/sheaf`

`http://math.wolfram.com/Sheaf.html`

We continue by introducing the abstract concept we shall need.

Definition. A **ringed space** is a pair $(X, \mathcal{F})$ consisting of a topological space X and a function $\mathcal{F}$ that assigns to every nonempty open subset $U \subset X$ a commutative ring with unit $\mathcal{F}(U)$ such that the following hold:

(1) For each pair of nonempty open subsets V and U such that $V \subset U$ there is a homomorphism of rings with unit $r_{VU} : \mathcal{F}(U) \to \mathcal{F}(V)$.

(2) The homomorphisms in the previous item satisfy the identities $r_{UU} = \mathrm{id}_{\mathcal{F}(U)}$ and $r_{WU} = r_{WV} \circ r_{VU}$ if $W \subset V \subset U$.

(3) Suppose that the open subset U is expressible as a union $\cup_\alpha U_\alpha$, and denote the homomorphisms $\mathcal{F}(U) \to \mathcal{F}(U_\alpha)$ and $\mathcal{F}(U_\alpha) \to \mathcal{F}(U_\alpha \cap U_\beta)$ by r_α and $r_{alpha,\beta}$ respectively. Given a family of elements $f_\alpha \in \mathcal{F}(U_\alpha)$ satisfying the "compatibility conditions" $r_{\alpha,\beta}(f_\alpha) = r_{\beta,\alpha}(f_\beta)$ for all α and β, there is a unique $f \in \mathcal{F}(U)$ such that $r_\alpha(f) = f_\alpha$ for all α.

For the sake of uniformity we often set $\mathcal{F}(\emptyset) = \{0\}$. If the rings $\mathcal{F}(U)$ are all algebras over the real numbers and the homomorphisms are maps of $\mathbf{R}$-algebras, then we say that $(X, \mathcal{F})$ is a *ringed space taking values in* $\mathbf{R}$*-algebras*.

Since it should be clear that such a complicated definition was created with some basic examples in mind, we proceed to an important one immediately.

Ringed spaces of continuous real valued functions. Let X be an arbitrary topological space, and let $\mathcal{F}(U)$ denote the real algebra of continuous real valued functions on an open subset U. In this example the maps r_{UV} are restriction homomorphisms taking a function f to $f|V$, and the lengthy third condition simply reflects the fact that one can find a continuous function restricting to locally defined functions f_α if and only if one has the basic consistency relations

$$f_\alpha |\, U_\alpha \cap U_\beta \quad = \quad f_\beta |\, U_\alpha \cap U_\beta$$

for all α and β.

Monumental results of I. Gelfand and M. Naimark from the middle of the twentieth century imply that *one can completely recover a compact Hausdorff space from its algebra of continuous functions*. In fact, one can also conclude that continuous mappings between such spaces are in 1–1 correspondence with continuous algebra homomorphisms of these structured function algebras. Some additional discussion of this topic appears in the file(s)

`http://math.ucr.edu/~res/math205A/solutions.*`

in the Mathematics 205A course directory (specifically, see the final problem). For our purposes here, the point of mentioning the results of Gelfand and Naimark is to indicate that function algebras often carry a great deal of information about topological and geometric structures. Other examples along these lines are theorems of L. E. Pursell and M. E. Shanks about retrieving a smooth manifold from its algebra of smooth functions and related objects. Here some references:

[1] P. W. Michor, J. Vanžura, *Characterizing algebras of $\mathcal{C}^\infty$-functions on manifolds*, Comment. Math. Univ. Carolinæ **37** (1996), 519–521. [Available online at `http://` [*continue*] `www.karlin.mff.cuni.cz/cmuc/pdf/cmuc9603/michor.pdf`]

[2] L. E. Pursell, Algebraic structures associated with smooth manifolds, Ph. D. Thesis, Purdue University, 1952.

[3] L. E. Pursell and M. E. Shanks, *The Lie algebra of smooth manifolds*, Proc. Amer. Math. Soc. **5** (1954) 468–472.

In view of the preceding discussion, the next example should also not be surprising: Let M be a smooth manifold, and for each open subset U let $\mathcal{F}(U)$ denote the algebra of all smooth $\mathcal{C}^\infty$ functions on U. Note that this defines a sub-ringed-space of the previous example (the terminology is clumsy because of our self-imposed constraints; in the language of sheaf theory this would be known as a sheaf of subrings). The main goal of this appendix is to state the key result describing a smooth structure on a topological manifold using such a sub-ringed-spaces of the ringed space of continuous functions.

THEOREM. *Suppose that M is a topological and $(M, \mathcal{F}$ is a ringed space with values in $\mathbf{R}$-algebras suc that the following hold:*

(*i*) *For each open subset $U \subset M$, the algebra $\mathcal{F}(U)$ is a real subalgebra on U.* [We shall denote the inclusion homomorphism by θ_U.]

(*ii*) *For each pair of open subsets U and V such that $V \subset U$, we have $\theta_V \circ r_{VU}(f) = \theta_U(f)|V$ (in other words, r_{VU} corresponds to restriction of a function to V).*

(*iii*) *For each point $x \in M$ there is an open neighborhood U and a homeomorphism from an subset $W \subset \mathbf{R}^n$ to U such that, for each open subset $V \subset U$, the $\mathbf{R}$-algebra isomorphism $h^* : \mathcal{C}(V) \to \mathcal{C}(\,h^{-1}V\,)$, defined by $h^*f(y) = f(h(y))$, sends $\mathcal{F}(V)$ to $\mathcal{C}^\infty(\,h^{-1}(V)\,)$.*

THEN *there is a unique smooth structure on M such that for each open subset $V \subset M$ the subalgebra $\mathcal{F}(V) \subset \mathcal{C}(V)$ is equal to the subalgebra of $\mathcal{C}^\infty$ functions on V.*

Proofs of this result and the one stated below are contained in the file(s) `ringedspacedef.`$*$ (where once again $* =$ `ps` or `pdf`) in the course directory.■

There is also a characterization of smooth maps in terms of ringed spaces.

THEOREM. *Let M and N be smooth manifolds, and let $h : M \to N$ be continuous. Then h is smooth if and only if for each open subset $V \subset N$ the algebra homomorphism $h^* : \mathcal{C}(V) \to \mathcal{C}(,h^{-1}(V)\,)$, defined by $h^*f(y) = f(h(y))$, sends $\mathcal{C}^\infty(V)$ to $\mathcal{C}^\infty(\,h^{-1}(V)\,)$.*

In fact, it is possible to establish the result on smooth maps without introducing any of the concepts or results discussed above. The proof of this theorem is left to the reader as an exercise.■

III.3 : Smooth approximations

(Conlon, §§ 3.5, 3.8)

We begin by stating a simple consequence of the Stone-Weierstrass Approximation Theorem from real analysis.

PROPOSITION. *Let U be open in $\mathbf{R}^{m+n}m$ and let L be a smooth n-manifold given by the level set of a regular value for some smooth function $f : U \to \mathbf{R}^n$. Assume that L is compact, $\varepsilon > 0$, and $g : L \to \mathbf{R}$ is continuous. Then there is a smooth function h such that $|h - g| < \varepsilon$.*

Proof. Let $\mathcal{C}(L)$ denote the Banach space of continuous functions on L. If $\mathcal{C}^\infty(L)$ denotes the subalgebra of $\mathcal{C}^\infty$ functions, then $\mathcal{C}^\infty(L)$ contains all restrictions of polynomial functions and accordingly it separates points and contains the constant function. These conditions suffice to apply the Stone-Weierstrass approximation Theorem, and therefore it follows that every function in $\mathcal{C}(L)$ can be uniformly approximated by a function in $\mathcal{C}^\infty(L)$.∎

Motivated by this result and the simplicity of the proof, it is natural to ask if one has similar approximation results if L and $\mathbf{R}$ are replaced by arbitrary smooth manifolds. One objective of this section is to establish results of this type; we shall not attempt to obtain the sharpest possible conclusions but instead will try to illustrate the general approach. The proposition and its proof should suggest that finding good approximations locally is fairly easy; our previous results on partitions of unity suggest that some smooth version of the latter might provide the means for obtaining local approximations from global ones. Therefore our first step will be to prove a result on smooth partitions of unity generalizing the previous ones for (a) continuous partitions of unity on topological manifolds, (b) smooth partitions of unity on open subsets of Euclidean spaces.

III.3.1 : *Smooth partitions of unity for manifolds*

Any reasonable definition of smooth manifolds should imply that such objects admit smooth partitions of unity, and we shall verify this for our definition right now.

To formulate the basic result, we shall again use notation similar to that of Sections I.2 and II.3. Let $(M, \mathcal{A})$ be a smooth manifold, and let $\mathcal{U}$ be an open covering of M. Then the methods of Sections I.2 and II.3 yield a countable locally finite refinement $\mathcal{V}$ such that the following hold:

(i) Each V_α in $\mathcal{V}$ is the image of some smooth chart of the form $(N_2(0), h_\alpha)$

(ii) The sets W_α, defined as the images $h_\alpha(N_1(0))$, also form an open covering of M.

As in the previous cases, $\overline{W_\alpha}$ is compact and $\overline{W_\alpha} \subset V_\alpha$ by construction.

EXISTENCE OF SMOOTH PARTITIONS OF UNITY, VERSION 2. *Let $(M, \mathcal{A})$ be a smooth manifold and let $\mathcal{M}$ and $\mathcal{N}$ be countable open coverings satisfying the properties of $\mathcal{V}$ and $\mathcal{W}$ as above. Then there is a family of smooth $\mathcal{C}^\infty$ functions $\varphi_j : M \to \mathbf{R}$ with values in $[0, 1]$ such that*

(i) *the* **support** *of φ_j — that is, the closure of the set on which $\varphi \neq 0$ — is a compact subset of M_j,*

(ii) *we have $\sum_j \varphi_j = 1$.*

As before, such a family of smooth functions is called a *smooth* $(\mathcal{C}^\infty)$ *partition of unity* subordinate to the open covering $\mathcal{M}$. As in the previous situations there is no convergence problem with the sum even if there are infinitely many sets in the open covering $\mathcal{M}$.

Proof. The argument is analogous to the proofs of the results in Sections I.2 and II.3, so we shall concentrate on the changes that are needed to make that proof work in the present situation.

Let ω be the smooth $\mathcal{C}^\infty$ bump function on the interval $[0, 2]$ such that $\omega = 1$ on $[0, 1]$, ω decreases linearly from 1 to 0 on $[1.\frac{3}{2}]$, and $\omega = 0$ on $[\frac{3}{2}, 2]$. Define a smooth function h_α

on $M_\alpha = h_\alpha(N_2(0))$ such that $f_\alpha(h_\alpha(x)) = \omega(|x|)$. This definition is justified by the fact that h_α defines a diffeomorphism from $N_2(0)$ onto the image of h_α, and as in the earlier proofs one can extend f_α to a smooth function on all of M by setting it equal to zero on the open subset $M - h_\alpha(N_{3/2}(0)$. If one chooses the maps f_j in this fashion, the proofs of the previously established existence theorems go through with no other changes.■

We can now generalize the previous applications of partitions of unity to smooth manifolds.

FIRST PROPOSITION, SMOOTH VERSION 2. *Suppose that $(M, \mathcal{A})$ is a smooth manifold Ω is an open neighborhood of $M \times \{0\}$ in $M \times \mathbf{R}$. Then there is a smooth $\mathcal{C}^\infty$ real valued function $f : X \to (0, \infty)$ such that the set*

$$\{\ (x,t) \in M \times [0,\infty) \mid t < f(x)\ \}$$

is contained in Ω.■

SECOND PROPOSITION, SMOOTH VERSION 2. *Let $(M, \mathcal{A})$ be a smooth manifold. Then there is a smooth $\mathcal{C}^\infty$ function $f : U \to \mathbf{R}$ with values in $[0, +\infty)$ such that for each $K > 0$ the inverse image $f^{-1}([0,K])$ is compact (in other words, f is a* **proper** *smooth map).*■

This is also a good place to record a generalization of another result from Section II.3.

GERM EXTENSION THEOREM, GENERAL VERSION. *Let $(M, \mathcal{A})$ be a smooth manifold, let $x \in M$, and let $U \subset M$ be an open subset containing x. Then there is an open neighborhood $U \subset W$ of x such that the restriction $f|V$ extends to a smooth $\mathcal{C}^r$ function from M to $\mathbf{R}^m$.*

Proof. Let $W \subset U$ be the image of a smooth coordinate chart in $\mathcal{A}$ at x. Then by the proof of the previously stated version of the Germ Extension Theorem, one can find an open subneighborhood $V \subset W$ and an intermediate neighborhood V_1 such that

$$V \subset \overline{V} \subset V_1 \subset \overline{V}_1 \subset W$$

and a smooth function $g_0 : U \to \mathbf{R}^n$ such that $g_0|V = f|V$ and $g_0|W - \overline{V}_1$ is the constant function whose value is $\mathbf{0}$. We can extend g_0 to a smooth function on M by setting it equal to zero on $M - \overline{V}_1$.■

III.3.2 : *Smooth perturbations of continuous maps*

We proceed to establish a strong generalization of the first result in this subsection.

SMOOTH PERTURBATION THEOREM. *Let $(M, \mathcal{A})$ be a compact smooth n-manifold, let W be open in $\mathbf{R}^q$, let $f : M \to U$ be continuous, and let $\varepsilon > 0$ be given. Then there is a smooth map $g : M \to U$ such that $|f(x) - g(x)| < \varepsilon$ for all x. Furthermore, it ε is sufficiently small, then f and g are homotopic.*

Proof. Let $K = f(M)$, so K is a compact subset of U. Then there is a $\delta > 0$ such that $N_\delta(K; \mathbf{R}^q) \subset U$; without loss of generality we may assume that $\varepsilon < \delta$ (if $\varepsilon' < \varepsilon$ then an ε'-approximation is also an ε-approximation). By the compactness of M there is a finite open covering of K by open disks of the form $N_\delta(y_i)$ where $y_i \in K$. By the continuity of f there is also a finite collection of smooth charts $(N_2(0), h_j)$ for M such that the image of each map $f \circ h_j$ lies in one of the sets $N_\delta(y_i)$ and the open subsets $h_j(N_1(0))$ also form an open covering of M.

Let $\frac{3}{2}D^n$ be the closed disk of radius $\frac{3}{2}$ in $\mathbf{R}^n$. By the Stone-Weierstrass Approximation Theorem, for each j there is a polynomial function g_j of n variables such that $|g_j - f| < \varepsilon$ on $\frac{3}{2}D^n$. Let $\{\varphi_i\}$ be a smooth partition of unity on M subordinate to the covering by the images of the maps h_i such that the support of φ_i contains $h_i(D^n)$ and is contained in $h_i\left(\frac{3}{2}D^n\right)$. Set g equal to the sum $\sum_i \varphi_i \cdot g_i$. If $x \in h_i(\frac{3}{2}D^n)$, then

$$|\varphi_i \cdot g_i \ - \ \varphi_i \cdot f| \ < \ \varphi_i \cdot \varepsilon$$

and the left hand side is zero elsewhere. Therefore it follows that

$$|g - f| \ = \ \left| \sum_i \varphi_i \cdot g_i \ - \ \sum_i \varphi_i \cdot f \right| \ \leq$$

$$\sum_i |\varphi_i \cdot g_i \ - \ \varphi_i \cdot f| \ < \ \sum_i \varphi_j \cdot \varepsilon \ = \ \varepsilon$$

so that g is a smooth ε-approximation to f. Since we are assuming $\varepsilon < \delta$ it also follows that $g(M) \subset U$. This proves everything except the final assertion in the theorem.

In order to prove that the maps are homotopic if ε is sufficiently small, we need the following topological result:

LEMMA. *Let U be open in $\mathbf{R}^n$, and let $K \subset U$ be compact. Then there is an $\eta > 0$ such that $N_\eta(x) \subset K$ for all $x \in K$.*

Proof of Lemma. For each $x \in K$ there is an $r(x) > 0$ such that $N_{r(x)}(x) \subset U$. Since this neighborhood is convex, if $(y, z) \in N_{r(x)}(x) \times N_{r(x)}(x)$ then the straight line segment joining y to z lies entirely in U. The union of these products over all $x \in K$ defines an open neighborhood of the diagonal Δ_K in $U \times U$. Let W be the intersection of this neighborhood with $K \times U$. Since the complement of W is closed in $K \times W$ and disjoint from Δ_K, it follows that the distance between these two sets is a positive constant η; here we are using the cartesian product distance, viewing $K \times W$ as a subset of $\mathbf{R}^{2n}$. Now if $x \in K$ and $|x - y| < \eta$, then it follows that for all points z on the line segment joining x to z we know that the distance from (x, x) to (x, z), which is just $|x - z|$, is less than η and hence (x, z) must belong to W. Therefore we have shown that $N_\eta(x) \subset U$.■

Conclusion of the proof of the Smooth Perturbation Theorem. If $\varepsilon < \eta$, then it follows that U contains the straight line segments joining the points $f(x)$ and $g(x)$ for all x. Therefore f and g are homotopic by the standard straight line homotopy

$$h(x, t) \ = \ t \cdot g(x) \ + \ (1 - t) \cdot f(x)$$

because the right hand side always belongs to U.■

Generalization. A similar result holds if M is noncompact and N arbitrary with some minor modifications. Since we shall not need this result later in the course and some additional complications arise, we shall not attempt give a proof here.■

The first part of [MUNKRES2] contains many further and more delicate results of this type, including results for approximating a smooth $\mathcal{C}^s$ function by a smooth $\mathcal{C}^s$ function for some $r > s$. Such results figure in the proofs of the theorems on raising differentiability classes that were mentioned in Section III.1.

III.3.3 : *Smooth homotopies* $(1\frac{1}{2}\star)$

According to Corollary 3.8.18 on page 119 of Conlon, two smooth maps f and g from one smooth manifold to another are continuously homotopic if and only if they are smoothly homotopic. We shall not prove this statement here, but it seems worthwhile to say something about the definition of a smooth homotopy, partly because the definition in Conlon uses the notion of manifolds with boundary that is not covered in these notes.

Definition. Let M and N be smooth manifolds and let $f, g : M \to N$ be smooth maps. An *admissible smooth homotopy* from f to g is a continuous map $H : M \times [0,1] \to N$ such that the following hold:

- (i) $H(x,0) = f(x)$ and $H(x,1) = g(x)$ for all $x \in M$
- (ii) $H|M \times (0,1)$ is smooth.
- (iii) For each $x \in M$ there is an open neighborhood U and an $\varepsilon > 0$ such that the restrictions of H to $U \times [0,\varepsilon)$ and $U \times (\varepsilon, 1]$ depend only on the first variable.

One then has the following analog of Exercise 3.8.3 on page 188 of Conlon.

PROPOSITION. *Let M and N be smooth manifolds and let $f, g : M \to N$ be smooth maps that are homotopic by an admissible smooth homotopy. Then there is a* strongly admissible homotopy *$K : M \times [0,1] \to N$ for which there is a $\delta > 0$ such that $K(x,t) = f(x)$ for all x and all $t \in [0,\delta)$ and $K(x,t) = g(x)$ for all x and all $t \in (\delta, 1]$.*

Proof. Let H be an admissible smooth homotopy and extend H to $M \times \mathbf{R}$ by setting $H(x,t) = f(x)$ for $t \leq 0$ and $H(x,t) = f(x)$ for $t \geq 1$. This determines a well-defined continuous mapping because the definitions agree on the two overlapping pieces $M \times \{0,1\}$. It also follows immediately that the function is smooth except perhaps on the latter set. Smoothness at the latter points follows because every such point (x,s) has a product neighborhood $U \times J$ for some open interval J such that $H|U \times J$ depends only upon the first coordinate and the associated function $H|U \times \{s\}$ is smooth. Let $\eta \in (0, \frac{1}{4})$, and define a linear map λ from $\mathbf{R}$ to itself that takes 0 and 1 to $-\eta$ and $1 + \eta$ respectively. One can then take $K(x,t) = H(x, \lambda(t))$.■

As noted in Conlon, a result of this sort is needed to prove the following expected result:

COROLLARY. *Suppose that $f, g, h : M \to N$ are smooth maps and there are admissible smooth homotopies from f to g and g to h. Then there is an admissible smooth homotopy from f to h.*

Sketch of proof. By the proposition we known that there are strongly admissible homotopies. Using these we can construct a continuous homotopy from f to h in the usual manner; by construction, this homotopy will be smooth on $M \times (0,1)$ and also strongly admissible. Perhaps the most significant point is that one needs to check directly that the constructed homotopy will be smooth on points of $M \times \{\frac{1}{2}\}$. This follows because the constructed homotopy is a smooth map depending only on the M coordinate on some open set of the form $M \times N_\gamma(\frac{1}{2})$.■

One common thread running through Conlon and these notes is that the notion of smooth homotopy defined in each source is equivalent to the notion of strongly admissible homotopy described here.

III.4 : Amalgamation theorems

(Conlon, § 1.3)

Many physical and geometrical objects are describable in terms of pieces that are somehow glued together. This principle applies particularly to the theory of smooth manifolds. The purpose of this section is to develop the mathematical concepts and results that are needed to assemble topological spaces and smooth manifolds from a collection of pieces. More formally, we must answer the following question at the topological level: Given a collection of topological spaces, what sorts of data do we need in order to glue them together and form a single "reasonable" space? Most of the time we also want a simple additional condition; namely, the space we construct should have an open, or finite closed, covering consisting of subsets homeomorphic to the objects in the original collection.

Since it is generally useful to have specific examples when setting up abstract mathematical machinery, here is one that is fairly simple and familiar but not entirely trivial: Physically it is clear that one can form a cube from six pairwise disjoint squares with sides of equal length by gluing the latter together in a suitable way along the edges. Whatever formalism we develop should provide a mathematical model for this well known process.

III.4.A. : Topological amalgamation

Since smooth manifolds are topological spaces with additional structure, we shall begin by discussing the underlying topological concepts and results. The first step is to introduce some constructions that are elementary and necessary for this course but do not appear in most point set topology texts (including [MUNKRES1]!).

III.4.A.1 : *Disjoint unions*

We shall need an elementary set-theoretic construction that is described in the ONLINE 205A NOTES. Namely, given two sets A and B we need to have a *disjoint union*, written $A \sqcup B$ or $A \coprod B$, which is a union of two disjoint subsets that are essentially xerox copies of A and B.

Most texts and courses on set theory and point set topology (*e.g.*, [MUNKRES1]) do not say much if anything about disjoint union constructions, one reason being that everything is fairly elementary when one finally has the right definitions (two references in print are Section 8.7 of Royden, *Real Analysis*, and Sections I.3 and III.4–III.7 of the text by K. Jänich mentioned at the beginning of these notes).

Since constructions of this sort play a crucial role beginning with the next section of these notes, a brief but comprehensive treatment seems worthwhile for the sake of precision and clarity.

Formally, **the disjoint union** (or **set-theoretic sum**) of two sets A and B is defined to be the set

$$A \coprod B \quad = \quad \big(A \times \{1\}\big) \bigcup \big(B \times \{2\}\big) \quad \subset \quad \big(A \cup B\big) \times \{1, 2\}$$

with **injection** maps $i_A : A \to A \coprod B$ and $i_B : A \to A \coprod B$ given by $i_A(a) = (a, 1)$ and $i_B(b) = (b, 2)$. The images of these injections are disjoint copies of A and B, and the union of the images is $A \coprod B$.

Definition. If X and Y are topological spaces, the *disjoint union topology* or *(set-theoretic) sum topology* on the set $X \coprod Y$ consists of all subsets having the form $U \coprod V$, where U is open in X and V is open in Y.

We claim that this construction defines a topology on $X \coprod Y$, and the latter is a union of disjoint homeomorphic copies of X and Y such that each of the copies is an open and closed subset. Formally, all this is expressed as follows:

ELEMENTARY PROPERTIES. *The family of subsets described above is a topology for $X \coprod Y$ such that the injection maps i_X and i_Y are homeomorphisms onto their respective images. These images are pairwise disjoint, and they are also open and closed subspaces of $X \coprod Y$. Each injection map is continuous, open and closed.*

Sketch of proof. This is all pretty elementary, but we include it because the properties are so fundamental and the details are not readily available in the standard texts.

Since X and Y are open in themselves and $\emptyset$ is open in both, it follows that $X \coprod Y$ and $\emptyset = \emptyset \coprod \emptyset$ are open in $X \coprod Y$. Given a family of subsets $\{ U_\alpha \coprod V_\alpha \}$ in the so-called disjoint union topology, then the identity

$$\bigcup_\alpha \left(U_\alpha \coprod V_\alpha \right) \quad = \quad \left(\bigcup_\alpha U_\alpha \right) \coprod \left(\bigcup_\alpha V_\alpha \right)$$

shows that the so-called disjoint union topology is indeed closed under unions, and similarly the if $U_1 \coprod V_1$ and $U_2 \coprod V_2$ belong to the so-called disjoint union topology, then the identity

$$\bigcap_{i=1,2} \left(U_i \coprod V_i \right) \quad = \quad \left(\bigcap_{i=1,2} U_i \right) \coprod \left(\bigcap_{i=1,2} V_i \right)$$

shows that the so-called disjoint union topology is also closed under finite intersections. In particular, we are justified in calling this family a topology.

By construction U is open in X if and only if $i_X(U)$ is open in $i_X(X)$, and V is open in Y if and only if $i_Y(V)$ is open in $i_Y(Y)$; these prove the assertions that i_X and i_Y are homeomorphisms onto their images. Since $i_X(X) = X \coprod \emptyset$, it follows that the image of i_X is open, and of course similar considerations apply to the image of i_Y. Also, the identity

$$i_X(X) = \left(X \coprod Y \right) \; - \; i_Y(Y)$$

shows that the image of i_X is closed, and similar considerations apply to the image of i_Y.

The continuity of i_X follows because every open set in $X \coprod Y$ has the form $U \coprod V$ where U and V are open in X and Y respectively and

$$i_X^{-1} \left(U \coprod V \right) \quad = \quad U$$

with similar conditions valid for i_Y. The openness of i_X follows immediately from the identity $i_X(U) = U \coprod \emptyset$ and again similar considerations apply to i_Y. Finally, to prove that i_X is closed, let $F \subset X$ be closed. Then $X - F$ is open in X and the identity

$$i_X(F) = F \coprod \emptyset = \left(X \coprod Y\right) - \left((X - F) \coprod Y\right)$$

shows that $i_X(F)$ is closed in $X \coprod Y$; once more, similar considerations apply to i_Y.■

IMMEDIATE CONSEQUENCE. *The closed subsets of $X \coprod Y$ with the disjoint union topology are the sets of the form $E \coprod F$ where E and F are closed in X and Y respectively.*■

If the topologies on X and Y are clear from the context, we shall generally assume that the $X \coprod Y$ is furnished with the disjoint union topology unless there is an explicit statement to the contrary.

Since the disjoint union topology is not covered in many texts, we shall go into more detail than usual in describing their elementary properties.

FURTHER ELEMENTARY PROPERTIES. *(i) If X and Y are discrete, then so is $X \coprod Y$.*

(ii) If X and Y are Hausdorff, then so is $X \coprod Y$.

(iii) If X and Y are homeomorphic to metric spaces, then so is $X \coprod Y$.

(iv) If $f : X \to W$ and $g : Y \to W$ are continuous maps into some space W, then there is a unique continuous map $h : X \coprod Y \to W$ such that $h \circ i_X = f$ and $h \circ i_Y = g$.

(v) The spaces $X \coprod Y$ and $Y \coprod X$ are homeomorphic for all X and Y. Furthermore, if Z is a third topological space then there is an "associativity" homeomorphism

$$\left(X \coprod Y\right) \coprod Z \quad \cong \quad X \coprod \left(Y \coprod Z\right)$$

(in other words, the disjoint sum construction is commutative and associative up to homeomorphism).

Sketches of proofs. (i) A space is discrete if every subset is open. Suppose that $E \subset X \coprod Y$. Then E may be written as $A \coprod B$ where $A \subset X$ and $B \subset Y$. Since X and Y are discrete it follows that A and B are open in X and Y respectively, and therefore $E = A \coprod B$ is open in $X \coprod Y$. Since E was arbitrary, this means that the disjoint union is discrete.

(ii) If one of the points p, q lies in the image of X and the other lies in the image of Y, then the images of X and Y are disjoint open subsets containing p and q respectively. On the other hand, if both lie in either X or Y, let V and W be disjoint open subsets containing the preimages of p and q in X or Y. Then the images of V and W in $X \coprod Y$ are disjoint open subsets that contain p and q respectively.

(iii) As noted in Theorem 20.1 on page 121 of [MUNKRES1], if the topologies on X and Y come from metrics, one can choose the metrics so that the distances between two points are ≤ 1. Let $\mathbf{d}_X$ and $\mathbf{d}_Y$ be metrics of this type.

Define a metric $\mathbf{d}^*$ on $X \coprod Y$ by $\mathbf{d}_X$ or $\mathbf{d}_Y$ for ordered pairs of points (p, q) such that both lie in the image of i_X or i_Y respectively, and set $\mathbf{d}^*(p, q) = 2$ if one of p, q lies in the image of i_X and the other lines in the image of i_Y. It follows immediately that $\mathbf{d}^*$ is nonnegative, is zero if and only if $p = q$ and is symmetric in p and q. All that remains to check is the Triangle Inequality:

$$\mathbf{d}^*(p, r) \quad \leq \quad \mathbf{d}^*(p, q) + \mathbf{d}^*(q, r)$$

The verification breaks down into cases depending upon which points lie in the image of one injection and which lie in the image of another. If all three of p, q, r lie in the image of one of the injection maps, then the Triangle Inequality for these three points is an immediate consequence of the corresponding properties for $\mathbf{d}_X$ and $\mathbf{d}_Y$. Suppose now that p and r lie in the image of one injection and q lies in the image of the other. Then we have $\mathbf{d}^*(p,r) \leq 1$ and

$$\mathbf{d}^*(p,q) \ + \ \mathbf{d}^*(q,r) \quad = \quad 2+2 \quad = \quad 4$$

so the Triangle Inequality holds in these cases too. Finally, if p and r lie in the images of different injections, then either p and q lie in the images of different injections or else q and r lie in the images of different injections. This means that $\mathbf{d}^*(p,r) = 2$ and $\mathbf{d}^*(p,q) + \mathbf{d}^*(q,r) \geq 2$, and consequently the Triangle Inequality holds for **all** ordered pairs (p,r).

(iv) Define $h(x,1) = f(x)$ and $h(y,2) = g(y)$ for all $x \in X$ and $y \in Y$. By construction $h \circ i_X = f$ and $h \circ i_Y = g$, so it remains to show that h is continuous and there is no other continuous map satisfying the functional equations. The latter is true for set theoretic reasons; the equations specify the behavior of h on the union of the images of the injections, but this image is the entire disjoint union. To see that h is continuous, let U be an open subset of X, and consider the inverse image $U^* = h^{-1}(U)$ in $X \coprod Y$. This subset has the form $U^* = V \coprod W$ for some subsets $V \subset X$ and $W \subset Y$. But by construction we have

$$V \quad = \quad i_X^{-1}(U^*) \quad = \quad i_X^{-1} \circ h^{-1}(U) \quad = \quad f^{-1}(U)$$

and the set on the right is open because f is continuous. Similarly,

$$W \quad = \quad i_Y^{-1}(U^*) \quad = \quad i_Y^{-1} \circ h^{-1}(U) \quad = \quad g^{-1}(U)$$

so that the set on the right is also open. Therefore $U^* = V \coprod W$ where V and W are open in X and Y respectively, and therefore U^* is open in $X \coprod Y$, which is exactly what we needed to prove the continuity of h.

(v) We shall merely indicate the main steps in proving these assertions and leave the details to the reader as an exercise. The homeomorphism τ from $X \coprod Y$ to $Y \coprod X$ is given by sending $(x,1)$ to $(x,2)$ and $(y,2)$ to $(y,1)$; one needs to check this map is 1–1, onto, continuous and open (in fact, if τ_{XY} is the map described above, then its inverse is τ_{YX}). The "associativity homeomorphism" sends $(\,(x,1),1)$ to $(x,1)$, $(\,(y,2),1)$ to $(\,(y,1),2)$, and $(z,2)$ to $(\,(z,2),2)$. Once again, one needs to check this map is 1–1, onto, continuous and open.■

COMPLEMENT. *There is an analog of Property* (iv) *for untopologized sets.*

Perhaps the fastest way to see this is to make the sets into topological spaces with the discrete topologies and then to apply (i) and (iv).■

Property (iv) is dual to the fundamental defining property of direct products. Specifically, oredered pairs of maps from a fixed object A to objects B and C correspond to maps from A into $B \times C$, while oredered pairs of maps going **TO** a fixed object A and coming **FROM** objects B and C correspond to maps from $B \coprod C$ into A. For this reason one often refers to $B \coprod C$ as the **coproduct** of B and C (either as sets or as topological spaces); this is also the reason for denoting disjoint unions by the symbol $\coprod$, which is merely the product symbol $\prod$ turned upside down.

III.4.A.2 : *Copy, cut and paste constructions* $(1\frac{1}{2}\star)$

Frequently the construction of spaces out of pieces proceeds by a series of steps where one takes two spaces, say A and B, makes disjoint copies of them, finds closed subspaces C and D that are homeomorphic by some homeomorphism h, and finally glues A and B together using this homeomorphism. For example, one can think of a rectangle as being formed from two right triangles by gluing the latter along the hypotenuse. Of course, there are also many more complicated examples of this sort.

Formally speaking, we can try to model this process by forming the disjoint union $A \coprod B$ and then factoring out by the equivalence relation

$$x \sim y \Longleftrightarrow x = y \quad \textbf{OR}$$

$$x = i_A(a),\ y = i_B(h(a)) \quad \text{for some } a \in A \quad \textbf{OR}$$

$$y = i_A(a),\ x = i_B(h(a)) \quad \text{for some } a \in A.$$

It is an elementary but tedious exercise in bookkeeping to to verify that this defines an equivalence relation (the details are left to the reader!). The resulting quotient space will be denoted by

$$A \bigcup_{h:C\equiv D} B.$$

As a test of how well this approach works, consider the following question:

Scissors and Paste Problem. *Suppose we are given a topological space X and closed subspaces A and B such that $X = A \cup B$. If we take $C = D = A \cap B$ and let h be the identity homeomorphism, does this construction yield the original space X?*

One would expect that the answer is yes, and here is the proof:

Retrieving the original space. Let Y be the quotient space of $A \coprod B$ with respect to the equivalence relation, and let $p : A \coprod B \to Y$ be the quotient map. By the preceding observations, there is a unique continuous map $f : A \coprod B \to X$ such that $f \circ i_A$ and $f \circ i_B$ are the inclusions $A \subset X$ and $B \subset X$ respectively. By construction, if $u \sim v$ with respect to the equivalence relation described above, then $f(u) = f(v)$, and therefore there is a unique continuous map $h : Y \to X$ such that $f = h \circ p$. We claim that h is a homeomorphism. First of all, h is onto because the identities $h \circ p \circ i_A = \text{inclusion}_A$ and $h \circ p \circ i_B = \text{inclusion}_B$ imply that the image contains $A \cup B$, which is all of X. Next, h is 1–1. Suppose that $h(u) = h(v)$ but $u \neq v$, and write $u = p(u')$, $v = p(v')$. The preceding identities imply that h is 1-1 on both A and B, and therefore one of u', v' must lie in A and the other in B. By construction, it follows that the inclusion maps send u' and v' to the same point in X. But this means that u' and v' correspond to the same point in $A \cap B$ so that $u = p(u') = p(v') = v$. Therefore the map h is 1–1. To prove that h is a homeomorphism, it suffices to show that h takes closed subsets to closed subsets. Let F be a closed subset of Y. Then the inverse image $p^{-1}(F)$ is closed in $A \coprod B$. However, if we write write $h(F) \cap A = P$ and $h(f) \cap B = Q$, then it follows that $p^{-1}(F) = i_A(P) \cup i_B(Q)$. Thus $i_A(P) = p^{-1}(F) \cap i_A(A)$ and $i_B(Q) = p^{-1}(F) \cap i_B(B)$, and consequently the subsets $i_A(P)$ and $i_B(Q)$ are closed in $A \coprod B$. But this means that P and Q are closed in A and B respectively, so that $P \cup Q$ is closed in X. Therefore it suffices to verify that $h(F) = P \cup Q$. But if $x \in F$, then the surjectivity of p implies that $x = p(y)$ for some $y \in p^{-1}(F) = i_A(P) \cup i_B(Q)$; if $y \in i_A(P)$ then we have

$$h(x) \quad = \quad h(p(y)) \quad = \quad f(y) \quad = \quad f \circ i_A(y) \quad = \quad y$$

for some $y \in P$, while if $y \in i_B(Q)$ the same sorts of considerations show that $h(x) = y$ for some $y \in Q$. Hence $h(F)$ is contained in $P \cup Q$. On the other hand, if $y \in P$ or $y \in Q$ then the preceding equations for P and their analogs for Q show that $y = h(p(y))$ and $p(y) \in F$ for $y \in P \cup Q$, so that $P \cup Q$ is contained in $h(F)$ as required.■

One can formulate an analog of the scissors and paste problem if A and B are open rather than closed subset of X, and once again the answer is that one does retrieve the original space. The argument is similar to the closed case and is left to the reader as an exercise.

Examples. Many examples for the scissors and paste theorem can be created involving subsets of Euclidean 3-space. For example, as noted before one can view the surface of a cube as being constructed by a sequence of such operations in which one adds a solid square homeomorphic to $[0,1]^2$ to the space constructed at the previous step. Our focus here will involve examples of objects in 4-dimensional space that can be constructed by a single scissors and paste construction involving objects in 3-dimensional space.

1. The *hypersphere* $S^3 \subset \mathbf{R}^4$ is the set of all points (x, y, z, w) whose coordinates satisfy the equation

$$x^2 + y^2 + z^2 + w^2 \quad = \quad 1$$

and it can be constructed from two 3-dimensional disks by gluing them together along the boundary spheres. An explicit homeomorphism

$$D^3 \bigcup_{\mathrm{id}(S^2)} D^3 \longrightarrow S^3$$

can be constructed using the maps

$$f_{\pm}(x, y, z) \quad = \quad \left(x, y, z, \sqrt{1 - x^2 - y^2 - z^2}\right)$$

on the two copies of D^3. The resulting map is well defined because the restrictions of $f_{\pm}$ to S^2 are equal.

2. We shall also show that the Klein bottle can be constructed by gluing together two Möbius strips along the simple closed curves on their edges. Let $g_{\pm} : [-1, 1] \to S^1$ be the continuous 1–1 map sending t to $\left(\pm\sqrt{1 - t^2}, t\right)$. It then follows that the images $F_{\pm}$ of the maps $\mathrm{id}_{[0,1]} \times [-1, 1]$ satisfy $F + \cup F_- = [0, 1] \times S^1$ and $F_+ \cap [0, 1] \times \{-1, 1\}$. If $\varphi : [0, 1] \times S^1 \to \mathbf{K}$ is the quotient projection to the Klein bottle, then it is relatively elementary to verify that each of the sets $\varphi(F_{\pm})$ is homeomorphic to the Möbius strip (look at the equivalence relation given by identifying two points if they have the same images under $\varphi \circ g_{\pm}$) and the intersection turns out to be the set $\varphi(F_+) \cap \varphi(F_-)$, which is homeomorphic to the edge curve for either of these Möbius strips.

III.4.A.3 : *Disjoint unions of families of sets*

As in the case of products, one can form disjoint unions of arbitrary finite collections of sets or spaces recursively using the construction for a pair of sets. However, there are also cases where one wants to form disjoint unions of infinite collections, so we shall sketch how this can be done, leaving the proofs to the reader as exercises.

Definition. If A is a set and $\{ X_\alpha \mid \alpha \in A \}$ is a family of sets indexed by A, the *disjoint union* (or set-theoretic sum)

$$\coprod_{\alpha \in A} X_\alpha$$

is the subset of all

$$(x, \alpha) \ \in \ \left(\bigcup_{\alpha \in S} X_\alpha \right) \times A$$

such that $x \in X_\alpha$.

This is a direct generalization of the preceding construction, which may be viewed as the special case where $A = \{1, 2\}$. For each $\beta \in A$ one has an injection map

$$i_\beta : X_\beta \to \coprod_{\alpha \in A} X_\alpha$$

sending x to (x, β); as before, the images of i_β and i_γ are disjoint if $\beta \neq \gamma$ and the union of the images of the maps i_α is all of $\coprod_\alpha X_\alpha$.

Notation. In the setting above, suppose that each X_α is a topological space with topology $\mathbf{T}_\alpha$. Let $\sum_\alpha \mathbf{T}_\alpha$ be the set of all disjoint unions $\coprod_\alpha U_\alpha$ where U_α is open in X_α for each α.

As in the previous discussion, this defines a topology on $\coprod_\alpha X_\alpha$, and the basic properties can be listed as follows:

[1] *The family of subsets $\sum_\alpha \mathbf{T}_\alpha$ defines a topology for $\coprod_\alpha X_\alpha$ such that the injection maps i_α are homeomorphisms onto their respective images. the latter are open and closed subspaces of $\coprod_\alpha X_\alpha$, and each injection is continuous, open and closed.*

[2] *The closed subsets of $\coprod X_\alpha$ with the disjoint union topology are the sets of the form $\coprod F_\alpha$ where F_α is closed in X_α for each α.*

[3] *If each X_α is discrete then so is $\coprod_\alpha X_\alpha$.*

[4] *If each X_α is Hausdorff then so is $\coprod_\alpha X_\alpha$.*

[5] *If each X_α is homeomorphic to a metric space, then so is $\coprod_\alpha X_\alpha$.*

[6] *If for each α we are given a continuous function $f : X_\alpha \to W$ into some fixed space W, then there is a unique continuous map $h : \coprod_\alpha X_\alpha \to W$ such that $h \circ i_\alpha = f_\alpha$ for all α.*

The verifications of these properties are direct extensions of the earlier arguments, and the details are left to the reader.∎

In linear algebra one frequently encounters vector spaces that are isomorphic to direct sums of other spaces but not explicitly presented in this way, and it is important to have simple criteria for recognizing situations of this type. Similarly, in working with topological spaces one frequently encounters spaces that are homeomorphic to disjoint unions but not presented in this way, and in this context it is also convenient to have a simple criterion for recognizing such objects.

INTERNAL SUM RECOGNITION PRINCIPLE. *Suppose that a space Y is a union of pairwise disjoint subspaces X_α, each of which is open and closed in Y. Then Y is homeomorphic to $\coprod_\alpha X_\alpha$.*

Proof. For each $\alpha \in A$ let $j_\alpha : X_\alpha \to Y$ be the inclusion map. By [6] above there is a unique continuous function

$$J : \coprod_\alpha X_\alpha \longrightarrow Y$$

such that $J \circ i_\alpha = j_\alpha$ for all α. We claim that J is a homeomorphism; in other words, we need to show that J is 1–1 onto and open. Suppose that we have $(x_\alpha, \alpha) \in i_\alpha(X_\alpha)$ and $(z_\beta, \beta) \in i_\beta(X_\beta)$ such that $J(x_\alpha, \alpha) = J(z_\beta, \beta)$. By the definition of J this implies $i_\alpha(x_\alpha) = i_\beta(z_\beta)$. Since the images of i_α and i_β are pairwise disjoint, this means that $\alpha = \beta$. Since i_α is an inclusion map, it is 1–1, and therefore we have $x_\alpha = z_\beta$. The proof that J is onto drops out of the identities

$$J\left(\coprod_\alpha X_\alpha\right) \quad = \quad J\left(\bigcup_\alpha i_\alpha(X_\alpha)\right) \quad = \quad \bigcup_\alpha J\left(i_\alpha(X_\alpha)\right) \quad = \quad \bigcup_\alpha j_\alpha(X_\alpha) \quad = \quad Y \; .$$

Finally, to prove that J is open let W be open in the disjoint union, so that we have

$$W \quad = \quad \coprod_\alpha U_\alpha$$

where each U_α is open in the corresponding X_α. It then follows that $J(W) = \cup_\alpha U_\alpha$. But for each α we know that U_α is open in X_α and the latter is open in Y, so it follows that each U_α is open in Y and hence that $J(W)$ is open.■

III.4.A.4 : *Constructing topological spaces out of pieces*

In this subsection we shall describe a method for constructing spaces out of relatively complicated data. This procedure is used repeatedly in differential topology and geometry; for example, it provides the framework for constructing spaces of tangent vectors to smooth manifolds in Section III.5 as well as numerous important generalizations.

Disassembly of a space via an open covering. Let X be a topological space, and let $\mathcal{U}$ be an open covering of X consisting of the sets U_α where α lies in some indexing set $\mathbf{A}$. The inclusion map of U_α into X will be denoted by j_α. By results from the preceding subsection, there is a unique continuous function

$$j : \coprod_\alpha U_\alpha \longrightarrow X$$

such that $j \circ i_\alpha = j_\alpha$ for all $\alpha \in \mathbf{A}$. CLAIM: j is an open mapping. — An arbitrary open subset of $\coprod_\alpha U_\alpha$ has the form $\coprod_\alpha V_\alpha$ where V_α is open in U_α (hence also in X). It follows that $j\left(\coprod_\alpha V_\alpha\right)$ is equal to $\cup_\alpha V_\alpha$, and this set is open in X because each V_α is open in X.■

Let $R(\mathcal{U})$ be the equivalence relation on $\coprod_\alpha U_\alpha$ that identifies a and b if and only if $j(a) = j(b)$, and let $\mathbf{p} : \coprod_\alpha U_\alpha \to X^*$ be the projection onto the set of equivalence classes for $R(\mathcal{U})$. Then there is a unique continuous map $J : X^* \to X$ such that $j = J \circ \mathbf{p}$, and by construction J is 1–1 and onto. Since j is open, results on quotient maps in [MUNKRES1] show that *the map J is a homeomorphism.*

III.4.A.5 : *Transition data associated to an open covering*

We are interested in the following problem: *Given an indexed family of topological spaces $\mathcal{U} = \{U_\alpha\}_{\alpha \in \mathbf{A}}$, what additional data are needed to construct an arbitrary topological space X with an open covering topologically equivalent to $\mathcal{U}$?*

The most effective way to analyze this problem is to start with a space X, an open covering $\mathcal{U}$, and the associated continuous open surjection j defined as above. One crucial aspect of understanding the construction of the space X is to study the intersections of two arbitrary open sets in the open covering. Given U_α and U_β in $\mathcal{U}$, define

$$V_{\beta\alpha} \quad = \quad j_\alpha^{-1}(U_\beta) \quad \subset \quad \coprod_\sigma U_\sigma \ .$$

By construction j_α maps $V_{\beta\alpha}$ homeomorphically onto $U_\alpha \cap U_\beta$; likewise, j_β maps $V_{\alpha\beta}$ homeomorphically onto $U_\alpha \cap U_\beta$. Of course this means that $V_{\beta\alpha}$ and $V_{\alpha\beta}$ are homeomorphic, and an explicit homeomorphism

$$\psi_{\beta\alpha} : V_{\beta\alpha} \to V_{\alpha\beta}$$

is given by the following composite:

$$V_{\beta\alpha} \quad = \quad (U_\alpha \cap U_\beta) \times \{\alpha\} \quad \cong \quad (U_\alpha \cap U_\beta) \times \{\beta\} \quad = \quad V_{\alpha\beta}$$

The homeomorphisms $\psi_{\beta\alpha}$ satisfy two basic relations of the form

$$\psi_{\alpha\alpha} \quad = \quad \mathrm{id}(U_\alpha)$$

$$\psi_{\alpha\beta} \quad = \quad \psi_{\beta\alpha}^{-1}$$

as well as a third relation that can be expressed informally as "$\psi_{\gamma\beta} \circ \psi_{\beta\alpha} = \psi_{\gamma\alpha}$." A little care is needed to formulate this precisely because the codomain of $\psi_{\beta\alpha}$ is usually not a subset of the domain of $\psi_{\gamma\beta}$, so it is necessary to be specific about when the composite is definable. This begins with the following observation:

For all α, β, γ in $\mathbf{A}$ the homeomorphism $\psi_{\beta\alpha}$ sends the open subset $V_{\beta\alpha} \cap V_{\gamma\alpha} \subset U_\alpha \times \{\alpha\}$ homeomorphically onto $V_{\alpha\beta} \cap V_{\gamma\beta} \subset U_\beta \times \{\beta\}$.

This is true because the images of the two intersections in X are merely $U_\alpha \cap U_\beta \cap U_\gamma$. Details of this verification are left to the reader.

A precise version of the third relation is then given by

$$\psi_{\gamma\beta}\left(\psi_{\beta\alpha}(x)\right) \quad = \quad \psi_{\gamma\alpha}(x) \quad \text{ if } \quad x \in V_{\beta\alpha} \cap V_{\gamma\alpha} \ .$$

The equivalence relation $R(\mathcal{U})$ has an alternate description in terms of the transition homeomorphisms $\psi_{\beta\alpha}$.

PROPOSITION. *Let a and b be points of $\coprod_\sigma U_\sigma$, and let α and β be the indices in $\mathbf{A}$ such that $a \in \mathrm{Image}\ i_\alpha \approx U_\alpha \times \{\alpha\}$ and $b \in \mathrm{Image}\ i_\beta \cong U_\beta \times \{\beta\}$. Then $(a, b) \in R(\mathcal{U})$ if and only if $b = \psi_{\beta\alpha}(a)$.*

Proof. ($\Longrightarrow$) By definition $(a, b) \in R(\mathcal{U})$ means that $j(a) = j(b)$. If this happens, then the common image point lies in $U_\alpha \cap U_\beta$. But this means that $a \in V_{\beta\alpha}$, $b \in V_{\alpha\beta}$ and $b = \psi_{\beta\alpha}(a)$. ($\Longleftarrow$) If $b = \psi_{\beta\alpha}(a)$ then the definition of $\psi_{\beta\alpha}$ implies that $j(a) = j(b)$.

The reason for dwelling on all these definitions and formulas is that they provide the framework for building a space out of pieces. The first step in establishing this is to formulate everything abstractly.

Definition. A *set of topological amalgamation data* is a pair

$$\big(\{Y_\alpha\}, \{\varphi_{\beta\alpha}\}\big)$$

where $\{Y_\alpha\}$ is an indexed family of topological spaces with indexing set $\mathbf{A}$ and $\{\varphi_{\beta\alpha}\}$ is an indexed family of homeomorphisms with indexing set $\mathbf{A} \times \mathbf{A}$ such that the following conditions hold:

(i) For every α and β the map $\varphi_{\beta\alpha}$ is a homeomorphism from an open subset $W_{\beta\alpha}$ of Y_α to an open subset $W_{\alpha\beta}$ of Y_β.

(ii) For every α the map $\varphi_{\alpha\alpha}$ is the identity map for Y_α, and for every α and β the map $\varphi_{\alpha\beta}$ is the inverse homeomorphism to $\varphi_{\beta\alpha}$.

(iii) For every α, β and γ the map $\varphi_{\beta\alpha}$ sends $W_{\beta\alpha} \cap W_{\gamma\alpha} \subset Y_\alpha$ homeomorphically onto $W_{\alpha\beta} \cap W_{\gamma\beta} \subset Y_\beta$, and $\varphi_{\gamma\beta}\left(\varphi_{\beta\alpha}(y)\right) = \varphi_{\gamma\alpha}(y)$ for all $y \in W_{\beta\alpha} \cap W_{\gamma\alpha}$.

The functional identities described above are called **cocycle formulas** or something similar in Conlon's book and numerous other places (the key word is "cocycle").

We have stated the definition so that the preceding construction defines a set of topological amalgamation data associated to an open covering of a topological space.

There is a corresponding concept of isomorphism; for convenience we shall assume that we have two sets of topological amalgamation data with the same indexing set (QUESTION: *What modifications are necessary if we do not make this assumption?*). Given two such structures

$$\big(\{Y_\alpha\}, \{\varphi_{\beta\alpha}\}\big) \qquad\qquad \big(\{U_\alpha\}, \{\psi_{\beta\alpha}\}\big)$$

an isomorphism between them consists of an indexed family of homeomorphisms $\{h_\alpha : Y_\alpha \to U_\alpha\}$ such that

(a) for every α and β the maps h_α and h_β send the domain and codomain of $\varphi_{\beta\alpha}$ homeomorphically onto the domain and codomain of $\psi_{\beta\alpha}$ respectively,

(b) for every α and β and for every y in the domain of $\varphi_{\beta\alpha}$ we have the following commutativity relation:

$$\varphi_{\beta\alpha}\left(h_\alpha(y)\right) \quad = \quad h_\beta\left(\varphi_{\beta\alpha}(y)\right)$$

We are now ready to state the result we want on building a space out of pieces:

TOPOLOGICAL REALIZATION THEOREM. *If* $\mathbf{Y} = \left(\{Y_\alpha\}, \{\varphi_{\beta\alpha}\}\right)$ *is a set of topological amalgamation data then there is a space* X *and an open covering* $\mathcal{U}$ *of* X *such that* $\mathbf{Y}$ *is isomorphic to the topological amalgamation data associated to* $\mathcal{U}$*. The space* X *is uniquely determined up to homeomorphism.*

Proof. ($\star$) Let Y be the disjoint union $\coprod_\sigma Y_\sigma$, and define a binary relation $R(\mathbf{Y})$ on Y by stipulating that (a, b) lies in the graph of $R(\mathbf{Y})$ if and only if $b = \varphi_{\beta\alpha}(a)$, where $a \in Y_\alpha$ and $b \in Y_\beta$.

The first order of business is to verify that $R(\mathbf{Y})$ is an equivalence relation. The relation is *reflexive* because the first part of property (ii) in the definition implies that $a = \varphi_{\alpha\alpha}(a)$. Similarly, the relation is *reflexive* because the second part of property (ii) in the definition shows that $a =$

$\varphi_{\alpha\beta}(b)$ if $b = \varphi_{\beta\alpha}(a)$. Finally, to verify that the relation is also *transitive*, let a, b and c satisfy $b = \varphi_{\beta\alpha}(a)$ and $c = \varphi_{\gamma\beta}(b)$. It follows that b lies in the intersection $W_{\alpha\beta} \cap W_{\gamma\beta}$, and therfore by the first part of property (iii) it follows that a lies in $W_{\beta\alpha} \cap W_{\gamma\alpha}$. Therefore $\varphi_{\gamma\alpha}(a)$ is defined, and by the second part of property (iii) we have

$$\varphi_{\gamma\alpha}(a) \quad = \quad \varphi_{\gamma\beta}(\varphi_{\beta\alpha}(a))$$

and using the assumptions on a, b and c we may rewrite the right hand side as

$$\varphi_{\gamma\beta}(b) \quad = \quad c$$

so that $c = \varphi_{\gamma\alpha}(a)$, which means that (a, c) lies in the graph of $R(\mathbf{Y})$ and consequently the latter is an equivalence relation as expected.

Let X be the set of equivalence classes of $R(\mathbf{Y})$ with the quotient topology, and for each α let k_α be the composite of the quotient map $p : Y \to X$ with the inclusion $i_\alpha : Y_\alpha \to Y$. We claim that for each α the map h_α is 1–1, continuous and open. Continuity follows immediately because h_α is a composite of two continuous functions. If a and a' lie in Y_α, then their images in X are equal if and only if $a' = \varphi_{\alpha\alpha}(a)$. But $\varphi_{\alpha\alpha}$ is the identity map, so we must have $a = a'$. Finally, to show that k_α is open, let N be an open subset of Y_α; to show that $k_\alpha(N)$ is open in X we need to show that $p^{-1}[k_\alpha(N)]$ is open in Y. But

$$p^{-1}[k_\alpha(N)] \quad \cong \quad \coprod_\beta \varphi_{\beta\alpha}^{-1}[N]$$

and the latter is open in Y by the continuity of the maps $\varphi_{\beta\alpha}$. Therefore k_α is open as claimed.

If we set $U_\alpha = k_\alpha(Y_\alpha)$ then $\mathcal{U} = \{U_\alpha\}$ is an open covering of X. It is left as an exercise for the reader to verify that the original set of topological amalgamation data is isomorphic to the corresponding data set associated to $\mathcal{U}$.

It remains to prove that X is unique up to homeomorphism. Suppose there are spaces X and X' with open coverings $\mathcal{U}$ and $\mathcal{U}'$ such that $\mathbf{Y}$ is isomorphic to the sets of topological amalgamation data associated to both $\mathcal{U}$ and $\mathcal{U}'$. By transitivity of isomorphisms it follows that the data sets associated to the two open coverings are isomorphic, so it suffices to show that X and X' are homeomorphic if the data sets are isomorphic.

For each α let $h_\alpha : U_\alpha \to U'_\alpha$ be the homeomorphism associated to the isomorphism of amalgamation data. We then have an corresponding homeomorphism:

$$\coprod_\alpha h_\alpha : \coprod_\alpha U_\alpha \longrightarrow \coprod_\alpha U'_\alpha$$

Let $\mathbf{p}$ and $\mathbf{p}'$ be the canonical quotient maps from $\coprod_\alpha U_\alpha$ and $\coprod_\alpha U'_\alpha$ to X and X' (respectively) as defined at the beginning of this writeup. By the commutativity relation from (b) in the definition of an isomorphism it follows that $\coprod_\alpha h_\alpha$ passes to a continuous map $h : X \to X'$ of these quotient spaces. We claim that h is a homeomorphism.

The continuity of h is already known, and the next step is to prove that h is onto. If $z \in X'$, choose α so that $z \in U'_\alpha$ (at least one exists because we have an open covering of X'). If $t \in Y'_\alpha$ maps to z under the quotient map $\mathbf{p}'$, then by construction we have that

$$z \quad = \quad h\left(j_\alpha\left(h_\alpha^{-1}(t)\right)\right)$$

showing that the arbitrary point z lies in the image of h.

To show that h is 1–1, it suffices to show that if $h(\mathbf{p}(a)) = h(\mathbf{p}(b))$ then $\mathbf{p}(a) = \mathbf{p}(b)$. Since h is the map of quotient spaces determined by $\coprod_\sigma h_\sigma$ we have the commutativity relation

$$h^\circ \mathbf{p} \quad = \quad \mathbf{p}'^\circ \left(\coprod_\sigma h_\sigma \right)$$

and thus if $a \in U_\alpha$ and $b \in U_\beta$ the hypothesis $h(\mathbf{p}(a)) = h(\mathbf{p}(b))$ can be rewritten as $\mathbf{p}'(h_\alpha(a)) = \mathbf{p}'(h_\beta(b))$. But this means that $\psi'_{\beta\alpha}(h_\alpha(a)) = h_\beta(b)$. On the other hand, by the commutativity relation in part (b) of the definition of an isomorphism we know that $\psi'_{\beta\alpha}{}^\circ h_\alpha = h_\beta{}^\circ \psi_{\beta\alpha}$, and this in turn implies that

$$h_\beta(b) \quad = \quad h_\beta\left(\psi_{\beta\alpha}(a)\right).$$

Since h_β is a homeomorphism this implies that $b = \psi_{\beta\alpha}(a)$, which means that $\mathbf{p}(b) = \mathbf{p}(a)$ and proves that h is 1–1.

The last step is to prove that h is open. This will be a special case of the following more general result:

LEMMA. *Let X and Y be topological spaces, let R and S be equivalence relations on X and Y respectively, let $p : X \to X/R$ and $q : Y \to Y/S$ be the corresponding quotient space projections, and suppose that f is a continuous map from X to Y that is $1-1$ onto and takes R-equivalent points in X to S-equivalent maps in Y. Denote the associated map of quotient spaces from X/R to Y/S by h. If f, p and q are open mappings then so is h.*

Proof. Suppose that U is open in X/R. Then $h(U)$ is open in Y/S if and only if $q^{-1}[h(U)]$ is open in Y. Since f is continuous, open and onto, it follows that $q^{-1}[h(U)]$ is open in Y if and only if

$$f^{-1}\left[q^{-1}[h(U)]\right] \quad = \quad p^{-1}\left[h^{-1}[h(U)]\right]$$

is open in X. Since h is 1–1 and onto it follows that $U = [h^{-1}[h(U)]$, and therefore the right hand side of the displayed equation is merely the set $p^{-1}[U]$, which is open by the continuity of p. It follows that the map h is open as asserted.■

As noted above, this completes the proof of the Realization Theorem.■

One recurrent question is whether the space constructed from amalgamation data is Hausdorff if all the pieces are. Perhaps the simplest examples yielding non-Hausdorff spaces are given by taking $Y_1 = Y_2 = \mathbf{R}^n$, $U_{21} = U_{12} = \mathbf{R}^n - \{0\}$, and $\varphi_{12} = \varphi_{21}$ to be the identity map. The space constructed from these data is the non-Hausdorff Forked Line that we first introduced in Section I.1.

In contrast, the next result essentially says that problems with pairs of subspaces are the only things can prevent the constructed space from being Hausdorff.

PROPOSITION. *Let $\mathbf{Y} = (\{Y_\alpha\}, \{\varphi_{\beta\alpha}\})$ be a set of topological amalgamation data, and let X be the space with an open covering $\mathcal{U}$ of X such that $\mathbf{Y}$ is isomorphic to the topological amalgamation data associated to $\mathcal{U}$. Then the space X is Hausdorff if and only if each Y_α is Hausdorff and for each β and γ the space*

$$X_{\beta\alpha} \quad = \quad Y_\alpha \cup_{\varphi_{\beta\alpha}: W_{\alpha\beta} \equiv W_{\beta\alpha}} Y_\beta$$

is Hausdorff.

Proof. The conditions are clearly necessary. To prove they are sufficient, let u and v be distinct points of X. If they both lie in some Y_α, then they have disjoint neighborhoods in Y_α because the latter is Hausdorff. If one lies in, say, Y_β and the other in Y_γ, then the points have disjoint neighborhoods in $X_{\beta\alpha}$ because this space is Hausdorff. In either case, two distinct points have disjoint neighborhoods as required.■

III.4.B : Smooth amalgamation

One can adapt much of the discussion in III.A to obtain a comparable theory for smooth manifolds.

III.4.B.1 : *Smooth structures on disjoint unions*

The first step is to give a smooth version of the disjoint union construction. This turns out to be extremely straightforward.

SMOOTH DISJOINT UNIONS. *Let $\{(M_\alpha, \mathcal{A}_\alpha)\}$ be a family of smooth manifolds, and let $\coprod_\alpha M_\alpha$ be their disjoint union. Then $\coprod_\alpha \mathcal{A}_\alpha$ defines a smooth atlas for $\Sigma = \coprod_\alpha M_\alpha$ such that*

(i) the injections $i_\alpha : M_\alpha \to \Sigma$ are smooth mappings,

(ii) if $(L, \mathcal{E}$ is a smooth manifold and $f : \Sigma \to L$ is continuous, then f is smooth if and only if each composite $f \circ i_\alpha$ is smooth.

In Section I.2 we made a standing hypothesis of second countability for the manifolds considered in this course. Since a disjoint union of of nonempty spaces is second countable only if the number of such spaces is $\leq \aleph_0$ (see page 3 of the file

`http://math.ucr.edu/∼res/math205A/solutions5.∗`

(where once again $* =$ `ps` or `pdf`) the standing hypothesis implies that the family $\{(M_\alpha, \mathcal{A}_\alpha)\}$ should be assumed to be countable.

Proof. The images of the charts in $\coprod_\alpha \mathcal{A}_\alpha$ form an open covering for $\coprod_\alpha M_\alpha$; to see that $\coprod_\alpha \mathcal{A}_\alpha$ determines a smooth atlas, note that the images of two charts (U, h) and (V, k) intersect nontrivially only if both belong to one of the subfamilies $\mathcal{A}_\ell$, and because of this all nontrivial transition maps (*i.e.*, those defined on nonempty open sets) will be smooth. Therefore $\coprod_\alpha \mathcal{A}_\alpha$ is a smooth atlas (however, as noted in the exercises it is usually not a maximal atlas).

Smoothness of the injection maps i_α may be established by noting that for each smooth chart (U, h) in $\coprod_\alpha \mathcal{A}_\alpha$ the local map "$(i_\alpha \circ h)^{-1} \circ h$" is equal to the id_U. To verify property (*ii*), first observe that the ($\Longrightarrow$) implication follows because composites of smooth maps are smooth. Conversely, if each of the composites $h \circ i_\alpha$ is smooth, then for all charts (U, h) in $\coprod_\alpha \mathcal{A}_\alpha$ and (V, k) in $\mathcal{E}$ satisfying $f \circ i_\alpha \circ h(U) \subset k(V)$ we know that "$k^{-1} \circ h \circ h$" is smooth; but this implies that h itself is also smooth.■

In Section III.A we gave a result called the Internal Sum Recognition Principle for recognizing spaces that are equivalent to disjoint unions; as noted there, similar situations arise naturally in linear algebra where it is often important to find internal direct sum structures on vector spaces. We would like to state and prove a corresponding recognition principle for smooth manifolds. Before doing so we shall a prove a preliminary result of independent interest.

LEMMA. *In the notation of the previous result, for each smooth manifold $(M_\beta, \mathcal{A}_\beta)$ in the collection $\{ (M_\alpha, \mathcal{A}_\alpha) \}$ the map i_β defines a diffeomorphism i'_β from M_β to $i_\beta(M_\beta)$.*

Proof. We already know that the associated map i'_β is a homeomorphism from M_β to $i_\beta(M_\beta)$, and the considerations of Section III.2 combine with conclusion (i) in the previous result to imply that i'_β is smooth. We shall prove that the inverse is smooth by presenting the inverse as a composite of smooth maps.

By the results for topological disjoint unions, one can define a continuous map $q : \coprod_\alpha \to M_\beta$ such that $q \circ i_\beta$ = identity and $q \circ i_\alpha$ = constant if $\alpha \neq \beta$. By construction each of the maps $q \circ i_\alpha$ is smooth, and therefore conclusion (ii) of the previous result implies that q is smooth. Direct computation then shows that the smooth map $q|i_\beta(M_\beta)$ is an inverse (hence **the** inverse) to $i_\beta(M_\beta)$.■

The lemma leads directly to the following simple criterion for recognizing smooth manifolds that are diffeomorphic to disjoint unions:

SMOOTH INTERNAL SUM RECOGNITION PRINCIPLE. *Suppose that a smooth manifold $(N, \mathcal{B})$ is a union of pairwise disjoint open subspaces M_α. Then N is homeomorphic to $\coprod_\alpha M_\alpha$.*

Proof. For each $\alpha \in A$ let $j_\alpha : M_\alpha \to N$ be the inclusion map. By the previous results for topological and smooth disjoint unions, there is a unique smooth function

$$J : \coprod_\alpha M_\alpha \longrightarrow N$$

such that $J \circ i_\alpha = j_\alpha$ for all α. We claim that J is a diffeomorphism; by the topological version of the internal sum recognition principle we know that J is a homeomorphism.

To show that J is a diffeomorphism, it only remains to show that J^{-1} is smooth. It will suffice to show that the restriction of J^{-1} to each open subset M_α is smooth. A direct examination of the definitions shows that $J^{-1}|M_\alpha$ is equal to i_α, which we know is smooth, and therefore it follows that J^{-1} must also be smooth.■

III.4.B.2 : *Constructing smooth manifolds out of pieces*

For smooth manifolds one can also give a condition for realizing the amalgamation data by a smooth atlas.

SMOOTH REALIZATION THEOREM. *In the setting of the Topological Realization Theorem above, suppose that the spaces Y_α are all open subsets of $\mathbf{R}^n$, the maps $\varphi_{\beta\alpha}$ are all diffeomorphisms, and the associated space X is Hausdorff and second countable. Then X is a second countable topological n-manifold, and it has a smooth atlas $\{(U_\alpha, h_\alpha)\}$ such that the transition maps "$h_\beta^{-1} h_\alpha$" are equal to $\varphi_{\beta\alpha}$ for all α and β.*

Notation. In the situation of this result, if $\mathcal{A}$ is the original set of amalgamation data then the corresponding smooth atlas on the constructed space X will be called the *associated smooth atlas* for the amalgamation data.

Proof of the theorem. We shall use the notation in the proof of the Topological Realization Theorem (*q.v.*). The space X is a topological manifold because it has an open covering consisting of topological manifolds (in fact, open subsets in $\mathbf{R}^n$). Consider the atlas consisting of the pairs (Y_α, k_α) described in the proof of the Topological Realization Theorem. The transition maps "$k_\beta^{-1} \circ k_\alpha$" for this atlas are equal to the diffeomorphisms $\varphi_{\beta\alpha}$ by construction.■

III.5 : Tangent spaces and vector bundles

(Conlon, §§ 3.3–3.4)

A basic idea underlying the theory of smooth manifolds is that such objects can be studied using a mixture of techniques from multivariable calculus and point set topology. We have already discussed some constructions for topological spaces for which there are similar constructions on smooth manifolds in at least some cases, including finite products, covering space projections, submanifolds, quotient constructions related to covering space projections and disjoint sums.

Despite these similarities, there are also clear differences between what one can do for topological spaces as opposed to smooth manifolds. In particular, there are numerous constructions on topological spaces that do not work at all for smooth manifolds, but on the other hand there are also some important constructions for smooth manifolds that cannot be carried out for topological spaces. The **tangent bundle** of a smooth manifold is a fundamental example of this sort.

III.5.1 : *Definitions and examples*

The definition of the tangent bundle requires some digressions, so it seems best to begin with a description of what we want. For an open subsubset U of $\mathbf{R}^n$ we defined the space of all tangent vectors to points of U to be the product $U \times \mathbf{R}^n$, the idea being that for each $x \in U$ the space $\{x\} \times \mathbf{R}^n$ can be viewed as a space of tangent vectors at x (or as a a physicist might say, vectors whose point of application is x). Similarly, if we are given a smooth n-manifold $(M, \mathcal{A})$ and a point p in M, we would like to describe a smooth manifold $T(M)$ such that for each $p \in M$ it contains an n-dimensional vector space $T_p(M)$ of *tangent vectors* to p in M, and such that $T(M)$ is the union of these vector spaces for all $p \in M$; for the record, we would also like these vector spaces to be pairwise disjoint. The space $T_p(M)$ should be defined so that its elements can be viewed as tangent vectors for smooth curves $\varphi : (-\varepsilon, \varepsilon)$ satisfying $\varphi(0) = p$; in other words, for each vector $\mathbf{v} \in T_p(M)$ one can find a φ of this sort so that it makes sense to say $\varphi'(0) = \mathbf{v}$.

If M is open in $\mathbf{R}^n$ our previous construction fulfills these requirements. As usual, the best test case for extending the definition is the standard 2-sphere in Euclidean 3-space.

There are two possible approaches, and they lead to the same answer. On one hand, in classical solid geometry one speaks about the tangent plane to a point on a sphere as the plane perpendicular to the radius at the point of contact. This is good for looking at a single tangent plane, but classical tangent planes generally intersect in a line and *we want our tangent planes at different points to* *be pairwise distinct*. One way of creating an object that fulfills this requirement and still leads to the classical notion of tangent plane is to view the tangent space for S^2 to be the set of all points $(x, v) \in S^2 \times \mathbf{R}^3$ such that $|x| = 1$ (*i.e.*, it lies on S^2) and y is perpendicular to x. The classical tangent plane to x will then be the set of all points of the form $x + y$ where y is perpendicular to x.

A second way of approaching this is through the following elementary result:

PROPOSITION. *Let $x \in S^2$ and $y \in \mathbf{R}^3$. Then there is a smooth curve $\varphi : (-\varepsilon, \varepsilon) \to S^2$ such that $\varphi(0) = x$ and $\varphi'(0) = y$ if and only if $\langle x, y\rangle = 0$, where $\langle\ ,\ \rangle$ denotes the usual inner product on $\mathbf{R}^3$.*

In fact, this all generalizes to level sets of regular values. If $f : \mathbf{R}^n \to \mathbf{R}^m$ is a smooth map (where $n > m$ as usual) and y is a nontrivial regular value of f, then the tangent space of level set $L = f^{-1}(\{y\})$ can be taken to be the set of all points $(u, \mathbf{v}) \in L \times \mathbf{R}^n$ such that $Df(u)\mathbf{v} = 0$. Since f is a regular value the dimension of the kernel of $Df(u)$ is $(n - m)$ for all $u \in \mathbf{R}^n$. The preceding proposition extends directly to such level sets with this definition of the tangent space. In particular, for the unit sphere we are looking at the set of all points where $f(x) = 1$, where $f(x) = |x|^2$, and in this case $Df(x)y = 2\langle x, y\rangle$.

By the Theorem on Level Sets in Section III.1, there is an atlas of smooth charts (U_α, h_α) for L such that each $j \circ h_\alpha$ is smooth. Suppose that $\in L$ is chosen so that $x \in h_\alpha(U_\alpha) \cap h_\beta(U_\beta)$, and let $\mathbf{v}$ be a vector in the kernel of $Df(x)$. Then one can use the coordinate charts to construct smooth curves $\Gamma_\alpha : (-\varepsilon, \varepsilon) \to U_\alpha$ and $\Gamma_\beta : (-\varepsilon, \varepsilon) \to U_\beta$ such that $h_\alpha \circ \Gamma_\alpha = h_\beta \circ \Gamma_\beta$, $h_\alpha(\Gamma_\alpha(0)) = h_\beta(\Gamma_\beta(0))$ and if $\Gamma : (-\varepsilon, \varepsilon) \to \mathbf{R}^n$ is the the associated smooth curve in Euclidean m-space then $\Gamma'(0) = \mathbf{v}$.

FUNDAMENTAL QUESTION. *What is the relationship between the tangent vectors $\Gamma'_\alpha(0)$ and $\Gamma'_\beta(0)$?*

ANSWER. By construction we have that Γ_β is equal to "$h_\beta^{-1} h_\alpha$" $\circ\, \Gamma_\alpha$, and therefore by the Chain Rule the tangent vector $\mathbf{w}$ at $u = \Gamma_\alpha(0)$ is identified with the tangent vector D"$h_\beta^{-1} h_\alpha$"$(u)\mathbf{w}$ at "$h_\beta^{-1} h_\alpha$"$(u) = \Gamma_\beta(0)$.■

All of these considerations are part of the following result:

THEOREM. *Let $n > m$ and let $f : \mathbf{R}^n \to \mathbf{R}^m$ be a smooth map such that y is a nontrivial regular value f (i.e., there is some x so that $f(x) = y$), and let $L = f^{-1}(\{y\})$. Then the tangent space to L, consisting of all $(x, y) \in L \times \mathbf{R}^n$ such that $Df(x)y = 0$, is a smooth manifold, and if $\mathcal{A} = \{(U_\alpha, h_\alpha)\}$ is a smooth atlas of the type described above, then there is a smooth atlas for the tangent space of L having the form $\{(U_\alpha \times \mathbf{R}^{n-m}, H_\alpha)\}$ where $H_\alpha(x, \mathbf{v}) = (h_\alpha(x), Dh_\alpha(x)\mathbf{v})$.*

The transition maps are smooth because they are given by the formula "$H_\beta^{-1} H_\alpha$"$(x, \mathbf{v}) =$ ("$h_\beta^{-1} h_\alpha$" $(x), D[$"$h_\beta^{-1} h_\alpha$"$](x)\mathbf{v})$.■

III.5.2 : *General construction for the tangent bundle*

Motivated by the level sets example, we would like to construct the tangent space of an arbitrary smooth manifold $(M, \mathcal{A})$ out of the following data:

For each chart (U_α, h_α) in the maximal atlas $\mathcal{A}$, define Y_α to be $U_\alpha \times \mathbf{R}^n$. Following standard practice we define $V_{\beta\alpha} \subset U_\alpha$ to be the open subset

$$h_\alpha^{-1}\big(\, h_\beta(U_\beta)\,\big)$$

and let $\psi_{\beta\alpha} : V_{\beta\alpha} \to V_{\alpha\beta}$ be the usual transition map "$h_\beta^{-1} \circ h_\alpha$" that is a diffeomorphism because $\mathcal{A}$ is a smooth atlas. We then take the open subset $W_{\beta\alpha}$ to be the product $V_{\beta\alpha} \times \mathbf{R}^n$ define mappings

$$\varphi_{\beta\alpha} : V_{\beta\alpha} \times \mathbf{R}^n \longrightarrow V_{\alpha\beta} \times \mathbf{R}^n$$

by the following formula:

$$\varphi_{\beta\alpha}(x, \mathbf{v}) \quad = \quad \big(\, \psi_{\beta\alpha}(x),\; D\psi_{\beta\alpha}(x)\mathbf{v}\,\big)$$

PROPOSITION. *The preceding data* $(\,\{Y_\alpha\},\, \{\,\varphi_{\beta\alpha}\,\}\,)$ *define a set of topological amalgamation data.*

Proof. In order to show that we have a set of amalgamation data it is necessary to

(i) verify that the maps $\varphi_{\beta\alpha}$ are homeomorphisms,

(ii) check that the cocycle formulas hold.

In fact, the first of these statements is implicitly contained in the second, so it $\varphi_{\beta\alpha} = \varphi_{\alpha\beta}^{-1}$), so it suffices to check that $\varphi_{\gamma,\gamma}$ = identity and "$\varphi_{\gamma\beta} \circ \varphi_{\beta\alpha}$'$= \varphi_{\gamma\alpha}$. These identities may be checked as follows:

(i) $\varphi_{\alpha\alpha}$ is the identity because $\psi_{\alpha\alpha}$ ="$h_\alpha^{-1}h_\alpha$" is the identity and the derivative of an identity map is just the identity matrix.

(ii) To see that $\varphi_{\beta\alpha}$ and $\varphi_{\alpha\beta}$ are inverse to each other, it suffices to calculate the composites explicitly using the fact that the inverse function identity, $\psi_{\beta\alpha}^{-1}$ =["$h_\beta^{-1}h_\alpha$"]$^{-1}$ equals $\psi_{\alpha\beta}$ ="$h_\alpha^{-1}h_\beta$", implies $D\psi_{\beta\alpha}(x)^{-1}$ =[D"$h_\beta^{-1}h_\alpha$"]$(x)^{-1}$ is equal to $D\psi_{\alpha\beta}(x)$ = D"$h_\alpha^{-1}h_\beta$"(x) for all x.

(iii) To see the composition relation, it suffices to calculate the composites explicitly using the fact that D"$h_\gamma^{-1}h_\alpha$" is the matrix product of equals D"$h_\gamma^{-1}h_\beta$" and D"$h_\beta^{-1}h_\alpha$" by the Chain Rule.

This completes the verification that we have a set of topological amalgamation data.∎

By the preceding result and the Topological realization theorems of Section III.4, there is a topological space $T(M)$ realizing the given data. If we choose n such that M is an n-manifold, it follows immediately that every point in $T(M)$ has an open neighborhood that is homeomorphic to an open subset of $\mathbf{R}^{2n}$. The transition maps $\varphi_{\beta\alpha}$ are all smooth (hence diffeomorphisms), and therefore $T(M)$ will be a topological manifold if it is Hausdorff. In addition to this, we really need to verify that $T(M)$ satisfies the standing hypothesis of second countability from Section I.2 (assuming that it holds for M itself!).

Our verification of the Hausdorff and second countability properties for $T(M)$ will depend upon the following fundamental result that is of considerable importance in its own right:

PROPOSITION. *In the setting described above, there is a continuous open surjection* τ_M : $T(M) \to M$ *such that for each smooth chart* (U_α, h_α) *in the maximal atlas* $\mathcal{A}$ *for* M *the, following conclusions hold:*

(*i*) *The inverse image* $\tau_M^{-1}(\,h_\alpha(U_\alpha)\,)$ *is homeomorphic to* $h_\alpha(U_\alpha)\times\mathbf{R}^n$.

(*ii*) *One can choose the homeomorphism* η *in* (*i*) *so that* $\tau_M\circ\eta(x,\mathbf{v})=x$ *for all* x *and* $\mathbf{V}$.

Proof. For each α in the indexing set for $\mathcal{A}$, define t_α on $Y_\alpha=U_\alpha\times\mathbf{R}^n$ by $t_\alpha(x,\mathbf{v})=h_\alpha(x)$. We claim that these maps fit together to define a continuous function on $T(M)$. This will hold if and only if the maps satisfy the following consistency condition with respect to the transition maps:

$$t_\alpha(x,\mathbf{v}) \quad = \quad t_\beta\big(\,\psi_{\beta\alpha}(x), D\psi_{\beta\alpha}(x)\mathbf{v}\,\big)$$

By construction the left hand side is equal to $h_\alpha(x)$ and the right hand side is equal to $h_\beta\,(\,\psi_{\beta\alpha}(x)\,)$. Since the latter is equal to $h_\alpha(x)$, it follows that the locally defined maps fit together to form a continuous map from $T(M)$ to M.

To see that the map τ_M is onto, note first that an arbitrary element of M is expressible as $h_\alpha(x)$ for some α and $x\in U_\alpha$; if $k_\alpha : Y_\alpha\to T(M)$ is the standard 1–1 continuous open map constructed in the realization theorem, then it follows immediately that $h_\alpha(x)=\tau_M\circ k_\alpha(x,\mathbf{0})$, so τ_M is onto. We shall next prove that τ_M is open. Since $\tau(\cup_\beta\, W_\beta)=\cup_\beta\,\tau_M(W_\beta)$ and the sets $k_\alpha(Y_\alpha)$ form an open covering of $T(M)$, it suffices to show that τ_M maps open subsets of $k_\alpha(Y_\alpha)$ to open subsets of $h_\alpha(U_\alpha)$ and since k_α is a homeomorphism onto the open subset $k_\alpha(Y_\alpha)$ it suffices to prove that t_α is open for each α. If $\pi : Y_\alpha = U_\alpha\times\mathbf{R}^n$ denotes projection onto the first coordinate then $t_\alpha=h_\alpha\circ\pi$. Both factors in this composite are open mappings, and therefore t_α is also open; it follows that τ_M is also open.

To prove conclusions (*i*) and (*ii*) it suffices to show that the inverse image of $h_\alpha(U_\alpha)$ is equal to $k_\alpha(Y_\alpha)$. By construction this set is contained in the inverse image. Suppose now that we are given some point $k_\beta(y,\mathbf{w})\in T(M)$ such that $\tau_M\circ k_\beta(y,\mathbf{w})\in h_\alpha(U_\alpha)$. By the definitions of the functions it follows that $h_\beta(y)\in h_\alpha(U_\alpha)$; the latter in turn implies that $\varphi_{\alpha\beta}(y,\mathbf{w})$ is defined and that

$$k_\beta(y,\mathbf{w})=k_\alpha\circ\varphi_{\alpha\beta}(y,\mathbf{w})$$

so that $k_\beta(y,\mathbf{w})$ lies in the image of k_α. It follows that $\tau_M^{-1}(h_\alpha(U_\alpha))$ is contained in $k_\alpha(Y_\alpha)$ and hence by the second sentence of this paragraph the two sets must be equal.∎

By construction the maps $\varphi_{\beta\alpha}$ are all diffeomorphisms, so we are reduced to showing two things; namely, the space $T(M)$ constructed from the preceding amalgamation data is second countable if M is, and it is always Hausdorff.

How does this help with proving that $T(M)$ is Hausdorff and second countable? It will suffice to combine the preceding observation with the following straightforward results in point set topology:

PROPOSITION. *Let* X *and* Y *be topological spaces, and let* $g : X\to Y$ *be a continuous map such that each point* $y\in Y$ *has an open neighborhood* V *for which* $g^{-1}(V)$ *is homeomorphic to a product* $V\times F$, *for some space* F, *by a homeomorphism* $h : V\times F\to g^{-1}(V)$ *satisfying* $g(h(v,z))=v$ *for all* v *and* z.

(A) *If* Y *and* F *are second countable then so is* X.

(B) *If* Y *and* F *are both Hausdorff then so is* X.

Sketch of Proof. (A) Since Y is second countable, there is a countable open covering $\{V_j\}$ where the open sets satisfy the local hypothesis. Each of the open subsets is also second countable, and a product of second countable sets is second countable, so X is a countable union of the second countable spaces $g^{-1}(V_j)$. But if a space can be expressed as a countable union of second countable open subsets, it must also be second countable (why?).

(B) Suppose that $x_1 \neq x_2$ in X. If $g(x_1) \neq g(x_2)$ then there are disjoint neighborhoods U_1 and U_2 of these image points in Y, and the inverse images $g^{-1}(U_1)$ and $g^{-1}(U_2)$ must be disjoint neighborhoods of x_1 and x_2. On the other hand, if $g(x_1) = g(x_2)$ let V be an open neighborhood of this point as described in the hypothesis of the theorem. The inverse image of this neighborhood is homeomorphic to $V \times F$ for some Hausdorff space F, and under this homeomorphism x_i corresponds to (v_i, z_i) where $v_1 = v_2$ but $z_1 \neq z_2$. Choose disjoint neighborhoods W_1 and W_2 for z_1 and z_2 in F such that $z_i \in W_i$ for $i = 1, 2$. Then the images $h(V \times W_1)$ and $h(V \times W_2)$ are open subsets of $g^{-1}(V)$ that are disjoint neighborhoods of x_1 and x_2 in X.■

COROLLARY. *The space $T(M)$ is a (second countable) smooth manifold and $\tau_M : T(M) \to M$ is a smooth map.*

Proof. We had reduced the proof that $T(M)$ was a second countable topological manifold to showing that it is Hausdorff and second countable. The preceding two propositions imply these facts. Since we had also shown that the transition maps for the amalgamation data are smooth, it follows that the amalgamation data yield a smooth atlas for $T(M)$.

The smoothness assertion follows because τ_M maps each set $k_\alpha(Y_\alpha)$ to $h_\alpha(U_\alpha)$ and the local map "$k_\alpha^{-1} \circ \tau_M \circ h_\alpha$" is just the projection map from $U_\alpha \times \mathbf{R}^n$ to U_α. Since this map is smooth, it follows that τ_M is smooth.■

We shall conclude this subsection with two remarks on atlases for $T(M)$.

The atlas we have constructed for $T(M)$ is **not** a maximal atlas for the tangent space. Consider the case $M = \mathbf{R}^n$. If we take $\mathcal{A}$ to be the atlas whose only chart is the identity map, then we see that $T(M) \cong M \times \mathbf{R}^n$ such that τ_M corresponds to projection onto the first factor (use the proposition). The charts in the standard atlas for $T(M)$ all map onto vertical strips of the form $W \times \mathbf{R}^n$, and of course there are many smooth charts on $T(M) \cong M \times \mathbf{R}^n \cong \mathbf{R}^{2n}$ that do not have this form.

In many situations the following observation on smooth atlases for $T(M)$ is useful:

PROPOSITION. *Let $(M, \mathcal{A})$ be a smooth manifold, and let $\mathcal{B}$ be a subatlas of $\mathcal{A}$. Given a smooth chart (U_α, h_α) in $\mathcal{A}$, let $(U_\alpha \times \mathbf{R}^n, k_\alpha)$ be the associated smooth chart for $T(M)$. Then the set $T(\mathcal{B})$ of all charts of the form $(U_\beta \times \mathbf{R}^n, k_\beta)$, where (U_β, h_β) belongs to $\mathcal{B}$, determines an equivalent smooth atlas for $T(M)$.*

Proof. Since $T(\mathcal{B})$ is contained in the standard smooth atlas, which we shall call $T(\mathcal{A})$, it suffices to show that the sets $k_\beta(U_\beta \times \mathbf{R}^n)$ form an open covering of $T(M)$. The proposition regarding the map τ_M provides a quick way of verifying the open covering assertion. Since $\mathcal{B}$ is an atlas for M, the sets $h_\beta(U_\beta)$ form an open covering of M; consequently, their inverse images with respect to τ_M form an open covering of $T(M)$. However, in the proof of the proposition on τ_M we have shown that $\tau^{-1}(h_\beta(U_\beta))$ is equal to $k_\beta(U_\beta \times \mathbf{R}^n)$, and thus we have shown that sets of the latter type form an open covering of $T(M)$.■

III.5.3 : *Vector space operations in the tangent bundle*

Now that we have constructed the tangent bundle, we need to show that it can be viewed as a union of n-dimensional vector spaces, with one for each point in the manifold; as noted previously, we would like this family of vector space to be continuously parametrized by the points of the manifold in some reasonable sense that must be defined.

Notation. If $x \in M$ where M is a smooth manifold, then $T_x(M)$ is defined to be the inverse image $\tau_M^{-1}(\{x\})$; this subspace is homeomorphic to $\mathbf{R}^n$ by construction and is called the **tangent space to** x **in** M, or more generically the **fiber** or x with respect to the map τ_M.

Formally, here is the structure that we want:

BASIC OBJECTS:

1. A parametrized zero map; specifically, a smooth map $z : T(M) \to T(M)$ such that $\tau_M \circ z = \tau_M$.

2. A parametrized scalar multiplication map; specifically, a smooth map $\mu : \mathbf{R} \times T(M) \to T(M)$ such that $\tau_M \circ \mu = \tau_M \circ \pi_{T(M)}$, where $\pi_{T(M)}$ denotes projection onto the $T(M)$ factor.

3. A smooth structure defined on the space $T(M) \times_M T(M)$, which is the inverse image of the diagonal $\Delta_M \subset M \times M$ under the squared projection map $\tau_M \times \tau_M$. If $\tau_2(M)$ denotes either of the maps

$$q \circ (\tau_m \times \tau_M) | \, T(M) \times_M T(M)$$

where q denotes projection onto the first or second factor (these are equal by the definition of $T(M) \times_M T(M)$!), then $\tau_2(M)$ is to be smooth with respect to this smooth structure.

4. A parametrized vector addition map; specifically, a smooth map $\Sigma : T(M) \times_M T(M) \to T(M)$ such that $\tau_M \circ \Sigma = \tau_2(M)$.

If z, μ and Σ are mappings as above, then it follows that z maps $T_x(M)$ to itself, μ maps $\mathbf{R} \times T_x(M)$ to $T_x(M)$, and Σ maps the fiber of x with respect to $\tau_2(M)$ — which is $T_x(M) \times T_x(M)$ — to $T_x(M)$. We shall denote the associated maps of fibers by z_x, μ_x and Σ respectively. With this notation we can state the final thing that we need fairly simply.

BASIC PROPERTY OF THESE OBJECTS: ∞. For each $x \in M$ the maps z_x, μ_x and Σ define an n-dimensional real vector space structure on $T_x(M)$, with Σ defining the vector addition, z_x defining the zero vector, and μ_x defining the scalar multiplication.

If U is open in $\mathbf{R}^n$ then we can do this very directly on $T(U) = U \times \mathbf{R}^n$ by simply taking the standard vector space operations that each set $\{x\} \times \mathbf{R}^n$ inherits from $\mathbf{R}^n$ with its usual vector space operations. One can also define vector space structures for the tangent spaces to points in an n-dimensional level set $L \subset \mathbf{R}^{m+n}$; in this case the tangent space to a point x in L is essentially an n-dimensional vector subspace of $\{x\} \times \mathbf{R}^{m+n}$. We shall proceed by using the first of these as a model, and later we shall see that our construction yields the vector space operations on the tangent spaces of level sets that we have described.

CONSTRUCTION OF THE ZERO VECTOR MAP. We define the map locally using charts and then prove that the definitions for different charts are compatible. Given a chart $(U\alpha \times \mathbf{R}^n, k_\alpha)$ for $T(M)$, define $z_\alpha : U_\alpha \times \mathbf{R}^n \to T(M)$ by the formula

$$z_\alpha(x, \mathbf{v}) \quad = \quad k_\alpha(x, \mathbf{0}) \; .$$

In order to show this yields a well-defined map on the tangent space we need to check that $z_\alpha \circ \varphi_{\alpha\beta} = z_\beta$ when the left hand side is defined. This is true by the following sequence of equations:

$$z_\alpha \circ \varphi_{\alpha\beta}(x, \mathbf{v}) \quad = \quad z_\alpha(\psi_{\alpha\beta}(x), D\psi_{\beta\alpha}(x)\mathbf{v}) \quad = \quad k_\alpha(\psi_{\alpha\beta}(x), \mathbf{0}) \quad =$$

$$k_\alpha(\psi_{\alpha\beta}(x), D\psi_{\beta\alpha}(x)\mathbf{0}) \quad = \quad k_\alpha \circ \varphi_{\beta\alpha}(x, \mathbf{0}) \quad = \quad k_\beta(x, \mathbf{0}) \quad = \quad z_\beta(x, \mathbf{v})$$

CONSTRUCTION OF THE SCALAR MULTIPLICATION MAP. In this case the local definition is

$$\mu_\alpha(t, x, \mathbf{v}) \quad = \quad k_\alpha(x, t\mathbf{v})$$

and the compatibility of these maps is true by a similar sequence of equations:

$$\mu_\alpha \circ [\mathrm{id}_{\mathbf{R}} \times \varphi_{\alpha\beta}]\,(t, x, \mathbf{v}) \quad = \quad \mu_\alpha(t,\, \psi_{\alpha\beta}(x),\, D\psi_{\beta\alpha}\mathbf{v}) \quad =$$

$$k_\alpha(\psi_{\alpha\beta}(x),\, t \cdot D\psi_{\beta\alpha}\mathbf{v})) \quad = \quad k_\alpha(\psi_{\alpha\beta}(x),\, D\psi_{\beta\alpha}(t\mathbf{v})\,) \quad =$$

$$k_\alpha \circ \varphi_{\alpha\beta}(x, t\mathbf{v}) \quad = \quad k_\beta(x, t\mathbf{v}) \quad = \quad \mu_\beta(t, x, \mathbf{v})$$

cONSTRUCTION OF A SMOOTH STRUCTURE ON $T(M) \times_M T(M)$. First of all, we note that each fiber $\tau_2(M)^{-1}(\mathrm{pt.}\}$ is homeomorphic to $\mathbf{R}^n \times \mathbf{R}^n$ and that the map $\tau_2(M)$ from $T(M) \times_M T(M)$ to $T(M)$ is continuous and open. One can prove that $T(M) \times_M T(M)$ is Hausdorff and second countable by the same sort of argument employed for $T(M)$; filling in the details is left to the reader as an exercise.

A smooth atlas may be defined as follows: Let $(U_\alpha \times \mathbf{R}^n, k_\alpha)$ be a coordinate chart for $T(M)$, note that the map

$$k_\alpha \times k_\alpha | \Delta_{U_\alpha} \times \mathbf{R}^n \times \mathbf{R}^n$$

is contained in $T(M) \times_M T(M)$, and define

$$\lambda_\alpha : U_\alpha \times \mathbf{R}^n \times \mathbf{R}^n \longrightarrow TM) \times_M T(M)$$

to be the map determined by $k_\alpha \times k_\alpha$ in this manner. This yields a smooth atlas because the transition maps

$$\Theta_{\beta\alpha} : V_{\beta\alpha} \times \mathbf{R}^n \times \mathbf{R}^n \longrightarrow V_{\alpha\beta} \times \mathbf{R}^n \times \mathbf{R}^n$$

are given by the formula

$$\Theta_{\beta\alpha}(x, \mathbf{v}, \mathbf{w}) \quad = \quad \left(\,\psi_{\beta\alpha}(x),\, D\psi_{\beta\alpha}(x)\mathbf{v},\, D\psi_{\beta\alpha}(x)\mathbf{w}\,\right)\,.$$

It follows from these definitions that the projection $\tau_2(M)$ is smooth.

CONSTRUCTION OF THE VECTOR ADDITION MAP. We now define addition by the formula

$$\Sigma_\alpha(x, \mathbf{v}, \mathbf{w}) \quad = \quad k_\alpha(x, \mathbf{v} + \mathbf{w}\,)\,.$$

Once again a lengthy computation is needed to prove the required consistency condition

$$\Sigma_\alpha\,(\,\Theta_{\alpha\beta}(x, \mathbf{v}, \mathbf{w})\,) \quad = \quad \Sigma_\beta(x, \mathbf{v}, \mathbf{w})$$

and as in the preceding two arguments the linearity of $D\psi_{\alpha\beta}(x)$ plays a crucial role in the verification. Details of this are left to the reader as an exercise.

In the preceding discussion we did not explicitly discuss the proofs of identities such as $\tau_m \circ z = \tau_m$ and the corresponding identities for μ and Σ. Once again it is left to the reader to verify that all these properties hold. Here is a hint in the case of the zero map: By construction the local maps z_α satisfy $\tau_M \circ z_\alpha(x, \mathbf{v}) = h_\alpha(x)$, and the same is also true for $\tau_M \circ k_\alpha(x, \mathbf{v})$.■

TANGENT BUNDLES FOR TOPOLOGICAL MANIFOLDS. (‡) Results of J. Milnor, J. M. Kister and B. Mazur from the nineteen sixties yield a partial generalization of the tangent bundle to arbitrary topological manifolds. More precisely, the topological tangent bundle for an n-manifold is a pair $(E, p : E \to M)$ such that E is a topological $2n$-manifold and p is a continuous map such that the following holds:

> Each $x \in M$ has an open neighborhood V such that V is an open subset of $\mathbf{R}^n$ and there is a homeomorphism $k : U \times \mathbf{R}^n \to p^{-1}(V)$ such that $p(k(x,y)) = x$ for all $(x,y) \in U \times \mathbf{R}^n$.

Further information on the construction of this object appears in the references cited below.

Note that there is no assumption about vector space operations on the fibers $p^{-1}(\{z\})$ where z runs through all the points of M. In fact, results from the previously cited book of Kirby and Siebenmann show that a manifold of dimension $\neq 4$ has a smooth structure if and only if one can impose reasonable continuously parametrized family of vector space structures on the fibers.

[x] J. M. Kister, *Microbundles are fibre bundles*, Ann. of Math. (2) **80** (1964), 190–199.

[x] J. W. Milnor, *Microbundles. I*, Topology **3** Suppl. 1 (1964), 53–80.

III.5.4 : *Naturality of the tangent bundle construction*

The aim of this subsection is to show that *the tangent space construction for smooth manifolds extends also yields a compatible construction for smooth maps of smooth manifolds.*

Here is a summary of the main construction:

THEOREM. *Let $f : M \to N$ be a smooth map of smooth manifolds (we suppress the atlases here to simplify the notation) where $\dim M = m$ and $\dim N = n$. Then there is a canonical smooth map $T(f) : T(M) \to T(N)$ such that the following hold:*

(i) For each $p \in M$, $T(f)$ sends $T_p(M)$ linearly to $T_{f(p)}(N)$.

(ii) If we have smooth charts (U_α, h_α) for M and (V_β, k_β) for N such that $f(h_\alpha(U_\alpha)) \subset k_\beta(V_\beta)$ and the maps for the associated charts in the tangent space atlases are denoted by H_α and K_β, then $T(f)$ maps $H_\alpha(U_\alpha \times \mathbf{R}^m)$ into $K_\beta(V_\beta \times \mathbf{R}^n)$ and "$K_\beta^{-1} \circ T(f) \circ H_\alpha$"$(x, \mathbf{v})$ is equal to ("$k_\beta^{-1} \circ f \circ h_\alpha$"$(x)$, D"$k_\beta^{-1} \circ f \circ h_\alpha$"$(x)\mathbf{v}$).

Proof. The second condition suggests that we define $T(f)$ on the image of a chart $H_\alpha(U_\alpha \times \mathbf{R}^n)$ by the given formula. This presupposes that f sends the image of h_α into the image of some chart for some atlas for N, but we know that we can find an atlas of charts for M with this property. If we let $f_1 : U_\alpha \to V_\beta$ be the local map determined by f — in other words, the map we have been describing as "$k_\beta^{-1} \circ f \circ h_\alpha$" most of the time — then we would like to say that

$$T(f) \circ H_\alpha(x, \mathbf{v}) \quad = \quad K_\beta(\, f_1(x),\, Df_1(x)\mathbf{v}\,) \ .$$

We need to show that this definition satisfies the basic consistency condition if we compare it with the corresponding formula for charts H_γ and K_δ (once again we assume that f sends the image of h_γ to the image of k_δ, and we denote the map corresponding to f by f_2. In terms of formulas, we need to show that if $(y, \mathbf{w}) = \varphi_{\gamma\alpha}(x, \mathbf{v})$, then

$$K_\beta\,(f_1(y),\, Df_1(y)\mathbf{w}) \quad = \quad K_\delta\,(f_2(x),\, Df_2(x)\mathbf{v}) \ .$$

The constructions of f_1 and f_2 from the original mapping f imply a consistency identity

$$f_1(\,\psi_{\gamma\alpha}(x)\,) \quad = \quad \psi_{\delta\beta}(\,f_2(x)\,)$$

whenever either side of the equation is defined. Direct calculation using this identity and the Chain Rule then yields the following sequence of equations:

$$K_\beta\,(f_1(y),\,Df_1(y)\mathbf{w}) \quad = \quad K_\beta\,(\,f_1(\psi_{\gamma\alpha}(x)),\,Df_1(\,\psi_{\gamma\alpha}(x)\,)\,[D\psi_{\gamma\alpha}(x)\mathbf{v}]\,) \quad =$$

$$K_\beta\,(\,f_1(\psi_{\gamma\alpha}(x)),\,D\,[f_1\circ\psi_{\gamma\alpha}(x)]\,\mathbf{v}) \quad = \quad K_\beta\,(\,\psi_{\beta\delta}(f_2(x)),\,D\,[\psi_{\delta\beta}\circ f_2(x)]\,\mathbf{v}) \quad =$$

$$K_\beta\,(\,\psi_{\delta\beta}(f_2(x)),\,D\psi_{\delta\beta}(f_2(x))\,[Df_2(x)\mathbf{v}]\,) \quad = \quad K_\beta\circ\varphi_{\delta\beta}\,(f_2(x),\,Df_2(x)\mathbf{v}\,)$$

By the defining construction for the tangent bundle, we know that the final expression is equal to $K_\delta\,(f_2(x),\,Df_2(x)\mathbf{v})$, and this completes the verification of the identity

$$K_\beta\,(f_1(y),\,Df_1(y)\mathbf{w}) \quad = \quad K_\delta\,(f_2(x),\,Df_2(x)\mathbf{v})$$

that we needed to conclude the existence of $T(f)$.■

In the language of category theory, the next result states that the constructions $M \longrightarrow T(M)$ and $f \longrightarrow T(f)$ define a **covariant functor** from the category of smooth manifolds to itself.

THEOREM. *The construction $f \to T(f)$ has the following properties:*

(*a*) $T(\mathrm{id}_M) = \mathrm{id}_{T(M)}$.

(*b*) *If $f : M \to N$ and $g : N \to P$ are smooth then $T(g\circ f) = T(g)\circ T(f)$.*

Proof. We shall first verify (*a*). — The definition of $T(\mathrm{id}_M)$ implies that if one takes a typical chart of the form $(U_\alpha \times \mathbf{R}^n, k_\alpha)$ then $T(\mathrm{id}_M)\circ k_\alpha(x,\mathbf{v}) = k_\alpha(x, D\mathrm{id}_\alpha(x)\mathbf{v})$, and since the derivative of an identity map is always the identity, it follows that the right hand side is equal to $k_\alpha(x,\mathbf{V})$. It follows that the restriction of $T(\mathrm{id}_M)$ to each open set of the form $k_\alpha(U_\alpha \times \mathbf{R}^n)$ is equal to the corresponding restriction of the identity map on $T(M)$. Since open sets of the form $k_\alpha(U_\alpha \times \mathbf{R}^n)$ form an open covering for $T(M)$, it follows that $T(\mathrm{id}_M)$ must be equal to $\mathrm{id}_{T(M)}$.

We shall now verify (*b*). — By construction, $T(f)$ and $T(g)$ are given as follows: First, one finds typical charts $(U_\alpha \times \mathbf{R}^n, k^M_\alpha)$, $(V_\beta \times \mathbf{R}^q, k^N_\beta)$, and $(W_\gamma \times \mathbf{R}^s, k^P_\gamma)$ for M, N and P respectively such that

(1) f maps the image of U_α in M to the image of V_β in N,

(2) g maps the image of V_β in N to the image of W_γ in P.

We shal denote the smooth maps from U_α to V_β and V_β to W_γ corresponding to f and g by f_1 and g_1 respectively. Then $T(f)$ and $T(g)$ are uniquely defined by the following identities:

$$T(f)\circ k^M_\alpha(x,\mathbf{v}) \quad = \quad k^N_\beta(\,f_1(x),\,Df_1(x)\mathbf{v}\,)$$

$$T(g)\circ k^N_\beta(y,\mathbf{w}) \quad = \quad k^P_\beta(\,g_1(y),\,Dg_1(y)\mathbf{w}\,)$$

Direct calculation then yields the following identity characterizing $T(g)\circ T(f)$:

$$T(g)\circ T(f)\circ k^M_\alpha(x,\mathbf{v}) \quad = \quad k^P_\gamma(\,g_1\circ f_1(x),\,[Dg_1(f_1(x))\cdot Df_1(x)]\,\mathbf{v}\,)$$

On the other hand, in the setting of the previous paragraph we also know that $g \circ f$ maps the image of U_α to the image of W_γ, and in fact the corresponding map from U_α to W_γ is just $g_1 \circ f_1$. Therefore the map $T(g \circ f)$ is uniquely defined by the following identity:

$$T(g \circ f) \circ k_\alpha^M(x, \mathbf{v}) \quad = \quad k_\gamma^P(\, g_1 \circ f_1(x),\, D\,[g_1 \circ f_1]\,(x)\mathbf{v}\,)$$

We now compare the final expressions in the two equations at the ends of the preceding paragraphs. The first coordinates are equal by construction, and the second are equal because the Chain Rule implies that

$$D\,[g_1 \circ f_1]\,(x) \quad = \quad [Dg_1(f_1(x)) \cdot Df_1(x)]$$

and therefore the restrictions of $T(g) \circ T(f)$ and $T(g \circ f)$ are equal on the set $k_\alpha^M(U_\alpha \times \mathbf{R}^n)$. Since these sets form an open covering for $T(M)$, it follows that $T(g \circ f) = T(g) \circ T(f)$ as required.■

Given a smooth map $f : M \to N$ and $p \in M$ it is often convenient to use $T_p(f)$ to denote the associated linear map from $T_p(M)$ to $T_{f(p)}(N)$.

III.5.5 : *Useful descriptions of some tangent spaces* $(\star)$

It is often useful to have simplified descriptions of tangent bundles when working with specific examples or abstract constructions. Here are some basic identities that arise fairly often in the subject.

THEOREM. *We have the following isomorphisms:*

(i) $T(\mathbf{R}^n) \cong \mathbf{R}^n \times \mathbf{R}^n$ *such that* τ *corresponds to projection onto the first factor and the vector space operations on* $\{pt.\} \times \mathbf{R}^n$ *are given by the standard* $1-1$ *correspondence between the latter and* $\mathbf{R}^n$.

(ii) If M *and* N *are smooth manifolds, then* $T(M \times N) \cong T(M) \times T(N)$ *such that* $\tau_{M \times N}$ *correspond to* $\tau_m \times \tau_N$.

(iii) If P *is a smooth manifold and* V *is an open subset of* P, *then* $T(V) \cong \tau_M^{-1}(V)$ *such that* τ_V *corresponds to* $\tau_M|T(V)$.

Note that the first and the third have the following consequence:

COROLLARY. *If* U *is an open subset of* $\mathbf{R}^n$, *then* $T(U) \cong U \times \mathbf{R}^n$ *such that* τ *corresponds to projection onto the first factor and the vector space operations on* $\{pt.\} \times \mathbf{R}^n$ *are given by the standard* $1-1$ *correspondence between the latter and* $\mathbf{R}^n$.■

Proofs of these identities are left to the exercises for this section.■

III.6 : Regular mappings and submanifolds

(Conlon, §§ 1.5, 2.5, 3.7)

This section has two related goals. The first is to formulate a general concept of smooth submanifold generalizing the two previously considered special cases: Open subsets of smooth manifolds and (regular) level sets of smooth functions from an open subset of some Euclidean space to a Euclidean space of lower dimension. The second goal is to generalize the notions of immersion and submersion from Section II.2 to arbitrary smooth maps. These are related by a simple idea: If M is a smooth submanifold of N, then M should of course be homeomorphic to a topological subspace of N, but in addition *the tangent space* $T(N)$ *should be homeomorphic to a topological subspace of* $T(N)$. Formally, the two themes are linked via the concept of smooth embedding. This will turn out to be a map $f : M \to N$ such that $f(M)$ is a smooth submanifold of N and f defines a diffeomorphism from M onto $f(M)$.

III.6.1 : *Immersions and submersions*

We have already discussed the mappings in the title for smooth maps of open subsets in Euclidean spaces. The tangent bundle construction allows us to define similar concepts for arbitrary smooth manifolds in a very direct and simple fashion.

Definition. If $f : M \to N$ is a smooth map of smooth manifolds (suppressing the atlases for notational simplicity), then f is said to be a (smooth) **submersion** at $x \in M$ if the linear map $T_p(f)$ is onto.

Definition. If $f : M \to N$ is a smooth map of smooth manifolds, then f is said to be an **immersion** (more correctly, a smooth immersion) at $x \in M$ if the linear map $T_p(f)$ is 1–1.

Following standard practice we shall simply say that f is a submersion or immersion if it is a submersion or immersion at x for each $x \in M$.

If V and W are finite dimensional vector spaces over some field, then elementary considerations from linear algebra imply that the rank of a linear transformation from V to W is less than or equal to the minimum of $\dim V$ and $\dim W$, and therefore we have the following elementary observations regarding immersions and submersions:

(1) *If there is an immersion from the smooth manifold* M *to the smooth manifold* N*, then* $\dim M \leq \dim N$.

(2) *If there is a submersion from the smooth manifold* M *to the smooth manifold* N*, then* $\dim M \geq \dim N$.

If the dimensions of M and N are equal, then one might have either type of map for a given choice of M and N, and in fact a smooth map between two manifolds of the same dimension is an immersion if and only if it is a submersion. In particular, *the identity map from a smooth manifold to itself is both an immersion and a submersion.*

As noted in the exercises, if M and N have the same dimension and $f : M \to N$ is an immersion/submersion, then f is an open mapping. On the other hand, if n is a positive integer that is greater than one, then the complex analytic map $f(x) = z^n$ from the field $\mathbf{C}$ of complex numbers to itself is open but it is not an immersion/submersion because the derivative vanishes when $z = 0$ (more generally, a complex analytic map from an open subset of $\mathbf{C}$ to $\mathbf{C}$ is open if it is not constant — a proof appears on pages 214–217 of Big Rudin).

Here are some elementary properties of immersions and submersions that are frequently very useful. Proofs of these results are left to the exercises for this section.

PROPOSITION. (i) *Let* $f : M \to N$ *be a smooth homoeomorphism of smooth manifolds. Then* f *is a diffeomorphism if and only if* f *is an immersion.*

(ii) Let $f : M \to N$ be a smooth homoeomorphism of smooth manifolds. Then f is a diffeomorphism if and only if f is a submersion.

(iii) Let $f : M \to N$ and $g : N \to P$ be smooth mappings of smooth manifolds. If f and g are immersions, then so is their composite $g \circ f$.

(iv) Let $f : M \to N$ and $g : N \to P$ be smooth mappings of smooth manifolds. If f and g are submersions, then so is their composite $g \circ f$.■

STRAIGHTENING OF IMMERSIONS AND SUBMERSIONS. Local characterizations for submersions and immersions were previously given when M and N were open subsets of Euclidean spaces, and it follows immediately that these also hold if M and N are arbitrary smooth manifolds.

IMMERSION STRAIGHTENING PROPOSITION, GLOBAL VERSION. *Let $f : M \to N$ be a smooth map of smooth manifolds. Then f is an immersion at $x \in M$ if and only if one can find smooth charts (U_0, h_0) and $(U_0 \times N_1(0; \mathbf{R}^{n-m}), k)$ at x and $f(x)$ respectively such that*

$$f \circ h(y) \quad = \quad k(y, 0)$$

for all $y \in U_0$.■

Important note on terminology. Conlon defines a topological immersion in Definition I.5.1 on page 20 to be a map that is locally 1–1. This is highly nonstandard; usually an immersion of topological manifolds is defined to be a a continuous map such that the conclusion of the Immersion Straightening Proposition holds for suitably chosen *continuous* coordinate charts. In the mathematical literature, a locally 1–1 map that does not satisfy such a hypothesis is generally called a *non-locally-flat immersion*.

SUBMERSION STRAIGHTENING PROPOSITION, GLOBAL VERSION. *Let $f : M \to N$ be a smooth map of smooth manifolds. Then f is a submersion at $x \in M$ if and only if one can find smooth charts (U_0, k) and $(N_1(0; \mathbf{R}^{m-n}) \times U_0, k)$ at $f(x)$ and x respectively such that*

$$f(\, h(y, z)\,) \quad = \quad k(z)$$

for all $(y, z) \in N_1(0; \mathbf{R}^{m-n}) \times U_0$.■

EXAMPLES OF IMMERSIONS.

Example 0. If U is an open subset of a smooth manifold M, then the inclusion is automatically an immersion.

Immersions are locally 1–1 by the results on straightening immersions, but they are not necessarily 1–1 globally. The next two examples illustrate this point quite clearly. One of the exercises contains a proof that an immersion is never onto if the dimension of the domain is strictly less than the dimension of the codomain.

Example 1. This actually yields a large class of examples. A smooth covering space projection is an immersion but it is globally 1–1 if and only if it is a diffeomorphism.■

Example 2. Here is another important, but quite different, example of an immersion that is not 1–1 globally: The **strophoid** is a classical plane curve that may be defined in any of the following equivalent manners:

In cartesian coordinates by the following equation:

$$y^2 \;=\; x^2(a-x)/(a+x)$$

In polar coordinates by the following equation:

$$r \;=\; a\cos(2\theta)/\cos(\theta)$$

Parametrically by the following equations:

$$\mathbf{x}(\theta) \;=\; \big(\,a\,\cos(2\theta),\, a\,\cos(2\theta)\tan(\theta)\,\big)\;, \qquad \text{where} \;\; -\frac{\pi}{2} < \theta < \frac{\pi}{2}$$

Graphs of this curve appear in the files `strophoid.`$*$ in the course directory, where $*$ is one of `gif`, `doc`, `ps` or `pdf`.

It is elementary to verify that

(*i*) $\mathbf{x}'(\theta) \neq \mathbf{0}$ for all θ,

(*ii*) $\mathbf{x}$ is 1–1 on the complement of the set $\{\,-\frac{\pi}{4},\, \frac{\pi}{4},\,\}$,

(*iii*) $\mathbf{x}(-\frac{\pi}{4}) = \mathbf{x}'(\frac{\pi}{4})$,

(*iv*) $\mathbf{x}'(-\frac{\pi}{4})$ and $\mathbf{x}'(\frac{\pi}{4})$ are linearly independent and hence form a basis for $\mathbf{R}^2$.

Properties (*ii*) – (*iv*) reflect important general phenomena in the theory of immersions. One can summarize (*ii*) and (*iii*) as saying that $\mathbf{x}$ *has no triple points* (*i.e.*, no distinct triples of parameter values t_1, t_2, t_3 such that $\mathbf{x}(t_1) = \mathbf{x}(t_2) = \mathbf{x}(t_3)$) — or quadruple points or quintuple points and so on — but $\mathbf{x}$ does have exactly one *double point*: Namely, $\mathbf{0} = \mathbf{x}(-\frac{\pi}{4}) = \mathbf{x}(\frac{\pi}{4})$. Multiple points are sometimes called *self-intersections* of a parametrized curve. The final condition (*iv*) is often summarized qualitatively by saying that the self-intersection is *transverse*; *i.e.*, if $t_1,\ \cdot\ , t_r$ are such that $\mathbf{x}(t_i) = \mathbf{x}(t_1)$, then the vectors $\mathbf{x}'(t_i)$ are linearly independent.

Here are some online references for more (historical and mathematical) information about strophoids:

`http://www.2dcurves.com/cubic/cubicst.html`

`http://www-groups.dcs.st-and.ac.uk/~history/Curves/Right.html`

`http://mathworld.wolfram.com/Strophoid.html`

`http://astron.berkeley.edu/~jrg/ay202/node192.html`

Many other standard plane and space curves yield instructive examples along the same lines.■

Example 3. Suppose that $f : M \to N$ is a smooth map of smooth manifolds, and define the *graph map* of f to be the smooth map $\mathbf{G}_f : M \to M \times N$ by the formula $\mathbf{G}_f(x) = (x, f(x))$. It follows immediately that $\mathbf{G}_f$ is 1–1. Here is a proof that it is an immersion: Let $\pi_M : M \times N \to M$ denote the projection map. Then $\pi_m \circ \mathbf{G}_f = \mathrm{id}_M$ and therefore

$$[T_{(x,f(x))}\,(\pi_M)] \circ T_x(\mathbf{G}_f) \;=\; \mathrm{identity}[T_x(M)]\;.$$

Now a linear transformation $T : V \to W$ is automatically 1–1 if there is another linear transformation $S : W \to V$ such that $S \circ T = \mathrm{id}_V$ (prove this!), and if we specialize this to the previous sentence we find that $T_x\,(\mathbf{G}_f)$ is1–1 for all $x \in M$. Therefore $\mathbf{G}_f$ is an immersion.■

Still further examples of immersions appear in the subsections on submanifolds and embeddings.

IMMERSIONS, MULTIVARIABLE CALCULUS AND ELEMENTARY DIFFERENTIAL GEOMETRY. In general, the "good" parametric equations for regular smooth curves in calculus and differential geometry correspond to smooth immersions from an interval in $\mathbf{R}$ to some Euclidean space $\mathbf{R}^n$. Similarly, the "good" parametrizations for (pieces of) surfaces in multivariable calculus and undergraduate differential geometry correspond to smooth immersions from open subsets in $\mathbf{R}^2$ to $\mathbf{R}^3$; one of the exercises for this section contains a more detailed discussion of this fact.

EXAMPLES OF SUBMERSIONS.

Example 0. If U is an open subset of a smooth manifold M, then the inclusion is automatically an immersion.

Submersions are locally onto by the results on straightening immersions, but they cannot even be locally 1–1 unless the dimensions of the domain and codomain are equal.

Example 1. As before, a smooth covering space projection is a submersion.

Example 2. If M and N are smooth manifolds, then the projections from $M \times N$ to either M or N are smooth submersions. Here is a proof for projection onto M; the proof in the other case is essentially identical. Given a point $(x, y) \in M \times N$, let $j_y : M \to M \times N$ be the map $j_y(z) = (z, y)$. Then the composite $p_M \circ j_y = \mathrm{id}_M$, and therefore it follows that the identity on $T_x(M)$ is the composite $[T_{(x,y)}(p_M)] \circ [T_x(j_y)]$. It follows that the linear map $T_{(x,y)}\,(p_M)$ is onto, and therefore p_M is a smooth submersion.

Example 3. If M is an arbitrary smooth manifold, then the tangent bundle projection $\tau_M : T(M) \to M$ is a smooth submersion. In fact, it is an example of an important object known as a **fiber bundle**.

Definition. Let $p : E \to B$ be a continuous map of topological spaces. Then p is said to be a **topological fiber bundle projection** if for each $x \in B$ there is an open neighborhood U of x and a homeomorphism $h : U \times F \to p^{-1}(U)$ such that

$$p\big(\,h(x,y)\,\big) \quad = \quad x$$

for all $(x, y) \in U \times F$.

The space F may be viewed as the inverse image of an arbitrary point in U, and it is called a *fiber* of the map; observe that if B is connected and locally connected, then the fibers $f^{-1}(\{y\})$ are homeomorphic to each other for all $y \in B$. Verification of this is left as an exercise.

If in addition we know that

(i) the map p is a smooth map of smooth manifolds,

(ii) the fiber F is a smooth manifold,

(iii) the homeomorphism h can be chosen to be a diffeomorphism,

then we say that p is a *smooth fiber bundle projection.* Every map of this form is a smooth submersion. If B is a connected manifold, then for all $y \in B$ the fibers $f^{-1}(\{y\})$ are all diffeomorphic to each other.

Example 4. This is a variant of Example 3. If M is an arbitrary smooth manifold and $\tau_2(M) : T(M) \times_M T(M) \to M$ is given as in the Section III.5, then $\tau_2(M)$ is also a smooth submersion.

In Units V and VI of these notes we shall work with a large number of smooth fiber bundles similar to τ_M and $\tau_2(M)$.

Example 5. Here is a slightly different class of smooth submersions involving a construction from the exercises for Section III.2. Given a space X and a homeomorphism $f : X \to X$, the *mapping torus of* X may be defined to be the quotient space X_f of $X \times [0,1]$ in which $(x, 0)$ is identified with $(f(x), 1)$ for all $x \in X$. If we take $X = (-1, 1)$ and $f(x) = -x$, then this construction yields the Möbius strip with its edge removed, and if we take $X = S^1$ and f to be complex conjugation then this construction yields the Klein bottle; if X is arbitrary and f is the identity, then the mapping torus is canonically homeomorphic to $X \times S^1$. In all cases, projection onto the first coordinate yields a well-defined map onto $S^1 \cong [0,1]_{\mathrm{id}}$.

By the results in the exercises to Section III.2, if X is a smooth manifold and f is a diffeomorphism, then X has a smooth structure and the canonical map into S^1 is smooth. In fact, this map is a smooth fiber bundle projection whose fiber is given by X itself.

The following book is the classic reference on fiber bundles:

[1] N. Steenrod, The Topology of Fibre Bundles. Princeton Mathematical Series, vol. 14. Princeton University Press, Princeton, N. J., 1951. [*Reprint of the 1957 edition.* Princeton Landmarks in Mathematics. Princeton Paperbacks. Princeton University Press, Princeton, NJ, 1999. viii+229 pp. ISBN: 0-691-00548-6.]

Here are a few additional background references:

[2] H. Cartan and S. Eilenberg, *Foundations of fibre bundles*, Symposium internacional de toplogía algebraica (International symposium on algebraic topology), pp. 16–23, Universidad Nacional Autónoma de México and UNESCO, Mexico City, 1958.

[3] S. Eilenberg, Foundations of fibre bundles, multicopied lecture notes, University of Chicago, 1957 [available from various libraries in the UC system].

`http://math.vassar.edu/faculty/McCleary/history.fibre.spaces.pdf`

Example 6. Not every smooth submersion is a smooth fiber bundle projection, an inclusion of an open subset or a covering space projection. In particular, a surjective mapping $f : \mathbf{R}^2 \to S^2$ of this sort is described in the file(s) `yarnball.*` in the course directory.

We have already noted that a submersion is locally surjective and open. By construction, smooth fiber bundle projections are globally surjective submersions, but of course inclusions of open subsets will never be globally surjective unless the domain and codomain are equal. On the other hand, there is one simple sufficient condition for global surjectivity.

PROPOSITION. *Suppose that $f : M \to N$ is a smooth submersion, and assume further that M is compact and N is connected. Then f is onto.*

In contrast to the results for submersions, a smooth immersion is **never** onto if the dimension of the domain is strictly less than the dimension of the codomain. In fact, such an image always has an empty interior. Proving this requires the notion of *subsets of measure zero* for smooth manifolds; details on all this appear in the exercises for this section.

Proof. This is extremely straightforward. Surjections have open images, so $f(M)$ is open. Since M is compact, so is $f(M)$. Therefore $f(M)$ is a nonempty open and closed subset of N, so by connectedness we must have $f(M) = n$.■

In fact, the hypotheses yield a much stronger conclusion. A basic result of Ch. Ehresmann implies that if a smooth submersion is **proper** (inverse images of compact subsets are compact), then the map is a smooth fiber bundle projection. Since a continuous map from a compact Hausdorff space is always proper, this implies in particular that *a smooth submersion from a compact manifold to a connected manifold is a smooth fiber bundle projection.* — A proof of Ehresmann's result appears in the file(s) `ehresmann.*` in the course directory.

Smooth submersions and motions of robot arms (‡) We have already noted a relation between mechanical properties of robot arms and questions about smooth manifolds; there is some further discussion of this on pages 9–13 of the file `dundasnotes.pdf` in the course directory. Let us say that a configuration of the arm is "nice" if each joint of the arm can move freely in every direction, so that the arm does not get stuck and can only move in a limited range of directions. One can then view the end point of the robot arm as a smooth function of all "nice" configuration, and the possibility of free movement in every direction implies that this map should be a smooth submersion. By construction robot arms are fairly rigid, and because of this some problems about the motion states of robot arms can be translated into questions about the existence of smooth submersions on compact manifolds. The following papers discuss some implications of Ehresmann's result for some of these mechanical problems:

[1] D. H. Gottlieb, *Robots and fibre bundles*, Bull. Soc. Math. Belgique Sér. A **38** (1987), 219-223.

[2] D. H. Gottlieb, *Topology and the Robot Arm*, Acta Appl. Math. **11** (1988), 117-121.

Factoring smooth maps into composites of mersions. If G and H are groups and $f : G \to H$ is a group homomorphism, then one can factor f into a composite of a surjective homomorphism and an injective homomorphism in two separate ways. The first and most naive is to view f as the composite of the projection $f_1 : G \to f(G)$ with the inclusion $f(G) \subset H$. Another way of doing this is to view f as the composite of the injective homomorphism $\Gamma_f : G \to G \times H$, which is given by the graph of f, with the surjective coordinate projection from $G \times H$ onto H. There is no analog of the first construction for smooth maps of smooth manifolds, one reason being that the image of a smooth map is not necessarily a smooth manifold. However, there is a strong analog of the second construction.

PROPOSITION. *If $f : M \to N$ is a smooth map of smooth manifolds, then f is a composite $g \circ h$, where h is a smooth $1-1$ immersion and g is a smooth fiber bundle projection.*

Proof. Let $h = \mathbf{G}_f$ and $g = \pi_N$, where $\mathbf{G}_f$ is the graph map of f that was defined in Example 3 for immersions and π_N is projection from $M \times N$ to N. The discussion for the example shows that the first map is a 1–1 immersion, and by definition the second is a trivial example of a smooth fiber bundle projection.■

III.6.2 : *Existence and classification of mersions* (2⋆)

Although it is easy to ask whether there is a smooth immersion and submersion from one smooth manifold to another, it is not immediately clear whether such questions can be answered in any reasonable fashion, and particularly whether this can be done using methods from topology. Several breakthrough results from the nineteen fifties and sixties gave very strong positive answers to suhc questions.

Before stating the results, it is useful to formulate a suitable notion of *regular homotopy* for immersions or submersions (which we shall denote generically by the neutral term **mersions**). Given two smooth manifolds M and N, a **regular homotopy** of mersions is a continuous homotopy $H : M \times [0,1] \to N$ such that the following conditions hold:

(i) The restriction of H to $M \times (0,1)$ is smooth.

(ii) If, as usual, $h_t : M \to N$ denotes the map determined by $H|M \times \{t\}$, then there is a $\delta > 0$ such that $h_t = h_0$ for $t < \delta$ and $h_t = h_1$ for $t > 1 - \delta$.

(iii) Each h_t is a smooth mersion.

With this terminology the main results can be stated quite simply. The result for immersions was first established by M. W. Hirsch when $\dim M < \dim N$ and by V. Poenaru when $\dim M = \dim V$. The result for submersions is due to A. Phillips. References are given below.

CLASSIFICATION OF SMOOTH IMMERSIONS. *Let M and N be connected topological manifolds such that either $\dim M < \dim N$ or $\dim M = \dim N$ and M is* **not** *compact. Then regular homotopy classes of immersions correspond bijectively to homotopy classes of continuous maps $F : T(M) \to T(N)$ satisfying the following conditions:*

(i) *There is a continuous map $f : M \to N$ such that $\tau_N \circ F = f \circ \tau_M$.*

(ii) *For each $x \in M$, the induced map F_x from $T_x(M)$ to $T_{f(x)}(N)$ is a linear injection.*

In this context, it is assumed that if $H : T(M) \times [0,1] \to T(N)$ is a homotopy, then for each value of t the map F_t satisfies the given conditions and that there is an associated homotopy of continuous maps from $M \times [0,1]$ to N. The bijective correspondence is defined by sending an immersion g to the tangent space map $T(g)$.■

In particular, if $\dim M < \dim N$, then there is a smooth immersion from M to N if and only if there is a map $F : T(M) \to T(N)$ as described in the theorem.■

CLASSIFICATION OF SMOOTH SUBMERSIONS. *Let M and N be connected topological manifolds such that either $\dim M \geq \dim N$ and M is* **not** *compact. Then regular homotopy classes of submersions correspond bijectively to homotopy classes of continuous maps $F : T(M) \to T(N)$ satisfying the following conditions:*

(i) *There is a continuous map $f : M \to N$ such that $\tau_N \circ F = f \circ \tau_M$.*

(ii) *For each $x \in M$, the induced map F_x from $T_x(M)$ to $T_{f(x)}(N)$ is a linear surjection.*

In this context, it is assumed that if $H : T(M) \times [0,1] \to T(N)$ is a homotopy, then for each value of t the map F_t satisfies the given conditions and that there is an associated homotopy of continuous maps from $M \times [0,1]$ to N. The bijective correspondence is defined by sending a submersion g to the tangent space map $T(g)$.■

Neither result extends to cases where $\dim M \geq \dim N$ and M is compact; one can use the previously cited result of Ehresmann (proper submersion $\Longrightarrow$ smooth fiber bundle) to construct explicit counterexamples.

Simple examples. Let $n \geq m$, and let $p : S^n \times S^1 \to S^m$ be a constant map. Then it is possible to construct a map $\Psi : T(S^n \times S^1) \to T(S^m)$ such that for each $x \in S^n \times S^1$ the map $\Psi|T_x(S^n \times S^1)$ maps the domain to $T_{p(x)}(S^m)$ by a linear isomorphism. However, the map p itself is not homotopic to a smooth submersion, for there are several different arguments which show that p is not homotopic to a smooth fiber bundle projection.

A proof of this fact and several general methods for constructing counterexamples are described in the file(s) `nonfiberingresults.tex` in the course directory.

Here are references for the material discussed above:

[1] A. Haefliger, and V. Poenaru, *La classification des immersions combinatoires*, Inst. Hautes Études Sci. Publ. Math. **23** (1964), 75–91.

[2] M. W. Hirsch, *Immersions of manifolds*, Trans. Amer. Math. Soc. **93** (1959), 242–276.

[3] A. V. Phillips, *Submersions of open manifolds*, Topology **6** (1967), 171–206.

[4] D. Spring, *The golden age of immersion theory in topology: 1959-1973. A mathematical survey from a historical perspective*, Bull. Amer. Math. Soc., to appear.

The last item is available online at the following site:

`www.ams.org/bull/0000-000-00/S0273-0979-05-01048-7/ S0273-0979-05-01048-7.pdf`

The paper by Hirsch has mainly historical value; the currently accepted "standard" approach to smooth immersion theory is similar to the one appearing in the papers of Haefliger-Poenaru and Phillips.

Mersions of topological manifolds. If one defines immersions and submersions of topological manifolds using the conclusions of the Straightening Propositions, then it is possible to prove strong analogs of the results for smooth mersions. The precise formulations and proofs are worked out in the following paper.

[5] D. Gauld, *Mersions of topological manifolds*, Trans. Amer. Math. Soc. **149** (1970), 539–560.

As in the smooth case, these results do not consider the cases where $\dim M \geq \dim N$ and M is not compact. The situation here is entirely parallel to the one for smooth manifolds. A result of L. C. Siebenmann shows that a proper topological submersion is a topological fiber bundle (whose fibers are topological manifolds). As before, one can combine this with results from algebraic topology to construct specific examples for which the conclusion of the classification theorem fails; in fact, the explicit, so-called "simple examples" are also not homotopic to topological submersions, and this fact is also noted in the online file(s) cited above. Here are the references to Siebenmann's papers:

[6] L. C. Siebenmann, *Deformation of homeomorphisms on stratified sets. I*, Comment. Math. Helv. **47** (1972), 123–136.

[7] L. C. Siebenmann, *Deformation of homeomorphisms on stratified sets. II*, Comment. Math. Helv. **47** (1972), 137–163.

Classifying submersions on compact manifolds. (‡) If f and g are regularly homotopic submersions (either smooth or topological) from a compact manifold E to another manifold B, then one can show there is a diffeomorphism (resp., homeomorphism) $\varphi : E \to E$ such that $f \circ \varphi = g$. This condition defines a weaker equivalence relation on submersions, and in principle the classification of smooth submersions up to this notion of equivalence may be viewed as an extended application of the bundle classification theory in Steenrod's book. However, this is much easier said than done except in a very limited number of special cases, for one quickly encounters a host of deep and difficult questions in algebraic and geometric topology even in some relatively simple cases; *e.g.*, when $E = S^p \times S^q$ and $B = S^q$. For a variety of reasons it is not feasible to give background references here.

III.6.3 : *Smooth submanifolds*

The preceding definitions provide the tools we need to formulate a definition of submanifold that includes open subsets of smooth manifolds and level sets of regular values.

Definition. Let $(M, \mathcal{A})$ and $(N, \mathcal{B})$ be smooth manifolds such that N is a subspace of M. Then $(N, \mathcal{B})$ is said to be a *smooth submanifold* of $(M, \mathcal{A})$ if the inclusion map $i : N \subset M$ is smooth and the associated map $T(i) : T(N) \to T(M)$ is also 1–1.

The following immediate consequences of the definition are worth noting at this point.

PROPOSITION. *If N is a smooth submanifold of M and i denotes the inclusion of N in M, then i is a smooth immersion.*

This follows because $T(i)$ maps the tangent space at each point of N injectively.∎

TRANSITIVITY PROPERTY. *If P is a smooth submanifold of N and N is a smooth submanifold of M, then P is a smooth submanifold of M.*

Proof. First of all, P is a subspace of M because "a subspace of a subspace is a subspace." Let $i : N \subset M$ and $j : P \subset N$ be the inclusion maps. Then by hypothesis both $T(j)$ and $T(i)$ are 1–1, and therefore $T(i \circ j) = T(i) \circ T(j)$ is also 1–1. Since $i \circ j$ is the inclusion of P in M, this proves that the conditions for a submanifold are satisfied.∎

The next result is extremely important conceptually and is used all the time when working with smooth submanifolds.

SMOOTH LOCAL FLATNESS PROPERTY. *Suppose that P is a smooth submanifold of M and $x \in P$, and assume that q and m denote the dimensions o P and M respectively. Then there exists a smooth chart $(U \times N_\delta(0; \mathbf{R}^{m-q}), h)$ for x in M such that*

(*i*) *U is an open subset of $\mathbf{R}^q$,*

(*ii*) *the restriction $h|U \times \{0\}$ determines a smooth chart for P,*

(*iii*) *the inverse image $h^{-1}(P)$ is equal to $U \times \{0\}$.*

Proof. Since the inclusion mapping i is an immersion it follows that there are smooth charts (U_0, h_0) and $(U_0 \times N_{\delta_0}(0), k_0)$ satisfying all the conditions except perhaps the final one. Without further consideration it is conceivable that the image of k_0 contains other points of N (in particular, see the examples below involving immersions!). We need to shrink U_0 and δ_0 in order to remove any such extraneous points.

Since $h(U_0)$ is open in P there is an open set $W \subset P$ such that $W \cap P = h(U_0)$. Consider the intersection

$$W \ \cap \ k_0\big(U_0 \times N_{\delta_0}(0)\big) \ .$$

One can then find an open set $U_0 \subset U$ containing x and some $\delta > 0$ satisfying $\delta < \delta_0$ for which

$$k_0\big(U \times N_\delta(0)\big) \quad \subset \quad W \ \cap \ k_0\big(U_0 \times N_{\delta_0}\big) \ .$$

The entire conclusion of the proposition now holds for $k = k_0|U \times N_\delta(0)$.∎

There is a converse to the preceding result; its proof is similar to the proof in Section III.1 that level sets have smooth atlases.

PROPOSITION. *Let $(M, \mathcal{A})$ be a smooth m-manifold, and let $P \subset M$ be a topological q-manifold such that for each $x \in P$ there exists a smooth chart of the form*

$$\Big(U \times N_\delta(0; \mathbf{R}^{m-q}), \ k\Big)$$

such that $x = k(u,0)$ for some $u \in U$ and

$$N \;\cap\; k\Big(U \times N_\delta(0;\mathbf{R}^{m-q}\Big) \quad = \quad N_\delta, \delta)^{m-n}) \quad = \quad k\big(U \times \{0\}\big) \ .$$

Then there is a smooth atlas $\mathcal{B}$ for N such that $(N,\mathcal{B})$ is a smooth submanifold of $(M,\mathcal{A})$.

Sketch of proof. We shall only mention a few basic points. A smooth atlas for P is given by charts of the form $(U \times \{0\}, k|U \times \{0\})$. From the construction it follows that the inclusion of P in M is an immersion, and since the inclusion i from P to M is 1–1 it follows that $T(i)$ is also 1–1.∎

A similar but more complicated argument yields a uniqueness result for smooth structures on a submanifold.

UNIQUENESS FOR SMOOTH SUBMANIFOLD STRUCTURES. *Let $(M,\mathcal{A})$ be a smooth manifold, let $P \subset M$ be a topological manifold, and let $\mathcal{B}$ and $\mathcal{B}'$ be smooth atlases for P that make the latter into a smooth manifold. Then $\mathcal{B}$ and $\mathcal{B}'$ define the same smooth structure on P.*

Proof.$(\star)$ The proof is similar to the argument for the theorem on level sets, so we shall concentrate on the steps that are different.

Each atlas $\mathcal{B}$ and $\mathcal{B}'$ for P contains a subatlas $\mathcal{B}_0$ and $\mathcal{B}_0'$ whose charts (W,ℓ_0) are constructible as follows: Given a smooth chart $(W \times N_\delta, \ell)$ in $\mathcal{A}$ such that

$$\mathrm{Image}(\ell) \;\cap\; P \quad = \quad \ell\big(U \times \{0\}\big)$$

take ℓ_0 to be the composite of $\ell|W \times \{0\}$ with the standard identification $W \cong W \times \{0\}$. In order to show that $\mathcal{B}$ and $\mathcal{B}'$ define the same smooth structure, it suffices to prove the same for $\mathcal{B}_0$ and $\mathcal{B}_0'$, and the latter in turn is equivalent to showing that for each pair of charts (U,k_0) and (U',k_0') in $\mathcal{B}_0$ and $\mathcal{B}_0'$ respectively, the transition map "$k_0^{-1} \circ k_0'$" is a diffeomorphism. By the previous discussion, the associated charts $(U \times N_\varepsilon, k)$ and $(U' \times N_\eta, k')$ belong to $\mathcal{A}$ and hence the transition map "$k^{-1} \circ k'$" is a diffeomorphism. As in the proof of the result on level sets, this transition map sends an open subset of $U \times \{0\}$ to an open subset of $U' \times \{0\}$ by "$k_0^{-1} \circ k_0'$" and this map will be a diffeomorphism for the same reasons given in the proof of the level sets theorem.∎

Examples of smooth submanifolds are generally constructed using smooth embeddings, which will be introduced in the next section. Therefore our discussion of examples will be rather limited for the time being.

Example. If M is a smooth manifold and U is an open subset of M, then U is a smooth submanifold. — The inclusion map has already been shown to be smooth, and the condition on tangent bundles follows from the identities at the end of the previous section.

We have already remarked that level sets also yield examples of smooth manifolds. Here is a generalized version of this fact.

PROPOSITION. *Let $f : M \to N$ be a smooth map of smooth manifolds and let $p \in N$ be a nontrivial regular value; i.e., $p = f(x)$ for some $x \in M$ and if $f(y) = p$ then $T_y(f)$ maps $T_y(M)$ onto $T_p(N)$. Then $V = f^{-1}(\{p\})$ is a smooth $(m-n)$-dimensional submanifold.*∎

The proof is nearly the same as for the special case considered in Section III.1, the only difference being that the local diffeomorphisms in that argument are replaced by smooth coordinate charts from atlases for M and N.

Nonsmoothable topological submanifolds. ‡ The local flatness condition for smooth submanifolds gives a strong necessary condition for a topological manifold $P \subset M$ to be a smooth submanifold. To illustrate this, we shall outline a method for constructing a subset of S^4 that is homeomorphic to S^2 but cannot be made into a smooth submanifold. This is actually a class of examples, and it depends upon the existence of nontrivially knotted curves in $\mathbf{R}^3$ and S^3. The most basic such example is the standard curve that one forms by first tying a piece of string into a simple knot, and then gluing the two loose ends of the string together. The files `trefoil.*` in the course directory discuss this further and mention one important property of this knot: If K is the knotted curve we have described, then the fundamental groups of $\mathbf{R}^3 - K$ and $S^3 - K$ are nonabelian (in contrast, if C is the standard circle in the plane, then the fundamental groups of $\mathbf{R}^3 - C$ and $S^3 - C$ are infinite cyclic, and this leads to a rigorous mathematical proof that one cannot untangle K without using a knife or scissors to cut it — this is modeled mathematically by removing one point from K). The fundamental group computation is the crucial information about the knot K that will be needed below.

Given any space X, one can construct its *unreduced suspension* ΣX as follows: Define an equivalence relation on $X \times [-1,1]$ whose equivalence classes are one point sets of the form $\{ (x,t) \}$ for $t \neq \pm 1$ along with the two larger subsets $X \times \{-1\}$ and $X \times \{+1\}$. If A is a subspace of X, then there is a natural inclusion of ΣA as a subspace of ΣX determined by the inclusion of $A \times [-1,1]$ in $X \times [-1,1]$. If Y is a sphere S^k then it is elementary to prove that $\Sigma Y = S^{k+1}$ (see the exercises for this section). Applying this to $K \subset S^3$, we obtain an inclusion of the form

$$S^2 \cong \Sigma S^1 \cong \Sigma K \subset \Sigma S^3 = S^4 .$$

A proof that this inclusion does not have the topological local flatness property (using the fact about fundamental groups in the previous paragraph) appears in the file(s) `suspknots.*` in the course directory.■

Definition. If M is a topological manifold and $N \subset M$ is also a topological manifold, then N is said to be *locally flat* if for each point one can find topological coordinate charts as in the conclusion of the (Smooth) Local Flatness Property above.

The following book by T. B. Rushing gives a comprehensive account of non-locally-flat embeddings of topological manifolds, including some that are even less well-behaved than the example considered above.

> T. B. Rushing, Topological embeddings. Pure and Applied Mathematics, Vol. 52. Academic Press, New York-London, 1973. ISBN: 0-12-603550-4.

We shall conclude this subsection with a basic result on smooth maps into submanifolds. This result is analogous to a standard fact about continuous maps into topological subspaces: *If $f : X \to Y$ is a continuous map and B is a subspace of Y such that $f(X) \subset B$, then there is a unique factorization of f as a composite $j \circ f'$ where $f' : X \to B$ is continuous and $j : B \to Y$ is the inclusion mapping.*

SMOOTH SUBMANIFOLD FACTORIZATION PROPERTY. *Suppose that N is a smooth submanifold of M, let $j : N \subset M$ be the inclusion map, let P be a smooth manifold, and let $f : P \to M$ be a continuous map such that $f(R) \subset N$. Then there is a unique factorization of f as a composite $j \circ f'$ where $f' : P \to N$ is smooth and $j : N \to M$ is the inclusion mapping.*

Proof. By the cited result for continuous mappings, there is a unique factorization of the desired type such that f' is continuous. Therefore it is only necessary to prove that f' is smooth.

Let $x \in P$ be arbitrary; then $f(x) \in N$ by hypothesis, and therefore there is a smooth chart at $f(x)$ in M of the form $(V \times N_\delta(0), k)$ such that the intersection of Image(k) and N is equal to $k(V \times \{0\})$. By continuity there is a smooth chart (U_x, h_x) at x in P such that f maps $h(U_x)$ into the image of k. Let $F : U_x \to V \times N_\delta(0)$ be the local map defined by f; this local map is smooth because f is smooth. The condition $f(P) \subset N$ implies that we may write F in coordinates as $F(u) = (F'(u), 0)$; the function F' will then be the corresponding local map for f', and thus we want to determine whether F' is smooth. But this follows easily because $F' = \pi_V \circ F$, where π_V denotes projection onto the first coordinate. Therefore we have shown that the restriction of f' to $h_x(U_x)$ is smooth. Since there is an open covering of P by sets of the form $h_x(U_x)$ for suitably chosen $x \in P$, it follows that f' is smooth everywhere.■

III.6.4 : *Smooth embeddings*

In Section I.4 we described an abstract notion of topological embedding; *i.e.*, a mapping that is 1–1, continuous, and defines a homeomorphism onto its image. Here is the analogous concept for smooth manifolds.

Definition. If $f : M \to N$ is a smooth map of smooth manifolds, then f is said to be a **smooth embedding** if it is a 1–1 immersion and maps M homeomorphically to $f(M)$.

EXAMPLES. If M is a smooth submanifold of N, then the inclusion map $i : M \subset N$ is a smooth embedding because it is an immersion and a homeomorphism onto its image. In particular, this applies if M is an open subset of M or a regular level subset for some smooth function defined on M. Also, if $f : M \to P$ is a smooth map, then the graph map $\mathbf{G}_f : M \to M \times P$ is an embedding (we have seen that it is a 1–1 immersion, and we also noted that it defines a homeomorphism onto its image).

The immersion condition in the definition of a smooth embedding reflects the preceding observation regarding our definition of smooth submanifold. One might also think of defining a smooth embedding to be a topological embedding $f : N \to M$ such that

(*i*) $f(N)$ is a smooth submanifold of M,

(*ii*) f determines a diffeomorphism from N to $f(N)$.

In fact, this formulation is equivalent to the definition we have given. The first step in showing this is to note that a map satisfying the two conditions above is a smooth embedding. Since we already know that f defines a homeomorphism onto its image, it is only necessary to show that f is an immersion. The map f factors as a composite $i \circ f'$ where $i : f(M) \to N$ is the inclusion map and f' is the homeomorphism (in fact, a *diffeomorphism*) from M to $f(M)$. Since both f' and i are smooth immersions, it follows that their composite $f = i \circ f'$ is also an immersion.■

The converse implication, which is the crucial relation between smooth embeddings and smooth submanifolds, is contained in the following result:

PROPOSITION. *Let $F : (N, \mathcal{B}) \to (M, \mathcal{A})$ be a smooth embedding. Then there is a smooth atlas $\mathcal{E}$ on $f(N)$ such that*

(*i*) *if $i : f(N) \to M$ is the inclusion map, then i is the inclusion of $f(N)$ as a smooth submanifold,* (*ii*) *if* $g : N \to f(N)$ *is the map of sets such that* $f = i \circ g$, *then* g *is a diffeomorphism.*

Proof. By hypothesis the map f defines a homeomorphism onto its image, and therefore the map g is a homeomorphism. As in the text, we can find a smooth atlas $\mathcal{E}$ on $f(N)$ such that g defines a diffeomorphism from $(N, \mathcal{B})$ to $(f(N), \mathcal{E})$; specifically, the charts in $\mathcal{E}$ all have the form $(U_\alpha, g \circ h_\alpha)$, where (U_α, h_α) is a smooth chart in $\mathcal{B}$. Since $i = f \circ g^{-1}$, it follows that i is also a smooth immersion.■

Example 1. Not every 1–1 immersion is an embedding. Consider the figure 8 curve $\varphi(t) = (\sin 2t, \sin t)$ for $t \in (0, 2\pi)$ that was first described in Section I.1. The image of this curve is a figure 8 where the crossing point is the origin, and therefore the image is not a manifold.

On the other hand, it is an elementary exercise to check that $\varphi'(t)$ is never zero and that φ is 1–1 on the open interval $(0, 2\pi)$. Of course, one can extend the definition of φ to all real values of t but then the function will not be 1–1.

Example 2. ($\star$) Here is another important but more complicated example of a smooth 1–1 immersion that is not an embedding. If $p : \mathbf{R}^2 \to T^2 = S^1 \times S^1$ is the map sending (s, t) to $(\exp(2\pi i\, s, (\exp(2\pi i\, t,)$, then p is a smooth covering space map. Given a positive irrational number α, consider the smooth map $f : \mathbf{R} \to T^2$ defined by $f(u) = p(\alpha u)$. By construction, f is a smooth map and f is a group homomorphism. Furthermore, under the identification of $T(\mathbf{R})$ with $\mathbf{R} \times \mathbf{R}$ given by the chart $k : \mathbf{R} \times \mathbf{R} \to T(\mathbf{R})$, we have that

$$[T(f)] \circ k(u, 1) \quad = \quad [T(p)] \circ k(\alpha u, \alpha)$$

and the right hand side is nonzero because the maps of tangent spaces $T(p)_{(s,t)}$ are linear isomorphisms for all (s, t). The latter implies that f is an immersion. Since f is a group homomorphism, it is 1–1 if and only if $f^{-1}(\{1\}) = \{0\}$; but $f(v) = 1 \implies \exp(2\pi i \alpha v) = \exp(2\pi i v) = 1$, and the latter equations imply that both v and αv are integers. Since α is irrational, this can only happen if $v = 0$.

CLAIM. The image $f(R)$ is **not** a smooth submanifold of T^2. The key to showing this is a basic result due to L. Kronecker. In order to state this result we need some simple terminology. Given a real number x, let $[x]$ denote the greatest integer $\leq x$ and let $\langle x \rangle$ denote the "fractional part" $x - [x]$.

KRONECKER APPROXIMATION LEMMA. *Let* $x \in \mathbf{R}$, *let* α *be a positive irrational real number and let* $\varepsilon > 0$. *Then there is an integer* n *such that*

$$\left| x - \langle n\alpha \rangle \right| \quad < \quad \varepsilon \; .$$

COROLLARY. *Let* α *be as above, and let* δ *be a positive real number. Then there is a positive integer* m *such that* $\langle m\alpha \rangle < \delta$.

Derivation of the corollary. Applying the Kronecker Approximation Lemma when $x = 0$, we see that there is an integer n such that $| \langle n\alpha \rangle | \quad < \quad \varepsilon$. Take m to be the absolute value of n.■

A proof of Kronecker's result appears in Sections 7.4 and 7.5 of the following text:

T. M. Apostol, Modular Functions and Dirichlet Series in Number Theory (Graduate Texts in Mathematics, Vol. 41, 2nd ed., Corr. 2nd printing), Springer, New York, 1997. ISBN: 0-387-97127-0.

A proof of the corollary is given in the file(s) `kroneckerappx.*` in the course directory. Here is an online reference for Kronecker's result:

`http://mathworld.wolfram.com/KroneckersApproximationTheorem.html`

We shall now use Kronecker's result to complete the discussion of our example.

PROOF OF CLAIM. If $f(\mathbf{R})$ is a submanifold of T^2, then one can find some $\varepsilon > 0$ and a neighborhood W of $(1,1) = f(0)$ such that

$$W \cap f(\mathbf{R}) \quad = \quad f\big((-\varepsilon, \varepsilon) \big) \ .$$

Without loss of generality we might as well assume that $\varepsilon < 1$. Now choose $\delta \in (0,1)$ such that p maps $(-\delta, \delta)^2$ diffeomorphically onto an open set $W_0 \subset W$. Choose m as in the preceding corollary. The mapping p satisfies

$$p(s,t) \quad = \quad p\big(\langle s \rangle , \, \langle t \rangle , \, \big)$$

and therefore we have

$$p(m, \alpha m) \quad = \quad p\big(0, \, \langle \alpha m \rangle \big) \ .$$

The point on the right hand side lies in both W_0 and by the right hand side of the equation it is also equal to $f(m)$; by the condition in the first sentence of this paragraph, we know that the a point in W_0 lies in the image of f only if it has the form $f(t)$ for $|t| < \varepsilon$. Since $m \geq 1 > \varepsilon$ and f is 1–1, this yields a contradiction. It follows that $f(\mathbf{R})$ cannot be a smooth submanifold of T^2.■

Note. One can push the preceding discussion further to show that $f(\mathbf{R})$ is dense in T^2.■

In Section I.4 we gave an abstract description of topological embeddings in terms of a universal mapping property. We shall conclude this subsection with a smooth analog of this property.

THEOREM. *Suppose that $g : N \to M$ is a smooth embedding and $f : P \to M$ is a smooth map such that $f(P) \subset g(N)$. Then there is a unique smooth map $f_0 : P \to N$ such that $f = g \circ f_0$.*

Proof. Let g' denote the induced diffeomorphism from N to $g(N)$ and let $j : g(N) \to M$ be the inclusion map. The condition $f(P) \subset g(N)$ implies that there is a unique continuous map $f' : P \to g(N)$ such that $f = j \circ f'$. If we set $f_1 = (g')^{-1} \circ f'$, then we have

$$f \quad = \quad j \circ f' \quad = \quad j \circ g' \circ (g')^{-1} \circ f' \quad = \quad g \circ f_1$$

so that a function with the desired properties exists. If f_2 is any such function, then we have $g \circ f_1 = f = g \circ g_2$, which implies that $g(f_1(y)) = g(f_2(y))$ for all $y \in P$. Since g is 1–1, this implies that $f_1 = f_2$, and therefore the uniqueness conclusion also follows.■

III.6.5 : *Embeddings in Euclidean spaces* $(1\frac{1}{2}\star)$

We have already mentioned that many undergraduate textbooks define smooth manifolds as objects satisfying the hypotheses of the Submanifold Recognition Principle (*cf.* the book by Edwards mentioned previously). In order to show that our definition is equivalent to the one in such texts, it is necessary to show that every submanifold is diffeomorphic to a submanifold of some Euclidean space; *i.e.*, we need to show that every smooth manifold admits a smooth embedding into some Euclidean space. The existence of such embeddings is also useful in a wide range of contexts; a few simple results appear in the Appendix to this section.

We begin by disposing of the compact case. In fact, we shall prove the following more general result; the argument is similar to one proof of the corresponding result for topological manifolds (*cf.* Theorem 36.2 on pages 226–227 of [MUNKRES1]).

ABSTRACT EUCLIDEAN EMBEDDING THEOREM. *Let M be a n-smooth manifold that admits a finite open covering by the images of smooth charts (U_i, h_i) for $1 \leq i \leq k$. Then M admits a smooth embedding into $\mathbf{R}^{k(n+1)}$.*

COROLLARY. *If M is a compact smooth manifold, then there is a smooth embedding of M into some Euclidean space.*

The corollary follows because M has a finite open covering by images of smooth coordinate charts.■

It will be convenient to make two preliminary observations before proceeding to the proof of the main result.

SHRINKING LEMMA. *Suppose that X is a $\mathbf{T}_4$ space and $\{U_i\}$ is a finite open covering of X. Then there are open sets $V_i \subset X$ such that $V_i \subset \overline{V_i} \subset U_i$ and $\{V_i\}$ is an open covering of X.*

Reference for proof. This result is established as Step 1 in the proof of Theorem 36.1 on page 225 of [MUNKRES1].■

SMOOTH VERSION OF URYSOHN'S LEMMA. *Let U be an open subset of $\mathbf{R}^n$, and let $V \subset U$ also be open. Then there is a smooth function $f : U \to \mathbf{R}$ with the following properties:*

(*i*) *The function f takes values in the closed interval $[0, 1]$.*

(*ii*) *For all $x \in \overline{V}$ (= the closure in U) we have $f(x) = 1$.*

(*iii*) *There is an open set W such that*

$$\overline{V} \quad \subset \quad W \quad \subset \quad \overline{W} \quad \subset \quad U$$

(once again, the closure in U) and $f(x) = 0$ for all $x \in U - W$.

Proof. $(2\star)$ Since $\mathbf{R}^n$ is metrizable ($\Longrightarrow$ normal), there is an open subset W such that

$$\overline{V} \quad \subset \quad W \quad \subset \quad \overline{W} \quad \subset \quad U$$

Let $\mathcal{U}$ be the open covering of U given by the two open sets W and $U - \overline{V}$. As in the subsection of Section II.3 treating smooth partitions of unity, there is a locally finite open refinement $\mathcal{V}$ such that the following hold:

(*i*) Each V_α in $\mathcal{V}$ is an open disk whose center and radius will be denoted by z_α and r_α respectively.

(*ii*) The sets W_α, defined as open disks whose centers are the same as those of the corresponding V_α and whose radii are half those of the corresponding V_α, also form an open covering of U.

(*iii*) Each set $\overline{W_\alpha}$ is compact and satisfies $\overline{W_\alpha} \subset V_\alpha$.

Let k_α be a smooth nonnegative real valued function on V_α such that $k_\alpha = 1$ on the disk of radius $\frac{1}{2}r_\alpha$ centered at z_α and $k_\alpha = 0$ off the disk of radius $\frac{3}{4}r_\alpha$ centered at z_α. As usual, this function extends smoothly to U by setting it equal to zero off V_α. Define $h = \sum_\alpha k_\alpha$, noting as usual that the sum is meaningful and smooth by the local finiteness of the covering $\mathcal{V}$, and it is positive because the sets W_α form an open covering of U. Next define $g = \sum_\beta k_\beta$ where β runs through all α such that $V_\alpha \cap \overline{V} \neq \emptyset$. By construction this sum is nonnegative, and we claim that g also has the following two properties:

(a) $g|\overline{V} \;=\; h|\overline{V}$

(b) If W is the union of all open sets W_α such that $W_\alpha \cap \overline{V} \neq \emptyset$, then $g|\,U - W \;=\; 0$.

To prove (a), note that $g \leq h$, and the only way strict inequality can hold at some point y is if $k_\alpha(y) > 0$ for some α such that V_α is disjoint from $\overline{V}$. However, if $x \in \overline{V}$ then $k_\alpha(x) = 0$ for all such α because $k_\alpha = 0$ off V_α. To prove (b), note that if $x \in U - W$, then the construction of $\mathcal{V}$ as a refinement of $\mathcal{U}$ implies that x only lies in disks V_α such that $V_\alpha \subset U - \overline{V}$. This means that $k_\beta(x) = 0$ for each β in the summation defining g and therefore that $g(x) = 0$ for such choices of x.

Therefore, if we define f to be the quotient g/h, it follows that $f = 1$ on $\overline{V}$ and $f = 0$ on $U - W$.■

Proof of the Abstract Euclidean Embedding Theorem. (2⋆) Several steps in the proof are very similar to portions of the argument presented in `smirnov.*` (see the course directory).

First of all, consider the open covering of M by the images of the given smooth charts (U_i, h_i). Using the Shrinking Lemma, define a second open covering given by sets of the form $h_i(V_i)$, where V_i is open in U_i and

$$\overline{h_i(V_i)} \;\subset\; h_i(U_i)\ .$$

Next, take open sets $W_i \subset U_i$ such that

$$\overline{h_i(V_i)} \;\subset\; h_i(W_i) \;\subset\; \overline{h_i(W_i)} \;\subset\; h_i(U_i)\ .$$

The smooth version of Urysohn's Lemma then implies the existence of a smooth function ω_i on U_i such that $\omega_i = 1$ on V_i and $\omega_i = 0$ on the complement of W_i.

Let J_i denote the inclusion of U_i in $\mathbf{R}^n$. Define a smooth map

$$G_i : U_i \longrightarrow \mathbf{R}^{n+1} \;\cong\; \mathbf{R}^n \times \mathbf{R}$$

by the formula

$$G_i(x) \;=\; \big(\omega_i(x)\cdot x,\ \omega_i(x)\big)$$

and convert this into a smooth map on $h_i(U_i)$ using the map $\varphi_i : h_i(U_i) \to U_i$ that is "inverse" to h_i. The composite $G_i \circ \varphi_i$ vanishes off $\overline{h_i(W_i)}$, and therefore it extends to a smooth map $H_i : M \to \mathbf{R}^{n+1}$. Now define a smooth map

$$H : M \longrightarrow \left(\mathbf{R}^{n+1}\right)^k \;\cong\; \mathbf{R}^{k(n+1)}$$

whose projection onto the i^{th} $\mathbf{R}^{n+1}$ factor is equal to H_i. CLAIM: H is a smooth embedding.

The argument on the second page of the file(s) `smirnov.*` goes through essentially unchanged to show that H defines a homeomorphism onto its image. Therefore the proof reduces to showing that H is a smooth immersion. It will suffice to show for each i the restriction of H to $h_i(V_i)$ is an immersion, and in fact it will suffice to show that the restriction of the i^{th} coordinate function H_i to $h_i(V_i)$ is an immersion for each i, or equivalently that the restriction of G_i to V_i is an immersion. Now $\omega_i = 1$ on V_i and therefore $G_i|V_i$ sends a point $x \in V_i$ to $(x, 1)$. This map is clearly an immersion (in fact, it is a constant plus an injective linear transformation!), and therefore it follows that for each i the maps $G_i|V_i$ and $H_i|h_i(V_i)$ are immersions as required. As noted before, this completes the proof that H defines a smooth embedding into some Euclidean space.■

Our next objective is to prove that every noncompact (second countable) smooth manifold admits a smooth embedding into some Euclidean space. The approach is based upon the following point-set-theoretic input from pages 20–21 of [MUNKRES2]:

LEMMA. *Let M be a topological n-manifold, and let $\mathcal{U}$ be an open covering of M. Then there is a countable locally finite open refinement $\mathcal{W}$ of $\mathcal{U}$ that is a union of pairwise disjoint families $\mathcal{W}_i$ for $0 \le i \le n$.*■

It will also be helpful to have the following observation involving smooth coordinate charts.

SUBLEMMA. *Let M be a smooth n-manifold, and let U be open in $\mathbf{R}^n$. Suppose that U is a union of the pairwise disjoint open subsets U_j and that $h : U \to M$ is a $1-1$ continuous open map such that for each j the pair $(U_j, h|U_j)$ is a smooth chart. Then (U, h) is also a smooth chart.*

Proof. We need to show that if (V, k) is a smooth chart for M then the transition map "$k^{-1} \circ h$" is a diffeomorphism, or equivalently that it is smooth and has nonvanishing Jacobian everywhere. This will hold if and only if it does for the restriction to each U_j. But the restriction to such a subset also has the form "$k^{-1} \circ h|U_j$" and these transition maps are smooth with nonvanishing Jacobians by the hypotheses.■

Using these we shall prove the desired embedding result.

WEAK EUCLIDEAN EMBEDDING THEOREM. *If M^n is a smooth n-manifold, then M admits a smooth embedding into $\mathbf{R}^{(n+1)^2}$.*

Proof. $(2\star)$ Start with an open covering of M by images of smooth coordinate charts (U_α, h_α), and let $\mathcal{U}$ be the family of sets $h_\alpha(U_\alpha)$. Let $\mathcal{W}$ be the locally finite refinement given by the lemma, let the subfamilies $\mathcal{W}_i$ be as in the conclusion of the lemma, and write $\mathcal{W}_i$ as $\{ W_{i,j} \}$. Since $\mathcal{W}$ is a refinement of $\mathcal{U}$ each set $W_{i,j}$ is the image of some smooth coordinate chart; specifically, let $W_{i,j}$ be the image of $(V_{i,j}, k_{i,j})$.

The map

$$f(x) \quad = \quad \frac{1}{2} \ + \ \left(\frac{2}{\pi} \arctan x \right)$$

defines a diffeomorphism from $\mathbf{R}$ onto the open unit interval $(0, 1)$, and we may similarly construct a diffeomorphism from $\mathbf{R}^n$ to $(0,1)^n$ by the formula

$$F(x_1, \ \cdots \ , x_n) \quad = \quad \big(f(x_1), \ \cdots \ , f(x_n) \ \big) \ .$$

Using this diffeomorphism and its translates, one can construct open subsets

$$\Omega_{i,j} \quad \subset \quad (\, 2j, \, 2j+1\,)^n$$

and diffeomorphisms $\varphi_{i,j} : \Omega_{i,j} \to V_{i,j}$. If $\Omega_i = \cup_j \, \Omega_{i,j}$ then the latter presents Ω_i as a union of pairwise disjoint subsets, and therefore we may define $\lambda_i : \Omega_i \to M$ by setting $\lambda_i|\Omega_{i,j} = k_{i,j} \circ \varphi_{i,j}$. By construction this map is continuous, 1–1 and open, and furthermore each restriction $\lambda_i \circ \Omega_{i,j}$ is a smooth chart for M. By the sublemma, it follows that λ_i is also a smooth chart.

The preceding argument yields $(m+1)$ smooth charts (Ω_i, λ_i) for M whose images cover M. Therefore the Abstract Euclidean Embedding Theorem implies that there is a smooth embedding of M into $\mathbf{R}^{(n+1)^2}$.■

Optimal embedding dimensions. (‡) We have called our general embedding result a weak embedding theorem because it there are similar results with $(n+1)^2$ replaced with much lower dimensions. As noted in Section 50 of [MUNKRES1], if a compact metric space X has topological dimension n in an appropriately defined sense, then there is a topological embedding of X in $\mathbf{R}^{2n+1}$ and this turns out to be the best possible general result (for example, Section 64 of [MUNKRES1] contains examples of 1-dimensional spaces that do not admit topological embeddings into $\mathbf{R}^2$). A basic result of H. Whitney yields a comparable smooth embedding result; namely, every smooth n-manifold can be smoothly embedded in $\mathbf{R}^{2n+1}$. In fact, a subsequent and much deeper result of Whitney shows that every smooth n-manifold can be smoothly embedded in $\mathbf{R}^{2n}$ (CAUTION: Page 22 of Conlon contains an incorrect assertion to the contrary just before the beginning of Section 1.6). However, the Hard Whitney Embedding Theorem is in fact the best possible general result. The classic text by J. Milnor and J. Stasheff (listed the references below) contains a proof that $\mathbf{RP}^n$ does not embed smoothly in $\mathbf{R}^{2n-1}$ if n is a power of 2. In fact, one can extend Whitney's techniques to prove that topological n-manifolds always embed topologically in $\mathbf{R}^{2n}$ using the work of Kirby and Siebenmann, and results of A. Haefliger show that $\mathbf{RP}^n$ does not even embed topologically in $\mathbf{R}^{2n-1}$ if n is a power of 2. Here are some references for the results discussed in this paragraph:

[1] A. Haefliger, *Differentiable imbeddings*, Bull. Amer. Math. Soc. **67** (1961), 109–112.

[2] A. Haefliger, *Plongements différentiables de variétés dans variétés*, Comment. Math. Helv. **36** (1961), 47–82.

[3] J. W. Milnor and J, D. Stasheff, Characteristic classes. Annals of Mathematics Studies, No. 76. Princeton University Press, Princeton, N. J., 1974. ISBN: 0-691-08122-0.

[4] H. Whitney, *The self-intersections of a smooth n-manifold in $2n$-space*, Ann. of Math. (2) **45** (1944), 220–246.

III.6.6 : *Applications of Euclidean embedding theorems*

The existence of smooth embeddings into Euclidean spaces is is important for practical as well as conceptual reasons, for it often can be used to simplify the proofs of fundamental results. Two related examples that appear in Sections 3.7 and 3.8 of Conlon. The first of these shows, among other things, that a smooth submanifold M of Euclidean space has a special type of neighborhood V such that M is a *smooth retract* of V; *i.e.*, there is a smooth map $r : V \to M$ such that $r|V$ is the identity on M. This result in turn makes it fairly easy to prove results on approximating a continuous map of smooth manifolds $P \to M$ by a smooth one. Roughly speaking, the idea is to approximate the original map by a smooth map into V and then to compose it with the retraction to get back into M. Further information may be found in the sections of Conlon that are cited above.

IV. Vector fields

In Section II.4 we considered vector fields for open subsets of Euclidean spaces previously, with particular attention to the systems of autonomous ordinary differential equations they generate. One objective of this unit is to extend the definitions to manifolds. Another is to examine the integral flow curves for the associated differential equations, including some important special properties of solution curves for autonomous differential equations.

In ordinary multivariable calculus there are various operations involving vector fields that play an important role in the subject and its applications to physics. These include the gradient of a smooth function and the divergence and curl of a vector field. Although we shall describe generalizations of such objects to arbitrary manifolds in this course, we shall need additional mathematical tools in order to do so, and therefore the discussion of such generalizations will be postponed to Unit V. On the other hand, there is an important operation on vector fields called the Lie bracket that can and will be described in this section. The bracket is named after the Norwegian mathematician Sophus Lie and his last name is pronounced "lee." Although this operation is not really needed in undergraduate multivariable calculus, it plays an important role in many mathematical and physical contexts. In particular, the Lie bracket of a pair of vector fields describes some ways in which their associated flow curves are related to each other.

IV.1 : Global vector fields

(Conlon, §§ 2.7–2.8, 4.1)

We begin by globalizing the notion of vector field to arbitrary smooth manifolds.

IV.1.1 : *Vector fields on smooth manifolds*

Intuitively, a (tangent) vector field on a manifold is supposed to specify a tangent vector for each point of the manifold. Here is the formal definition.

Definition. Let M be a smooth manifold and let $\tau_M : T(M) \to M$ be its tangent bundle. A **(tangent) vector field** on M is a continuous map $X : M \to T(M)$ such that $\tau_M \circ X = \mathrm{id}_M$. Unless specifically stated otherwise, we shall assume that all vector fields under consideration are smooth.

If U is open in $\mathbf{R}^n$, so that $T(U) \cong U \times \mathbf{R}^n$, then a (smooth!) tangent vector field has the form

$$X(u) = (u, \mathbf{F}(u))$$

where $\mathbf{F} : U \to \mathbf{R}^n$ is smooth mapping. This is clearly equivalent to the notion of vector field that one sees in undergraduate physics and multivariable calculus courses.

Example 1. The preceding paragraph gives a vast collection of examples that is exhaustive if M is an open subset of $\mathbf{R}^n$. On an arbitrary smooth manifold M the zero map from M to $T(M)$ is a smooth vector field.

Example 2. If X and Y are vector fields on M, then their sum $X + Y$ can be defined and it is also a vector field. Likewise, if $g : M \to \mathbf{R}$ is smooth then the pointwise scalar product $g \cdot X$ is also a vector field. The operations of addition and multiplication by functions in $C^\infty(M)$ make the set of vector fields $\mathbf{X}(M)$ into a *module* over $C^\infty(M)$; *i.e.*, the operations satisfy all the rules for vector addition and scalar multiplication that one has for a vector space (although $C^\infty(M)$ is merely a commutative ring with unit rather than a field).

Example 3. If M is a smooth manifold with tangent bundle $\tau_M : T(M) \to M$ and U is open in M with inclusion map $i : U \subset M$, then $T(U)$ is isomorphic to $\tau_M^{-1}(U)$ and the accordingly restriction $X|U$ determines a smooth vector field i^*X on U. The construction $i^* : \mathbf{X}(M) \to \mathbf{X}(U)$ is a linear transformation of real vector spaces (verify this!), and the product with smooth functions satisfies the identity

$$i^*(g \cdot X) = (g|U) \cdot X.$$

Example 4. If M is a smooth manifold and $f : M \to N$ is a diffeomorphism, then an isomorphism of real vector spaces $f_* : \mathbf{X}(M) \to \mathbf{X}(N)$ is defined by $f_*Y(p) = T(f) \circ Y \circ f^{-1}$. This is a direct generalization of the previously defined construction for open subsets of Euclidean spaces. One natural question is what can be said about $f_*(g \cdot Y)$ if $g \in C^\infty(M)$; this is left to the reader as an exercise. One important property of the f_* construction is that it is covariantly functorial: $(h \circ f)_* = h_* \circ f_*$ and $(\mathrm{id}_M)_* = \mathrm{id}_{\mathbf{X}(M)}$.

In the classical approach to tensor analysis, tangent vector fields for a smooth manifold M are described in terms of a smooth atlas $\mathcal{A} = \{(U_\alpha, h_\alpha)\}$ for M. It will be convenient for us to write $\psi_{\beta\alpha}$ for the transition maps "$h_\beta^{-1} \circ h_\alpha$" $= \psi_{\beta\alpha}$ in this note, both for the sake of notational conciseness and for its consistency with the standard notation of tensor analysis. The following result is essentially the classical characterization of vector fields that is basic to tensor analysis (in the terminology of tensor analysis, this is a *contravariant tensor field of rank* 1.

GLOBAL VECTOR FIELD CONSTRUCTION PRINCIPLE. *Given a smooth n-manifold M with smooth atlas $\mathcal{A} = \{(U_\alpha, h_\alpha)\}$, suppose that for each α we are given a smooth map $f_\alpha : U_\alpha \to \mathbf{R}^n$ and that these maps satisfy the condition*

$$f_\beta \circ \psi_{\beta\alpha}(u) \quad = \quad [D\psi_{\beta\alpha}(u)] f_\alpha(u)$$

for all α and β. Then there is a unique vector field Y on M such that

$$k_\alpha(u, f_\alpha(u)) \quad = \quad Y \circ h_\alpha(u)$$

for all U_α and $u \in U_\alpha$.

Proof. Define maps $\chi_\alpha : U_\alpha \to T(M)$ by the formula $\chi_\alpha(x) = k_\alpha(x, f_\alpha(x))$. This defines a vector field over $h_\alpha(U_\alpha)$ because $\tau_M \circ \chi_\alpha = h_\alpha$. To show this defines a vector field, we need to prove the consistency identity $\chi_\alpha \circ \psi_{\beta\alpha} = \chi_\beta$.

As in Section III.5 we verify the consistency condition using a sequence of equations:

$$\chi_\alpha \circ \psi_{\alpha\beta}(x) \quad = \quad k_\alpha(\psi_{\alpha\beta}(x),\, f_\alpha \circ \psi_{\alpha\beta}(x)\,) \quad =$$

$$k_\alpha \circ \psi_{\alpha\beta}(\,x,\, [D\psi_{\beta\alpha}(x)] f_\alpha(x)\,) \quad = \quad k_\alpha \circ \varphi_{\alpha\beta}(x, f_\beta(x)\,) \quad =$$

$$k_\beta(x, f_\beta(x)\,) \quad = \quad \chi_\beta(x)$$

The standard results now imply the existence of a vector field Y with the required properties.■

The next result is just the same result restated in different terminology; verification is left to the reader as an exercise.

GLOBAL VECTOR FIELD CONSTRUCTION, ALTERNATE VERSION. *Given a smooth n-manifold M with smooth atlas $\mathcal{A} = \{(U_\alpha, h_\alpha)\}$, denote the transition maps "$h_\beta^{-1} \circ h_\alpha$" by $\psi_{\beta\alpha}$, and let $V_{\beta\alpha}$ and $V_{\alpha\beta}$ be the domain and image of $\psi_{\beta\alpha}$ respectively. Suppose that for each α we are given a smooth vector field Y_α on U_α and that these vector fields satisfy the compatibility condition*

$$[\psi_{\beta\alpha}]_*(Y_\alpha | V_{\beta\alpha}) \quad = \quad Y_\beta | V_{\alpha\beta} \ .$$

Then there is a unique vector field Y on M such that

$$(h_\alpha)_* Y_\alpha \quad = \quad Y | h_\alpha(U_\alpha)$$

$$k_\alpha(u, f_\alpha(u)) \quad = \quad Y \circ h_\alpha(u)$$

for all α.■

IV.1.2 : *Globally independent vector fields* $(1\frac{1}{2}\star)$

If U is an open subset in $\mathbf{R}^n$ then we can easily construct a set of n vector fields on U that are everywhere linearly independent. Specifically, if $\{\, \mathbf{e}_j \,\}$ denotes the standard unit vector bases for $\mathbf{R}^n$, we have the following smooth vector fields:

$$\frac{\partial}{\partial x_i}(\, u \,) \quad = \quad (u, \mathbf{e}_i)$$

The reason for this terminology will become apparent in Section IV.3.

More generally, one can use a variant of the Germ Extension Theorems in Sections II.3 and III.3 to prove the following local result for an arbitrary smooth manifold:

PROPOSITION. *Let M be a smooth n-manifold, and let $y \in M$. Then there exist smooth vector fields $\{\, X_i \,\}$ on M and an open neighborhood W of M such that for all $z \in U$ the vectors $X_i(W)$ form a basis for $T_z(M)$.*

Proof. Let (U, h) be a smooth chart at y, and define vector fields Y_i on $h(U)$ by the formula

$$Y_i \quad = \quad (h')_* \left(\frac{\partial}{\partial\, x_i} \right)$$

where h' is the diffeomorphism from U to $h(U)$ determined by h, the operation $(h')_*$ on vector fields is defined as in Section IV.1, and the unit vector fields $\partial/\partial x_i$ are defined as above. Let ω be a smooth function on U such that $\omega = 1$ near y and $\omega = 0$ off some open neighborhood V of y whose closure is compact and contained in $h(U)$. Then the vector fields $\omega \cdot Y_i$ vanish off the closure of V and thus one can extend these to smooth vector fields on all of M by setting them equal to zero on $M - \overline{V}$. We shall call these extended vector fields X_i. Let $W \subset V$ be an open subneighborhood of y in V such that $\omega = 1$ on W. By construction $Y_i(z) = X_i(z)$ for all i and all $z \in W$, and hence the smooth vector fields $X_i(z)$ form a basis of $T_z(M)$ for all $z \in W$.■

In contrast, it is not always possible to define a collection of smooth vector fields on a given smooth manifold M such that $X_i(z)$ form a basis for $T_z(M)$ at each point $z \in M$. The following basic result in topology shows that problems arise already when one tries to find a **single** vector field on S^2 that is everywhere nonzero.

THEOREM. *If X is a smooth vector field on S^2, then there is some point $p \in S^2$ such that $X(p) = 0$.* $\diamondsuit$

In fact, a similar result holds for all even-dimensional spheres. A proof of the result for S^2 using a minimum of heavy machinery appears in the paper by J. Milnor cited below, and a quick proof for all even-dimensional spheres using some algebraic topology is posted in the online PlanetMath site which is also listed below:

[1] J. W. Milnor, *Analytic proofs of the "hairy ball theorem" and the Brouwer fixed-point theorem*, Amer. Math. Monthly **85** (1978), 521–524.

`http://planetmath.org/encyclopedia/HairyBallTheorem.html`

RESULTS FOR ODD-DIMENSIONAL SPHSERES. (‡) As noted in the PlanetMath site, every odd-dimensional sphere has a nowhere zero tangent vector field, and it is given as follows: View $\mathbf{R}^{2n+2}$ as $\mathbf{C}^{n+1}$, and consider the smooth map J from the latter to $T(\mathbf{R^{2n+2}}) \cong (\mathbf{C}^{n+1})^2$ sending a complex n-dimensional vector $\mathbf{v}$ to the pair $(\mathbf{v},\, i\,\mathbf{v})$. This map sends a point on the unit sphere S^{2n+1} into a vector $(\mathbf{v},\, \mathbf{w}) \in \mathbf{R}^{4n+4}$ such that $\langle \mathbf{v},\, \mathbf{w} \rangle = 0$, and therefore it defines a smooth map from S^{2n+1} to $T(S^{2n+1})$ that is a vector field. Since $\mathbf{v} \neq \mathbf{0} \Longrightarrow i\,\mathbf{v} \neq \mathbf{0}$, it follows that this vector field is nowhere zero.

Using quaternions one can similarly prove that if $n = 4k+3$ for some $k \geq 0$ then S^n has three globally defined smooth vector fields that are everywhere linearly independent (and one can even go a step further, using the Cayley numbers or octonions to define seven everywhere linearly independent vector fields on S^{8k+7} for all $k \geq 0$). On the other hand, breakthrough results from the middle of the twentieth century showed that if $n = 4k+1$ then one cannot find a pair of vector fields X, Y on S^n that are everywhere linearly independent. There are several different proofs of this, and they illustrate some major themes in topology and its relation to algebra and analysis; in particular, the paper listed below by Steenrod and Whitehead involves the interaction between topology and algebra, while the brief monograph by Atiyah reflects some important relationships between topology and the theory of elliptic partial differential equations.

[1] M. F. Atiyah, Vector fields on manifolds. Arbeitsgemeinschaft für Forschung des Landes Nordrhein-Westfalen, Heft 200, Westdeutscher Verlag, Cologne, 1970, 26 pp.

[2] N. E. Steenrod and J. H. C. Whitehead, *Vector fields on the n-sphere*, Proc. Nat. Acad. Sci. U. S. A. **37** (1951), 58–63.

The preceding discussion generates two questions.

(1) Given a smooth n-manifold M, under what conditions is it possible to find a set of n smooth vector fields X_i such that the tangent vectors $X_i(z)$ form a basis of $T_z(M)$ for each $z \in M$? A manifold admitting such a collection of vector fields is said to be *parallelizable*.

(2) What is the maximum number of everywhere linearly independent vector fields that a given smooth manifold can support? In particular, what is this maximum number for S^n?

The answers to these questions are completely known for spheres. In particular, S^n is parallelizable if and only if $n = 0, 1, 3, 7$, and the maximal number of linearly independent vector fields is given as follows: Given a positive integer n, there is a unique expression for $n+1$ of the form

$$n+1 \quad = \quad 2^{4s+r}\, q$$

where q is an odd integer and r and s are nonnegative integers with $r < 4$. The maximum number of everywhere linearly independent vector fields on S^n is then given by

$$2^r \ + \ 8s \ - \ 1 \ .$$

The existence of this many linearly independent vector fields is essentially an algebraic fact that comes out of research by A. Hurwitz and J. Radon in the late nineteenth and early twentieth centuries. Subsequent work near the middle of the twentieth century produced cleaner, more accessible and more elegant proofs. The following online article by M. Shapiro provides explicit references and an excellent description of both the results and their historical background.

`www.math.ohio-state.edu/ shapiro/lec1.pdf`

Proving that there cannot be any more everywhere linearly independent vector fields on S^n requires techniques from algebraic topology, and the final result was established by J. F. Adams around 1960. Further discussion and references are given on pages 5–6 of Conlon (see also the bibliography on pages 403–404).

IV.2 : Global flows and completeness

(Conlon, §§ 2.7–2.8, 4.1)

We have already listed several online sites with useful discussions and graphics (some interactive) involving vector fields and their integral flows. We shall begin this section by listing one more:

`http://www.vias.org/simulations/simusoft_vectorfields.html`

In Section II.4 we noted that a vector field on an open subset of $\mathbf{R}^n$ determines a special type of ordinary differential equation (or system of such equations) $\mathbf{x}'(t) = \mathbf{F}(\mathbf{x})$ that were called *autonomous*; *i.e.*, the expression on the right hand side of the equation is independent of t. We also mentioned that the integral curves of such differential equations have important special properties. In this section we shall discuss this further.

IV.2.1 : *Integral flows and local* 1*-parameter groups*

The solution curves for an autonomous ordinary differential equation (or system of such equations) satisfy the following simple identity:

PROPOSITION. *Let U be open in $\mathbf{R}^n$, and let $F : U \to \mathbf{R}^n$ be a smooth function. Suppose that Φ defines the integral curves of F. Given an ordered pair (t, x) which lies in the open set $\mathcal{D}(\mathbf{R}\times U)$ on which Φ is defined (see Section* II.4 *for background and definitions), let $y = \Phi(x, t)$, and suppose that the maximal integral curve for F with initial condition y is defined on an interval containing*

(a,b). Then the maximal integral curve for x is defined on an interval containing $(a+t, b+t)$ and we have

$$\Phi(s,y) \quad = \quad \Phi\big(s,\,\Phi(t,x)\big) \quad = \quad \Phi(s+t,x)$$

for all $s \in (a,b)$.

Proof. Let β be the maximal integral curve for F with initial condition x. The desired equation is equivalent to saying that $\alpha(s) = \beta(s+t)$ is an integral curve for F with initial condition $y = \Phi(t,x)$. The second part follows immediately because $\alpha(0) = \beta(t) = y$. To verify the first part, note that by definition we have

$$\alpha'(s) \quad = \quad \frac{\partial}{\partial s}\Phi(s+t,x) \quad = \quad \frac{du}{ds}\cdot\frac{\partial}{\partial u}\Phi(u,x)$$

where the second equation follows by the Chain Rule if we make the substitution $u = s+t$. We may simplify this further using $u'(s) = 1$ and the definition of solution curves to show that the right hand side of the previous display is equal to

$$\beta'(u) \quad = \quad \beta'(s+t) \quad = \quad F\big(\Phi(s+t,x)\big) \quad = \quad F\circ\beta(s+t) \quad = \quad F\circ\alpha(s)$$

and combining this with the previous display we conclude that $\alpha(s)$ is the solution curve with the desired properties.■

This identity turns out to be fundamentally important in the study of autonomous (ordinary) differential equations. It is easy to construct examples of nonautonomous systems for which this identity does not hold. Some examples are described in the exercises for this section.

The preceding result implies that Φ is a special case of the following abstract topological concept:

Definition. Given a space Y and a continuous map $\Phi : \mathcal{D} \to Y$ defined on an open neighborhood $\mathcal{D}$ of $Y \times \{0\}$ in $Y \times \mathbf{R}$, we say that Φ is a *local 1-parameter group* if it satisfies the following conditions:

(*i*) $\Phi(y,0) = y$ for all $y \in Y$.

(*ii*) If U is an open subset of Y such that $U \times \{t\} \subset \mathcal{D}$ for some real number t, then $\varphi_t(u) = \Phi(u,t)$ maps U homeomorphically onto an open subset of Y.

(*iii*) If $v = \Phi(u,t) = \varphi_t(u)$ and $w = \Phi(v,s) = \varphi_s(v)$ are defined, then $(u, t+s) \in \mathcal{D}$ and $w = \varphi_{t+s}(u)$.

The first condition can be rewritten $\varphi_0 = \mathrm{id}_Y$, and the third can be rewritten informally as $\varphi_{t+s} = \varphi_s\circ\varphi_t$. For every point $y \in Y$ there is an open neighborhood U of y and an interval $(-\varepsilon,\varepsilon)$ such that $U \times (-\varepsilon,\varepsilon) \subset \mathcal{D}$, and on this neighborhood one can informally write $\varphi_t^{-1} = \varphi_{-t}$.

Since the maps φ_t are diffeomorphisms if Φ is given by the flow curves solving an autonomous differential equation, in our situation we actually have a **smooth** local 1-parameter group.

Globalization to smooth manifolds. Most of the basic theory is the same as in the special case of open subsets of Euclidean spaces. It will be convenient to start by isolating an elementary but crucial step in the proof.

LEMMA. *Let $U, V \subset \mathbf{R}^n$ be open, let X be a smooth vector field on U, and let $h : U \to V$ be a diffeomorphism. If $\Phi : \mathcal{D} \to U$ denotes the integral flow of X, then the integral flow of h_*X is given by*

$$\Psi(t,y) \quad = \quad h\circ\Phi\big(t,\,h^{-1}(y)\big)$$

for all $(t, y) \in \mathrm{id}_{\mathbf{R}} \times h(\mathcal{D})$.

Proof. Let $F : U \to \mathbf{R}^n$ be the second coordinate of the vector field X. Then the second coordinate of the vector field $h_* X$ is given by

$$G(v) \quad = \quad \bigl[Dh\bigl(\, h^{-1}(v)\,\bigr)\,\bigr]\, F\bigl(\, h^{-1}(v)\,\bigr) \ .$$

The proof of the flow assertion reduces to showing that if γ is an integral curve for X with initial condition u, then the curve $\beta = h \circ \gamma$ is an integral curve for $h_* X$ with initial condition $h(u)$. By construction $\beta(0) = h \circ \gamma(0) = h(u)$, so it remains to verfy that β' has the right form. Once again, we do this by a string of equations; the equality of the third and fourth expressions follows from the Chain Rule, while the equality of the fourth and fifth follows because $\beta = h \circ \gamma \Longrightarrow \gamma = h^{-1} \circ \beta$.

$$\beta(t) \quad = \quad [h \circ \gamma]'\,(t) \quad = \quad \bigl[Dh\bigl(\,\gamma(t)\,\bigr)\bigr]\,\gamma'(t) \quad = \quad \bigl[Dh\bigl(\,\gamma(t)\,\bigr)\bigr]\,F(\,\gamma(t)\,) \quad =$$

$$\bigl[Dh\bigl(\, h^{-1} \circ \beta(t)\,\bigr)\bigr]\,F(\, h^{-1} \circ \beta(t)\,) \quad = \quad G\bigl(\,\beta(t)\,\bigr)$$

It follows that $h \circ \Phi(t, x) = \Psi(t,\, h(x)\,)$, which is equivalent to the formula in the conclusion of the lemma.■

Notation. Given a smooth curve $\gamma : J \to M$ defined on some interval J with values in a smooth manifold M, we define the map $\gamma' : J \to T(M)$ to be the smooth curve $T(\gamma) \circ \partial_t$, where ∂_t is the smooth vector field $\partial/\partial t$ described in the previous subsection. The definition is structured so that $\gamma'(t)$ gives the tangent vector to the curve γ at parameter value t. Note that by construction we have $\tau_M \circ \gamma' = \gamma$ (verify this!).

THEOREM. *Let M be a smooth manifold, and let X be a smooth vector field on M. Then there is a smooth local 1-parameter group Φ defined on an open set $\mathcal{D} \subset \mathbf{R} \times M$ such that for each $x \in M$ the smooth curve $\gamma_x(t) = \Phi(t, x)$ is the maximal integral curve for X with initial condition x.*

Sketch of proof. Given a smooth chart (U_α, h_α), let Y_α be the vector field

$$\bigl(h_\alpha^{-1}\bigr)_*\, X|\, h_\alpha(U_\alpha) \ .$$

Suppose we are given $x_0 \in M$. For each α let Γ_α be the local 1-parameter group associated to Y_α. Then $\Phi_\alpha(t, y) = \Gamma_\alpha(t,\, h_\alpha^{-1}(y)\,)$ defines a smooth local 1-parameter group such that for each y, the tangent vectors to the curves $\xi(t) = \Phi_\alpha(t, y)$ are given by $X(y)$. We claim that locally ξ does not depend upon the choice of α. This follows from the local case because on $h_\alpha(U_alpha) \cap h_\beta(U_\beta)$ the local vector fields Y_α and Y_β are related by the identity $(\psi_{\beta\alpha})_* Y_\alpha = Y_\beta$, and therefore by the lemma the local 1-parameter groups are related by an identity of the form

$$\psi_{\beta\alpha} \circ \Gamma(t, x) \quad = \quad \Gamma_\beta\bigl(\, t,\, \psi_{\beta\alpha}(x)\,\bigr) \ .$$

The preceding discussion yields a local existence and uniqueness result for solution curves to the global differential equation $\gamma' = X \circ \gamma$; we needed the lemma to show that different choices of charts lead to the same local solution on M. One can piece them together, exactly as in the case for an open subset $U \subset \mathbf{R}^n$, to form the desired smooth local 1-parameter group on M.■

We can also come full circle and use the theorem to generalize the previous lemma to smooth vector fields on arbitrary smooth manifolds.

COROLLARY. *Let M and N be smooth manifolds, let X be a smooth vector field on U, and let $h : M \to N$ be a diffeomorphism. If $\Phi : \mathcal{D} \to M$ denotes the integral flow of X, then the integral flow of $h_* X$ on N is given by*

$$\Psi(t,y) \quad = \quad h^\circ \Phi\big(t,\, h^{-1}(y)\big)$$

for all $(t,y) \in \mathrm{id}_{\mathbf{R}} \times h(\mathcal{D})$.

The preceding corollary has another consequence that will be useful in Section IV.3.

INVARIANCE CRITERION. *Let X be a smooth vector field on a smooth manifold M, let Φ be the local 1-parameter group of X, and let $h : M \to M$ be a diffeomorphism. Then $h_* X = X$ if and only if the flow curves φ_t all satisfy equations of the form $\varphi_t {}^\circ h = h^\circ \varphi_t$ (for all t).*■

Stationary points of flows. If X is a smooth vector field on the smooth manifold M and $X(x) = 0$ for some $x \in M$, then the constant curve through x trivially satisfies the differential equation for trivial reasons, and by the local uniqueness properties of solutions it follows that *the integral curve $\Phi(t,x)$ must be the constant curve through x.* In other words, the point X is not moved by the diffeomorphisms in the local 1=parameter group – *i.e.*, it is *stationary*. Therefore our preceding observations about vector fields on even-dimensional spheres have the following consequence:

PROPOSITION. *If X is a smooth vector field on S^{2n} for some n, then X has a stationary point.*■

It might be worthwhile to say a little about the physical meaning of this. We have already mentioned that vector fields may be used to view certain laws of physics in an abstract mathematical setting. In particular, one might try to model the motion of wind on the surface of the earth is describable by some smooth vector field. The proposition then implies that there must be a point on the earth's surface at which the wind is calm. Strictly speaking the discussion of this paragraph only applies to situations where the velocity at each point is independent of time, but one can also extend the conclusion to a time-dependent vector field; *i.e.*, a smooth map $X : ((-\varepsilon, \varepsilon) \times \mathbf{M} \to \mathbf{T}(\mathbf{M})$ such that $\tau_M {}^\circ X$ is projection onto the first coordinate. The details for this are outlined in the exercises.

IV.2.2 : *Completeness of integral flows*

For many basic examples of differential equations, one knows that $\mathcal{D}(M)$ is equal to all of $\mathbf{R} \times M$. In particular, this is true for the linear differential equation (system)

$$\mathbf{x}' \quad = \quad A \cdot \mathbf{x}$$

where A is a square matrix. However, it is also not difficult to find examples of differential equations for which $\mathcal{D}(M) \neq \mathbf{R} \times M$. One specific example in Conlon (specifically, 2.8.13) is given by the 1-dimensional equation $x' = e^x$. Standard methods from elementary differential equations courses show that the general solution of this curve has the form $x(t) = -\log(t - K)$ for some integration constant K, and the associated curves can only be defined when $t > K$.

One important reason for wanting to know whether $\mathcal{D}(M) = \mathbf{R} \times M$ is that the latter implies each φ_t is a **global** diffeomorphism from M to itself. More generally, if Φ is an arbitrary local 1-parameter group, then each φ_t is a global homeomorphism from the associated space Y to itself if $\mathcal{D} = \mathbf{R} \times Y$.

Definition. If Y is a topological space and Φ is 1-parameter group defined on Y, then Φ is said to be a *global* 1*-parameter group* if $\mathcal{D} = \mathbf{R} \times Y$. A smooth vector field X on a smooth manifold M is said to be **complete** if its associated flow Φ defines a global 1-parameter group.

The following is a straightforward generalization of Lemma 4.1.10 on page 133 of Conlon:

PROPOSITION. *Let Φ be a local 1-parameter group on the topological space Y that is defined on the open set $\mathcal{D} \subset \mathbf{R} \times Y$. If $\mathcal{D}$ contains $Y \times (-\varepsilon, \varepsilon)$ for some $\varepsilon > 0$, then $\mathcal{D} = Y \times \mathbf{R}$.*∎

COROLLARY 1. *If Y is compact then $\mathcal{D} = Y \times \mathbf{R}$.*∎

COROLLARY 2. *If M is a compact manifold and X is a vector field on M, then X is complete.*∎

Compare this with Corollary 4.1.12 on page 134 of Conlon.

The preceding corollary illustrates one feature of the global situation that does not arise for open sets in Euclidean space. All vector fields on a compact manifold are complete. In contrast, for open subsets of Euclidean spaces it is always possible to construct vector fields that are not complete (we have already given examples of noncomplete vector fields for $\mathbf{R}$, and it is not difficult to extend this example to $\mathbf{R}^n$; in other cases the constructions require more work).

Physically speaking, a noncomplete vector field (one where $\mathcal{D} \neq \mathbf{R} \times M$) corresponds to a dynamical system that breaks down or blows up in a finite amount of time. The next proposition can be viewed as a mathematical formulation of this principle:

PROPOSITION. *Suppose that X is a vector field on M and that the maximal integral curve of X with initial condition p is only definable for $t \in (a^-, a^+)$ where $-\infty < a^-$ or $a^+ < +\infty$. If Γ is the image of this curve, then Γ is not contained in any compact subset of M.*

Proof. We first consider the case where $a^+ < +\infty$. Let K be the closure of Γ. By construction we know that Φ maps $\mathcal{D} \cap (\Gamma \times \mathbf{R})$ to itself, so by continuity it must also map $\mathcal{D} \cap (K \times \mathbf{R})$ to itself. It follows that $\Phi | \mathcal{D} \cap K \times \mathbf{R}$ is a local 1-parameter group of transformations on K. If Γ were contained in a compact set, then K would also be compact, and by the preceding corollary it would follow that $\mathcal{D} \cap K \times \mathbf{R}$ would be all of $(K \times \mathbf{R})$. The same would also hold if K were replaced by Γ. But this contradicts the hypothesis on Γ, and therefore Γ cannot lie in any compact subset of M.

On the other hand, if $-\infty < a^-$, then one can retrieve the result from the previous case by considering the reverse vector field $-X$, whose flow is given by $\Psi(u, t) = \Phi(u, -t)$.∎

IV.3 : Lie brackets

(Conlon, §§ 2.7, 2.8, 4.3)

Explicit formulas for vector fields are indispensable for computations, but for many conceptual and theoretical purposes it is better to characterize them in other ways. In particular, one can define directional derivatives for smooth functions along a vector field and study the latter in terms of the associated differentiation operator. It is not difficult to show that different vector fields define different different differentiation operators, and one basic result of this section establishes a 1–1 correspondence between between vector fields and a purely algebraically defined class of operations.

One particular advantage of the algebraic characterization is that it yields important algebraic structures and identities; the **Lie bracket** is a particularly significant and fundamental example of this type.

The first step is to consider a class of algebraic operations that have the basic formal properties of differentiation operations.

IV.3.1 : *Algebraic abstractions*

If $\mathbf{k}$ is a field, then an *algebra over* $\mathbf{k}$ consists of a vector space A over $\mathbf{k}$ and a multiplication map $m : A \times A \to A$ with the following properties:

$a(b+c) = ab + ac$ for all a, b, c in A.

$(a+b)c = ac + bc$ for all a, b, c in A.

$(ra)b = r(ab) = a(rb)$ for all a, b in A and $r \in \mathbf{k}$.

An algebra is said to be *associative* if $(ab)c = (ab)c$ for all a, b, c in A; various other ring-theoretic conditions are also meaningful to formulate over and algebra (for example, a multiplicative identity), and there are corresponding definitions of algebras with identities, commutative algebras, division algebras, and so forth.

Perhaps the simplest example of an algebra that does not satisfy the associativity condition is $\mathbf{R}^3$ with the usual cross product. This algebra is in fact an example of a system called a **Lie algebra**. Not that Lie is pronounced "lee" and is named after the Norwegian mathematician Sophus Lie.

Definition. An algebra A over the field $\mathbf{k}$ is said to be a *Lie algebra* if it satisfies the following two conditions:

ANTI-COMMUTATIVITY. $ab = -ba$ for all a and b in A.

JACOBI IDENTITY. $a(bc) + b(ca) + c(ab) = 0$ for all a, b, c in A.

There is a standard method for generating Lie algebras from associative algebras:

PROPOSITION. *Let A be an associative algebra over a field $\mathbf{k}$, and define $[a, b]$ to be the commutator $ab - ba$. Then A with respect to the binary operation $[\ ,\ ]$ is a Lie algebra.*

Sketch of proof. Verification of the result involves checking that the commutator satisfies the following properties:

(1) It is distributive with respect to both variables.

(2) It is homogeneous with respect to scalars.

(3) It is anticommutative.

(4) It satisfies the Jacobi identity.

In each case it is an elementary (but sometimes fairly tedious) exercise to check that the given identity is valid.∎

This produces many examples of Lie algebras. In particular, it determines a Lie algebra structure on the $n \times n$ matrices over $\mathbf{k}$. The subset of matrices whose traces are zero forms a Lie subalgebra in the obvious sense (a vector subspace and closed under forming commutators).

Similarly, the set of all skew-symmetric matrices turns out to form a Lie subalgebra (check this out!). The same is true for the set of skew-symmetric matrices.

IV.3.2 : *Derivations*

None of the preceding discussion suggests any reason why Lie algebras might be relevant to the study of smooth manifolds. The crucial ideas are contained in the following definition and proposition:

Definition. If A is an algebra over $\mathbf{k}$, the a *derivation* of A is a linear transformation D from A to itself that satisfies the *Leibniz identity*:

$$D(ab) \quad = \quad a\,(Db) \; + \; (Da)\,b \qquad \text{for all } a \text{ and } b \text{ in } A$$

The most basic examples of derivations are given by partial differentiations on the algebra $\mathcal{C}^\infty(U)$ of functions on an open subset $U \subset \mathbf{R}^n$ with continuous partial deriviatives of all orders. In fact, if D_i denotes partial differentiation with respect to the i^{th} variable and g_i is a smooth function on U then $\sum_i \, g_i \, D_i$ is also a derivation.

The set of derivations on an algebra A is a vector subspace of the algebra $\mathcal{L}(A, A)$ of $\mathbf{k}$-linear transformatons from A to itself. However, one can also draw a much stronger conclusion:

ABSTRACT DERIVATION THEOREM. *If A is an associative algebra over the field $\mathbf{k}$ and $\mathcal{D}(A)$ is the set of derivations on A, then A is a Lie subalgebra with respect to the vector space operations on linear transformations from A to itself and the commutator product.*

Sketch of proof. We need to show that $\mathcal{D}(A)$ is a vector subspace of the space of all linear transformations from A to itself and that the commutator of two derivations is also a derivation. In each case the crucial point is to verify that the Leibniz identity holds; *i.e.*, if D_1 and D_2 are derivations and $r \in \mathbf{k}$ is a scalar, then each of the linear transformations $D_1 + d_2$, $r\,D_1$ and $[D_1,\, D_2]$ satisfies the Leibniz identity. In each case this is a routine and elementary calculation; not surprisingly, the proof for the commutator is the longest one.

If we apply this to our example $\mathcal{C}^\infty(U)$ we conclude that the space of derivations on $\mathcal{C}^\infty(U)$ is a Lie algebra with respect to the commutator product.

Warning. Given a derivation D on $\mathcal{C}^\infty(U)$ and a function g in $\mathcal{C}^\infty(U)$, one can define a new derivation $g \cdot D$ by the formula $[g \cdot D](f) = g\,(Df)$, where the right hand side is merely the product of two functions. This defines an analog of scalar multiplication by $\mathcal{C}^\infty(U)$ on the space of derivations — formally, a *module* structure — but in general the functions $[gD_1,\, D_2]$, $[D_1,\, gD_2]$ and $g\,[D_1, D_2]$ are unequal. In fact, one has the following identity:

$$[fX,\, gY] \quad = \quad fg\,[X,Y] \; + \; f\,(Xg)\,Y \; - \; g\,(Yf)\,X$$

We shall need the following generalization of the fact that the (partial) derivatives of constant functions are zero.

PROPOSITION. *If the algebra A has a two-sided unit element 1 and $r \in \mathbf{k}$ is arbitrary, then $D(r\,1) = 0$.*

Proof. Since $D(r\,1) = r\,D(1)$ it suffices to show that $D(1) = 0$. But $1 \cdot 1 = 1$ implies that $D(1) = D(1 \cdot 1) = 2\,D(1)$. Subtraction of $D(1)$ from both sides of the equation yields $D(1) = 0$.■

If U is an open subset of $\mathbf{R}^n$ then a smooth vector field is essentially a smooth n-dimensional vector valued function on U. In view of the preceding examples, if we are given a smooth vector field $X(y)) = (y, \mathbf{F}(y))$ which can be expressed in coordinate form as $\sum_j f_j(x)\,\mathbf{e}_j$, where $\mathbf{e}_j$ is the standard j^{th} unit vector, then there is an associated derivation $D_{\mathbf{F}}$ defined by the formula

$$D_{\mathbf{F}}(g) \quad = \quad \sum_j f_j \frac{\partial g}{\partial x_j} \ .$$

This directional derivative is also called the **Lie derivative** of f with respect to the vector field X and usually one writes Xg or $\mathcal{L}_X(g)$ to denote the function we have called $D_{\mathbf{F}}(g)$.

Now the derivation corresponding to the j^{th} unit vector is partial differentiation with respect to the j^{th} variable, and THIS *is why we have denoted the the vector field whose value is always* $\mathbf{e}_j$ *by the symbol*

$$\frac{\partial}{\partial x_j} \ .$$

It follows immediately that *every vector field on* U *can be expressed uniquely as a linear combination of these basic vector fields with coefficients in* $\mathcal{C}^\infty(U)$.

For our purposes, it is important to know there is a converse to this statement:

FACT. *Every derivation of* $\mathcal{C}^\infty(U)$ *has the form* $D_{\mathbf{F}}$ *for some unique vector field* $\mathbf{F}$.

This result is established in Conlon; although the statement of the result is clean, the details of the proof are somewhat delicate in several places, so in these notes we shall prove an alternate formulation of the result that is slightly more complicated to state but easier to prove and also adequate for the purposes of this course.$\diamondsuit$

If U and V are open subsets of $\mathbf{R}^n$ such that $V \subset U$, then there is a restriction map

$$i^*_{V,U} : \mathcal{C}^\infty(U) \longrightarrow \mathcal{C}^\infty(V)$$

sending a smooth function f on U to its restriction $f|V$, and likewise if X is a smooth vector field on U one can define the restriction $X|V$. We shall be interested in systems of derivations with compatibility properties similar to those of vector fields and their restrictions:

Definition. Let U be an open subset of $\mathbf{R}^n$. A *compatible system of derivations* on the family of algebras $\mathcal{C}^\infty(V)$, where V runs through all the open subsets of U, consists of a collection of derivations D^V on the function algebras $\mathcal{C}^\infty(V)$ such that

$$\big(D^V(f)\big)\,|W \quad = \quad D^W\,(f|W)$$

for all open subsets such that $W \subset V$.

LOCAL REALIZATION THEOREM. *Every compatible system of derivations on the open subsets of* U *is defined by a unique vector field on* U.

Proof. $(\star)$ We shall begin by verifying that if $D_{\mathbf{F}} = 0$ then $\mathbf{F}$ is the zero vector field. As before, if we write

$$D_{\mathbf{F}} \quad = \quad \sum_j f_j \frac{\partial}{\partial x_j}$$

then $D_{\mathbf{F}}$ of the k^{th} coordinate function x_k is equal to f_k, and therefore $D_{\mathbf{F}} = 0$ implies that $f_k = 0$ for all k, which in turn implies that $\mathbf{F} = \mathbf{0}$.

Next, suppose that E^V is an arbitrary system of compatible derivations on the open subsets of U. Let x_j be the j^{th} coordinate function as usual, and set $g_j = E^U(x_j)$. We claim that

$$E^W \quad = \quad \sum_j \, g_j \cdot \frac{\partial}{\partial x_j}$$

on every open subset $W \subset U$. This amounts to showing that

$$E^W f(p) \quad = \quad \sum_j g_j(p) \cdot \frac{\partial f}{\partial x_j}(p)$$

for each smooth function f defined on an open subset W and each point $p \in W$.

Let f and W be as in the preceding sentence, and choose $\varepsilon(p) > 0$ such that $V = N_{\varepsilon(p)}(p) \subset W$. As in Lemma 2.2.20 on page 48 of Conlon, the function $f|V$ may be written as a sum

$$\sum_j \; h_j(x) \cdot (x_j - p_j)$$

where h_j is a smooth function on V and p_j is the j^{th} coordinate of p. Applying E^V to this and remembering that a derivation sends a constant function to zero, we find that

$$E^V f|V \quad = \quad \sum_j \; E(h_j) \cdot (x_j - p_j) \; + \; h_j(x) \cdot E(x_j - p_j) \quad = \quad \sum_j \; E(h_j) \cdot (x_j - p_j) \; + \; h_j \cdot g_j \; .$$

If we evaluate this at p we find that

$$E^W f(p) \quad = \quad E^V f|V(p) \quad = \quad \sum_j \; E(h_j)(p) \cdot (p_j - p_j) \; + \; h_j(p) \cdot g_j(p) \; .$$

Now the first term of each summand clearly vanishes, and therefore we have that

$$E^W f(p) \quad = \quad \sum_j \; h_j(p) \cdot g_j(p) \; .$$

On the other hand, we also have that

$$E_{\mathbf{F}} f(p) \quad = \quad \sum_j \; g_j \, \frac{\partial f}{\partial x_j}(p)$$

and the formula for $f|V$ shows that the value of the j^{th} partial derivative of f at p is equal to $h_j(p)$. Therefore we have shown that $E^W f(p) = E_{\mathbf{F}} f(p)$, and since W, f and p are arbitrary this proves that the system E^V is given by $\mathbf{F}$.

COROLLARY. *If U is open in $\mathbf{R}^n$ and $\mathbf{F}$ and $\mathbf{G}$ define smooth vector fields on U, then there is a unique vector field $\mathbf{H}$ such that $D_{\mathbf{H}} = [D_{\mathbf{F}}, D_{\mathbf{G}}]$.*

Proof. The commutator of $D_{\mathbf{F}}$ and $D_{\mathbf{G}}$ defines a system of derivations on the algebras of continuous functions, and therefore this commutator has the of $D_{\mathbf{H}}$ for some unique vector field $\mathbf{H}$.

Here is a general formula for the Lie bracket of two vector fields in terms of its factors:

COORDINATE FORMULA. If we have

$$D_{\mathbf{F}} \;=\; \sum_j f_j \frac{\partial}{\partial x_j} \quad , \quad D_{\mathbf{G}} \;=\; \sum_j g_j \frac{\partial}{\partial x_j}$$

then the commutator is given by the following identity:

$$[D_{\mathbf{F}}, D_{\mathbf{G}}] \;=\; \sum_j \left(\sum_i f_i \frac{\partial g_j}{\partial x_i} - g_i \frac{\partial f_j}{\partial x_i} \right) \cdot \frac{\partial g_j}{\partial x_i}$$

IV.3.4 : *Globalization to smooth manifolds*

The notions of Lie derivative and Lie bracket generalize to arbitrary smooth manifolds, and one also has a characterization of vector fields over arbitrary manifolds in terms of systems of derivations on the smooth function algebras. We shall now explain how one globalizes all these definitions and results.

The first step in globalizing the Lie bracket is to show that it behaves well with respect to smooth changes of coordinates. Suppose that U and V are open in $\mathbf{R}^n$, let $\varphi : U \to V$ be a diffeomorphism, and let $\mathbf{X}(U)$ and $\mathbf{X}(V)$ denote the spaces of vector fields on U and V respectively. We had previously constructed a vector space isomorphism

$$\varphi_* : \mathbf{X}(U) \longrightarrow \mathbf{X}(V)$$

which takes a vector field X to the composite $T(\varphi) \circ X \circ \varphi^{-1}$, and the desired invariance property of the Lie bracket is given by the following result:

THEOREM. *The vector space isomorphism φ is an isomorphism of Lie algebras; formally,*

$$[\varphi_* X, \varphi_* Y] \;=\; \varphi_*[X, Y]$$

for all vector fields X and Y in $\mathbf{X}(U)$.

Proof. $(\star)$ We shall need the following result relating Lie derivatives and φ_*:

$$[\varphi_* X]\, f \;=\; [X(f \circ \varphi)] \circ \varphi^{-1}$$

Verification of this identity is a routine exercise that is left to the reader.

The proof of the theorem is an immediate consequence of the following sequence of equations:

$$\mathcal{L}_{\varphi_*[X,Y]} f \;=\; \left[\mathcal{L}_{[X,Y]} f \circ \varphi\right] \circ \varphi^{-1} \;=\; \left[\mathcal{L}_X \mathcal{L}_Y f \circ \varphi\right] \circ \varphi^{-1} \;-\; \left[\mathcal{L}_Y \mathcal{L}_X f \circ \varphi\right] \circ \varphi^{-1} \;=$$

$$\left[\mathcal{L}_X \left(\mathcal{L}_Y f \circ \varphi\right) \varphi^{-1} \varphi\right] \circ \varphi^{-1} \;-\; \left[\mathcal{L}_Y \left(\mathcal{L}_X f \circ \varphi\right) \varphi^{-1} \varphi\right] \circ \varphi^{-1} \;=$$

$$\left[\mathcal{L}_X \left(\mathcal{L}_{\varphi_* Y} f\right) \varphi\right] \varphi^{-1} \;-\; \left[\mathcal{L}_Y \left(\mathcal{L}_{\varphi_* X} f\right) \varphi\right] \varphi^{-1} \;=$$

$$\mathcal{L}_{\varphi_* X} \left(\mathcal{L}_{\varphi_* Y} f\right) \;-\; \mathcal{L}_{\varphi_* Y} \left(\mathcal{L}_{\varphi_* X} f\right) \;=\; \left[\mathcal{L}_{\varphi_* X}, \mathcal{L}_{\varphi_* Y}\right] f \; .$$

The final expression in this sequence is merely the Lie derivative of f with respect to $[\varphi_* X, \varphi_* Y]$.■

The next order of business is to define Lie derivatives globally. The Lie derivative $\mathcal{L}_X f$ of a smooth function f on the open set U along the smooth vector field X is another smooth function on U, and there are two basic formulas for defining the value of the Lie derivative $\mathcal{L}_X f$ at a point $p \in U$:

> $\mathcal{L}_X f(p) = \pi_2 T(f)[\,X(p)\,]$, where $\pi_2 : T(\mathbf{R}) \to \mathbf{R}$ is the projection map that is a linear isomorphism on the tangent space of each point, or equivalently projection onto the second coordinate under the standard identification of the tangent space projection with the first coordinate projection from $\mathbf{R} \times \mathbf{R}$ to $\mathbf{R}$.
>
> $\mathcal{L}_X f(p) = [f \circ \Phi_p]'(0)$, where Φ_p is the unique integral curve of X with initial condition $\Phi_p(0) = p$.

This equivalence is mentioned in Definition 2.8.15 on page 76 of Conlon.

It is a routine exercise to check that this reduces to the earlier definition if U is an open subset of $\mathbf{R}^n$. Furthermore, if M is a smooth manifold and X is a smooth vector field on M then the Lie derivative associates a derivation $\mathcal{L}_X$ of $\mathcal{C}^\infty(M)$ to a smooth vector field M, and one can define compatible systems of derivations just as in the special case of open subsets of $\mathbf{R}^n$.

GLOBAL REALIZATION THEOREM. *Every compatible system of derivations on the open subsets of M is defined by a unique vector field on M.*

Proof. $(\star)$ Let $\mathcal{U} = \{\, W_\alpha \,\}$ be an open covering of M by smooth coordinate charts, and let k_α be a diffeomorphism from U_α to an open subset U_α of $\mathbf{R}^n$. Denote the derivation on $\mathcal{C}^\infty(W_\alpha)$ by E_α, and define a new derivation on $\mathcal{C}^\infty(U_\alpha)$ by

$$E^*_\alpha(f) \quad = \quad \big[E_\alpha(f \circ k_\alpha)\big] \circ k_\alpha^{-1} \ .$$

By the local realization theorem we know that E^*_α is given by a vector field X_α on U_α. We claim that we can piece together these local vector fields to obtain a global vector field and that this global vector field defines the entire system of derivations.

Let $h_\alpha : U_\alpha \to W_\alpha$ be a diffeomorphism inverse to k_α. Then direct calculation shows that $(h_\alpha)_* X_\alpha$ and E define the same system of derivations on W_α. Given a second set W_β in the original open covering, the same considerations show the existence of a vector field $(h_\beta)_* X_\beta$ such that $(h_\beta)_* X_\beta$ and E define the same system of derivations on W_β. In particular, the restrictions of both vector fields to $W_\alpha \cap W_\beta$ define the same system of derivations as E on this intersection, so by the uniqueness portion of the Local Realization Theorem we know that the restrictions of both $(h_\alpha)_* X_\alpha$ and $(h_\beta)_* X_\beta$ to $W_\alpha \cap W_\beta$ must be identical. But this implies that there is a global vector field Y on M whose restriction to W_α is $(h_\alpha)_* X_\alpha$ for all α.

We now need to prove that E and Y determine the same system of derivations on M, or equivalently that for each open subset $V \subset M$ they determine the same derivation on V. Suppose that $f : V \to \mathbf{R}$ is a smooth function. Then the functions Ef and Yf are equal if and only if their restrictions to each of the subsets $W_\alpha \cap V$ are equal. However, the restriction of Y to W_α is Y_α, so it follows that Y and Y_α determine the same derivation as E and E_α on $V \cap W_\alpha \subset W_\alpha$. Therefore we conclude that the restrictions of Ef and Yf to W_α are equal for each α, and as noted before this means that $Ef = Yf$ for all f, which in turn implies that the system of derivations E is given by the vector field Y.

The preceding argument shows existence; at this point uniqueness is very straightforward to prove. Suppose Y and Z are two smooth vector fields determining the same system of derivations. Then for all α they determine the same derivation on W_α, and by the Local Realization Theorem it follows that the restrictions of Y and Z to W_α are equal for all α. Since the open sets W_α cover M, it follows that $Y = Z$.■

Comparison of results. The comparable result in Conlon states that every derivation on $\mathcal{C}^\infty(M)$ is given by Lie differentiation with respect to some vector field; thus Conlon's hypothesis is weaker than ours. However, a crucial step in proving the Conlon's version is to show that if U is open in M then a derivation on $\mathcal{C}^\infty(M)$ gives rise to a compatible derivation on $\mathcal{C}^\infty(U)$, and this takes a fair amount of delicate work. In this course (and Conlon) there are many other important realization theorems that we want to prove, and in each case if one starts with a comparably weak hypothesis the proof will be comparably more difficult. In most contexts where one works with smooth manifolds, there is no problem seeing that the stronger types of hypotheses adopted in these notes are satisfied, so the trade-off yields results with effectively the same usefulness as those in Conlon and notably simpler proofs.

IV.3.5 : *Differentiation formula for the Lie bracket* $(\star)$

In the discussion of Lie derivatives of functions with respect to vector fields, we noted a formula for the Lie derivative as an ordinary derivative related to the local 1-parameter group of the integral flow. Because of its relation to the material of this subsection, we shall state the result in a form which reflects our objectives:

LIE DERIVATIVE FORMULA FOR SMOOTH FUNCTIONS. *Let M be a smooth manifold, let X be a smooth vector field on M, let φ_t denote the flow map(s) associated to the local 1-parameter group Φ, and let $f : M \to \mathbf{R}$ be a smooth function. Then at each point $y \in M$ the Lie derivative Xf is given by the following formula:*

$$\lim_{t \to 0} \frac{1}{t} \cdot \big(f - (\varphi_t)_* f \big)$$

In this expression, $(\varphi_t)_* f(y)$ is equal to $f(\varphi_t^{-1}(y))$.

Once again the proof is left to the reader as an exercise; perhaps the main step is to recall the relationship between X and the integral flow curves.■

The important thing to notice about the displayed expression inside the limit sign is that it remains meaningful it we replace f by a smooth vector field Y. Two questions then arise naturally. First, is there a limit as $t \to 0$? And if so, what is this limit? Given that we are discussing this in a section devoted to the Lie bracket, one might guess that any answers to these questions somehow involve the Lie bracket of X and Y, and in fact one can answer both questions very definitively in such terms:

LIE DERIVATIVE FORMULA FOR VECTOR FIELDS. *Let M be a smooth manifold, let X and Y be smooth vector fields on M, and let φ_t denote the flow map(s) associated to the local 1-parameter group Φ for X. Then at each point $y \in M$ the commutator $[X, Y]$ satisfies the following equation:*

$$[X, Y] \quad = \quad \lim_{t \to 0} \frac{1}{t} \cdot \big(Y - (\varphi_t)_* Y \big)$$

Note. Strictly speaking, the notation $(\varphi_t)_* Y$ is an abuse of language if X is not a complete vector field (in which case φ_t might not be defined everywhere), but we can adjust for this by taking restrictions to open subsets on which the map φ_t is defined. Recall that for all points y and values of t sufficiently close to zero, one can find an open neighborhood on which φ_t is defined.

Before proving the Lie derivative formula it will be convenient to establish two lemmas.

LEMMA 1. *Let $J_\varepsilon = (-\varepsilon, varepsilon)$, let U be an open subset of the smooth manifold M, and suppose that $f : J_\varepsilon \times U \to \mathbf{R}$ is a smooth map such that $f(0,x) = 0$ for all $x \in U$. Then there is a smooth map $g : J_\varepsilon \times U \to \mathbf{R}$ such that $f(t,p) = t\, g(t,p)$ and*

$$g(t,p) = \frac{\partial}{\partial t} f(0,p) \ .$$

Proof. Define g by the formula

$$g(t,p) \quad = \quad \int_0^1 \frac{\partial}{\partial t} f(ts,p)\, ds \ .\blacksquare$$

LEMMA 2. *Let X be a smooth vector field on the smooth manifold M, and denote the diffeomorphisms from the local 1-parameter group by φ_t. Given a smooth function $f : M \to \mathbf{R}$ there is a function $g_t(x) = g(t,x)$ with the following properties on M:*

(*i*) $\quad f \circ \varphi_t \quad = \quad f \ + \ t \cdot g_t$

(*ii*) $\quad g_0 \quad = \quad Xf$

Proof. Let $h(t,x) - f(\,\varphi_t(x)\,) - f(x)$. By the preceding lemma, there is a smooth function g such that $h(t,p) = t\, g(t,p)$ and

$$g(t,p) = \frac{\partial}{\partial t} h(0,p) \ .$$

It then follows that $f \circ \varphi_t \ = \ f + t \cdot g_t$ and

$$\lim_{t\to 0} \frac{1}{t} \big(\, f(\,\varphi_t(x)\,) - f(x)\,\big) \quad = \quad lim_{t\to 0} \frac{f(t,x)}{t} \quad = \quad \lim_{t\to 0} g(t,x) \quad = \quad g_0(x) \ .$$

On the other hand, the left hand side is equal to Xf, and hence $g_0 = Xf$ as required.$\blacksquare$

Proof of the Lie derivative formula for vector fields. Given a smooth function $f : M \to \mathbf{R}$, let g be given as in Lemma 2. We then have

$$\big\{\,[\,(\varphi_t)_* Y\,]\, f\big\}\,(x) \quad = \quad \big[\,Y\,(f \circ \varphi_t)\,)\big]\ \big(\,\varphi_t^{-1}(x)\,\big) \quad = \quad \big[\,Yf \ + \ t \cdot (Y g_t)\big]\ \big(\,\varphi_t^{-1}(x)\,\big)$$

and consequently we have

$$\lim_{t\to 0} \ \frac{1}{t}\ \big\{\,\big[\,Y \ - \ (\varphi_t)_* Y\,\big]\, f\big\}\,(x) \quad = \quad \lim_{t\to 0}\ \frac{1}{t}\ \big[Yf(x) \ - \ Yf(\,\varphi_t^{-1}(x)\,)\big] \quad - \quad \lim_{t\to 0} Y g_t(\,\varphi_t^{-1}(x)\,) \quad =$$

$$\big[X(Yf)\big(x\big) \ - \ \big[Y g_0\big](x) \quad = \quad \big[X(Yf)\big(x\big) \ - \ \big[Y(Xf)\big(x\big) \quad = \quad [X,Y]f(x) \ .$$

Since the two expressions involving vector fields have the same effect on an arbitrary smooth function, they must be equal.$\blacksquare$

COROLLARY. *In the notation above we also have*

$$(\varphi_s)_*[X,Y] \quad = \quad \lim_{t\to 0} \frac{1}{t} \cdot \Big((\varphi_s)_* Y \quad - \quad (\varphi_{s+t})_* Y \Big)$$

for all s.■

These formulas imply that a Lie bracket $[X,Y]$ reflects the interaction between the local 1-parameter groups associated to X and Y. One can make this precise in several different ways. We shall limit ourselves to one of the simplest.

COMMUTING VECTOR FIELDS LEMMA. *Let X and Y be two smooth vector fields on the smooth manifold M. Then the following are equivalent:*

(*i*) *The Lie bracket $[X,Y]$ is the zero vector field.*

(*ii*) *If φ_s and ψ_t are given by the local 1-parameter groups for X and Y respectively, then $\psi_t \circ \varphi_s = \varphi_s \circ \psi_t$ for all (s,t) sufficiently close to $(0,0)$.*

Proof. ($\star$) We begin with the ($\Longleftarrow$) implication. By previous observations, we know that the composite

$$\varphi_s \circ \psi_t \varphi_s^{-1}$$

yields the local 1-parameter group for $(\varphi_s)_* Y$. But the commutativity assumption implies that the displayed expression is equal to ψ_t, and therefore both $(\varphi_s)_* Y$ and Y have the same local flows, which means they must be equal. If we substitute this into the Lie derivative formula for $[X,Y]$, we see that we are taking the limit as $t \to 0$ of something that is equal to zero, and therefore $[X,Y]$ must also be equal to 0.

Turning to the ($\Longrightarrow$) implication, fix a point $p \in M$ and consider the curve in the tangent space of p given by evaluating $(\varphi_s)_* Y$ at p. Since the tangent space is a smooth manifold of M (see the exercises for this section), it follows that this is in fact a smooth curve $\rho(s)$ in $T_p(M)$ with $\rho(0) = Y(p)$. The hypothesis that $[X,Y] = 0$ and the preceding corollary then imply that $\rho'(t)$ is identically zero, and therefore ρ must be a constant; *i.e.*, for each sufficiently small value of s the vector fields Y and $(\varphi_s)_* Y$ have the same value at p. Since p was arbitrary this implies that the vector fields must be equal and therefore must have the same flows. In the previous paragraph we noted that these flows are ψ_t and $\varphi_s \circ \psi_t \varphi_s^{-1}$ respectively, and therefore we have

$$\psi_t \quad = \quad \varphi_s \circ \psi_t \varphi_s^{-1}$$

which is equivalent to the desired identity $\psi_t \circ \varphi_s = \varphi_s \circ \psi_t$.■

IV.4 : Introduction to Lie groups ($2\star$)

(Conlon, §§ 5.1–5.3)

This part of the notes can be found in the file(s) `liegps.*` (where $*$ =`ps` or `pdf`) in the course directory. Although this is an extremely important topic with ties to many branches of mathematics, there is not enough time to cover it in this course.

V. Cotangent spaces and tensor algebra

Two objectives of this unit have already been mentioned briefly. In Unit IV we noted that generalizations of some standard constructions of vector analyis — specifically, the gradient vector field of a smooth function and the divergence and curl of a smooth vector field — require additional mathematical formalism. Part of this is essentially algebra and provides the framework for defining **tensor fields** on manifolds. These are generalizations of vector fields and play a fundamental role in studying the geometrical properties of smooth manifolds and their applications to other areas of mathematics and physics.

The abstract definition of tensors involves concepts from basic linear algebra that are not overwhelmingly difficult but seldom appear in undergraduate courses (there already is plenty of excellent material that deserves inclusion in such courses). Therefore a substantial part of this unit will be devoted to the algebraic construction of tensor spaces and certain related objects that will be needed in Unit VI.

We shall now attempt to describe the contents of this unit more specifically. Given a vector space V over a field $\mathbf{k}$, or more generally a finite collection of vector spaces over a given field, one has associated vector spaces that are called tensor spaces. In particular, if we are given two nonnegative integers r and s and a vector space V, then there is another vector space $\mathrm{Tensor}^r_s(V)$ which is called *the space of tensors with* **contravariant rank** r *and* **covariant rank** s.

If $\dim V = n < \infty$, then by construction we shall have

$$\dim\,[\,\mathrm{Tensor}^r_s(V)\,] \quad = \quad n^{r+s} \ .$$

Frequently the elements of this vector space are called tensors of type (r, s) on V; for every vector space V we set $\mathrm{Tensor}^0_0(V)$ equal to $\mathbf{k}$ by definition. By construction $\mathrm{Tensor}^1_0(V)$ will be equal to V.

Suppose now that M is a smooth n-manifold and $x \in M$; as usual let $T_x(M)$ denote the tangent space to x with respect to M. Then the tensor construction can be applied to each of the vector spaces $T_x(M)$ to yield associated tensor spaces

$$[\mathbf{T}^r_s]_x(M) \quad = \quad \mathrm{Tensor}^r_s\big(\,T_x(M)\,\big)$$

that is indexed by the points of M. Since the tangent spaces $T_x(M)$ can be viewed as a smoothly parametrized family of vector spaces over M, it is natural to ask if the same is true for the family of vector spaces $[\mathbf{T}^r_s]_x(M)]$. We shall prove that these vector spaces fit together smoothly to form a smooth manifold $[\mathbf{T}^r_s]\ (M)]$ called a tensor bundle in a manner quite similar to the way the tangent spaces fit together smoothly to for the tangent bundle of M. All these smoothly parametrized families will be special cases of objects known as **vector bundles**. The latter will be introduced in Section V.1. There is also a concept of **tensor field** of type (r, s) for $[\mathbf{T}^r_s]_x(M)]$ that generalizes the notion of vector field for $T(M)$. Tensor bundles are fundamental examples of a recurrent theme from linear algebra, point set topology, and Unit III of these notes; namely, the construction of new objects of a given type out of old ones. In Section V.2 we shall introduce a few relatively simple methods for constructing new vector bundles out of known examples, and we shall also discuss another theme from linear algebra and point set topology. A second theme in each of the latter subjects is the importance of some additional metric structure. We have also seen the first theme in connection with the theory of smooth manifolds in Unit III of these notes. In Section V.2 we shall introduce some basic methods for constructing new vector bundles out of old ones and also introduce parametrized families of inner products on parametrized families of vector spaces. Such structures, usually known as **riemannian metrics**, play an extremely important role in the study of vector bundles and their applications to smooth manifolds.

Section V.3 is devoted to the most basic example of a tensor bundle aside from $T(M) = [\mathbf{T}^1_0]_x(M)]$ itself; namely, the **cotangent bundle**

$$T^*(M) \quad = \quad [\mathbf{T}^0_1]_x(M)] \ .$$

In fact, it is possible the vector space $\mathrm{Tensor}^0_1(V)$ without mentioning tensors at all, for the latter is the *dual space* that is studied in some undergraduate linear algebra texts. Since there are also many such texts that do not include a discussion of dual spaces, we shall summarize their definitions and basic properties in a preliminary Section V.A immediately following this overview. Two reasons for studying the cotangent bundle first are its relative simplicity and its particular importance in several contexts. One of these is that $(0,1)$ tensor fields — which are also known as **differential 1-forms** — provide an extremely convenient way for expressing the integrands of line integrals over smooth (or piecewise smooth) curves on a smooth manifold. In particular, one can use differential 1-forms to formulate general versions of the usual conditions under which line integrals of a given expression are independent of path. This will be explained in Section V.3.

The algebraic material needed to define higher order tensors appears in Section V.4; in many respects the linear algebra is fairly straightforward, but we shall formulate much of the theory in a coordinate-free manner using universal mapping properties (compare Sections I.4 and III.2 above) for the sake of conceptual and computational simplicity. The standard algebraic descriptions of tensors by coordinates and other expressions will then be derived from the abstract setting.

Finally, in Section V.5 we shall describe the coordinate free methods for defining tensor fields in mathematics and physics that have become standard in tensor analysis. In older work the definitions of tensor fields were often extremely messy and required confusing arrays of terms involving multiple indices like R^i_{jkl} or or g_{ij} or g^{ij} or δ^i_j. The methods described this section have simplified the theory and use of tensor analysis tremendously.

V.A : Dual spaces

(Conlon, § 6.1)

There are a few general remarks about dual spaces on page 183 and 184 of Conlon, but the discussion assumes the reader is already familiar with the concept of dual space. However, many — probably most — current undergraduate linear algebra texts omit this topic (although certain examples are always implicit in the discussion of matrix algebra). Therefore we shall include a summary of dual spaces and their fundamental properties.

Definition. Given a field $\mathbf{k}$ and a vector space V over $\mathbf{k}$, the **dual space** V^* is the set of all $\mathbf{k}$-linear transformations from V to $\mathbf{k}$. The elements of V are often called *linear functionals*.

More generally, if V and W are vector spaces over the field $\mathbf{k}$, standard methods from undergraduate linear algebra show that the set of $\mathbf{k}$-linear transformations fro V to W

$$\mathcal{L}_{\mathbf{k}}(V,W) \quad := \quad \mathrm{Hom}_{\mathbf{k}}(V,W)$$

is a vector space over $\mathbf{k}$ with respect to pointwise vector addition and scalar multiplication of W-valued functions. The proof for the case $\mathbf{k} = \mathbf{R}$ goes through word for word.

This might also be a good place to mention a general principle for elementary linear algebra over a field $\mathbf{k}$: *Except for the discussion of inner products, all of undergraduate linear algebra through the theory of determinants goes through if* $\mathbf{k}$ *replaces* $\mathbf{R}$ *as the field of scalars.*

In fact, except for the discussion of determinants, everything goes through if we only assume that $\mathbf{k}$ is a *division ring*, which is a system that has all the properties of a field except the multiplication is not necessarily commutative (strictly speaking, this means we need to assume both left and right distributivity as separate axioms and the condition on multiplicative inverses must be interpreted as saying that every nonzero element has a unique two-sided inverse). The quaternions are the most straightforward example of a division ring that is not a field.

V.A.1 : *The dimension of a dual space*

We begin with the following absolutely fundamental observation.

PROPOSITION. *If* V *is a vector space over the field* $\mathbf{k}$ *and* $\dim V = n < \infty$*, then* $\dim V^*$ *is also equal to* n*.*

Sketch of proof. Given an ordered basis $\{v_1, \cdots, v_n\}$ for V one constructs a *dual basis* $\{f_1, \cdots, f_n\}$ for V^* as follows: By the linearity identity

$$f\left(\sum_j x_j v_j\right) = \sum_j x_j f(v_j)$$

it suffices to define the functionals f_i on the basis vectors for V. Set $f_i(v_j) = 0$ if $i \neq j$ and $f_i(v_i) = 1$. It is then a straightforward exercise to verify that the set $\{f_1, \cdots, f_n\} \subset V^*$ defines a basis for V^*; the explicit verification is left to the reader as an exercise.■

Warning. If V is infinite dimensional, then as in Unit I of the ONLINE 205A NOTES it turns out that V has a basis and any two bases have the same cardinality, so it is meaningful to define the dimension of V to be an infinite cardinal number in such cases. However, the proposition above does not extend to the infinite dimensional case; in fact, it is sometimes possible to prove $\dim V < \dim V^*$ by simply computing the cardinalities of the respective vector space. For example, if $\mathbf{k} = \mathbf{Z}_2$ and $\dim V = \aleph_0$, then the cardinalities of V and V^* are $\aleph_0$ and $2^{\aleph_0}$ respectively.

COROLLARY. *If* V *is an* n*-dimensional vector space over* $\mathbf{k}$ *and* $\mathcal{B} = \{v_1, \cdots, v_n\}$ *is an ordered basis for* V*, then there is a* $\mathbf{k}$*-linear isomorphism* $T_{\mathcal{B}} : V \to V^*$ *such that* T *sends* v_i *to the* i^{rmth} *vector in the dual basis for* V^**.*■

Notation. Frequently we shall denote the dual ordered basis to $\mathcal{B} = \{v_1, \cdots, v_n\}$ by $\mathcal{B}^* = \{v_1^*, \cdots, v_n^*\}$, and similarly for other ordered bases.

It is important to realize that $T_{\mathcal{B}}$ depends very much upon the choice of $\mathcal{B}$. If $\mathcal{A} = \{w_1, \cdots, w_n\}$ is another ordered basis for V with dual basis $\mathcal{A}^*$, then in general $T_{\mathcal{A}}$ and $T_{\mathcal{B}}$ will be quite different and in particular $T_{\mathcal{B}} w_i$ will unusually not be equal to the corresponding vector w_i^* in the dual basis $\mathcal{A}^*$. In fact, one has the following relationship:

THEOREM. *In the setting described above, suppose that $\mathcal{A}$ and $\mathcal{A}^*$ are expressed as linear combinations of $\mathcal{B}$ and $\mathcal{B}^*$ as follows:*

$$w_j \;=\; \sum_i p_{ij}\, v_i \qquad\qquad w_j^* \;=\; \sum_i q_{ij}\, v_i^*$$

Then the matrices $P = (\,p_{ij}\,)$ and $Q = (\,q_{ij}\,)$ are transposed inverses of each other.

The proof amounts to verifying that ${}^{\mathbf{T}}Q \cdot P = I$ and is left to the reader as an exercise.∎

In particular, it follows that the matrix of $T_{\mathcal{B}}$ with respect to the ordered bases $\mathcal{A}$ and $\mathcal{A}'$ is equal to $Q^{-1}\,P = {}^{\mathbf{T}}P\,P$, while the matrix of $T_{\mathcal{A}}$ with respect to these ordered bases is the identity. Over the real numbers, this means that $T_{\mathcal{B}} = T_{\mathcal{A}}$ if and only if P is an orthogonal matrix.

V.A.2 : *Formal properties of the dual space construction*

The dual space construction extends to linear transformations in a fairly direct manner.

THEOREM. *Let V and W be vector spaces over the field $\mathbf{k}$, and let $A : V \to W$ be a linear transformation. Then there is a unique linear transformation $A^* : W^* \to V^*$ such that $[A^*f](v) = f \circ A(v)$ for all $f \in W^*$ and $v \in V$. This construction has the following additional properties:*

(1) *If A_1 and A_2 are linear transformations from V to W, then $(A_1 + A_2)^* = A_1^* + A_2^*$.*

(2) *If $A : V \to W$ is linear and c is a scalar, then $(c\,A)^* = c \cdot A^*$.*

(3) *If I_V is the identity linear transformation on V, the I_V^* is the identity linear transformation on V^*.*

(4) *If $B : U \to V$ and $A : V \to W$ are linear transformations, then $(AB)^* = B^*\,A^*$.*

Sketch of proof. The existence of A^* begins by noting that if $f \in W^*$ then $f \circ A \in V^*$. To prove that A^* is linear, one notes that if $f, g \in V^*$ and c is a scalar then we have

$$A^*(f \;+\; c\,g) \;=\; (f \;+\; c\,g)\circ A \;=\; (f\circ A) \;+\; c\,(g\circ A) \;=\; A^*f \;+\; c\,A^*g\;.$$

The verifications of $(1) - (4)$ are all routine calculations that are left to the reader as exercises.∎

Note that there are no finite dimensionality hypotheses in the preceding result.

We also have an important and useful relation between the construction $A \to A^*$ and the transposition operation on matrices.

THEOREM. *Let V and W both be finite dimensional vector spaces over the field $\mathbf{k}$, let $\mathcal{B} = \{v_1,\ \cdots\ , v_n\}$ and $\mathcal{A} = \{w_1,\ \cdots\ , w_n\}$ be ordered bases for V and W respectively, let $T : V \to W$ be a linear transformation, and let P denote the matrix of T with respect to the ordered bases $\mathcal{B}$ and $\mathcal{A}$. Then the matrix of T^* with respect to the dual ordered bases $\mathcal{A}^*$ and $\mathcal{B}^*$ is equal to the transpose of P.*

Verification of this is another straightforward computation that is left to the reader.∎

One can iterate the dual space construction to get double dual spaces V^{**} and associated linear transformations of double dual spaces. If $\dim V = n$, then we also have $n = \dim V^* = \dim V^{**}$, so that V^{**} is also isomorphic to V. However, in contrast to the case of V and V^*, there **IS** a

natural isomorphism from V to V^{**} that does not depend upon a choice of ordered basis. The key to seeing this is the following result:

THEOREM. *Let V be a vector space over the field $\mathbf{k}$ (no assumption of finite dimensionality here). Then there is a linear transformation $e_V : V \to V^{**}$ such that for all $x \in V$ the map $e_V(x) : V^* \to \mathbf{k}$ is defined by*

$$[[\, e_V(f)\,]\, f \quad = \quad f(x) \ .$$

Furthermore, if $T : V \to W$ is a linear transformation then we have the commutativity identity

$$T^{**} \circ e_V \quad = \quad e_W \circ T \ .$$

This is yet another sequence of routine calculations that is left to the reader as an exercise.■

COROLLARY. *If in the preceding result V is finite dimensional and $\mathcal{B} = \{v_1, \ \cdots \ , v_n\}$ is an ordered basis for V, then e_V maps the vectors of $\mathcal{B}$ to the corresponding vectors of the double dual basis $\mathcal{B}^{**}$.*■

In other words, for such examples there is an isomorphism that does not depend upon the choice of basis.

V.A.3 : *A geometric application*

If V is the n-dimensional real vector space $\mathbf{R}^n$, then there is a simple 1–1 correspondence between the p-dimensional subspaces of $\mathbf{R}^n$ and the $(n-p)$-dimensional subspaces that is defined by sending a subspace W to its orthogonal complement $W^{\perp}$; since $(W^{\perp}){*}\perp = W$, it follows that this map is 1–1 onto and in fact is its own inverse. A similar statement and proof aply to the complex vector space $\mathbf{C}^n$ and its complex subspaces. We shall use dual spaces to prove a similar result for p-dimensional subspaces of an n-dimensional vector space V over an arbitrary field $\mathbf{k}$.

THEOREM. *Let V be an n-dimensional vector space over the field $\mathbf{k}$, and let p be an integer between 0 and n. Then there is a $1-1$ correspondence between the p-dimensional subspaces of V and the subspaces of dimension $(n-p)$.*

Proof. Since V^* is also n-dimensional we have a linear isomorphism $V \cong V^*$. More generally, if we are given an isomorphism T from an n-dimensional vector space V to a vector space W, then $S \subset V$ is a q-dimensional subspace if and only if $T(S) \subset W$ is a q-dimensional subspace, and therefore there is a 1–1 correspondence between q-dimensional subspaces of V and W. Therefore it will be enough to prove that there is a 1–1 correspondence between the p-dimensional subspaces of V and the $(n-p)$-dimensional subspaces of V^*.

Given a subspace $U \subset V$, define the *annihilator* of U to be the set $U^{\dagger}$ of all linear functionals in V^* that are zero on U. CLAIM: If U is p-dimensional, then $U^{\dagger}$ is an $(n-p)$-dimensional subspace of V^*.

The proof of the claim is fairly elementary, so we shall merely sketch the argument and leave some details to the reader. First, one needs to show that $U^{\dagger}$ is a subspace. One may then construct a basis for $U^{\dagger}$ by starting with a basis $\mathcal{B}_0 = \{v_1, \ \cdots \ , v_p\}$ for U, expanding it to a basis $\mathcal{B}$ of V by adding $\mathcal{B}_1 = \{v_{p+1}, \ \cdots \ , v_n\}$, and then forming the dual basis $\mathcal{B}^*$. Let $\mathcal{B}_1^* \subset \mathcal{B}^*$ be the set

$\{v_{p+1}^*, \cdots, v_n^*\}$. By construction this set lies in $U^\dagger$, and we claim it is a basis (since it is linearly independent we only need to show it spans). If we write a typical element of V^* as

$$f = \sum_i y_i v_i^*$$

then f cannot belong to $U^\dagger$ if there is some $i \leq p$ such that $y_i \neq 0$. Therefore $\mathcal{B}_1$ must span $U^\dagger$, and this yields the assertion about $\dim U^\dagger$.

The preceding discussion shows the existence of a map Φ sending p-dimensional subspaces of V to $(n-p)$-dimensional subspaces of V^* that is given by $\Phi(U) = U^\dagger$. To see that this map is 1–1 onto consider the corresponding map from $(n-p)$-dimensional subspaces of V^* to subspaces of $V^(**)$ with dimension

$$p = [n - (n-p)] .$$

This map sends U to the subspace $U^{\dagger\dagger}$. It is elementary to verify that $e_V(U)$ is a subspace of $U^{\dagger\dagger}$, and since e_V is an isomorphism it follows that $e_V(U)$ is a p-dimensional subspace. On the other hand, the preceding discussion also shows that $\dim U^{\dagger\dagger} = p$, and hence we must have $e_V(U) = U^{\dagger\dagger}$. Therefore a one sided inverse Ψ to Φ is given by sending a subspace S to $e_V^{-1}(S^\dagger)$. Since $\Psi \circ \Phi$ is the identity, it follows that Φ is 1–1. To see that Φ is onto, it suffices to check that Ψ is 1–1. We may use the argument thus far to show that the annihilator construction sending $(n-p)$-dimensional subspaces of V^* to p-dimensional subspaces of V^{**} is 1–1, and the construction sending $X \subset V^{**}$ to $e_V^{-1}(X)$ is also 1–1, and if we combine these to observations we see that Ψ is indeed 1–1 as claimed.■

V.A.4 : *Dual pairings*

In many contexts it is desirable to have an algebraic criterion for recognizing dual spaces.

Definition. If V and W are finite dimensional vector spaces over a field $\mathbf{k}$, then a *dual pairing* is a function

$$\varphi : V \times W \to \mathbf{k}$$

with the following properties:

(1) For each $v \in V$ the function $\varphi_v^{(1)} : W \to \mathbf{k}$ defined by $\varphi_v^{(1)}(w) = \varphi(v, w)$ is linear.

(2) For each $w \in W$ the function $\varphi_w^{(2)} : V \to \mathbf{k}$ defined by $\varphi_w^{(2)}(v) = \varphi(v, w)$ is linear.

(3) If $0 \neq v \in V$ then there is a nonzero vector $w \in W$ such that $\varphi(v, w) \neq 0$.

(4) If $0 \neq w \in W$ then there is a nonzero vector $v \in V$ such that $\varphi(v, w) \neq 0$.

EXAMPLES.

1. If V is a finite dimensional vector space over $\mathbf{k}$ then the map $\varepsilon_V : V^* \times V \to \mathbf{k}$ given by $\varepsilon(f, v) = f(v)$ has the desired properties, and in fact it provides the motivation for the name.

2. If V is the space of $1 \times n$ row vectors over $\mathbf{k}$ and W is the space of $n \times 1$ column vectors and we identify the 1×1 matrices with $\mathbf{k}$ as usual, then matrix multiplication defines a dual pairing.

3. If $\mathbf{k} = \mathbf{R}$ and V is a finite dimensional inner product space over $\mathbf{R}$, then the inner product defines a dual pairing; the first and second properties are part of the definition for an inner product, while the last two follow because $v \neq 0 \Longrightarrow \langle v, v \rangle > 0$.

4. If $\mathbf{k} = \mathbf{C}$ and $V = \mathbf{C}^n$, then a dual pairing is defined by taking $\varphi(v, w) = \langle v, \overline{w} \rangle$ where the right hand side denotes the inner product of v with the vector $\overline{w}$ formed by taking the complex conjugates of the coordinates of w. As in the previous item, the four properties of a dual pairing are direct consequences of the defining properties for a complex inner product, but one must take into account the definition of the inner product as $\sum_j \ v_j \, \overline{w_j}$ when working out the details (hence φ is $\sum_j \ v_j \, w_j$).

The next result states that all pairings are essentially equivalent to the first example.

PROPOSITION. *Let V and W be finite dimensional vector spaces over the field $\mathbf{k}$, and supppose that φ is a dual pairing on $V \times W$. Then there is an isomorphism $A_\varphi : V \to W^*$ such that*

$$\varphi(v, w) \quad = \quad \varepsilon_W\big(\, A_\varphi(v), \, w \,\big)$$

for all $(v, w) \in V \times W$.

Proof. One defines $A_\varphi(v)$ to be the linear functional $\varphi_v^{(1)}$ described previously. Direct computation shows that A_φ is a linear transformation. We claim it is an isomorphism; *i.e.*, its kernel is $\{0\}$ and its image is all of V^*.

Suppose first that $A_\varphi(v) = 0$. This means that $\varphi_v = 0$. However, the conditions in the definition of a dual pairing imply that $\varphi_v = 0$ if and only if $v = 0$. Note that this implies in particular that $\dim V \leq \dim W^* = \dim W$.

We claim a similar inequality holds if the roles of V and W are reversed. To see this, let $\varphi^{\mathrm{op}} : W \times V \to \mathbf{k}$ be the reverse map defined by

$$\varphi^{\mathrm{op}}(w, v) \quad = \quad \varphi(v, w) \ .$$

Then φ^{rmop} also satisfies conditions $(1) - (4)$ above, and therefore the argument in the previous paragraph implies $\dim W \leq \dim V$. Combining this with the previous inequality we have $\dim W = \dim V = \dim V^*$.

We now know that A_φ is a 1–1 linear transformation between spaces of the same dimension, so therefore it must be onto and invertible by standard results from undergraduate linear algebra.■Many applications of the preceding result involve special cases where $W = V$.

Definition. Let V be a vector space over a field $\mathbf{k}$. A function $\varphi : V \times V \to \mathbf{k}$ is said to be a *bilinear form* if it for each $v \in V$ the functions

$\varphi_v^{(1)} : V \to \mathbf{k}$ defined by $\varphi_v^{(1)}(x) = \varphi(v, x)$, and

$\varphi_v^{(2)} : V \to \mathbf{k}$ defined by $\varphi_v^{(2)}(x) = \varphi(x, v)$

are both linear. Such a form is said to be *left nondegenerate* if $v \neq 0 \Longrightarrow \varphi_v^{(1)} \neq 0$. There is a similar definition for right nondegeneracy. If $\mathbf{k} = \mathbf{R}$, then inner products are basic examples of left nondegenerate bilinear forms.

FURTHER EXAMPLES. If $V = \mathbf{k}^n$ viewed as column vectors and A is an invertible $n \times n$ matrix, then the expression

$$\varphi_A(x, y) \quad = \quad {}^{\mathbf{T}}y \, A \, x$$

is easily checked to define a left nondegenerate bilinear form on V, and in fact all such forms are given in this manner for suitable choices of the matrix A. Details appear in the exercises.

The next result shows the close relationship between left nondegenerate bilinear forms and dual pairings.

PROPOSITION. *In the above notation, if φ is left or right nondegenerate, then it defines a dual pairing on $V \times V$.*

Proof. We only need to prove that φ is right nondegenerate if it is left nondegenerate and vice versa. Since the arguments are nearly identical, we shall only consider the case where φ is left nondegenerate.

As in previous arguments we obtain a map $A : V \to V^*$ that is 1–1. Since V^* and V have the same dimension the map A must also be onto. Therefore if $0 \neq w \in V$ one can find some $z \in V^*$ such that $\varepsilon_V(z, w) \neq 0$. Since A is onto we know $z = Ay$ for some $y \in V$, and with this choice of y we have $\varphi(y, w) \neq 0$.■

COROLLARY. *If V is a finite dimensional vector space over the field* $\mathbf{k}$ *and φ is a left nondegenerate bilinear form on V, then there is an isomorphism $A : V \to V^*$ such that $\varphi(x, y) = \varepsilon_V(Ax, y)$ for all $(x, y) \in V \times V$.*■

Final remarks. Having spent so much time and space discussing dual spaces and isomorphisms between a finite dimensional vector space and its dual, it is reasonable to expect that all this is needed for something in this course that is geometrically significant. In fact, if V is a finite dimensional real vector space and the isomorphism between V and V^* is given by an inner product, then the preceding corollary reflects a fundamental pair of constructions in tensor algebra that are classically called **raising and lowering of indices**. We shall discuss this further in Section V.4.

V.1 : Vector bundles

(Conlon, §§ 3.3–3.4)

Tangent bundles are not the only parametrized families of vector spaces that are important in the study of smooth manifolds. The purpose of this section is to generalize the the procedure for constructing the tangent bundle so that it will yield the other examples that arise naturally in geometry and topology.

One special feature of the tangent bundle is that it is a family of n-dimensional vector spaces over an n-dimensional manifold. We shall also be interested in families of m-dimensional vector spaces over an n-dimensional manifold where m is not necessarily equal to n, and in order to motivate this generalization our first objective is to describe some fundamental examples for which $m = 1$ and n can be as large as we please. However, in order to do this we must first discuss the manifolds over which these families lie.

V.1.1 : *Projective spaces*

When we see parallel lines that extend for a long distance, it looks as if they meet at some point on the horizon. During the fifteenth century this observation was studied in great detail by various artists beginning with F. Brunelleschi (1377–1446), and the resulting theory of perspective drawing yielded improved techniques for painting and drawing pictures that more accurately reflect the images produced by the human eye. These ideas and other considerations lead directly to a provocative idea:

> Perhaps our standard concept of parallel lines as not having any points in common should be modified to state that they meet at some point at infinity. All lines parallel to a given line should have the same point at infinity, but if two lines are not parallel then they should not have the same point at infinity.

Since visual experience indicates that all the points at infinity lie on the horizon line, one might also speculate that the set of all points at infinity should be viewed as a line at infinity. Here are some online references that (often literally) illustrate the ideas in this paragraph.

`http://mathforum.org/sum95/math_and/perspective/perspect.html`

`http://www.math.utah.edu/~treiberg/Perspect/Perspect.htm`

`http://www.dartmouth.edu/~matc/math5.geometry/unit11/unit11.html`

`http://www.math.nus.edu.sg/aslaksen/projects/perspective/alberti.htm`

`http://www.ski.org/CWTyler_lab/CWTyler/Art%20Investigations/···`
`PerspectiveHistory/Perspective.BriefHistory.html`

`http://www.mcm.edu/academic/galileo/ars/arshtml/renart1.html`

`http://www.collegeahuntsic.qc.ca/Pagesdept/Hist_geo/Atelier/···`
`Parcours/Moderne/perspective.html`

`http://gaetan.bugeaud.free.fr/pcent.htm`

If we incorporate these ideas into ordinary plane geometry, it will follow that every pair of lines meets at some point which is either an ordinary point or a point at infinity. Likewise, an ordinary line and the line at infinity meet at one point; namely, the point at infinity on the original line. Similarly, two points in this extended plane (which includes points at infinity) will always determine a unique line; if both points are ordinary this will be the usual line plus its point at infinity, if one is ordinary and the other is at infinity this will be the line in the given parallel family through the given point, and if both are at infinity this will be the line at infinity. Thus we have a system where every two points lie on a unique line and every two lines meet in a unique point. This is the starting point of synthetic projective geometry, and the system obtained as above is called the **projective plane**. It is natural to ask what mathematical value such a system might have. During the sixteenth and seventeenth century various mathematicians discovered that the addition of points at infinity led to (1) new discoveries in ordinary plane and solid geometry, (2) new proofs of geometrical facts that were often dramatically simpler than more traditional ones, (3) unified formulations of results that would otherwise involve long and unenlightening lists of special cases.

It is natural to ask the following question, and for our purposes the answer is fundamentally important:

COORDINATIZATION PROBLEM. *Is it possible to extend the usual notion of cartesian coordinates to the projective plane in some reasonable manner? If so, how can this be done?*

Since all ordered pairs of real numbers are used up by points in the cartesian plane, it should be clear that we shall need at least three real numbers to locate a point in the projective plane. However, since we are dealing with an object that is 2-dimensional, it is also clear that we also do not want a 1–1 correspondence between points of the projective plane and ordered triples of real numbers. There are two ways of addressing this problem:

(1) Impose some rigid constraints on the coordinate values; for example, require that the third coordinate be either 1 or 0.

(2) Find a decomposition of $\mathbf{R}^3$, or some reasonable subset, into equivalence classes such that the points of the projective plane correspond to the various equivalence classes of triples.

Both approaches work quite well, but eventually the first must borrow ideas from the second, so it is best to proceed with the second approach and see how it reflects the first one. For several reasons it is convenient to carry out the construction with $\mathbf{R}$ replaced by an arbitrary field $\mathbf{k}$; one can even weaken the hypothesis and allow $\mathbf{k}$ to be a division ring by being sufficiently careful about the order of multiplication (see the exercises). Note that in $\mathbf{R}^2$ or $\mathbf{R}^3$ ordinary lines and planes correspond to translates of subspaces; *i.e.*, subsets of the form $x + W$ where W is either a 1- or 2-dimensional vector subspace of $\mathbf{R}^2$ or $\mathbf{R}^3$. One can define lines and planes in $\mathbf{k}^2$ and $\mathbf{k}^3$ — or more generally $\mathbf{k}^n$ — in exactly the same fashion. As usual, a *hyperplane* in $\mathbf{k}^n$ will be the translate of an $(n-1)$-dimensional subspace.

ELEMENTARY FACT. *Every hyperplane in $\mathbf{k}^n$ is equal to the set of solutions for some (generally nonhomogeneous) linear equation of the form*

$$\sum_j a_i\, x_i \quad = \quad b$$

where at least one coefficient a_i is nonzero. Two ordered $(n+1)$-tuples $(a_1, \ \ldots\ , a_n; b)$ and $(a'_1, \ \ldots\ , a'_n; b')$ determine the same hyperplane if and only if there is a nonzero constant c such that $b = c\, b'$ and $a'_i = c\, a_i$ for all i.

This is a fairly straightforward exercise in undergraduate linear algebra (generalized so that the scalars are given by an arbitrary field $\mathbf{k}$).■

For our purposes it is important to know that the Euclidean Parallel Postulate generalizes to $\mathbf{k}^n$ if lines, planes, *etc.* are defined as above.

EUCLIDEAN PARALLELISM PROPERTY. *Let L be a line in $\mathbf{k}^2$, and let $\mathbf{x}$ be a point not on L. Then there is a unique line M in L such that $\mathbf{x} \in M$ and M is* parallel *to L; i.e., there is a plane Π containing both L and M.*

Sketch of proof. Write $L = \mathbf{y} + W$, where W is a 1-dimensional subspace and $\mathbf{y}$ is suitably chosen, and consider the line $M = \mathbf{x} + W$. We claim that L and M are coplanar but $L \cap M = \emptyset$.

We claim that $\mathbf{y} - \mathbf{x}$ does not belong to W, for if it did then we would have $\mathbf{x} \in \mathbf{y} + W = L$. If U is the span of a nonzero vector in W and $\mathbf{y} - \mathbf{x}$, then U is 2-dimensional and contains both L and M. Finally, to show the lines are disjoint, suppose that we had some $\mathbf{z} \in L \cap M$. Take $\mathbf{w}$ to be a nonzero vector in W ($\Longrightarrow$ the set $\mathbf{w}$ forms a basis for W). Then there are scalars r and s such that

$$\mathbf{z} \ = \ \mathbf{x} + r\,\mathbf{w} \ = \ \mathbf{y} + s\,\mathbf{w}$$

and therefore we have

$$\mathbf{y} - \mathbf{x} \ = \ (r - s)\,\mathbf{w}$$

which contradicts the conclusion of the first sentence in this paragraph.■

We are now ready to define our coordinates on the projective plane associated to $\mathbf{k}$. As in ordinary coordinate geometry, if we are given a line L in $\mathbf{k}^2$ and it is defined by an equation of the form

$$a_1 x_1 \ + \ a_2 x_2 \ = \ b$$

then we shall say that (a_1, a_2) determines a pair of *direction numbers* for L. Such direction numbers are not unique, but by the preceding observation any two pairs of direction numbers are related by a nonzero proportionality constant.

Motivated by the preceding paragraph, if L is a line and (a_1, a_2) is a set of direction numbers for L, we shall say that the triple $(a_1, a_2, 0)$ is a set of **homogeneous coordinates** for the point at infinity on L. Given an ordinary point $\mathbf{x} = (x_1, x_2) \in \mathbf{k}^2$, we shall similarly say that for each $c \neq 0$ in $\mathbf{k}$ every triple of the form $(c\,x_1,\ c\,x_2,\ c)$ is a set of homogeneous coordinates for $\mathbf{x}$.

The following observations may be checked by direct calculation:

THEOREM. *Every nonzero element of $\mathbf{k}^3$ is a set of homogeneous points for either an ordinary or ideal point. No such element can be a set of homogeneous coordinates for more than one point. An ordinary line in $\mathbf{k}^2$ defined by an equation of the form $ax + by = c$ corresponds to the set of points whose homogeneous coordinates satisfy the linear homogeneous equation $a\,U_1 + b\,u_2 - c\,u_3 = 0$ where the first two coefficients are not zero, and the line at infinity corresponds to all points whose homogeneous coordinates satisfy $u\,3 = 0$. In particular, every line in the projective plane is defined by a homogeneous linear equation of the form $a\,U_1 + b\,u_2 - c\,u_3 = 0$ where at least one of a, b, c is nonzero. Finally, two such triples define the same line in the projective plane if and only if and only if one set is proportional to the other (where the constant of proportionality must be nonzero).*

This is essentially an elementary but somewhat lengthy and tedious exercise in linear algebra.■

We may proceed similarly in $\mathbf{k}^n$ for $n > 2$, the main difference being that the set of points at infinity will define a *hyperplane at infinity* or *ideal hyperplane*; in the same spirit, one often describes the points at infinity for lines in $\mathbf{k}^n$ as *ideal points*. All this leads us to define **the n-dimensional projective space $\mathbf{kP}^n$** over the field $\mathbf{k}$ to be the set of equivalence classes for $\mathbf{k}^{n+1} - \{\,\mathbf{0}\,\}$ under the equivalence relation $\mathbf{u} \ \sim \ \mathbf{v} \iff \mathbf{u}$ is a nonzero scalar multiple of $\mathbf{v}$. Ordinary points correspond to those equivalence classes for which the last coordinate is nonzero, and in this setting one set of homogeneous points for an ordinary point $(x_1, \ \dots, x_n)$ is given by $(x_1, \ \dots, x_n; 1)$.

We do not need to go much further into projective geometry, but the following point seems worth mentioning.

THEOREM. *Suppose that $\mathbf{x}$ and $\mathbf{y}$ are distinct points of $\mathbf{k}^n$, and let $\alpha(\mathbf{x})$ and $\alpha(\mathbf{y})$ be the vectors in $\mathbf{k}^{n+1}$ whose first n coordinates are those of $\mathbf{x}$ and $\mathbf{y}$ respectively and whose last coordinates are equal to 1. Then the span U of $\alpha(\mathbf{x})$ and $\alpha(\mathbf{y})$ is 2-dimensional, and the unique line joining $\mathbf{x}$ and $\mathbf{y}$ is given by the set of ordinary points whose homogeneous coordinates lie in U.*

Sketch of proof. One first needs to check that $\alpha(\mathbf{x})$ and $\alpha(\mathbf{y})$ are linearly independent so that their span is indeed 2-dimensional.

The ordinary points on the line joining $\mathbf{x}$ and $\mathbf{y}$ are precisely those points $\mathbf{z}$ expressible as a linear combination of the form $t\,\mathbf{x} + (1-t)\,\mathbf{y}$ for some scalar t. Homogeneous coordinates for such a point are given by $t\,\alpha(\mathbf{x}) + (1-t)\,\alpha(\mathbf{y})$ and hence lie in U; since the latter is closed under scalar multiplication, all sets of homogeneous coordinates for $\mathbf{z}$ also belong to U. In the reverse direction, if $\mathbf{z}$ has a set of homogeneous coordinates in U, then the corresponding vector $\alpha(\mathbf{z})$ is a linear combination of $\alpha(\mathbf{x})$ and $\alpha(\mathbf{y})$, and since the third coordinates of all three vectors are equal to one it follows that the coefficients of $\alpha(\mathbf{x})$ and $\alpha(\mathbf{y})$ must add up to 1. As noted above, this means that $\mathbf{z}$ lies on the line joining $\mathbf{x}$ and $\mathbf{y}$.■

Smooth atlases for real and complex projective spaces. The first thing we need to do is construct a Hausdorff topology on $\mathbf{kP}^n$ is $\mathbf{k}$ denotes the real or complex numbers and to verify that this space is a topological manifold. Its dimension will be n in the real case and $2n$ in the complex case; if one is careful about the order of multiplication, it is also possible to carry out all of the above for the division ring $\mathbf{K}$ of quaternions, and in this case one will find that the associated projective n-space will be a $4n$-manifold, but we shall be content with $\mathbf{R}$ and $\mathbf{C}$ here.

The simplest way to define topologies on real and complex projective n-spaces is to take the quotient topology on these spaces that is associated to $\mathbf{k}^{n+1} - \{\,\mathbf{0}\,\}$ and the equivalence relation determined by nonzero scalar multiplication. One problem with this is that some important topological properties of $\mathbf{RP}^n$ and $\mathbf{CP}^n$ are not easily seen from this definition. Now every point in $\mathbf{R}^{n+1} - \{\,\mathbf{0}\,\}$ or $\mathbf{C}^{n+1} - \{\,\mathbf{0}\,\}$ is uniquely expressible as a product of a positive real number and a vector with unit length, and it follows that the sets of equivalence classes comprising $\mathbf{RP}^n$ and $\mathbf{CP}^n$ are also given by taking the unit sphere in $\mathbf{R}^{n+1}$ or $\mathbf{C}^{n+1}$ and identifying two points if one is a scalar multiple of the other; such a scalar must have absolute value (or modulus in the complex case) equal to 1. The exercises for Mathematics 205A show that one obtains the same quotient spaces for these constructions and the corresponding earlier ones (these also appear in the exercises for the present section).

We have already shown that $\mathbf{RP}^n$ is a smooth manifold and defined a smooth atlas for it in Sections I.1 and III.2, but it will be convenient to take an approach that works for both real and complex projective spaces; after this has been completed we shall show that the new construction yields a smooth structure on $\mathbf{RP}^n$ that is identical to the old one.

Let $\mathbf{k} = \mathbf{R}$ or $\mathbf{C}$, and let $d = 1$ for $\mathbf{R}$ and 2 for $\mathbf{C}$. As in the case where $n = 2$, we have a 1–1 map from $\mathbf{k}^n$ onto the set of all ordinary points in $\mathbf{kP}^n$ with nonvanishing last homogeneous coordinates. Given a permutation σ of the first $(n+1)$ positive integers, the corresponding permutation of homogeneous coordinates determines a homeomorphism from $\mathbf{kP}^n$ to itself, and if we take the transposition permutation interchanging $n+1$ with some fixed $j < n$, we obtain an identification of $\mathbf{k}^n$ with the set of all points in $\mathbf{kP}^n$ with nonvanishing j^{th} coordinates. Since every point y in $\mathbf{kP}^n$ there is some j such that all j^{th} homogeneous coordinates for y, are nonzero, we have apparently shown that each point in $\mathbf{kP}^n$ lies in a subset that somehow looks like $\mathbf{k}^n$. In order to show that we have a topological dn-manifold, we need to show that these maps from $\mathbf{k}^n$ into $\mathbf{kP}^n$ define homeomorphisms onto their images and that the space $\mathbf{kP}^n$ is Hausdorff. We shall verify these in the next result.

PROPOSITION. *The space* $\mathbf{kP}^n$ *is a compact topological* dn*-manifold.*

Proof. The space $\mathbf{kP}^n$ is compact because it can be presented as the quotient of the compact space S^{dn+1}. The next step will be to prove that every point has an open neighborhood homeomorphic to $\mathbf{k}^n$ using the candidates for coordinate charts defined above.

Suppose first that we have an ordinary point of $\mathbf{kP}^n$ whose last homogeneous coordinates are nonzero. Let $F : \mathbf{k}^n \to \mathbf{k}^{n+1} - \{\,\mathbf{0}\,\}$ be the map sending $(x_1, \ldots, x_n)$ to $(x_1, \ldots, x_n; 1)$, and let $h : U \to \mathbf{kP}^n$ be the map sending $(x_1, \ldots, x_n)$ to the point with homogeous coordinates $(x_1, \ldots, x_n; 1)$. It follows immediately that h is continuous and 1–1 and its image is the set of ordinary points. We claim that it is an open subset of $\mathbf{kP}^n$. This follows because the inverse image of the complement is the set of all points in $\mathbf{k}^{n+1}$ whose last coordinate is zero; since the latter is closed, it follows that the complement of the image of h is closed and hence that the image of h is open. To see that h defines a homeomorphism onto its image, let $X \subset \mathbf{k}^{n+1} - \{\,\mathbf{0}\,\}$ be the set of all points whose last coordinate is nonzero; it will suffice to define a map $\ell : X \to \mathbf{k}^n$ such that $\ell(z \cdot \mathbf{v}) = \ell_j(\mathbf{v})$ for all nonzero complex numbers z and $\ell \circ h$ is the identity. The first condition will imply that ℓ factors through a continuous map $\overline{\ell}$ defined on the image of h, and the second will imply that the map $\overline{\ell}$ is an inverse to h. To finish the argument, we may simply take ℓ to be $1/x_{n+1}$ times projection from $\mathbf{k}^{n+1} - \{\,\mathbf{0}\,\}$ onto the first n coordinates.

As in the discussion preceding the statement of the proposition, for each point $y \in \mathbf{kP}^n$ there is at least one value of j such that the j^{th} homogeneous coordinates are nonzero, and by taking a self-homeomorphism φ of $\mathbf{kP}^n$ that interchanges the j^{th} and last homogeneous coordinates we obtain an ordinary point $\varphi(y)$. It then follows that y lies in the image of "$\varphi^{-1} \circ \pi \circ h$" where π is the quotient space projection onto $\mathbf{kP}^n$. Thus we have shown that every point in $\mathbf{kP}^n$ has an open neighborhood that is homeomorphic to $\mathbf{k}^n \cong \mathbf{R}^{dn}$.

Finally, we need to show that $\mathbf{kP}^n$ is Hausdorff. If both points are ordinary points, then this follows because $\mathbf{k}^n$ is Hausdorff. Therefore it will suffice to show that if x and y are distinct points of $\mathbf{kP}^n$ then there is a homeomorphism θ of $\mathbf{kP}^n$ to itself such that $\theta(x)$ and $\theta(y)$ are both ordinary points. This will be a consequence of a more general fact:

SYMMETRY LEMMA FOR PROJECTIVE SPACES. *If T is a linear transformation from $\mathbf{k}^{n+1}$ to itself, then there is a homeomorphism $\widehat{T}$ from $\mathbf{kP}^n$ to itself such that if x is a point with homogeneous coordinates ξ then $\widehat{T}(x)$ is a point with homogeneous coordinates $T\xi$.*

Sketch of proof. In order to get a well defined continuous map, it is only necessary to show that if $\pi(\xi) = \pi(\xi')$ then $\pi(T\xi) = \pi(T\xi')$, where π gives the equivalence class of a vector in the projective space. But this follows immediately because $\xi' = c\xi \Longrightarrow T\xi' = cT\xi$. To see that $\widehat{T}$ is a homeomorphism, it suffices to check that if S is inverse to T then $\widehat{S}$ is inverse to $\widehat{T}$. But this is just a routine calculation.■

Proof of proposition concluded. Let ξ and η be homogeneous coordinates for x and y respectively. We may as well assume that $n > 0$ for otherwise $\mathbf{kP}^n$ consists of a single point. Since $x \neq y$ the vectors ξ and η are linearly independent. There are infinitely many pairs of vectors α and β in $\mathbf{k}^{n+1}$ such that α and β are linearly independent but have nonvanishing last coordinates. Choose an invertible linear transformation T of $\mathbf{k}^{n+1}$ such that $T\xi = \alpha$ and $T\eta = \beta$. Then the homeomorphism $\widehat{T}$ sends x and y to the ordinary points $\pi(\alpha)$ and $\pi(\beta)$. Since the latter have disjoint open neighborhoods U and V, it follows that $\widehat{T}^{-1}(U)$ and $\widehat{T}^{-1}(V)$ are disjoint open neighborhoods of x and y.■

The next order of business is to construct smooth atlases.

PROPOSITION. *For each j such that $1 \leq j \leq n+1$ let τ_j be the permutation on the first $n+1$ positive integers that interchanges j and $n+1$, and let φ_j be the homeomorphism of $\mathbf{kP}^n$ with itself determined by τ_j on homogeneous coordinates (by convention φ_{n+1} is the identity). Let h be the $1-1$ correspondence from $\mathbf{k}^n$ to ordinary points of $\mathbf{kP}^n$. Then the pairs $(\mathbf{k}^n, \varphi \circ h)$ form a smooth atlas for $\mathbf{kP}^n$. Furthermore, if $\mathbf{k} = \mathbf{R}$ then the atlas constructed in this fashion defines the same smooth structure as the one defined previously.*

Proof. Direct computation shows that the coordinates for the transition maps ψ_{ij} are all quotients of the form $1/x_i$ and x_j/x_i where $j \neq i$ (check this first with a few examples). These maps are all smooth, and the relation $\psi_{ji} = \psi_{ij}^{-1}$ show that all Jacobians are nonzero. Therefore the charts in the proposition define a smooth atlas.

It remains to show that one obtains the same smooth structure as before when $\mathbf{k} = \mathbf{R}$. For this purpose it is helpful to choose an atlas for the previously defined smooth structure on $\mathbf{RP}^n$, and the right choice is to take the hemispherical charts $q_j : N_1(0) \to S^n \to \mathbf{RP}^n$ whose images in S^n are the sets of all points on the unit sphere with positive first coordinates. In this case the associated inverse map "q_j^{-1}" is just the projection that forgets the j^{rmth} coordinate. It follows that the transition maps "$q_j^{-1} \circ h_i$" are given by composing the smooth map

$$\frac{1}{\sqrt{1+|x|^2}} \cdot \tau_i \circ h(x)$$

with the projection forgetting the $j^{\rm th}$ coordinate. Such maps are clearly smooth. Now consider the inverse transition maps "$h_i^{-1} \circ g_j$" to the ones considered above. If z_i denotes the $i^{\rm th}$ coordinate of g_j, then this transition map is given by composing the map

$$\frac{1}{z_i} \cdot g_j$$

with the coordinate projection that forgets the $i^{\rm th}$ coordinate, and therefore this map is also smooth. Thus we have shown that the union of the new atlas with an older one is still a smooth atlas for $\mathbf{RP}^n$, and thus the two atlases define the same smooth structure on this manifold.■

Later in this section we shall use the fact that $\mathbf{kP}^1$ is a very familiar object.

THEOREM. *If $\mathbf{k} = \mathbf{R}$ or $\mathbf{C}$ and $d = \dim_{\mathbf{R}} \mathbf{k}$, then $\mathbf{kP}^1$ is diffeomorphic to S^d.*

Sketch of proof. Since it is very easy to prove the manifolds are homeomorphic and considerably more difficult to prove they are diffeomorphic, we shall first prove the manifolds are homeomorphic; note that the diffeomorphism proof is marked with two stars.

Homeomorphism proof. By construction we have identified $\mathbf{k}$ with the set of all points in $\mathbf{kP}^n$ whose second coordinate is nonzero. Therefore every point in the complement has homogeneous coordinates of the form $(a, 0)$ for some nonzero scalar a. Since all vectors of this form are nonzero scalar multiples of $(1, 0)$ it follows that the complement of the image of $\mathbf{k}$ consists of a single point. The latter implies that $\mathbf{kP}^1$ must be homeomorphic to the one point compactification of $\mathbf{k} \cong \mathbf{R}^d$, which is S^d.

Diffeomorphism proof. $(\star\star)$ This proof uses stereographic coordinates, which were mentioned in the Secton III.1 of these notes (see Example 1.2.3 on page 3 of Conlon; also see the material on stereographic projections in the ONLINE 205A NOTES.). We shall let $g_\pm : \mathbf{R}^d \longrightarrow S^d$ denote the stereographic projection maps whose images are $S^d - \{\pm\,\mathbf{e}_{d+1}\}$ where $\mathbf{e}_{d+1}$ is the unit vector whose last coordinate is nonzero. We did not actually show that the maps $g_\pm$ we smooth coordinate charts, but we shall do so now as follows: If $j : S^d \to \mathbf{R}^{d+1}$, then the explicit formulas for $g_\pm$ show that the composites $j \circ g_\pm$ are smooth and have maximum rank. It follows that the same is true for the maps $g_\pm$ whose images are contained in the submanifold S^d. This is enough to imply that the charts $g_\pm$ are compatible with those in the standard submanifold atlas for S^d. Note that this argument works for all $d \geq 1$.

The transition maps "$g_+^{-1} g_-$" and "$g_-^{-1} g_+$" can be computed explicitly using basic linear algebra, and the computations show that both maps are given by

$$\psi(\mathbf{x}) \quad = \quad \frac{4}{|\mathbf{x}|^2} \cdot \mathbf{x} \ .$$

As indicated in the ONLINE 205A NOTES, this map is often described as an inversion about a sphere (or circle or pair of points) of radius 2 centered at the origin. The corresponding tansition maps for the standard atlas on $\mathbf{kP}^1$ are very close to this, for they are given by

$$\varphi(z) \quad = \quad \frac{1}{z} \quad = \quad \frac{1}{|z|^2} \cdot \chi(z)$$

where $\chi(z)$ denotes complex conjutation (hence is the identity if $\mathbf{k} = \mathbf{R}$). We shall use the similarity between transition functions to construct a diffeomorphism.

Let h_1 and h_2 be the standard coordinate charts for $\mathbf{kP}^1$ so that the transition maps are given by φ as above. It will suffice to construct smooth mappings

$$f_{\pm} : \mathbf{k} = \mathbf{R}^d \longrightarrow S^d$$

such that each is a smooth cnart, the union of the images is S^d, and $f_+ = f_- \circ \varphi$. Specifically, let

$$f_+(z) \quad = \quad g_+(2\,\chi(z)) \qquad \text{and} \qquad f_-(z) \quad = \quad g_-(2\,z)$$

so that all the conditions except $f_+ = f_- \circ \varphi$ follow immediately. To verify this final condition, use the sequence of equations

$$f_*(z) \quad = \quad g_+(2\chi(z)) \quad = \quad g_- \circ \psi(2\chi(z)) \quad = \quad g_- \left(\frac{4}{4|z|^2} \cdot 2\chi(z) \right) \quad =$$

$$f_- \left(\frac{1}{|z|^2} \cdot z \right) \quad = \quad f_- \circ \varphi(z)$$

and this completes the proof that $\mathbf{kP}^1$ is diffeomorphic to S^d.■

V.1.2 : *Canonical line bundles*

If $\mathbf{k}$ is a field, then to each point in $\mathbf{kP}^n$ one has a naturally associated 1-dimensional subspace of $\mathbf{k}^{n+1}$ consisting of the zero vector and all sets of homogeneous coordinates for that point (recall that if we have two sets of homogeneous coordinates then one is a nonzero multiple of the other. As in the case of the tangent space of S^n we may think of this parametrized family as a subset $E(\mathbf{kP}^n)$ of $\mathbf{kP}^n \times \mathbf{k}^{n+1}$: Specifically, it is the set of all pairs (x, y) such that $y = 0$ or y is a set of homogeneous coordinates for x. If $\mathbf{k} = \mathbf{R}$ or $\mathbf{C}$ then this family can be topologized using the subspace topology inherited from $\mathbf{kP}^n \times \mathbf{k}^{n+1}$. The restriction of the first factor projection

$$\mathbf{kP}^n \times \mathbf{k}^{n+1} \longrightarrow \mathbf{kP}^n$$

defines a map η from $E(\mathbf{kP}^n)$ that takes the 1-dimensional vector space of homogeneous coordinates for a point x to the point x itself. Furthermore, this map is continuous if $\mathbf{k} = \mathbf{R}$ or $\mathbf{C}$.

One important property of tangent spaces is that they locally look like products $U \times \mathbf{R}^n$ such that the vector space operations correspond to the usual vector addition and scalar multiplication on $\mathbf{R}^n$. We want to show that $E(\mathbf{kP}^n)$ has a similar property.

If $1 \le j \le n+1$, and $x \in \mathbf{kP}^n$, then the following observation is elementary:

FACT. *If* **a** *and* **b** *are homogeneous coordinates for x and the j^{th} coordinate of* **a** *is zero, then the j^{th} coordinate of* **b** *is also zero.*

This follows because **a** and **b** are nonzero scalar multiples of each other.∎

One immediate consequence of this result is that if the j^{th} coordinate is nonzero in one set of homogeneous coordinates for point in $\mathbf{kP}^n$, then the same is true for every set of homogeneous coordinates. Therefore it is meaningful to define the set $U_j \subset \mathbf{kP}^n$ of all points such that "the j^{th} homogeneous coordinate is nonzero." Consider the map

$$\lambda_j : U_j \times \mathbf{k} \longrightarrow E(\mathbf{kP}^n) \quad \subset \quad \mathbf{kP}^n \times \mathbf{k}^{n+1}$$

which is defined on (x, a) by choosing homogeneous coordinate $\mathbf{v}$ for x and and given by the formula

$$\lambda_j(x, a) \quad = \quad \left(x, \frac{a}{v_j}\mathbf{v} \right) \ .$$

This map is well-defined because if $\mathbf{w} = b\,\mathbf{v}$ is another set of homogeneous coordinates for x we have

$$\frac{a}{v_j}\mathbf{v} \quad = \quad \frac{a}{w_j}\mathbf{w} \ .$$

Direct examination shows that λ_j is 1–1 and its image is onto $U_j \times \mathbf{k}^{n+1} \cap E(\mathbf{kP}^n)$. Furthermore, this map takes the natural vector space structure on $\{x\} \times \mathbf{k}$ to the natural vector space structure on its image in $\{x\} \times \mathbf{k}^{n+1}$. All of this is very similar to the standard charts which one has for the tangent bundle.

For our purposes it is also important to understand the transition maps $\Phi_{m,j}$ = "$\lambda_m^{-1} \circ \lambda_j$" which send $(U_m \cap U_j) \times \mathbf{k}$ to itself. It follows immediately that

$$\Phi_{m,j}(x, a) \quad = \quad \left(x, \frac{v_j}{v_m}\mathbf{v} \right)$$

where $\mathbf{v}$ is a set of homogeneous coordinates for x; note that the ratios v_j/v_m do not depend upon the choice of homogeneous coordinates and thus may be viewed as a map from $U_i \cap U_j$ to $\mathbf{k} - \{\mathbf{0}\}$. Furthermore, if $\mathbf{k}$ is the reals or complex numbers then the map $g_{m,j}$ corresponding to v_j/v_m is continuous from $U_j \cap U_m$ to $\mathbf{k} - \{\mathbf{0}\}$, and in fact this map is smooth with respect to the standard coordinate charts for $\mathbf{kP}^n$. Finally, we note that the maps $g_{m,j}$ satisfy compatibility conditions like resembling certain identities for the tangent bundle:

$$g_{j,j} \quad = \quad \mathrm{id}, \qquad g_{j,m} \quad = \quad g_{m,j}^{-1}, \qquad g_{p,m} \cdot g_{m,j} \quad = \quad g_{p,j}$$

The smoothness of the maps $g_{m,j}$ implies that the transition maps $\Phi_{m,j}$ define a smooth atlas for $\mathbf{kP}^n$ if $\mathbf{k} = \mathbf{R}$ or $\mathbf{C}$ if we identify the sets U_j with $\mathbf{k}^n$ using the charts for $\mathbf{kP}^n$. Furthermore, we claim that the projection map η is smooth with respect to this smooth structure on $E(\mathbf{kP}^n)$ and the previously defined smooth structure for $\mathbf{kP}^n$. The latter is true because we have $\eta \circ \lambda_j(v, a) = v$.

V.1.3 : *Basic definitions*

We begin with an abstract definition of a parametrized family of vector spaces.

Definition. Let B be a topological space. A *family of n-dimensional real or complex vector spaces parametrized by B* is a triple $\xi = (E, p, \Sigma, \mu)$ consisting of a topological space E, a continuous map $p : E \to B$, and continuous mappings $\Sigma : E \times_B E \to E$, (where $E \times_B E$ is the set of all $(x, y) \in E \times E$ such that $p(x) = p(y)$) and $\mu : \mathbf{k} \times E \to E$ (where $\mathbf{k} = \mathbf{R}$ or $\mathbf{C}$) such that the following hold:

(1) If $q : E \times_B E \to B$ is the map $q(x, y) = p(x) = p(y)$, then $p \circ \Sigma = q$,

(2) If $\pi : \mathbf{k} \times E \to E$ is projection onto the second factor, then $p \circ \mu = p \circ \pi$.

(3) If $E_x \subset E$ denotes the inverse image of $x \in B$, then for each such x the mappings $\Sigma_x : E_x \times E_x \to E_x$ and $\mu_x : \mathbf{k} \times E_x \to E_x$ defined by (1) and (2) make E_x into an n-dimensional vector space.

The space B is said to be the *base space* of the family, and E is said to be the *total space*.

At this point we have three fundamental examples.

Example 1. If M is a smooth manifold, then the tangent bundle $T(M)$ with its natural structure maps is a continuously parametrized family of n-dimensional real vector spaces over M, where $n = \dim M$.

Example 2. If X is any topological space, there is always the (standard) *trivial family* of vector spaces over $\mathbf{k} = \mathbf{R}$ or $\mathbf{C}$ such that $E = X \times \mathbf{k}^n$, the map $\Sigma : X \times \mathbf{R}^n \times \mathbf{R}^n \to X \times \mathbf{R}^n$ sends (x, v, w) to $(x, v + w)$, and the map μ sends (c, x, v) to (x, cv)

Example 3. If $M = \mathbf{k}P^n$ as above, then the preceding consstruction defines a 1-dimensional parametrized family of $\mathbf{k}$-vector spaces over M.

In the discussion of tangent bundles we also included a map $z : B \to E$, which sends a point $x \in B$ to the zero vector in the vector space $E_x = p^{-1}(\{x\})$. The following result implies that one often gets such a map from the remaining structure.

PROPOSITION. *Suppose that ξ is a family of vector spaces over a space B and that $p : E \to B$ is an open map. Then there is a continuous map $z : B \to E$ such that for each $x \in B$ the the zero vector in E_x is vector $z(x)$.*

Proof. Consider the map $\widetilde{z} : E \to E$ sending $v \in E$ to $\mu(0, v)$. The image of this map is the set of all zero vectors, and $p(v) = p(w) \Longrightarrow \widetilde{z}(v) = \widetilde{z}(w)$. Therefore there is a map $z : B \to E$ with the required properties if B has the quotient topology determined by E and p, and since p is open we know this is the case.■

Two families of vector spaces $(E_1, \dots)$ and $(E_2, \dots)$ over the same space X are said to be *isomorphic* if there is a homeomorphism $\varphi : E_1 \to E_2$ such that $p_2 \circ \varphi = p_1$ and for each x the map φ sends the vector space $(E_1)_x$ to $(E_2)_x$ by a $\mathbf{k}$-linear isomorphism.

In the study of topological spaces and smooth manifolds, one important theme is the construction of new examples out of old ones. For families of vector spaces, the most basic construction takes a family over some space B and yields a family over a subspace $A \subset B$:

PROPOSITION. *Let B is a topological space, let (E, p, Σ, μ) be a family of n-dimensional real or complex vector spaces parametrized by B, and let $A \subset B$. Let $E_A = p^{-1}(A)$, let $p_A : E_A \to A$ be the map defined by p, similarly let*

$$\Sigma_A : E_A \times_A E_A \to E_A \quad , \quad \mu_A : \mathbf{k} \times E_A \to E_A$$

be determined by Σ and μ respectively. Then $(E_A, p_A, \Sigma_A, \mu_A)$ is a family of n-dimensional real or complex vector spaces parametrized by A.■

We shall call the family $\xi|A = (E_A, p_A, \Sigma_A, \mu_A)$ the *restriction* of (E, p, Σ, μ) to A.

A family of h-dimensional vector spaces will said to be *trivial* if it is isomorphic to the standard trivial family.

Definition. A family of vector spaces (E, p, Σ, μ) is said to be an n-dimensional real or complex **vector bundle** if for each $x \in B$ there is an open neighborhood U containing x such that the restriction $(E_U, p_U, \Sigma_U, \mu_U)$ is trivial. Frequently this condition is summarized by the statement that ξ is *locally trivial.* Sometimes one also uses the term *real or complex n-plane bundle* over the base space B as a synonym for an n-dimensional real or complex vector bundle over B.

Previous examples revisited. By construction, each of the Examples 1–3 above is a vector bundle. Furthermore, if ξ is a vector bundle over B and $A \subset B$ then $\xi|A$ is a vector bundle over A. Finally, in these cases the projection map p is open; by local triviality this is true for $\xi|U_\alpha$ for some family of open subsets U_α that cover B, so that the restriction of p to each open set $p^{-1}(U_\alpha)$ is open mapping, and since the latter form an open covering of E it follows that p itself is an open mapping. In particular, this implies that for our examples one always has a continuous map $z : B \to E$ sending x to the zero vector in E_x.■

If all vector bundles were isomorphic to the standard trivial examples, then there would not be much point in defining vector bundles abstractly, so at this point it is useful to show that some of the preceding examples are not trivial. Here are some useful criteria.

PROPOSITION. *Suppose ξ is an n-dimensional vector bundle over B.*

(i) If ξ is the tangent bundle for S^2, then ξ is nontrivial.

(ii) If ξ is a real or complex 1-dimensional vector bundle over a space C, then ξ is trivial only if the complement of the set of zero vectors is homeomorphic to $B \times (\mathbf{k} - \{0\})$.

Proof. The proof of the second statement reduces quickly to the special case of trivial bundles, where it is immediate.

To prove the first statement, note that if the tangent bundle to any n-manifold is trivial, then there is a set of n linearly independent continuous vector fields which defines a basis at every point of M. Since every continuous vector field on S^2 is zero at some point of S^2, it follows that the tangent bundle to S^2 cannot be trivial.■

COROLLARY. *The canonical 1-dimensional vector bundle over $\mathbf{kP}^n$, where $\mathbf{k} = \mathbf{R}$ or $\mathbf{C}$, is nontrivial.*

Proof. Let $E_0 \subset E$ be the set of all nonzero vectors. If we can show that E_0 is homeomorphic to $\mathbf{k}^{n+1} - \{0\}$ then we can conclude that E_0 is nontrivial for the following reasons:

(1) If $\mathbf{k} = \mathbf{R}$ and E is trivial, then by the proposition we know that E_0 is disconnected, but $\mathbf{R}^{n+1} - \{0\}$ is connected so the bundle over $\mathbf{RP}^n$ must be nontrivial.

(2) If $\mathbf{k} = \mathbf{C}$ and E is trivial, then by the proposition we know that

$$\pi_1(E_0) \quad \cong \quad \pi_1\big(\mathbf{CP}^n \times (\mathbf{C} - \{0\})\big) \quad \cong \quad \pi_1(\mathbf{kP}^n) \times \mathbf{Z}$$

is nontrivial, but $E_0 \cong S^{2n+1} \times \mathbf{R}$ is simply connected so the bundle over $\mathbf{CP}^n$ must be nontrivial.

To prove the assertion about E_0 recall that E is defined to be a subset of $\mathbf{kP}^n \times \mathbf{k}^{n+1}$ such that the zero vectors in E correspond to the set $\mathbf{kP}^n \times \{0\}$. Therefore by definition E_0 is a subset of

$$\mathbf{kP}^n \times \left(\mathbf{k}^{n+1} - \{0\}\right)$$

and thus the composition of inclusion and projection onto the second coordinate yields a continuous (in fact, smooth) map $g : E_0 \to \mathbf{k}^{n+1} - \{0\}$. We claim this map is a homeomorphism, and we shall do so by constructing an inverse explicitly.

Consider the map $F : \mathbf{k}^{n+1} - \{0\} \to \mathbf{kP}^n \times \mathbf{k}^{n+1}$ sending a vector v with associated point $[v] \in \mathbf{kP}^n$ to the pair $([v], v)$. By construction the image of this map lies in E, and therefore it determines a continuous map $f : \mathbf{k}^{n+1} - \{0\} \to E$. By construction the map $g \circ f$ is the identity, and therefore both maps will be homeomorphisms if either f is onto or g is 1–1. Since every point in E has the form (x, v) where $x = [v]$, it follows immediately that f is onto, and thus both maps are homeomorphisms as claimed.■

Important special case. If $B = \mathbf{RP}^1 \cong S^1$ then we claim that E is just an open Möbius strip; removing the set of zero vectors will correspond to cutting this strip in the middle. One simple way to see the identification of E with the open Möbius strip is as follows: The semicircular arc $\gamma(t)$ in $\mathbf{R}^2 - \{0\}$ defined by $\exp(\pi\, i\, t)$, where $t \in [0, 1]$, passes to a closed curve in $\mathbf{RP}^1$, and if one takes the associated curve Γ in E defined by $\Gamma(t) = ([\gamma(t)],\ \gamma(t)\,)$, then by construction one the vectors $\Gamma(0)$ and $\Gamma(1)$ are negatives of each other.

Sections of a vector bundle. The map sending a point in the base of a vector bundle to the zero in the fiber satisfies the property $p \circ z = \mathrm{id}_B$ analogous to the corresponding property for the zero vector field on the tangent bundle of a smooth manifold. More generally, if $\sigma : B \to E$ is an arbitrary continuous map such that $p \circ \sigma = \mathrm{id}_B$, then we shall say that σ is a continuous **cross section** of the vector bundle.

V.1.4 : *Vector bundle atlases*

We are now faced with two questions:

Problem 1. Given an open covering $\mathcal{U} = \{U_\alpha\}$ of a space B, how can one assemble the spaces $U_\alpha \times \mathbf{k}^n$ to form an n-dimensional $\mathbf{k}$-vector bundle over B?

Problem 2. If B is a smooth manifold, what sort of extra structure is needed to construct a smooth vector bundle?

As in Unit III, questions of these sorts will be answered using a suitable notion of atlas (which is slightly different for each case). In each case one requires homeomorphisms or diffeomorphisms of special types, and the following result provides the motivation for the conditions that are needed.

CHARACTERIZATION OF AUTOMORPHISMS. (*i*) *Let X be a topological space, let $\mathbf{k}$ denote the real or complex numbers, and let Φ be a continuous automorphism of the trivial vector bundle $X \times \mathbf{k}^n$, where isomorphisms are defined as above. Then there is a continuous function $\varphi : X \to GL(n, \mathbf{k})$ such that*

$$\Phi(x, v) \quad = \quad (x,\ \varphi(x) v)$$

for all (x, v).

(*ii*) *Let U be an open subset of $\mathbf{R}^m$, let $\mathbf{k}$ be as above, and let Φ be a smooth automorphism of the trivial vector bundle $X \times \mathbf{k}^n$, where isomorphisms are defined as above. Then there is a smooth function $\varphi : X \to GL(n, \mathbf{k})$ such that*

$$\Phi(x, v) \quad = \quad (x,\ \varphi(x) v)$$

for all (x,v).

Explanation. We have already noted that $GL(n,\mathbf{R})$ is open in the space of $n\times n$ matrices over the real numbers, which is equivalent to $\mathbf{R}^{n^2}$, and continuity and smoothness for $\mathbf{k}=\mathbf{R}$ are interpreted in this sense. One can proceed similarly if $\mathbf{k}=\mathbf{C}$; the group $GL(n,\mathbf{C})$ is a subspace of the set of of $n\times n$ matrices over the complex numbers, which is equivalent to $\mathbf{C}^{n^2}\cong\mathbf{R}^{2n^2}$, and the invertible matrices again form an open subset because they are the matrices whose complex determinants are nonzero. — In the second case, note that the matrix product $\varphi(x)v$ will be a smooth function of x and v if φ is smooth.

Proof. By hypothesis, for each $x\in X$ the map Φ sends $\{x\}\times\mathbf{k}^n$ to itself, with (x,v) being sent to $(x,\,\varphi(x)v\,)$ for some invertible $n\times n$ matrix φ over $\mathbf{k}$. We need to show that the entries $a_{ij}(x)$ of this matrix are continuous in x if Φ is continuous and smooth if Φ is smooth.

Let π_1 and π_2 denote the projections of $X\times\mathbf{k}^n$ to the first and second factors, and let $\{\mathbf{e}_1,\cdots,\mathbf{e}_n\}$ be the standard basis of unit vectors for $\mathbf{k}^n$. We then have

$$a_{ij}(x) \quad=\quad \langle\,\pi_2\circ\Phi(x,\mathbf{e}_i),\,\mathbf{e}_j\,\rangle$$

where $\langle\ ,\ \rangle$ as usual denotes the standard real or complex inner product. Since the right hand side is continuous or smooth if Φ has the appropriate property, it follows that the matrix valued function φ is also continuous or smooth if the same holds for Φ.∎

CONVERSE TO AUTOMORPHISM CHARACTERIZATION. *In the setting of the preceding result, if φ is a continuous or smooth function from X to the appropriate group of invertible matrices, then the map*

$$\Phi(x,v) \quad=\quad (x,\,\varphi(x)v)$$

is a continuous or smooth automorphism of $X\times\mathbf{k}^n$ or $U\times\mathbf{k}^n$ respectively.

Verification of this converse is a straightforward computational exercise.∎The preceding results lead directly to the appropriate notion of atlas for a vector bundle.

Definition. Let B be a topological space, let $\mathcal{V}=\{\,V_\beta\,\}$ be an open covering of B, let $\xi=(E,p,\Sigma,\mu)$ be a vector bundle over B such that each restriction $\xi|V_\alpha$ is trivial, and let $\mathcal{A}=\{\,(U_\beta,h_\beta)\,\}$ be a collection of topological spaces with homeomorphisms onto the respective open subsets V_β. A *vector bundle atlas* for ξ over $\mathcal{A}$ then consists of *topological vector bundle charts* $F_\alpha:U_\alpha\times\mathbf{k}^n\to E$ such that

- (*i*) for each $(U_\alpha\times\mathbf{k}^n,F_\alpha)$ there is an associated chart $(U_{|\alpha|},\,h_{|\alpha|})$ in $\mathcal{A}$ such that $U_{|\alpha|}=U_\alpha$ and $p\circ F_\alpha(x,v)=h_{|\alpha|}(x)$ for all x and v (*i.e.*, F_α is a chart over $h_{|\alpha|}$),
- (*ii*) each F_α is a homeomorphism onto $p^{-1}(V_{|\alpha|})$
- (*iii*) for each $x\in U_\alpha$ the map $F_\alpha|\{x\}\times\mathbf{k}^n$ is a vector space isomorphism from the domain to E_y, where $y=h_{|\alpha|}(x)$.

If we let $\psi_{|\beta|\,|\alpha|}$ be the usual transition map "$h_{|\beta|}^{-1}h_{|\alpha|}$" then by the previous results on automorphisms of trivial vector bundles the corresponding transition maps $\Phi_{\beta\alpha}=$"$H_\beta^{-1}H_\alpha$" from $\left(h_{|\alpha|}^{-1}(h_{|\beta|}(U_{|\beta|}))\right)\times\mathbf{k}^n$ to $\left(h_{|\beta|}^{-1}(h_{|\alpha|}(U_{|\alpha|}))\right)\times\mathbf{k}^n$ have the form

$$\Phi_{\beta\alpha}(x,v) \quad=\quad \left(\,\psi_{|\beta|\,|\alpha|}(x),\,g_{\beta\alpha}(x)v\,\right)$$

for some continuous map

$$g_{\beta\alpha} : \left(h_{|\alpha|}^{-1}(h_{|\beta|}(U_{|\beta|})) \right) \longrightarrow GL(n, \mathbf{k}) \ .$$

The maps $g_{\beta\alpha}$ satisfy the previously stated conditions: $g_{\alpha\alpha}$ = identity, $g_{\alpha\beta} = g_{\beta\alpha}^{-1}$ and $g_{\gamma\alpha} = g_{\gamma\beta} \cdot g_{\beta\alpha}$ whenever the right hand side is meaningful.

V.1.5 : *Smooth vector bundles*

Suppose now that B is the underlying space of some smooth manifold, and let $\mathcal{B}$ be a smooth atlas for B. Given an m-dimensional real or complex vector bundle $\xi = (\pi : E \to B,\ etc.)$ it follows immediately that E is a topological $(n + dm)$-manifold, where $d = 1$ or 2 depending upon whether the scalars $\mathbf{k}$ are the real or complex numbers. Our next objective is to describe the extra structure needed to give a smooth vector bundle over the smooth manifold B. The preceding discussion of vector bundle atlases and the construction and analysis of the tangent bundle in Section III.5 will provide good models for our approach to the definition of smooth structures.

We shall begin with a simple but important observation:

PROPOSITION. *Let B be a smooth manifold and let ξ be an m-dimensional vector bundle over B with scalars $\mathbf{k}$, where $\mathbf{k} = \mathbf{R}$ or $\mathbf{C}$. Then there is a smooth atlas $\mathcal{B}_0$ for B such that for each smooth chart (U_α, h_α) in $\mathcal{B}_0$ the restricted bundle $\xi|h_\alpha(U_\alpha)$ is trivial.*

Proof. Let $\mathcal{B}$ be the maximal atlas for B, and let $\mathcal{W} = \{ W_\gamma \}$ be an open covering for B such that each restriction $\xi|W_\gamma$ is trivial. By the Chart Restriction Lemma there is a family of smooth charts $\mathcal{B}_0 \subset \mathcal{B}$ that covers B such that the image of every chart in $\mathcal{B}_0$ lies in one of the open subsets W_γ. Since the restriction of a trivial bundle is trivial (why?), it follows that for each smooth chart (U_α, h_α) in $\mathcal{B}_0$ the restricted bundle $\xi|h_\alpha(U_\alpha)$ is trivial.■

Definition. Let B be a smooth manifold with maximal atlas $\mathcal{B}$, let $\xi = (E, p, \Sigma, \mu)$ be a vector bundle over B, and let $\mathcal{B}_0 \subset \mathcal{B}$ be a subatlas such that for each smooth chart (U_α, h_α) in $\mathcal{B}_0$ the restricted bundle $\xi|h_\alpha(U_\alpha)$ is trivial. A *smooth vector bundle atlas* for ξ over $\mathcal{B}_0$ is a vector bundle atlas

$$\mathcal{V}_\xi \quad = \quad \{ \ (\, F_\alpha, \, U_\alpha \times \mathbf{k}^n \,) \ \}$$

over $\mathcal{B}_0$ such that the transition maps $\Phi_{\beta\alpha} =$"$F_\beta^{-1} F_\alpha$" defined as above are diffeomorphisms (equivalently, smooth).

If we let $\psi_{\beta\alpha}$ be the usual transition map "$h_\beta^{-1} h_\alpha$" then $\psi_{\beta\alpha}$ is a diffeomorphism because $\mathcal{B}_0$ is a smooth atlas, and in the previous description of Φ as

$$\Phi_{\beta\alpha}(x, v) \quad = \quad \left(\, \psi_{\beta\alpha}(x), \, g_{\beta\alpha}(x) v \, \right)$$

the maps $g_{\beta\alpha}$ are smooth.

The existence of a smooth vector bundle atlas implies a substantial amount of extra structure generalizing basic properties of the tangent bundle.

THEOREM. *Suppose that B is a smooth manifold, ξ is a vector bundle over B, and $\mathcal{V}$ is a smooth vector bundle atlas for ξ over some smooth atlas for B. Then the following hold:*

(i) $\mathcal{V}$ is a smooth atlas for E such that the projection map $p : E \to B$ is a smooth fiber bundle projection.

(*ii*) *With respect to this smooth structure the zero section map* $z : B \to E$ *and scalar multiplication map* $\mu : \mathbf{k} \times E \to E$ *are smooth.*

(*iii*) *There is a smooth atlas for* $E \times_B E$ *such that the projection to* B *is a smooth fiber bundle projection and the vector addition map* $\Sigma : E \times_B E \to E$ *is smooth.*

The proofs of these properties are completely analogous to the arguments for the tangent bundle of a smooth manifold.■

Definition. Given a smooth vector bundle atlas for ξ (hence smooth structure on E) we shall say that a cross section $\sigma : B \to E$ is a smooth cross section if it is smooth as a map from B to E. Note that if one has a smooth vector bundle atlas then the zero section is always smooth.

As in the case of ordinary smooth atlases, we want to describe standard examples of smooth vector bundle atlases in order to have relatively efficient definitions of smooth vector bundles. We shall do this in a sequence of steps.

Definition. Given a smooth manifold B and a vector bundle ξ over B, we shall say that a subatlas $\mathcal{B}_0$ of the maximal atlas is a *trivializing subatlas* for ξ if for each smooth chart (U_α, h_α) in $\mathcal{B}_0$ the restriction of ξ to the image of h_α is trivial.

LEMMA 1. *Every trivializing atlas is contained in a unique maximal trivializing subatlas.*

Proof. Simply take the set of all smooth charts such that the restriction of the vector bundle to each image is trivial.■

Example. A maximal trivializing subatlas for the base B is not necessarily a maximal atlas for B. We know that $\mathbf{kP}^1 \cong S^d$ where $d = \dim_{\mathbf{R}} \mathbf{k}$, and if we take the canonical 1-dimenional vector bundle over $\mathbf{kP}^1$ we know it is nontrivial. Suppose we take the cartesian product of everything in sight with the real line. Then $S^d \times \mathbf{R}$ is open in $\mathbf{R}^{d+1}$, but we claim that the bundle $\eta' = (p \times \mathrm{id}_{\mathbf{R}},\ etc\ )$ is not trivial. One simple way to see this is to observe that

(*i*) if $A \subset B$ and ξ is a trivial vector bundle over B, then $\xi|A$ is also trivial,

(*ii*) if we restrict η' to $S^d \times \{0\}$ we obtain the original (nontrivial) vector bundle.

The following result on maximal atlases can be established using the same methods employed in Section III.1 to prove the existence of a unique maximal atlas containing a given smooth atlas. It is not particularly difficult to carry out this argument, but it is time-consuming and the details are of a familiar nature and not particularly enlightening, so we shall leave them to the reader.

LEMMA 2. *Let* ξ *be a vector bundle over a smooth manifold* B *with a maximal trivializing atlas* $\mathcal{B}_0$*, and let* $\mathcal{V}$ *be a smooth vector bundle atlas over* $\mathcal{B}_0$*. Then* $\mathcal{V}$ *is contained in a unique maximal smooth vector bundle atlas* $\mathcal{V}^*$ *over* $\mathcal{B}_0$*.*■

Definition. Let ξ be a vector bundle over a smooth manifold B, and let $\mathcal{A}_0$ be a maximal trivializing smooth atlas for B with respect to ξ. A *smooth vector bundle* is a pair consisting of a vector bundle ξ and a maximal smooth vector bundle atlas over $\mathcal{B}_0$.

By the previous results concerning smooth atlases, it follows that E and $E \times_B E$ are smooth manifolds such that each of the basic structure maps p, μ, z and Σ is smooth and the projections onto B are smooth fiber bundle projections.■

V.1.6 : *Vector bundle amalgamation data*

In practice vector bundles are often constructed from the sort of data yielding the tangent bundle, so we shall describe the construction in a general manner. As usual $\mathbf{k}$ will denote either $\mathbf{R}$ or $\mathbf{C}$ and d will denote its real dimension.

Let X be a topological space, let $\mathcal{U} = \{U_\alpha\}$ be an open covering of X and let $\{\psi_{\beta\alpha}\}$ denote the associated transition data; *i.e.*, $\psi_{\beta\alpha}$ is the homeomorphism identifying $V_{\beta\alpha} \cong U_\alpha \cap U_\beta \subset U_\alpha$ with $V_{\alpha\beta} \cong U_\beta \cap U_\alpha \subset U_\beta$. An m-dimensional $\mathbf{k}$-*vector bundle preatlas over* $\mathcal{U}$ is a set of topological amalgamation data $(\{Y_\alpha\}, \{\Phi_{\beta\alpha}\})$ such that

(*i*) $Y_\alpha = U_\alpha \times \mathbf{k}^m$,

(*ii*) $\varphi_{\beta\alpha}$ maps $V_{\beta\alpha} \times \mathbf{k}^m$ to $V_{\alpha\beta} \times \mathbf{k}^m$ such that

$$\Phi_{\beta\alpha}(x, \mathbf{y}) = (\psi_{\beta\alpha}(x), F_{\beta\alpha}(x, \mathbf{y}))$$

where $F_{\beta\alpha}$ is continuous and every slice mapping $F_x : \{x\} \times \mathbf{k}^m \to \{\psi_{\beta\alpha}(x)\} \times \mathbf{k}^m$ is a vector space isomorphism.

The conditions describing a vector bundle preatlas are parallel to those on the amalgamation data for the tangent bundle, and in fact the methods used to construct and establish properties of the tangent bundle also allow one to construct a vector bundle over X from a vector bundle preatlas.

PREATLAS REALIZATION THEOREM. *Given an m-dimensional $\mathbf{k}$-vector bundle preatlas as above, there is an m-dimensional $\mathbf{k}$-vector bundle ξ such that*

(*i*) *for every open set U_α in the open covering $\mathcal{U}$, there is a homeomorphism $H_\alpha : U_\alpha \times \mathbf{k}^m \to \pi^{-1}(U_\alpha)$ such that $\pi(H_\alpha(x, \mathbf{y})) = x$ for all x,*

(*ii*) *for each pair of open sets U_α and U_β the transition map "$H_\beta^{-1} H_\alpha$" is equal to $\Phi_{\beta\alpha}$.*

The space E is Hausdorff if X is Hausdorff, and E is second countable if X is second countable.■

Alternate description. In the construction of the tangent bundle the counterparts of the maps $F_{\beta\alpha}$ have the form

$$F_{\beta\alpha}(x, \mathbf{y}) = [L_{\beta\alpha}(x)]\mathbf{y}$$

for continuous (in fact, smooth) maps from $V_{\beta\alpha}$ to the group $GL(k, \mathbf{R})$ of all invertible $k \times k$ matrices (recall that this group is in fact an open subset of the m^2-dimensional Euclidean space of all $m \times m$ matrices, and the matrix multiplication and inverse maps are smooth). The basic compatibility conditions for amalgamation data followed from identities of the form

(*i*) $L_{\alpha\beta}(x) = [L_{\beta\alpha}(\psi_{\alpha\beta}(x))]^{-1}$,
(*ii*) $L_{\gamma\alpha}(x) = L_{\gamma\beta}(\psi_{\beta\alpha}(x)) \cdot L_{\beta\alpha}(x)$.

A family of maps satisfying these identities is essentially a $GL(m, \mathbf{k})$-cocycle in the notation of Conlon, Section 3.4.

The construction method for the tangent bundle shows that a $GL(m, \mathbf{k})$-cocycle in the preceding sense yields a vector bundle, for using these maps one can form an associated preatlas with transition maps

$$\Phi(x, \mathbf{y}) = (\psi_{\beta\alpha}(x), L_{\beta\alpha}(x)\mathbf{y}).$$

In fact, given a vector bundle preatlas one can retrieve a $GL(m, \mathbf{k})$-cocycle from the results characterizing automorphisms of a trivial vector bundle that were proved earlier in this section.

Of course, there is a corresponding notion of *smooth vector bundle preatlas*, but a few changes are needed in order to formulate this. Given a smooth manifold M and a smooth atlas $\mathcal{A} = \{(U_\alpha, h_\alpha)\}$ for M, a k-dimensional smooth vector bundle atlas over $\mathcal{A}$ is a set of amalgamation data $(\{Y_\alpha\}, \{\Phi_{\beta\alpha}\})$ such that

(*i*) $Y_\alpha = U_\alpha \times \mathbf{k}^m$,

(*ii*) $\varphi_{\beta\alpha}$ maps $h_\alpha^{-1}(h_\beta(U_\beta)) \times \mathbf{k}^m$ to $h_\beta^{-1}(h_\alpha(U_\alpha)) \times \mathbf{k}^m$ by a diffeomorphism such that $\Phi_{\beta\alpha}(x, \mathbf{y})$ is equal to $(\text{“}h_\beta^{-1} h_\alpha\text{”}(x), F_{\beta\alpha}(x, \mathbf{y}))$, where $F_{\beta\alpha}$ is smooth and each slice map $F_x : \{x\} \times \mathbf{k}^m \to \{\psi_{\beta\alpha}(x)\} \times \mathbf{k}^m$ is a vector space isomorphism.

Once again the point of the definition is that the structure leads to a vector bundle.

SMOOTH PREATLAS REALIZATION THEOREM. *Given an m-dimensional smooth vector bundle preatlas as above, there is a smooth m-dimensional* $\mathbf{k}$*-vector bundle* (ξ, *etc.*) *such that*

(*i*) *for every smooth chart (U_α, h_α) in the atlas $\mathcal{U}$, there is a homeomorphism $H_\alpha : U_\alpha \times \mathbf{k}^m \to \pi^{-1}(h_\alpha(U_\alpha))$ such that $\pi(H_\alpha(x,\ \mathbf{y})) = h_\alpha(x)$ for all x,*

(*ii*) *for each pair of open sets U_α and U_β the transition map “$H_\beta^{-1} H_\alpha$” is equal to $\Phi_{\beta\alpha}$.*■

V.1.7 : *Smoothing vector bundles over smooth manifolds*

One question generated by the discussion of smooth bundle atlases is the following:

Smoothing problem. *Given a vector bundle ξ over a smooth manifold B, is there a smooth vector bundle atlas for ξ over a suitable smooth atlas for B. If so, how unique is this structure?*

In fact, it is always possible to construct a smooth atlas under the given conditions, and the smooth vector bundle structure is unique up to a structure preserving diffeomorphism; *i.e.*, if $\mathcal{V}$ and $\mathcal{V}'$ are maximal smooth vector bundle atlases over a maximal trivializaing smooth atlas $\mathcal{B}_0$ for B and $\mathcal{W}$ and $\mathcal{W}'$ are the corresponding maximal smooth atlases for E, then there is a diffeomorphism $\Phi : (E, \mathcal{W}) \to (E, \mathcal{W}')$ such that $p \circ \Phi = p$ and for each $x \in B$ the restriction Φ_x of Φ to a fiber E_x is a linear isomorphism from E_x to itself. This is a fairly straightforward consequence of general classification theorems for vector bundles, but it is not particularly easy to find in the literature. Further information appears in the file(s) `vbsmoothings.*` in the course directory.

V.2 : Constructions on vector bundles

(Conlon, § 3.4)

Not surprisingly, there is a variety of techniques for constructing new vector bundles out of old ones. Some reflect constructions in linear algebra and others reflect constructions in topology.

V.2.1 : *Pullbacks of vector bundles*

We have already noted that if ξ is a family of vector spaces over a space B and A is a subspace of B then there is an induced family of vector spaces $\xi|A$; in particular, if $p : E \to B$ is the projection for ξ, then the corresponding bundle projection for $\xi|A$ is $p|p^{-1}(A)$. If ξ is a vector bundle, then $\xi|A$ is also a vector bundle, for if $x \in A$ and U is an open neighborhood of x such that $\xi|U$ is trivial, then $U \cap A$ is an open neighborhood of x in A such that $U|U \cap A$ is trivial. The bundle $\xi|A$ is also called the *pullback* of ξ with respect to the inclusion map $i : A \subset B$.

This construction can be extended to an arbitrary continuous map. Here is a fast way of doing so. Given a continuous map $f : X \to Y$ we can write f as a composite $p_Y \circ \Gamma(f)$, where p_Y is the projection $X \times Y \to Y$ and $\Gamma(f) : X \to Y$ is the graph of f $(\Gamma(f)(x) = (x, f(x)))$. If ξ is a vector bundle over Y then one can construct a vector bundle $\xi \times X$ over $X \times Y$ whose projection is the product map $\mathrm{id}_X \times p : X \times E \to X \times B$ and whose other structure maps are similarly defined by taking products with $\mathbf{id}_X$.

If M and N are smooth manifolds, (ξ, *etc.*) is a smooth vector bundle over N and $f : M \to N$ is a smooth map, then it is possible to make the pullback bundle into a smooth vector bundle in a similar fashion. In analogy with the preceding paragraph, if one can do this for embeddings of smooth submanifolds, then the general case follows by looking at the graph of a smooth mapping. One step in this is almost trivial; if we have a smooth structure on ξ then one can construct a smooth structure on $M \times \xi$ without much trouble. Finding smooth structures for restrictions to submanifolds requires more work. Since we shall not need such smooth structures at any subsequent point, we shall simply state the main results without proof.

PROPOSITION. *Let ξ be a smooth vector bundle over the smooth manifold B, and suppose that $A \subset B$ is a smooth submanifold. Then there is a smooth vector bundle structure on $\xi|A$ such that E_A is a smooth submanifold of E and $E_A \times_A E_A$ is a smooth submanifold of $E \times_B E$.*■

V.2.2 : *External products*

If ξ and ξ' are families of $\mathbf{k}$-vector spaces over B and B', then their external product has projection map $p \times p' : E \times E' \to B \times B'$, so that the fiber above a typical point $(x, y) \in B \times B'$ is equal to the direct product vector space $E_x \times E'_y$. One can then define vector space structure maps in a straightforward manner; perhaps the most substantial observation one needs is the identification

$$(E \times E') \times_{B \times B'} (E \times E') \quad \cong \quad (E \times_B E) \times (E' \times_{B'} E') \ .$$

If ξ and ξ' are locally trivial, then the same is true for $\xi \times \xi'$, for if $(x, y) \in B \times B'$ and U and V are open neighborhoods of x and y such that $\xi|U$ and $\xi'|V$ are trivial, then $U \times V$ is an open neighborhood of (x, y) and

$$\xi'|U \times V \quad \cong \quad (\xi|U) \ \times \ (\xi'|V)$$

is also trivial. Furthermore, if we are given smooth vector bundle atlases for ξ and ξ' then it is possible to construct a smooth vector bundle atlas for the product by taking products of smooth vector bundle coordinate charts. The details of verifying this procedure are again left to the reader as an exercise.

One of the most basic constructions in linear algebra is the direct sum. For vector bundles over the same space B there is a corresponding notion of direct sum:

Definition. set of all $(a, b) \in E \times E'$ such that $p(a) = p'(b)$. Let $\sigma : E \times_B E' \to B$ send (a, b) to $p(a) = p'(b)$. The *direct sum* $\xi \oplus \xi'$ is defined to be the pullback of $\xi \times \xi'$ under the diagonal mapping $\Delta_B : B \to B \times B$. The resulting object is an $(m + n)$-dimensional $\mathbf{k}$-vector bundle over B.

Suppose now that B is a smooth manifold and we have smooth structures on ξ and ξ'. We shall provide some details on the construction of a smooth vector bundle atlas for $\xi \oplus \xi'$ because it similar methods will be needed later.

LEMMA. *Let ξ and ξ' be m- and n-dimensional $\mathbf{k}$-smooth vector bundles over the smooth manifold B. Then there is a smooth atlas $\{ (U_\alpha, h_\alpha) \}$ for B such that there are smooth bundle charts*

$$F_\alpha : U_\alpha \times \mathbf{k}^m \longrightarrow E$$

$$G_\alpha : U_\alpha \times \mathbf{k}^n \longrightarrow E'$$

such that $p \circ F_\alpha = p' \circ G_\alpha = h_\alpha \circ \pi_1$, where π_1 denotes projection onto the first factor.

Sketch of proof. Let $x \in B$. We claim there is a neighborhood W of x such that $p^{-1}(W) = F_\alpha(U_\alpha \times \mathbf{k}^m)$ and ${p'}^{-1}(W) = G'_\beta(U_\beta \times \mathbf{k}^m)$ for suitable vector bundle charts in the maximal smooth vector bundle atlases for ξ and ξ' respectively. This is true because one can find charts F_γ and G'_β whose image contain $p^{-1}(\{x\})$ and these charts can be restricted so that their images are the same open neighborhood of $p^{-1}(\{x\})$ because we are working with maximal smooth vector bundle atlases. If h_α and h_β are the smooth charts for B which are covered by F_α and G'_β respectively, then by construction we then have $h_\alpha(U_\alpha) = W = h_\beta(U_\beta)$. By our hypotheses we have $p \circ F_\alpha = h_\alpha \circ \pi_1$ and we have $p' \circ G'_\beta = h_\beta \circ \pi_1$, where π_1 denotes projection onto the first factor.

Let $\psi_{\beta\alpha}$ be the transition map from U_α to U_β and let

$$G_\alpha \quad = \quad G'_\beta \circ \big(\psi_{\beta\alpha} \times \mathrm{id}(\mathbf{k}^{\mathrm{n}}) \big) \ .$$

Since $(\psi_{\beta\alpha} \times \mathrm{id}(\mathbf{k}^{\mathrm{n}}))$ is smooth and linear on each vector space $\{pt.\} \times \mathbf{k}^n$ it follows that G_α also belongs to the maximal vector bundle atlas for ξ'. By construction we also have $p \circ G_\alpha = h_\alpha \circ \pi_1$.■

We can now form a smooth atlas for the total space $E \times_B E'$ of $\xi \oplus \xi'$ by the same process we employed to put a smooth atlas on $E \times_B E$ previously. Specifically, take vector bundle charts F_α and G_α as above, and consider the product maps

$$F_\alpha \times G_\alpha : U_\alpha \times \mathbf{k}^m \times U_\alpha \times \mathbf{k}^n \longrightarrow E \times E'$$

notice that the images of their restrictions to the sets of all (U, x, v, y) with $u = v$ are contained in $E \times_B E'$, and let $H_\alpha : U_\alpha \times \mathbf{k}^m \times \mathbf{k}^n \to E \times_B E'$ be the maps defined in this fashion. It then follows that the transition maps associated to these charts are smooth vector bundle automorphisms and therefore define a smooth vector bundle atlas.

Important remark. From the viewpoint of geometric cocycles the direct sum construction has a simple formal interpretation. Suppose that geometric cocycles for ξ and ξ' are given by families of smooth functions $g_{\beta\alpha} : U_{\beta\alpha} \to GL(m, \mathbf{k})$ and $g'_{\beta\alpha} : U_{\beta\alpha} \to GL(n, \mathbf{k})$. Then a geometric cocycle for $\xi \oplus \xi'$ is given by $\mathbf{B}(g_{\beta\alpha}, g'_{\beta\alpha})$ where

$$\mathbf{B} : GL(m, \mathbf{k}) \times GL(n, \mathbf{k}) \longrightarrow GL(m+n, \mathbf{k})$$

is the block sum construction sending a pair of matrices (A, C) to their block sum $A \oplus C$. Since $\mathbf{B}$ is a smooth group homomorphism it follows that by $\mathbf{B}(g_{\beta\alpha}, g'_{\beta\alpha})$ is also a smooth geometric cocycle, and as noted before the vector bundle it defines is just the direct sum.

We shall encounter many other constructions of this type that yield important examples of vector bundles. The important point is that one has a smooth homomorphism from one general linear group Γ, or perhaps a product of several general linear groups, to some other general linear group Γ'.

V.2.4 : *Riemannian metrics*

Inner products are another basic concept in linear algebra that extend to vector bundles.

Definition. Let $\xi = (\pi,\ etc.)$ be a vector bundle with scalar field $\mathbf{k} = \mathbf{R}$ or $\mathbf{C}$; for each $x \in B$ denote the vector space $p^{-1}(\{x\})$ by E_x as usual and denote the canonical projection $E \times_B E \to E$ by π_2. A (continuous) **riemannian metric** on ξ is a continuous map

$$g : E \times_B E \to \mathbf{k}$$

such that for each $x \in B$ the map g defines an inner product

$$g_x : E_x \times E_x \cong \pi_2^{-1}(\{x\}) \to \mathbf{k}$$

If we are given a (maximal) smooth atlas on B and an associated richer structure of a smooth vector bundle for Π, the metric will be said to be **smooth** if g is smooth.

Frequently in the literature the term "riemannian metric" is reserved for real vector bundles and inner products on complex vector bundles are called *Hermitian metrics.*

Of course inner products play an important role in the study of real and complex vector spaces, and riemannian (and Hermitian) metrics play at least an equally important role in the study of vector bundles and smooth manifolds (and they are arguably even more important for manifolds).

The following elementary observation is important for our purposes.

PROPOSITION. *Let V be a vector space over $\mathbf{k} = \mathbf{R}$ or $\mathbf{C}$, let $g_i : V \times V \to \mathbf{k}$ be an inner product on V for $i = 1$ and 2, and let s be a real number in the closed interval $[0, 1]$. Then $s \cdot g_1 + (1-s) \cdot g_2$ is also an inner product on V.*■

This proposition has an extremely far-reaching consequence:

EXISTENCE THEOREM FOR RIEMANNIAN METRICS. *Let ξ be a vector bundle such that the base space B is paracompact Hausdorff (e.g., take B to be a second countable topological manifold). Then there is a riemannian metric on ξ. If B has a smooth atlas and we are given an associated richer structure of a smooth vector bundle for ξ, then a smooth riemannian metric exists.*

Proof. We begin with the topological conclusion. Take a locally finite open covering $\mathcal{U}$ of B such that the bundle looks like a product over each open set U_α in the open covering. For each U_α a riemannian metric f_α can be constructed on $\pi^{-1}(U_\alpha) \cong U_\alpha \times \mathbf{k}^m$ simply by taking the standard inner product on each vertical slice $\{u\} \times \mathbf{k}^m$. Let φ_α be a partition of unity subordinate to $\mathcal{U}$. Then we can extend each product function $\varphi_\alpha \cdot f_\alpha$ to a continuous map g_α from $E \times_B E$ to $\mathbf{k}$, and if $W_\alpha \subset B$ is the set of points where $\varphi_\alpha \neq 0$, then then the restriction of g_α to $\pi_2^{-1}(W_\alpha)$ is a riemannian metric. By local finiteness of the open covering the sum $g = \Sigma_\alpha g_\alpha$ is meaningful; but for each $x \in B$ there is an α such that $\varphi_\alpha(x) \neq 0$, and therefore by the proposition it follows that g defines a riemannian metric on all of $E \times_B E$.

The existence of a smooth metric follows by taking a smooth atlas $\mathcal{A} = \{(V_\beta, k_\beta)\}$ for B such that each image $k_\beta(V_\beta)$ lies inside some U_α from $\mathcal{U}$ and the resulting open covering $\mathcal{V} = \{k_\beta(V_\beta)\}$ is locally finite (verify that one can find an atlas so that both conditions hold). One can then choose a *smooth* partition of unity φ_b subordinate to $\mathcal{V}$, and the corresponding metrics f_β over the open sets in this open covering are smooth by construction. Therefore the sum $g = \Sigma_\beta \varphi_\beta \cdot f_\beta$ is a smooth riemannian metric.∎

To illustrate the usefulness of riemannian metrics we shall use them to define the length of a smooth curve in a smooth manifold.

Notation. If g is a smooth riemannian metric on the tangent bundle τ_M and $\Gamma : (c,d) \to M$ is a smooth curve and $t \in (c,d)$, then $\Gamma'(t) \in T_{\Gamma(t)}(M)$ is the image of $(t,1)$ under the canonical identification

$$(c,d) \times \mathbf{R} \cong T\left((c,d)\right)$$

followed by the associated map of tangent spaces

$$T(\Gamma) : T\left((c,d)\right) \to T(M).$$

If $c < a < b < d$ then the **length** of $\Gamma|[a,b]$ is given by the integral

$$\int_a^b \sqrt{g\left(\Gamma'(t), \Gamma'(t)\right)}\, dt$$

which exists by the smoothness of g and Γ.

One can also define the length of a smooth curve if it is only defined on the closed interval $[a,b]$, but this will be left to the reader in order to keep the discussion relatively brief.

Many important examples of riemannian metrics arise from inclusions of submanifolds. For example, suppose that M is a smooth submanifold of $\mathbf{R}^n$ and i denotes the inclusion map. Then $T(i)$ defines a smooth embedding of $T(M)$ in $T(\mathbf{R}^n) \cong \mathbf{R}^n \times \mathbf{R}^n$, and if P_2 denotes projection onto the coordinate then for each $p \in M$ the composite $P_2 \circ T(i)|T_p(M)$ is a 1–1 linear transformation. Therefore if $\mathbf{v}$ and $\mathbf{w}$ lie in $T_p(M)$ then

$$\langle P_2 \circ T(i)\mathbf{v}, P_2 \circ T(i)\mathbf{w} \rangle$$

defines a riemannian metric on $T(M)$. In the classical theory of surfaces where $n = 3$ and $\dim M = 2$, this riemannian metric is called the **First Fundamental Form** of the embedded surface.

In mathematical physics (or older books on tensor analysis and differential geometry) a riemannian metric is often described locally over each chart in an atlas, with a compatibility requirement for the definitions over different charts. We shall state a version of this approach from the perspective of these notes, but first we need an observation and some notation. If $\mathbf{Symm}(n)$ is the set of all symmetric $n \times n$ matrices, then $\mathbf{Symm}(n)$ is a subspace of the space of all $n \times n$ matrices, and its dimension is $\frac{1}{2}\,(n^2 + n)$. As usual, given a matrix A we shall denote its transpose by ${}^{\mathbf{T}}A$.

PROPOSITION. *Let M^n be a smooth manifold, let $\mathcal{A} = (U_\alpha, h_\alpha)$ be a smooth atlas for M, and let $(U \times \mathbf{R}^n), H_\alpha)$ be the chart for the tangent space $T(M)$ associated to (U_α, h_α). Suppose that for each α we have a smooth map $G_\alpha : U_\alpha \to \mathbf{Symm}(n)$ such that for all $x \in U_\alpha$ and nonzero $\mathbf{v} \in \mathbf{R}^n$ we have ${}^{\mathbf{T}}\mathbf{v}G_\alpha(x)\mathbf{v} > 0$, and that for all α and β we have $G_\beta(\,\text{“}h_\beta^{-1}h_\alpha\text{”}(x)\,) = {}^{\mathbf{T}}L_{\alpha\beta}(x)G_\alpha(x)L_{\alpha\beta}(x)$, where $L_{\alpha\beta}(x)$ is equal to $D\,\text{“}h_\beta^{-1}h_\alpha\text{”}(x)$. Then there is a smooth riemannian metric g on the tangent bundle of M such that*

$$g\,(H_\alpha(x,\ \mathbf{v}), H_\alpha(x,\ \mathbf{w})) = {}^{\mathbf{T}}\mathbf{w}G_\alpha(x)\mathbf{v}$$

for all x, $\mathbf{v}$ and $\mathbf{w}$.

Notation. If the entries of G_α are denoted by $g_{i,j}$ then the latter are smooth functions on U_α and the classical presentation of the metric locally is an expression of the form

$$\sum_{i,j} g_{i,j}(u)\,dx^i\,dx^j$$

(and frequently the summation sign is suppressed, the convention being that if a variable appears twice then one sums over it).

V.2.5 : *Riemannian metrics and distance functions* $(2\star)$

In elementary plane, solid and spherical geometry, one often thinks of the distance between two points as the length of the shortest curve in the given set joining the two points. Since a riemannian metric provides a method for defining lengths of piecewise smooth curves, it is meaningful to ask if one can use this structure to define a distance function on a manifold in a similar fashion. Modulo one relatively minor complication, this turns out to be the case.

GEOMETRIC DISTANCE THEOREM. *Let M be a smooth manifold with a riemannian metric g, and define the lengths of piecewise smooth curves on M using g as above. Given two points $x, y \in M$, let $\mathbf{d}_g(x, y)$ be the greatest lower bound of the lengths of all piecewise smooth curves Γ that joint x and y. Then $\mathbf{d}_g$ defines a metric on M.*

Proof. Most of the conditions for a metric are very straightforward to check. For example, $\mathbf{d}_g \geq 0$ because lengths of curves are always nonnegative, and $\mathbf{d}_g(x, x) = 0$ because the constant curve at x joins x to itself and has length zero. The symmetric property $\mathbf{d}_g(y, x) = \mathbf{d}_g(x, y)$ follows because if Γ is a piecewise smooth curve joining x to y then the opposite curve Γ^{op} is a piecewise smooth curve joining y to x and the lengths of Γ and Γ^{op} are equal; it follows immediately that $\mathbf{d}_g(y, x) \leq \mathbf{d}_g(x, y)$, and by reversing the roles of x and y it follows that $\mathbf{d}_g(y, x) = \mathbf{d}_g(x, y)$. To prove the triangle inequality for x, y and z note that if Γ is a piecewise smooth curve joining x to y and Γ' is a piecwise smooth curve joining y to z, then the concatenation $\Gamma + \Gamma'$ is a picewise smooth that joins x to z such that

$$\mathrm{Length}(\Gamma\ +\ \Gamma')\ \ =\ \ \mathrm{Length}(\Gamma)\ +\ \mathrm{Length}(\Gamma')\ .$$

Therefore we have

$$\mathbf{d}_g(x,z) \quad \leq \quad \mathrm{Length}(\Gamma) \; + \; \mathrm{Length}(\Gamma')$$

for all choices of Γ and Γ'. It follows that

$$\mathbf{d}_g(x,z) \; - \; \mathrm{Length}(\Gamma) \quad \leq \quad \mathrm{Length}(\Gamma')$$

for all Γ' and therefore

$$\mathbf{d}_g(x,z) \; - \; \mathrm{Length}(\Gamma) \quad \leq \quad \mathbf{d}_g(y,z) \; .$$

The latter inequality implies

$$\mathbf{d}_g(x,z) \; - \; \mathbf{d}_g(y,z) \quad \leq \quad \mathrm{Length}(\Gamma)$$

for all Γ, which in turn implies

$$\mathbf{d}_g(x,z) \; - \; \mathbf{d}_g(y,z) \quad \leq \quad \mathbf{d}_g(x,y)$$

and one obtains the Triangle Inequality from the latter by adding $\mathbf{d}_g(y,z)$ to both sides.

The only remaining condition to verify is that $\mathbf{d}_g(x,y) = 0$ implies $x = y$. This turns out to be somewhat delicate and requires several preliminary observations. The basic idea is to prove the result first for riemannian metrics on open subsets of $\mathbf{R}^n$ and then to extend the result using some simple topological methods.

We begin the proof for disks with some general observations about arbitrary riemannian metrics on such open sets. By definition, a riemannian metric on an open subset U of $\mathbf{R}^n$ is completely determined by a smooth matrix valued function G on U such that for each $x \in \mathbf{R}^n$ the matrix $G(x)$ is the *Gram matrix* for the inner product on $\{x\} \times \mathbf{R}^n$ whose entries are defined by

$$g_{ij}(x) \quad = \quad g\left((x, \mathbf{e}_i), \, (x, \mathbf{e}_j) \right)$$

where $\{ \mathbf{e}_1, \; \cdots \; , \mathbf{e}_n \}$ is the standard set of unit vectors in $\mathbf{R}^n$.

A standard result in linear algebra states that an $n \times n$ matrix G is the Gram matrix for an abstract inner product on $\mathbf{R}^n$ k if and only if $G = {}^{\mathbf{T}}P \cdot P$ for some invertible matrix P; in fact, if $Q = P^{-1}$ then the columns of Q give an orthonormal basis for $\mathbf{R}^n$ with respect to the inner product defined by G (PROOF: If Q is given as above then orthonormality implies $I = {}^{\mathbf{T}}Q \, G \, Q$ and therefore $P = Q^{-1}$ implies that

$${}^{\mathbf{T}}P \, P \quad = \quad {}^{\mathbf{T}}P \, I \, P \quad = \quad {}^{\mathbf{T}}P \, {}^{\mathbf{T}}Q \, G \, Q \, P \quad = \quad I \cdot G \cdot I \quad = \quad G \; .)$$

We shall need a parametrized version of this result.

LEMMA 1. *If U is open in $\mathbf{R}^n$ and $G(x)$ is an $n \times n$ smooth matrix valued function corresponding to a smooth riemannian metric on U, then there is a smooth function $P : U \to GL(n, \mathbf{R})$ such that $G(x) = {}^{\mathbf{T}}P(x) \cdot P(x)$ for all $x \in U$.*

Sketch of proof. The idea is to use the Gram-Schmidt process. More precisely, if we start with the standard set of basic vector fields $\frac{\partial}{\partial x_i}$ on U then we can apply the Gram-Schmidt process to these vector fields to obtain a set of n smooth vector fields $\mathbf{p}_j$ on U that are othonormal with respect to the riemanian metric g at every point $x \in U$. One can then define $P(x)$ to be the matrix valued function whose columns are given by the vector fields $\mathbf{p}_j(x)$.■

We also need the following result giving an upper bound for the norm of a matrix:

LEMMA 2. *If A is an $m \times n$ matrix over the real numbers then its norm satisfies the inequality*

$$\|A\| \quad \leq \quad \left(\sum_{i,j} a_{ij}^2 \right)^{1/2}$$

where the entries of A are the coefficients a_{ij}.

Proof. If $\mathbf{v} \in \mathbf{R}^n$ is written as a linear combination of unit vectors $\sum_j \ x_j \, \mathbf{e}_j$ then by definition we have

$$A\,\mathbf{v} \quad = \quad \sum_{i,j} a_{ij} \, x_j \, \mathbf{e}_i$$

so we are interested in estimating the quantity

$$\sum_i \left(\sum_j a_{ij} \, x_j \right)^2 .$$

By the Cauchy-Schwarz Inequality the summands satisfy

$$\left(\sum_j a_{ij} \, x_j \right)^2 \quad \leq \quad \left(\sum_j a_{ij} \right)^2 \cdot |\mathbf{v}|^2$$

and if we sum these inequalities over all values of i we see that

$$|A\,\mathbf{v}|^2 \quad \leq \quad \left(\sum_{i,j} a_{ij}^2 \right) \cdot |\mathbf{v}|^2$$

which immediately yields the bound for $\|A\|$.∎

We shall now apply these observations to riemannian metrics on an open disk.

PROPOSITION. *Let g be a smooth riemannian metric on $N_r(0; \mathbf{R}^n)$. Then there is a positive constant K such that if γ is a piecewise smooth curve joining 0 to a point y for which the image of γ is entirely contained in $N_{3r/4}(0; \mathbf{R}^n)$, then the length $L_g(\gamma)$ with respect to g and the Euclidean length $L_0(\gamma)$ satisfy $L_g(\gamma) \geq K \cdot L_0(\gamma)$.*

Proof. Let G be the smooth function which gives the Gram matrix of g at a point of $N_r(0)$, and using the previous results write $G(x) = {}^{\mathbf{T}}P(x) \cdot P(x)$ where we know $P(x)$ is a continuous function of x. Since $P(x)$ is invertible we know that

$$|\mathbf{v}| \quad = \quad |P(x)^{-1} P(x) \mathbf{v}| \quad \leq \quad \|P^{-1}(x)\| \cdot |P(x)\mathbf{v}|$$

which implies that

$$|P(x)\mathbf{v}| \quad \geq \quad \|P(x)^{-1}\|^{-1} \cdot |\mathbf{v}|$$

and by the previous lemma we know that the right hand side is bounded from below by $\beta(x) \cdot |\mathbf{v}|$, where $\beta(x)$ is a smooth positive valued continuous function of x. Let K be the minimum value of $\beta(x)$ for $|x| \leq \frac{3}{4}\,r$. Then we have

$$L_g(\gamma) \quad = \quad \int_0^a \sqrt{\langle G\circ\gamma(t)\,\gamma'(t),\ \gamma'(t)\rangle}\;dt \quad = \quad \int_0^a |P\circ\gamma(t)\,\gamma'(t)|\,dt \quad \geq$$

$$K \cdot \int_0^a |\gamma'(t)|\,dt \quad = \quad K \cdot L_0(\gamma)$$

as required.∎

COROLLARY. *In the preceding discussion we have* $L_g(\gamma) \geq K\,|y|$.

This is true because $|y|$ is the length of the line segment joining 0 to y, and this line segment is the shortest curve joining the two points.∎

Completion of the proof of the Geometric Distance Theorem. Let $x \in M$ and suppose we are given a smooth coordinate chart $(N_r(0), h)$ such that $h(0) = x$. Suppose that $y \neq x$ and γ is a piecwise smooth curve joining x to y.

Let S be the image of the sphere of radius $r/2$ under h. Then $h(\,N_{r/2}(0)\,)$ is an open and closed subset of $M - S$ which contains x. If $y \notin h(\,N_{r/2}(0)\,)$ then connectedness considerations imply that the image of γ must contain a point of S.

Suppose now that $y \notin h(\,N_{r/2}(0)\,)$ and let $z \in S$ be a point in the image of γ. Since $\gamma^{-1}(S)$ is compact it follows that this subset of a closed interval contains a least element t_0. We might as well assume that $z = \gamma(t_0)$ at this point. Let γ_0 denote the restriction of γ to $[0, t_0]$. It follows that

$$L_g(\gamma) \quad \leq \quad L_g(\gamma_0) \quad \leq \quad K \cdot \frac{r}{2}$$

where (1) the first inequality is true because $L_g(\gamma_0)$ and $L_g(\gamma)$ have the same positive integrand but the integral for $L_g(\gamma_0)$ is taken over a subinterval, (2) the second inequality is a consequence of the previous proposition. Since the right hand side is a positive bound that is independent of γ, it follows that $\mathbf{d}_g(x, y) \geq Kr/2 > 0$ if $y \notin h(\,N_{r/2}(0)\,)$.

Suppose now that $y \in h(\,N_{r/2}(0)\,)$. If the image of γ is entirely contained in $h(\,N_{r/2}(0)\,)$ then the preceding corollary implies that $L_g(\gamma) \geq K\,|h^{-1}(y)|$. On the other hand, if the image of γ is not entirely contained in this open set, then the argument in the previous paragraph implies that $L_g(\gamma) \geq Kr/2$. Since $|h^{-1}(y)| \leq \frac{1}{2}r$ it follows that $L_g(\gamma) \geq K|h^{-1}(y)|$ in both cases so that of γ, it follows that $\mathbf{d}_g(x, y) \geq K|h^{-1}(y)| > 0$. Therefore we have shown that $x \neq y \Longrightarrow \mathbf{d}_g(x, y) > 0$ as required.∎

Example. Frequently it is possible to find a smooth curve of length $\mathbf{d}_g(x, y)$ joining x to y, but this is not always the case. One simple way of constructing an example is to take the standard riemannian metric on the tangent bundle of $\mathbf{R}^2 - \{0\}$. Then the distance between $(+1, -0)$ and $(-1, 0)$ in the above sense is equal to 2, but there is no piecewise smooth curve of length 2 joining these points that lies entirely inside of the set $\mathbf{R}^2 - \{0\}$. Verification of this is left to the exercises for this section. This example may look somewhat artificial because it is obtained by removing one point from a "good" example, but it is not difficult to construct other examples that cannot be fixed in this way.

Footnote. A celebrated theorem proved by J. Nash in the nineteen fifties states that *every smooth riemannian metric can be realized as a metric coming from some smooth embedding of* *the manifold in a Euclidean space.* Here is an online site describing the result further and giving references:

`http://en.wikipedia.org/wiki/Nash_embedding_theorem`

On the other hand, another celebrated theorem proved by D. Hilbert at the beginning of the twentieth century shows that one cannot realize the hyperbolic metric on the unit disk $\mathbf{H}^2$ in $\mathbf{R}^2$, which is defined by the formula

$$\frac{dx^2 + dy^2}{\left(1 - (x^2 + y^2)\right)^2}$$

as the First Fundamental Form for a smooth $\mathcal{C}^2$ embedding of $\mathbf{H}^2$ as a closed subset of $\mathbf{R}^3$. The following online document discusses Hilbert's Theorem along with many related topics from hyperbolic geometry, and it also gives references to complete proofs of Hilbert's result; the discussion of the latter begins at the bottom of page 9 in the article.

`http://www.math.utah.edu/~treiberg/Hilbert/Hilbert.pdf`

V.2.6 : *Generalizations of riemannian metrics* (2⋆)

Inner products are special cases of *symmetric nondegenerate bilinear forms*, for which the positive definiteness condition

$$\mathbf{v} \neq 0 \qquad \Longrightarrow \qquad \langle \mathbf{v}, \mathbf{v} \rangle \ > \ 0$$

is replaced by a nondegeneracy condition: *For each nonzero* $\mathbf{v}$ *there is a vector* $\mathbf{w}$ *such that* $\langle \mathbf{v},\ \mathbf{w} \rangle \neq 0$.

Such forms on a finite-dimensional real vector space V are classified up to equivalence by their *type*, which can be viewed as an ordered pair of nonnegative integers (r, s) such that $r + s = \dim V$ and there are subspaces V_+ and V_- such that

(*i*) we have $V_+ \cap V_- = \{\mathbf{0}\}$, $V_+ + V_- = V$, $\dim V_+ = r$ and $\dim V_- = s$,

(*ii*) if $\mathbf{x} \in V_+$ and $\mathbf{y} \in V_-$ then $\langle \mathbf{x},\ \mathbf{y} \rangle = 0$,

(*iii*) if $\mathbf{x} \in V_+$ is nonzero then $\langle \mathbf{x}, \mathbf{x} \rangle \ > 0$ and if $\mathbf{y} \in V_-$ is nonzero then $\langle \mathbf{y}, \mathbf{y} \rangle \ < 0$.

A riemannian metric is merely the special case where $s = 0$. We shall use the term *indefinite metric of type* (r, s) to denote a map g as above such that for each $z \in B$ the restriction to $E_z \times E_z$ is a nondegenerate symmetric bilinear form of type (r, s) where $s > 0$.

There are important contexts in which one wishes to consider analogs of riemannian metrics given by maps $g : T(M) \times_M T(M) \to \mathbf{R}$ such that the restriction to each $T_p(M) \times T_p(M)$ is a nondegenerate bilinear form that is not an inner product. For example, in relativity one considers *Lorentz metrics* on 4-dimensional manifolds for which the type is $(3, 1)$. Closely related objects are needed in classical mechanics, where considers maps $\Omega : T(M) \times_M T(M) \to \mathbf{R}$ that are bilinear, nondegenerate and *skew-symmetric* (*i.e.*,

$$\Omega\,(\mathbf{y}, \mathbf{x}) \ = \ -\Omega\,(\mathbf{x}, \mathbf{y})$$

for all $\mathbf{x}$ and $\mathbf{y}$); such a structure is called a *pre-symplectic structure* on M. It is not possible to construct indefinite metrics on arbitrary tangent bundles using partitions of unity as for riemannian metrics because the analog of the convexity proposition does not hold. In fact, indefinite metrics of a prescribed type and pre-symplectic structures do not necessarily exist on an arbitrary manifold.

V.2.7 : *The Second Fundamental Form* (2$\star$)

The **Second Fundamental Form** of an oriented hypersurface (smooth submanifold) $M^{n-1} \subset \mathbf{R}^n$ is another example of a smooth map

$$T(M) \times_M T(M) \to \mathbf{R}$$

whose restriction to each $T_p(M) \times T_p(M) \cong \tau_2^{-1}(\{p\})$ is symmetric and bilinear. However, this map may be degenerate at some points (or even zero everywhere!). The orientation on an oriented hypersurface is essentially given by a unit normal vector field

$$\mathbf{N} : M \to \mathbf{R}^n$$

such that $\mathbf{N}(p)$ is perpendicular to $T_p(M)$, where the latter is viewed as a subspace of $\{p\} \times \mathbf{R}^n$ via the linear injection from $T_p(M)$ to $\{p\} \times \mathbf{R}^n$ induced by inclusion.

If $\mathbf{v} \in T_p(M)$ then it is not difficult to show that the image of $T_p(N)\mathbf{v}$ in $\{p\} \times \mathbf{R}^n$ is perpendicular to $\mathbf{N}(p)$ and therefore lies in $T_p(M)$. This means that $T(\mathbf{N})$ defines a smooth map S from $T(M)$ to itself such that for each $p \in M$ the map S send $T_p(M)$ to itself linearly. If we let $\mathbf{F}^{\mathbf{I}}_M$ denote the first fundamental form, then the Weingarten map has the self adjointness property

$$\mathbf{F}^{\mathbf{I}}_M\,(S(\mathbf{v}),\ \mathbf{w}) = \mathbf{F}^{\mathbf{I}}_M\,(\mathbf{v},\ S(\mathbf{w}))$$

for all

$$(\mathbf{v},\ \mathbf{w}) \quad \in \quad T(M) \times_M T(M)$$

and the second fundamental form is defined to be

$$\mathbf{F}^{\mathbf{II}}_M\,(S(\mathbf{v}),\ \mathbf{w}) \quad = \quad \mathbf{F}^{\mathbf{I}}_M\,(S(\mathbf{v}), \mathbf{w})\,.$$

If $n = 2$ the Weingarten map provides a very neat way to handle some classical concepts from the differential geometry of oriented surfaces in $\mathbf{R}^3$. Since S is self-adjoint, it has an orthonormal basis of (real) eigenvectors with real eigenvalues. The two eigenvalues of S at the point p are the *principal sectional curvatures* at p, half the trace of S at p is the *mean curvature* at p, and the determinant of S at p is the *Gaussian curvature* at p. An important result in differential geometry, known as the **Theorema Egregium** of Gauss, states that the Gaussian curvature only depends upon the **First** Fundamental Form. Among other things, this result leads one to a concept of curvature for arbitrary riemannian metrics.

V.3 : Cotangent spaces and differential 1-forms

(Conlon, §§ 6.1–6.3)

One important goal of this section is to study the following question that was raised earlier in these notes:

Gradient vector field construction. *Is there some way of extending the construction of gradients from smooth real valued functions on open subsets of Euclidean spaces to smooth real valued functions defined on an arbitrary smooth manifold?*

One immediate question is what happens to the gradient when one changes coordinates using a transition function $\psi_{\beta\alpha} : U_{\beta\alpha} \to U_{\alpha\beta}$ where the domain and codomain are open subsets in $\mathbf{R}^n$ and $\psi_{\beta\alpha}$ is a diffeomorphism. Under the transition function a smooth map $f : U_{\alpha\beta} \to \mathbf{R}$ should correspond to $g = f \circ \psi_{\beta\alpha}^{-1} : U_{\alpha\beta} \to \mathbf{R}^n$, and by the chain rule the coordinates for the gradients of f and g are related as follows:

CHANGE OF COORDINATES FORMULA. *In the setting above we have*

$$\nabla g\,(\,\psi_{\beta\alpha}(x)\,) \quad = \quad {}^{\mathbf{T}}[D\psi_{\beta\alpha}(x)\,]^{-1}\,\nabla f$$

for all $x \in U_{\beta\alpha}$.

Sketch of proof. If $u = \psi_{\beta\alpha}(x)$, then the coordinates on the left hand side of the equation are the functions $\partial g/\,\partial u_i$ and the coordinates of ∇f are the functions $\partial f/\partial x_j$. By the Chain Rule these are related by the equations

$$\frac{\partial f}{\partial x_j} \quad = \quad \sum_i \frac{\partial g}{\partial u_i} \cdot \frac{\partial u_i}{\partial x_j}$$

and one can rewrite this as the following equation involving column vectors:

$$\nabla f \quad = \quad {}^{\mathbf{T}}D\psi_{\beta\alpha}\,\nabla g$$

This is clearly equivalent to the equation in the formula.■

This result is initially disappointing because the change of coordinates formula is not what one needs to form a global vector field. However, in the next subsection we shall see that one actually has something which is not quite in the form we initially would have hoped for but gives us what we need for many purposes.

V.3.1 : *Dual vector bundles*

We have already seen that the direct sum construction on vector spaces can be extended to a construction on vector bundles. Similarly, the dual space construction summarized in Section V.A can also be extended, and in fact it plays an important role in the theory of smooth manifolds (and the applications of this theory to other subjects as well).

If V is a finite dimensional vector space over a field $\mathbf{k}$ then V and V^* are isomorphic as vector spaces, but there are numerous contexts where it is still useful to have both available. One way of distinguishing between the n-dimensional vector space $\mathbf{k}^n$ and its dual is to view the former as the set of all $n \times 1$ column vectors and the latter as the set of all $1 \times n$ row vectors. With this convention the evaluation of a linear functional $\mathbf{w}^*$ in the dual space of $\mathbf{k}^n$ at a vector $\mathbf{v} \in \mathbf{k}^n$ is given by the matrix product $\mathbf{w}^* \cdot \mathbf{v}$ (note that this is a 1×1 matrix).

If $\mathbf{k} = \mathbf{R}$ or $\mathbf{C}$ and ξ is an n-dimensional vector bundle over a space B, then there is a natural concept of dual vector bundle such that the fibers F_x of the dual bundle are dual spaces to the fibers of the original vector bundle. Formally, one proceeds as follows: Suppose we are given a vector bundle ξ with a vector bundle atlas with charts $F_\alpha : U_\alpha \times \mathbf{k}^n \to E$ and transition maps $\Phi_{\beta\alpha}$ defined by a geometric cocycle $g_{\alpha\beta}$. The dual bundle is formed from amalgamation data given by $U_\alpha \times \mathbf{k}^n$ with new transition maps $\Phi^*_{\beta\alpha}$ which are defined as

$$\Phi^*_{\beta\alpha}(x, \mathbf{v}) \quad = \quad \left(\psi_{\beta\alpha}(x),\, {}^{\mathbf{T}}g^{-1}_{\beta\alpha}(x)\mathbf{v} \,\right)$$

where $\psi_{\beta\alpha}$ denotes the transition function associated to the open covering $\{ U_\alpha \}$. Since the transposed inverse (or "contragredient") construction defines a smooth group automorphism of $GL(n, \mathbf{k})$ it follows that the functions above define vector bundle amalgamation data. Moreover, if we started with a smooth vector bundle atlas then these data are smooth amalgamation data. The resulting dual space bundle is denoted by ξ^* and we often denote the projection by $p^* : E^* \to B$.

If we specialize this to the tangent bundle of a smooth manifold we obtain the **cotangent bundle** $\tau^*_M : T^*(M) \to M$.

One important motivation for introducing the cotangent bundle is that it provides a framework for defining a global version of the gradient of a smooth function. We shall need the following principle for constructing global cross sections out of pieces.

SECTION CONSTRUCTION LEMMA. *Suppopse we are given an atlas of charts (U_α, h_α) for a space B and a q-dimensional $\mathbf{k}$-vector bundle ξ over B whose restriction to each set $h_\alpha(U_\alpha)$ is trivial. Let $\{ \psi_{\beta\alpha}$ be the transition functions and let $g_{\beta\alpha} : V_{\beta\alpha} \to GL(q, |bfk)$ be the associated functions in a vector bundle atlas for ξ. If for each α we have a local cross section $s_\alpha : U_\alpha \to |bfk^q$ and these satisfy the compatibility condition*

$$s_\beta \circ \psi_{\beta\alpha}(u) \quad = \quad g_{\beta\alpha}(u)]s_\alpha(u)$$

then there is a cross section $s : B \to E$ such that for all u the point $s \circ h_\alpha(u)$ is the image of $s_\alpha(u)$ in E. If all atlases in sight are smooth and the functions s_α are all smooth, then s is also smooth.

This can be shown by the same methods used to construct vector fields in Unit IV.■

THEOREM. *Let M be a smooth manifold, and let $f : M \to \mathbf{R}$ be a smooth function. Then there is a unique smooth cross section df of the cotangent bundle τ^*_M such that if (U, h) is an arbitrary smooth chart for M and the associated vector bundle chart for $T^*(M)$ is H^*_α, then*

$$df \circ h(x) \quad = \quad H^*_\alpha\big(x, \nabla[f \circ h](x)\big) \ .$$

Sketch of proof. The identities in the conclusion of the theorem imply that df must be unique if it exists because they yield its values at every point of M. Therefore we need to prove existence.

The problem here is analogous to the global vector field construction principle in Section IV.1, for we have the section of the cotangent bundle defined locally and we need to show that these local definitions are compatible under changes of coordinates.

Given the smooth n-manifold M with smooth atlas $\mathcal{A} = \{(U_\alpha, h_\alpha)\}$, suppose that for each α we are given a smooth map $g_\alpha : U_\alpha \to \mathbf{R}^n$. These maps will determine a cross section of the cotangent bundle if and only if they satisfy the condition

$$g_\beta \circ \psi_{\beta\alpha}(u) \quad = \quad {}^{\mathbf{T}}[D\psi_{\beta\alpha}(u)]^{-1} g_\alpha(u) \ .$$

If we take $g_\alpha = \nabla(f \circ h_\alpha)$, the Change of Coordinates Formula at the beginning of this section implies that the compatibility condition is satisfied.■

V.3.2 : *The parametrized dual pairing*

For vector spaces the evaluation map $e : V^* \times V \to \mathbf{k}$ defined by $e(f, v) = f(v)$ is a bilinear map. We claim that there is a well behaved global version of this map for the tangent and cotangent bundles. More precisely, there is a smooth map $e : T^*(M) \times_M T(M) \to \mathbf{R}$ such that for each $p \in M$ the restriction of e to $T_p^*(M) \times T_p(M)$ is just this canonical evaluation map.

Locally this is easy to do. If we are given an open subset of Euclidean space, then the manifold $T^*(M) \times_M T(M)$ is simply $U \times \mathbf{R}^n \times \mathbf{R}^n$, and the map we want is simply the one sending $(u,\ \mathbf{v},\ \mathbf{w})$ to ${}^{\mathbf{T}}\mathbf{v} \cdot \mathbf{w}$.

We must now verify that this definition is compatible with the transition maps defining $T^*(M) \times_M T(M)$. If $\mathcal{A} = \{(U_\alpha, h_\alpha)\}$ is a smooth atlas for M, then charts for $T^*(M) \times_M T(M)$ have the form $(U_\alpha \times \mathbf{R}^n \times \mathbf{R}^n,\ etc.)$ and the associated transition maps send $(u,\ \mathbf{v},\ \mathbf{w})$ to

$$\left(\psi_{\beta\alpha}(u), {}^{\mathbf{T}}[D\psi_{\beta\alpha}(u)]^{-1}\mathbf{v}, [D\psi_{\beta\alpha}(u)]\mathbf{w}\right) .$$

Therefore everything reduces to checking whether ${}^{\mathbf{T}}\mathbf{v} \cdot \mathbf{w}$ is equal to

$${}^{\mathbf{T}}\left({}^{\mathbf{T}}[D\psi_{\beta\alpha}(u)]^{-1}\mathbf{v}\right) \cdot \left([D\psi_{\beta\alpha}(u)]\mathbf{w}\right) .$$

But the latter simplifies to

$$\left({}^{\mathbf{T}}\mathbf{v}[D\psi_{\beta\alpha}(u)]^{-1}\right) \cdot \left([D\psi_{\beta\alpha}(u)]\mathbf{w}\right)$$

which further simplifies to ${}^{\mathbf{T}}\mathbf{v} \cdot \mathbf{w}$ as desired.■

Notational convention. If

$$e_M : T^*(M) \times_M T(M) \longrightarrow \mathbf{R}$$

is the map described above, then we frequently denote $e_M(w, v)$ by $\langle w, v \rangle$. Following standard terminology from tensor analysis, we shall also say that e_M is the *contraction* mapping or pairing.

We shall now prove a result that raises questions about the reasons for carrying out the entire construction of the cotangent bundle.

THEOREM. *If M is a smooth manifold, then the tangent and cotangent bundles are isomorphic vector bundles.*

Proof. Let g be a riemannian metric on (the tangent bundle of) M. We can define a map Γ of sets from $T(M)$ to $T^*(M)$ such that for each p the map sends $T_p(M)$ to $T_p(M)^* \cong T_p^*(M)$ by the formula $[\Gamma_p(\mathbf{v})]\mathbf{w} = g(\mathbf{v},\ \mathbf{w})$. For each p the map Γ_p is well defined, it maps $T_p(M)$ top a subspace of the same dimension, and it is 1–1 because $0 = \Gamma_p(v) \Rightarrow 0 = [\Gamma_p(\mathbf{v})]\mathbf{v} = g_p(\mathbf{v}, \mathbf{v})$ and the latter is zero if and only if $\mathbf{v} = 0$. Thus Γ_p is an isomorphism for each p.

We need to show that this map is a diffeomorphism. It suffices to work locally. Suppose that U is open in $\mathbf{R}^n$ and for each $u \in U$ let $G(u)$ be the Gram matrix of the riemannian metric g with respect to the standard unit basis (hence the (i, j) entry is the value of g at $(u,\ \mathbf{e}_i,\ \mathbf{e}_j)$ where $\{\mathbf{e}_k\}$ denotes the standard unit vectors in $\mathbf{R}^n$). Then Γ takes the form

$$\Gamma(u,\ \mathbf{v}) = \big(u, {}^{\mathbf{T}}\mathbf{v} \cdot G(u)\big)$$

and the smoothness of this follows because G is smooth. Therefore we have shown that we have a smooth map $\Gamma : T(M) \to T^*(M)$ with the desired properties. Since $G(u)$ is an invertible symmetric matrix for all u one can show directly that Γ^{-1} is given locally by

$$\Gamma^{-1}(u,\ \mathbf{w}) = \big(u,\ G(u)^{-1}({}^{\mathbf{T}}\mathbf{w})\big)$$

and therefore Γ is a diffeomorphism.■

We return to the question: Why do we need both the tangent and the cotangent spaces? The reason is that each is better for some purposes and each has different uses. Cross sections of the tangent bundle provide the right way to look at ordinary differential equations and the Lie bracket. Cross sections of the cotangent bundle provide the right way to generalize the gradient, and we shall see that such cross sections are also useful for other purposes, including defining line integrals on manifolds.

One can generalize the preceding argument to show that every real vector bundle is isomorphic to its dual. However, over the complex numbers a vector bundle is not necessarily isomorphic to its dual.

On a more abstract note, the crucial point is that the isomorphism between a finite dimensional vector space and its dual space is unnatural in the sense that it requires one to pick some extra structure in order to construct an isomorphism; the structure may be a basis or an inner product or certain generalizations of either, and it is often very clumsy to manipulate objects using these isomorphisms that depend on extrinsic data.

Implications for defining gradients. Given a riemannian metric the theorem above yields an isomorphism Γ from $T(M)$ to $T^*(M)$. One can use this isomorphism to define **gradient vector fields with respect to the given metric.** Specifically, if $f : M \to \mathbf{R}$ is a smooth function one takes $\mathbf{Grad}_g(f)$ to be the vector field $\Gamma \circ df$. Therefore it is possible to generalize the ordinary gradient for smooth functions on open subsets of Euclidean spaces, but it is necessary to have metric structure in order to do so. In the ordinary case the metric structure is so simple that it is essentially invisible in the definition.

Having dualized the tangent bundle, it is natural to ask if there is a corresponding dualization of the tangent space map $T(f) : T(M) \to T(N)$ associated to a smooth map $f : M \to N$. The following result describes such a dualization. Note that we must state this result in terms of pullback bundles; one obvious difficulty with trying to define a map from $T^*(N)$ to $T^*(M)$ is that there is no reasonable way of canonically defining a smooth map $f^* : N \to M$ that is dual to $f : M \to N$ unless f is a diffeomorphism.

THEOREM. *Suppose that $f : M \to N$ is a smooth map. Then there is a unique smooth map $T^\bullet(f) : f^*T^*(N) \to T(M)$ such that for each $x \in M$ the map $T^*(f)$ sends the fiber $f^*T^*(N)_{f(x)} = T_x(N)^*$ to $T^*(M)_x = T_x(M)^*$ by the dual $T_x(f)^*$ of the usual map of tangent spaces:*

$$T_x(f) : T_x(M) \longrightarrow T_{f(x)}(N)$$

The construction sending f to $T^*(f)$ satisfies the identities $T^\bullet(\mathrm{id}_M) = \mathrm{id}_{T^*(M)}$ and $T^\bullet(f \circ g) = T^\bullet(g) \circ T^\bullet(f)$.

Sketch of proof. The description of $T^\bullet(f)$ yields a unique set-theoretic map from $f^*T^*(N)$ to $T(M)$ as well as the identities in the final sentence of the theorem. Therefore we need only show that the mapping in question is smooth. As usual we shall do this locally.

Suppose that we are given smooth charts (U, h) at x and (V, k) at $f(x)$ such that f maps $h(U)$ into $k(V)$; assume that U and V are open in $\mathbf{R}^n$ and $\mathbf{R}^m$ respectively. Denote the associated map "$k^{-1}fh$" by $g : U \to V$. Everything then reduces to describing the map

$$T^\bullet(g) : V \times \mathbf{R}^m \longrightarrow U \times \mathbf{R}^n$$

defined as in the statement of the theorem. The definition implies this map is given by the formula

$$T^\bullet(g)(x, w^*) \quad = \quad (x, \, {}^{\mathbf{T}}Dg(x)w^*)$$

and by the smoothness of g we know that this mapping is smooth.■

COMPLEMENT. *Suppose that $f : M \to N$ is a diffeomorphism. Then there is a unique smooth map $T^*(f) : T^*(N) \to T^*(M)$ that maps each cotangent space $T^*_y(N)$ linearly to the cotangent space $T^*_{f^{-1}(y)}(M)$ by the map of dual spaces associated to the linear map from $T_{f^{-1}(y)}(M)$ to $T_y(N)$. This construction satisfies the identities $T^*(\mathrm{id}_M) = \mathrm{id}_{T^*(M)}$ and $T^*(f \circ g) = T^*(g) \circ T^*(f)$.*

Sketch of proof. The idea is basically the same; we have uniquely described the map, the identities follow from the definition, and we only need to check that it is smooth. This is done locally, where one has the formula

$$T^*(g)(x, w^*) \quad = \quad \left(x, \, {}^{\mathbf{T}}Dg(x)w^* \right)$$

which immediately yields the smoothness of the mapping under consideration.■

Given an arbitrary vector bundle $\pi : E \to B$ one can imitate the construction above to construct a good fiberwise evaluation map $\varepsilon_E : E^* \times_B E \to \mathbf{R}$ in analogy with the tangent bundle. If E is a real vector bundle and one has a riemannian metric on E it is also possible to construct a vector bundle isomorphism $E \to E^*$ like the isomorphism Γ (depending upon the metric) that we constructed for the tangent bundle.

V.3.3 : *Differential* 1*-forms*

We define a *differential* (or *exterior*) 1*-form* on M to be a (smooth) cross section of the cotangent bundle; *i.e.*, a smooth map $\omega : M \to T^*(M)$ such that $\tau^*_M \circ \omega = \mathbf{1}_M$.

As in the case of vector fields, if ω is a differential 1-form on M and U is open in M, then the restriction $\omega|U$ defines a differential 1-form on U.

Following standard notational conventions, if U is open in $\mathbf{R}^n$ we write dx^i to denote the 1-form sending u to $(u, \mathbf{e}_i^*)$ where $\mathbf{e}_i^*$ is the i^{th} vector in the dual basis to the standard unit vectors. Clearly we then have

$$e\left(dx^i, \frac{\partial}{\partial x^j}\right) = \delta_j^i$$

where δ_j^i is 0 if $i \neq j$ and 1 if $i = j$ (i.e., the values are given by the *Kronecker delta function*). On U every form can be expressed as a $C^\infty(M)$ linear combination

$$\omega(p) \quad = \quad \sum_i \; g^i(p)\, dx^i$$

for suitable smooth functions $g^i : U \to \mathbf{R}$.

In the language of tensor analysis, differential 1-forms correspond to *covariant vector fields of rank* 1.

The set of all differential 1-forms on a manifold M is a module over the algebra $\mathcal{C}^\infty(M)$, and it is denoted by $\wedge^1(M)$. Since the tangent and cotangent bundles are isomorphic we know that $\mathbf{X}(M)$ and $\wedge^1(M)$ are isomorphic as $\mathcal{C}^\infty(M)$-modules.

Given a smooth differential 1-form ω on a smooth manifold M and a smooth vector field X on M one can use the evaluation or **contraction** mapping $e : T^*(M) \times_M T(M) \longrightarrow \mathbf{R}$ to define a real valued function $e_M(\omega, X)$. For our purposes it will be helpful to know that the value of this pairing has reasonable invariance properties with respect to some diffeomorphism $f : M \to N$. In Unit IV we described the direct image f_*X of a vector field X on M under a diffeomorphism f. The corresponding definition of direct image for 1-forms is as follows:

$$f_*\omega(p) \quad = \quad T^*(f^{-1}) \circ \omega \circ f^{-1}(p)$$

One can immediately check this is linear and satisfies the naturality properties $\mathrm{id}_* = \mathrm{id}$ and $(g \circ f)_* = g_* \circ f_*$. In particular, it follows that the direct image construction defines an isomorphism between 1-forms on M and 1-forms on N. With this definition of direct image for forms, we have the desired invariance property:

$$e_N\,(\,f_*\omega,\; f_*X\,)\,[f(p)] \quad = \quad e_M\,(\,\omega,\; X\,)\,[p]$$

Verification of this formula is left to the reader as an exercise.

Vector fields have a natural extra structure given by the Lie bracket, and the differential forms have a much different extra structural property: *Given a smooth map $f : M \to N$, there is a natural way of pulling back a 1-form on N to a 1-form on M.*

We shall prove the existence of pullback forms indirectly, and the method will be similar to the construction of the Lie bracket in an important respect: The latter used the characterization of vector fields in terms of derivations on smooth functions, and here we shall use a characterization of differential forms as suitable operators on vector fields. Here is the basic result:

DIFFERENTIAL 1-FORM RECOGNITION PRINCIPLE. *Let ω be a differential 1-form on M, let U be open in M and let $\mathbf{X}(U)$ be a $\mathcal{C}^\infty(U)$-module of vector fields on U. Then ω defines* *a $\mathcal{C}^\infty(U)$-linear map $E\omega_U$ from $\mathbf{X}(U)$ to $\mathcal{C}^\infty(U)$ via the contraction mappings e_U, and if $V \subset U$ these maps $E\omega_U$, $E\omega_V$ satisfy the compatibility relation*

$$[E\omega_U(Y)]\,|V \quad = \quad E\omega_V(Y|V)\ .$$

Conversely, If $\{\, L_U \,\}$ is any system of $\mathcal{C}^\infty(U)$-linear maps from $\mathbf{X}(U)$ to $\mathcal{C}^\infty(U)$ satisfying the compatibility condition $[L_U(Y)]\,|V \;=\; L_V(Y|V)$, then there is a unique differential 1-form ω such that $L = E\omega$.

Proof. It follows immediately that one can define $E\omega_U$ so that it has the desired properties.

Suppose now that we are given a system of maps L_U with the specified properties. Consider first the case where M is open in $\mathbf{R}^n$. Let f_i be the value of L_M at the basic vector field $\partial/\partial x_i$. Then by $\mathcal{C}^\infty(M)$-linearity and the compatibility under restrictions we conclude that $L = E\omega$ where $\omega = \sum_i\ f_i\, dx^i$. Furthermore there is only one such 1-form with this property because if $E\omega = E = \lambda$ then their values on the basic vector fields would be the same and that would imply that the coefficients of each dx^i would be equal.

The conclusion now also follows for manifolds that are diffeomorphic to open subsets of $\mathbf{R}^n$ by the invariance property for direct images.

To dispose of the general case take an open covering $\{\, W_\alpha \,\}$ of M by images of smooth coordinate charts. This yields differential 1-forms ω_α on each W_α, and by the uniquess and compatibility properties we have $\omega_\alpha|W_\alpha \cap W_\beta = \omega_\beta|W_\alpha \cap W_\beta$. Therefore there is a unique 1-form ω on M such that $\omega|W_\alpha = \omega_\alpha$ for all α.

We now need to show that

$$e_V\big(\, E\omega_V, X\,\big) \quad = \quad L_V(X)$$

for each vector field X on an open subset V. Now both sides of this equation are smooth functions, so it is enough to show that their restrictions to each open set $V \cap W_\alpha$ agree. Let U be a typical set of this type. Since the restriction of L to open subsets U of W_α is given by ω_α and $\omega|W_\alpha = \omega_\alpha$ it follows that $E\omega_U(X|U) = L_U(X|U)$ for $U = V \cap W_\alpha$; by the compatibility properties the left and right hand sides are equal to the restrictions of $E\omega_V(X)$ and $L_V(X)$ to U. Therefore we have shown that $E\omega_V(X) = L_V(X)$ as required.■

Before proceeding we note that the previous result yields a simple characterization of the exterior derivative. Namely, if f is a smooth function on M, then df is the unique 1-form such that $e_U(df, X)$ is the Lie derivative $X\cdot(f|U)$ for all vector fields X defined on open subsets $U \subset M$. Note that (1) the family of operators L_U defined by the Lie derivative satisfies all the conditions of the theorem, (2) for open subsets of $\mathbf{R}^n$ one can check directly that this yields the exterior derivative, (3) if φ is a coordinate chart from some open set W to M and $f : M \to \mathbf{R}$ is smooth, then by our previous construction the restriction of df to $\varphi(W)$ is $\varphi_* d(f\circ\varphi)$.

Definition. Let $f : M \to N$ be a smooth manifold, and let ω be a differential 1-form on N. The *pullback form* $f^*\omega$ on M is then the unique form θ such that

$$E\theta_U(X)[p] \quad - \quad e_N\big(\,\omega(f(p)),\, T(f)X(p)\,\big)$$

for all vector fields X defined on open subsets of M. It is a routine exercise to verify that the right hand side satisfies the condition in the recognition result, and this guarantees the existence and uniqueness of the pullback. It also follows immediately that the pullback construction

$$f^* : \wedge^1(N) \to \wedge^1(M)$$

is a linear transformation. It is also a routine exercise to verify that the pullback construction has the basic naturality properties $(g \circ f)^* = g^* \circ f^*$ and $\mathrm{id_M}^* = \mathrm{id}[\wedge^1(M)]$.

What does the pullback construction look like in local coordinates? Suppose V is open in $\mathbf{R}^n$, U is open in $\mathbf{R}^m$ and $f : U \to V$ is smooth. Locally a differential 1-form ω on V is given by a vector valued function $\mathbf{g} : V \to \mathbf{R}^n$, and by the identity characterizing $f^*\omega$ the corresponding vector valued function $\mathbf{h} : U \to \mathbf{R}^m$ satisfies

$$\mathbf{h}(x) \quad = \quad {}^{\mathbf{T}}[Df(x)]\mathbf{g}(f(x)) \ .$$

Alternatively, if we write ω locally as $\sum_i \, g_i \, du_i$ and $x_1, \ \cdots \ , x_m$ are the standard Cartesian coordinates for U then

$$f^*\omega \quad = \quad \sum_j \left(\sum_i \, g_i \, \frac{\partial f_i}{\partial x_j} \right) \cdot dx_j \ .$$

the local form

$$du = \sum \frac{\partial u}{\partial x^i} \, dx^i \ .$$

Important notational point. In the literature of mathematics and physics one often sees dx^i used in place of dx_i. It is important to recognize that in such notation **the variable i is a superscript and not an exponent.**

Remark. The form that is traditionally called dx^i is in fact the exterior derivative of the i^{th} coordinate function $x^i : \mathbf{R}^n \to \mathbf{R}$.

The exterior derivative and pullback construction have an important compatibility property:

COMPATIBILITY FORMULA. *If $h : N \to \mathbf{R}$ is smooth and $f : M \to N$ is smooth, then $f^*dh = d(f \circ h)$.*

Derivation. It suffices to prove this over an open covering of M, and it is convenient to take an open covering of M by images of coordinate charts in M such that f sends each of these to the image of some coordinate chart in N. This means that we can reduce to the situation of a smooth map $f : U \to V$ where U and V are open in Euclidean spaces as in the preceding discussion. The preceding local formula for the pullback yields the equation

$$f^*dh \quad = \quad \sum_j \left(\sum_i \frac{\partial h}{\partial u_i} \cdot \frac{\partial f_i}{\partial x_j} \right) \cdot dx^j \ .$$

On the other hand we have

$$df^*h \quad = \quad d(h \circ f) \quad = \quad \sum_j \ \frac{\partial (hf)}{\partial x_j} \, dx^j$$

and by the Chain Rule the right hand side of this equation is equao to the right hand side of the preceding one.■

V.3.4 : *Line integrals*

In multivariable calculus a line integral of a vector field has an integrand of the form $\sum_i \, f_i \, dx^i$, which looks very much like a differential 1-form in local coordinates. Thus it should not be surprising that one can generalize line integrals to curves in smooth manifolds if one takes the integrand to be a differential 1-form. There is an extensive discussion of this in Section 6.3 of Conlon. In these notes we shall just summarize some of the main points in terms of our setting and notation.

In order to simplify the discussion we shall limit attention to the following special cass of curves.

Hypothesis. *For the rest of this section, we shall only consider continuous curves* $\gamma : [a,b] \to M$ *that are piecewise smooth in the following sense: There is a partition*

$$\Delta \quad = \quad \{ \; a \; = \; t_0 \; < \; t_1 \; < \; \cdots \; < \; t_q \; = \; b \; \}$$

such that for each i *the restriction* γ_i *to the* i^{th} *closed subinterval extends to a smooth curve on an open interval containing the subinterval* $[t_{i-1},\, t_i]$.

We would like to define the line integral

$$\int_\gamma \omega$$

of a smooth 1-form as follows. For each i the 1-form $\gamma_i^*\omega$ has the form $g_i(t)$ for some smooth function g_i; the line integral should then be given by the sum

$$\sum_i \int_{t_{i-1}}^{t_i} \gamma^*\omega \quad = \quad \sum_i \int_{t_{i-1}}^{t_i} g_i(s)\, ds \; .$$

We need to show this is independent of the choice of Δ, and we use an approach that also occurs elsewhere in real analysis:

(1) The additivity properties of integrals show that value of the sum does not change if we take a refinement of the partition; *i.e.*, we insert additional points.

(2) Given two partitions one can always construct a common refinement.

If U is open and $\omega = \sum_i F_i\, dx^i$ then this definition reduces to the usual definition of the line integral

$$\int_\gamma \textstyle\sum_i F_i\, dx^i \; .$$

In multivariable calculus it is well known that a line integral of the form $\int_\gamma df$ is independent of path. This generalizes to arbitrary smooth manifolds as follows: Suppose that we have a partition Δ, so that the line integral $\int_\gamma df$ is a sum of the integrals $\int \gamma_i^* df$ where each γ_i is smooth. The Fundamental Theorem of Calculus then implies that the summands are given by

$$f \circ \gamma(t_i) \;-\; f \circ \gamma(T_{i-1})$$

and if we add these we see that the line integral is equal to the path independent expression $f(\gamma(b)) - f(\gamma(a))$.∎