Amiya Mukherjee

Differential Topology

Second Edition

HINDUSTAN BOOK AGENCY

Birkhäuser

HINDUSTAN BOOK AGENCY

Birkhäuser

Amiya Mukherjee

Differential Topology

HINDUSTAN BOOK AGENCY

Birkhäuser

Amiya Mukherjee
Indian Statistical Institute
Statistics and Mathematics Unit
Calcutta, India

A copublication with Hindustan Book Agency, New Delhi, licensed for sale in all countries, in print only outside of India. Sold and distributed in print within India by Hindustan Book Agency, P 19 Green Park Extn., New Delhi 110 016, India.

HBA ISBN 9789380250762

ISBN 978-3-319-19044-0 ISBN 978-3-319-19045-7 (eBook)
DOI 10.1007/978-3-319-19045-7

Library of Congress Control Number: 2015944813

Mathematics Subject Classification (2010): 58A05, 57R12, 57R22, 58A10, 58C35, 70G45, 57R17, 57R70 57Q65, 58A20, 57R65, 57R80

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.birkhauser-science.com)

To my teachers at Mathematical Institute,
24-29 St. Giles', Oxford, England,
who taught me mathematics.

PREFACE

The book is the outcome of lectures and seminars on various aspects of differentiable manifolds that I have given over the years at the Indian Statistical Institute, Calcutta, and at other universities in India. The purpose of these lectures was to provide the necessary background for, and to train the students in the use of some fundamental tools of differential topology. The book may be used as an orientation course for advanced-level research students or for independent study. The prerequisites are an elementary knowledge of linear algebra, multivariate calculus, general topology, analysis, and some algebraic topology. I have tried to make demands on the reader's knowledge of background materials as modest as possible by brushing them up adequately whenever they are needed.

The book provides a systematic and comprehensive account of the theory of differentiable manifolds. More explicitly, the book has two objectives. The first is to serve as an introduction to the subject. The aim of the second part is to acquaint the reader with some epochal discoveries in the field of manifolds, mainly the earlier works of Stephen Smale for which he was awarded the Fields Medal. The topics covered include (1) Thom transversality, (2) Morse theory, (3) Theory of handle presentation, (4) h-cobordism theory and generalised Poincaré's conjecture. However, while trying to achieve these two objectives, I have not made any sharp division between them. They intermix in a natural way.

Chapter 1 begins with the basics of differentiable manifolds. These are introduced mainly by means of examples. Chapter 2 addresses the problems of various approximations leading to Whitney's embedding theorem. The main ingredient is Sard's theorem. Chapter 3 introduces tangent spaces, vector fields and flows, and the exterior algebra. The main results of this chapter are the Darboux-Weinstein theorem on symplectic structures, and an analogous theorem for contact structures. Chapter 4 discusses Riemannian manifolds, geodesics, and the exponential map. The main result of this chapter is the Hopf-Rinow theorem. Chapter 5 is a brief introduction to the concepts of differentiable vector bundles. We include Atiyah's construction of vector bundles, the homotopy property, and orientations. Chapter 6 is devoted to elementary transversality theory with simple applications. The main result here is Hopf's degree theorem.

The subject matter of Chapter 7 is tubular neighbourhoods and collar neighbourhoods. Their existence and uniqueness are obtained by using the isotopy extension theorem. Also included in this chapter are discussions of straightening corners of manifolds, and constructions of manifolds by the gluing process. Chapter 8 outlines topics and results concerning spaces of differentiable maps, and spaces of jets. We prove here Thom's transversality theorem, and the multi-jet transversality theorem. The main applications are Whitney's immersion and embedding theorems. Chapter 9 is concerned with Morse theory, its applications in computation of homology groups, and triangulation of differentiable manifolds. Chapter 10 offers the theory of handle presentation and its simplification leading to the h-cobordism theorem, and the solution of the generalised Poincaré's conjecture. The writing of this chapter has been influenced mainly by the book of J. Milnor on h-cobordism, and also by the lecture notes of C.T.C. Wall on differentiable manifolds.

I do not presume to give all the basic information in the field, or a survey touching on every topic. Rather, I treat various selected topics in differential topology, which interest me, from a point of view that I hope the reader will find appealing. There are also matters which I have simply mentioned without elaboration. I welcome comments, suggestions, and corrections from readers of the book so that a later edition may benefit from experience with this one.

The book may be divided into three parts. The first part comprising Chapters 1 to 4 is foundational. It will be useful to general students of pure mathematics even if they are not going to take up research in mathematics. It can be used to design a course at the M.Sc. level in Indian universities. The second part consists of Chapters 5 to 7. It caters to researchers in the areas of Topology, Differential or Algebraic Geometry and Global Analysis. It touches on advanced topics which are of general interest to serious people in these areas. These topics help in an in-depth understanding of these particular areas. Finally the third part, which is the remainder of the book, is meant for those desirous of working in the field of Differential Topology itself.

The present book is a revised version of an earlier book with the title "Topics in Differential Topology". The significant difference between this edition and the earlier one is that here we have omitted the last chapter (Chapter 11) on Gromov theory of homotopy principle of certain partial differential relations. Because this topic is too technical for beginners, and also because it deserves a separate volume. I have added a new section in Chapter 5 on integration of differential forms on manifolds and Stokes' theorem, although we have not used this topic in the rest of the book.

I am thankful to the Department of Science and Technology, Government of India, for providing a grant under its USERS scheme for writing this book. I would like to take this opportunity to thank the Institute of Mathematical Sciences (IMSc), Chennai, and the Harish-Chandra Research Institute (HRI), Allahabad, for inviting me to spend some time in these research centres. I would like to thank Prof. R. Balasubramanian of IMSc, Prof. R.S. Kulkarni of

HRI, and all of the faculty and staff of these institutions for the effort they put into making my stay there so memorable. Parts of the text were used in the lectures given at these institutions. Thanks are also due to those who attended my lectures and for their interest in my book. The book was written in the very friendly and stimulating academic environment of the Stat-Math Unit of the Indian Statistical Institute, Calcutta. I would like to thank my colleagues and staff of the unit who were extremely helpful. Finally, I wish to express my deep gratitude and appreciation to the Director of Indian Statistical Institute for his unstinting help and encouragement while the book was being written.

Amiya Mukherjee

Stat-Math Unit
Indian Statistical Institute
Calcutta

Contents

CHAPTER 1

BASIC CONCEPTS OF MANIFOLDS

1.1. Two definitions of a manifold and examples

There are two ways one can look at a differentiable manifold. Firstly, it is a topological space with a structure which helps us to define differentiable functions on it, just as a topological structure on a set is designed to define continuous functions on that set. Secondly, it is a topological space which can be obtained by gluing together open subsets of some Euclidean space in a nice way; think, for example, of the surface of a ball or a torus covered with small paper disks pasted together on overlaps without making any crease or fold.

Both the approaches to a differentiable manifold are based on the standard differentiable structure on a Euclidean space $\mathbb{R}^n$. Let us therefore recall from calculus the notion of differentiable functions[1] on $\mathbb{R}^n$. Let $u_1, \dots, u_n$ denote the coordinate functions, where $u_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ is the function mapping a point $p = (p_1, \dots, p_n)$ onto its i-th coordinate p_i. A function f from an open subset U of $\mathbb{R}^n$ into $\mathbb{R}$ is **differentiable of class** C^r, or simply a C^r **function**, if it has continuous partial derivatives of all orders $\leq r$ with respect to $u_1, \dots, u_n$. A C^0 **function** is just a continuous function. A C^∞ **function** is C^r for every $r \geq 0$. A function of class C^r is also of class C^s for all $s < r \leq \infty$. The important point to note here is that if $f : U \longrightarrow \mathbb{R}$ is locally C^r (i.e. C^r on an open neighbourhood of each point of U), then f is C^r on U, where r may be any non-negative integer, or ∞.

Let $C^r(U)$ denote the set of all real valued C^r functions on U. This is an algebra over $\mathbb{R}$, in the sense that it is a vector space over $\mathbb{R}$ together with a bilinear product which is associative and distributive over the sum. The operations on $C^r(U)$ are given by pointwise addition, pointwise multiplication, and the multiplication of functions by scalars:

$$(f+g)(p) = f(p) + g(p), \quad (f \cdot g)(p) = f(p)g(p), \quad (\lambda f)(p) = \lambda f(p),$$

where $f, g \in C^r(U)$, $p \in U$, and $\lambda \in \mathbb{R}$. Then every $C^r(U)$ is a subalgebra $C^0(U)$.

A map $\phi : U \longrightarrow \mathbb{R}^m$, U open in $\mathbb{R}^n$, can be written as $\phi = (\phi_1, \dots, \phi_m)$, where $\phi_i = u_i \circ \phi : U \longrightarrow \mathbb{R}$ are the components of ϕ. The map ϕ is C^r if

[1] The terms 'function' and 'map' are synonymous. However, we use the term 'function' for a map whose range is $\mathbb{R}$.

A. Mukherjee, *Differential Topology*, DOI 10.1007/978-3-319-19045-7_1

each ϕ_i is C^r. It follows from the chain rule that ϕ is C^r if and only if $f \circ \phi \in C^r(\phi^{-1}(V))$ for every $f \in C^r(V)$, where V is open in $\mathbb{R}^m$ and $\phi^{-1}(V) \neq \emptyset$. A map ϕ between two open subsets of $\mathbb{R}^n$ is called a C^r **diffeomorphism** if it is a homeomorphism and both ϕ and ϕ^{-1} are C^r maps. We shall call a C^∞ diffeomorphism simply a diffeomorphism. For example, any linear isomorphism $\mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a diffeomorphism.

We shall use the words "smooth", "differentiable", and the symbol "C^∞" interchangeably. Our standard practice in this book will be to work with smooth functions, maps, and manifolds.

Consider a topological space $(X, \mathcal{U})$, where $\mathcal{U}$ is the topology or family of open sets in X.

Definition 1.1.1. A **sheaf of continuous functions** on X is a map $\mathcal{F}_X$ on $\mathcal{U}$ which assigns to each open subset $U \in \mathcal{U}$ a subalgebra $\mathcal{F}_X(U)$ of the algebra $C^0(U)$ such that

(1) $\mathcal{F}_X(\emptyset) = 0$,
(2) if $U, V \in \mathcal{U}$ with $V \subseteq U$, and $f \in \mathcal{F}_X(U)$, then $f|V \in \mathcal{F}_X(V)$,
(3) if $U \in \mathcal{U}$, and $f : U \longrightarrow \mathbb{R}$ is a function such that each $p \in U$ has an open neighbourhood $V \subset U$ on which f coincides with a $g \in \mathcal{F}_X(V)$, then $f \in \mathcal{F}_X(U)$.

The condition (2) implies that $\mathcal{F}_X$ is a contravariant functor from the category of open sets of X and inclusions of open sets into the category of algebras. Note that if $U = \cup\, U_i, U_i \in \mathcal{U}$, and $\{f_i\}$ is a family of functions such that $f_i \in \mathcal{F}_X(U_i)$ and $f_i|(U_i \cap U_j) = f_j|(U_i \cap U_j)$ for all i, j for which $U_i \cap U_j \neq \emptyset$, then there exists a unique continuous function $f : U \longrightarrow \mathbb{R}$ whose restriction to U_i is f_i, for all i. This $f \in \mathcal{F}_X(U)$ by the condition (3).

We call the pair $(X, \mathcal{F}_X)$ an a-**space** (or a space with a sheaf of algebras on it). An example is provided by $X = \mathbb{R}^n$ with the sheaf $\mathcal{F}_X$ as the sheaf of smooth functions, $C^\infty : U \longrightarrow C^\infty(U)$. The a-space $(\mathbb{R}^n, C^\infty)$ will play a key role in the next definition. Other examples of a-spaces appear in Remarks 1.1.6 below.

A **morphism** $\phi : (X, \mathcal{F}_X) \longrightarrow (Y, \mathcal{F}_Y)$ between a-spaces is defined by a continuous map $\phi : X \longrightarrow Y$ such that, for every open set U in Y, $f \in \mathcal{F}_Y(U)$ implies that $f \circ \phi \in \mathcal{F}_X(\phi^{-1}(U))$. Clearly, the identity map is a morphism, and so is the composition of two morphisms. A morphism ϕ is called an **isomorphism** if its inverse ϕ^{-1} exists as a morphism.

Example 1.1.2. Let $(X, \mathcal{F}_X)$ be an a-space, Y a topological space, and $\phi : X \longrightarrow Y$ a homeomorphism. Then the sets

$$\mathcal{F}_Y(U) = \{f : U \longrightarrow \mathbb{R} \mid f \circ \phi \in \mathcal{F}_X(\phi^{-1}(U))\},$$

U open in Y, define a sheaf of functions on Y, and ϕ induces an isomorphism between the a-spaces $(X, \mathcal{F}_X) \longrightarrow (Y, \mathcal{F}_Y)$.

The first definition of a manifold is as follows.

Definition 1.1.3. A **smooth manifold** of dimension n is a second countable Hausdorff a-space $(M, \mathcal{F}_M)$ which is locally isomorphic to the a-space $(\mathbb{R}^n, C^\infty)$. This means that each point of M has an open neighbourhood U and a homeomorphism ϕ of U onto an open subset of $\mathbb{R}^n$ such that $f \in \mathcal{F}_M(U)$ if and only if $f \circ \phi^{-1} \in C^\infty(\phi(U))$.

Then $\mathcal{F}_M$ is called a **smooth** or **differentiable sheaf** on M. The functions in $\mathcal{F}_M(U)$ are called **smooth functions** on U.

$\Diamond$ **Exercises 1.1.** Convince yourself that the following sets are smooth manifolds, and determine their dimensions.

(1) The empty set $\emptyset$ (the definition is vacuously satisfied).
(2) Any countable discrete set.
(3) Any disjoint union of a countable family of manifolds, each of dimension n.
(4) The Euclidean space $\mathbb{R}^n$ with smooth functions taken in the ordinary sense.
(5) Any open subset of a smooth manifold.
(6) The Cartesian product of two smooth manifolds.

We shall consider (4), (5), (6) again in Examples 1.1.7(1), 1.1.8(1), and 1.1.13 respectively.

We now turn to the second definition of a manifold.

Definition 1.1.4. A **smooth manifold** M of dimension n is a second countable Hausdorff space together with a smooth structure on it. A **smooth structure** consists of a family $\mathcal{D}^\infty$ of pairs (U_i, ϕ_i), i is in some index set Λ, where U_i is an open set of M and ϕ_i is a homeomorphism of U_i onto an open set of $\mathbb{R}^n$ such that

(1) the open sets U_i, $i \in \Lambda$, cover M,
(2) for every pair of indices $i, j \in \Lambda$ with $U_i \cap U_j \neq \emptyset$ the homeomorphisms
$$\phi_i \circ \phi_j^{-1} : \phi_j(U_i \cap U_j) \longrightarrow \phi_i(U_i \cap U_j),$$
$$\phi_j \circ \phi_i^{-1} : \phi_i(U_i \cap U_j) \longrightarrow \phi_j(U_i \cap U_j)$$
are smooth maps between open subsets of $\mathbb{R}^n$,
(3) the family $\mathcal{D}^\infty$ is maximal in the sense that it contains all possible pairs (U_i, ϕ_i) satisfying the properties (1) and (2).

The restriction $U_i \cap U_j \neq \emptyset$ in the condition (2) may be omitted provided we agree to assume that the empty map on the empty set is always smooth.

A pair $(U, \phi) \in \mathcal{D}^\infty$ with $p \in U$ is called a **coordinate chart** at p, U is called a **coordinate neighbourhood** of p, and $\phi = (x_1, \dots, x_n)$, where $x_i = u_i \circ \phi : U \longrightarrow \mathbb{R}$ is the i-th component of ϕ, is called a (local)**coordinate system** at p. Two charts (U_i, ϕ_i) and (U_j, ϕ_j) satisfying the conditions in

(2) are said to be C^∞ **related** or **compatible**, and each of $\phi_i \circ \phi_j^{-1}$ and $\phi_j \circ \phi_i^{-1}$ is called a **transition map** or a **change of coordinates**. A family of coordinate charts on M satisfying (1) and (2) is called a **smooth atlas**[1]. A smooth structure is a smooth atlas satisfying (3).

To understand the maximality condition (3) more clearly, consider the family of all smooth atlases on M. Say that two atlases $\mathcal{A}$ and $\mathcal{B}$ are compatible if each chart in $\mathcal{A}$ is compatible with each chart in $\mathcal{B}$, or equivalently, if $\mathcal{A} \cup \mathcal{B}$ is an atlas on M. It is easy to check that this is an equivalence relation. Then the union of all atlases in an equivalence class is a maximal atlas or a smooth structure on M. Thus any atlas can be enlarged to a unique smooth structure by adjoining all smoothly related charts to it.

The maximality condition allows us to restrict coordinate charts. If (U, ϕ) is a chart, U' is an open set in U, and $\phi' = \phi|U'$, then the charts (U, ϕ) and (U', ϕ') are compatible by the transition map $\phi \circ {\phi'}^{-1} = \mathrm{id}$, where id denotes the identity map.

Next observe that the charts (U, ϕ) and $(U, \alpha \circ \phi)$, where $\alpha : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a diffeomorphism, are always compatible. In particular, taking α to be the translation which sends $\phi(p)$ to 0, we can always suppose that every point $p \in M$ admits a coordinate chart (U, ϕ) such that $\phi(p) = 0$. We may also suppose that $\phi(U)$ is a convex set, or the whole of $\mathbb{R}^n$.

Proposition 1.1.5. *The two definitions of a manifold are equivalent.*

Proof. Suppose that $(M, \mathcal{F}_M)$ is a manifold according to Definition 1.1.3, then the local isomorphism $(M, \mathcal{F}_M) \longrightarrow (\mathbb{R}^n, C^\infty)$ defines an atlas on M. If $\phi_i : U_i \longrightarrow \mathbb{R}^n$ and $\phi_j : U_j \longrightarrow \mathbb{R}^n$ are two coordinate charts, then $f \in C^\infty(\phi_j(U_i \cap U_j))$ implies $f \circ \phi_j \in \mathcal{F}_M(U_i \cap U_j)$, and therefore $f \circ \phi_j \circ \phi_i^{-1} \in C^\infty(\phi_i(U_i \cap U_j))$. This shows that $\phi_j \circ \phi_i^{-1}$ is a smooth map. Similarly, $\phi_i \circ \phi_j^{-1}$ is smooth. Therefore M is a manifold according to Definition 1.1.4.

Conversely, if M is a manifold in the sense of Definition 1.1.4, then, for a coordinate chart (U_i, ϕ_i) of M where $\phi_i : U_i \longrightarrow V_i \subset \mathbb{R}^n$ is a homeomorphism, define

$$\mathcal{F}_M(U_i) = \{f \circ \phi_i : f \in C^\infty(V_i)\}.$$

This gives a sheaf of functions $\mathcal{F}_M$ on M (note that we need only to define $\mathcal{F}_M$ for coordinate neighbourhoods, in view of (3) of Definition 1.1.1), and the collection $\{\phi_i\}$ gives a local isomorphism from $(M, \mathcal{F}_M)$ to $(\mathbb{R}^n, C^\infty)$. Thus M is a manifold in the sense of Definition 1.1.3. $\square$

Remarks 1.1.6. (1) **Analytic manifold**. The notion of an analytic manifold is obtained from Definition 1.1.3 by working with the *a*-space $(\mathbb{R}^n, C^\omega)$, instead of $(\mathbb{R}^n, C^\infty)$, where C^ω is the sheaf which associates with each open set

[1]The terminology is probably due to Carl Friedrich Gauss (1777-1855) who formulated in mathematical terms the method of drawing maps of earth's surface.

U of $\mathbb{R}^n$ the algebra $C^\omega(U)$ of real valued analytic functions on U. Recall that a function $f : U \longrightarrow \mathbb{R}$ is **analytic**, or a C^ω **function**, if it admits an infinite Taylor's series expansion in a neighbourhood of each point of U which converges absolutely to f in that neighbourhood. Equivalently, one has to replace "smooth" in Definition 1.1.4 by "analytic" (note that a map f between open sets of $\mathbb{R}^n$ is analytic if and only if each component of f is so). An analytic structure or a maximal C^ω atlas is denoted by $\mathcal{D}^\omega$.

In the sequel an analytic manifold will hardly ever be mentioned, even if a smooth manifold is also analytic.

(2) C^r **manifold**. In the same way, we could have defined a C^r manifold for an $r < \infty$. The corresponding C^r differentiable structure $\mathcal{D}^r$ is the maximal atlas of C^r related charts (the definition is obvious, just replace "smooth", by "C^r"). If $0 \le r < s$, then a $\mathcal{D}^s$ is also a C^r atlas, so it extends to a $\mathcal{D}^r$ which is the unique minimal structure containing $\mathcal{D}^s$ (in the sense that any smooth structure containing $\mathcal{D}^r$ also contains $\mathcal{D}^s$). Conversely, it can be proved that for $r \ge 1$ every $\mathcal{D}^r$ contains a $\mathcal{D}^s$ for each s, $r < s \le \infty$, which is not unique, of course, but is unique up to an equivalence relation (see Definition 1.3.4 below). A proof may be given using an embedding theorem of Whitney, which says, roughly speaking, that any C^r manifold may be considered as lying in a Euclidean space with its usual smooth structure. In the next chapter, we shall prove this result for the special case when $r = \infty$. A more general embedding theorem of Whitney implies that any smooth structure contains an analytic structure. In view of these facts, it appears that the degree of differentiability of a manifold is not a very significant factor, and for most purposes it is sufficient to consider smooth manifolds. Besides, the category of smooth manifolds has several advantages over the C^r category with finite r. For example, the algebra $C^\infty(U)$ is closed under the operation of differentiation, while the algebra $C^r(U)$ is not when $0 < r < \infty$.

The largest structure $\mathcal{D}^0$ comprises of all local homeomorphisms $M \longrightarrow \mathbb{R}^n$. A second countable Hausdorff space with a $\mathcal{D}^0$ structure is called a **topological manifold** or a **locally Euclidean space**. It follows from the definition that a C^0 structure on a space is unique. But a space may have several different C^r structures, if $r > 0$ (see the results which appear at the end of §1.3 in p. 16). Next note that a C^∞ differentiable structure $\mathcal{D}^\infty$ can be written as $\mathcal{D}^\infty = \cap_{r \ge 0} \mathcal{D}^r$, where each $\mathcal{D}^r$ is uniquely determined by $\mathcal{D}^\infty$.

(3) **Complex manifold**. Another variant of Definition 1.1.4 is a complex manifold M, where $\mathbb{R}^n$ is replaced by the complex n-space $\mathbb{C}^n$, and the transition maps $\phi_i \circ \phi_j^{-1}$ and $\phi_j \circ \phi_i^{-1}$ are required to be holomorphic or complex analytic (instead of smooth). Recall that if we write the coordinates in $\mathbb{C}^n$ as $(z_1, \ldots, z_n)$, where $z_i = x_i + \sqrt{-1}\, y_i$, $x_i, y_i \in \mathbb{R}$, then we may identify $\mathbb{C}^n \simeq \mathbb{R}^{2n}$ by the $\mathbb{R}$-linear isomorphism

$$(z_1, \ldots, z_n) \mapsto (x_1, y_1, \ldots, x_n, y_n).$$

A map $f : \mathbb{C}^n \longrightarrow \mathbb{C}^m$ with components $f_1, \ldots, f_m$, where

$$f_i = u_i + \sqrt{-1}v_i,\ u_i = u_i(x_1, y_1, \ldots, x_n, y_n),\ v_i = v_i(x_1, y_1, \ldots, x_n, y_n),$$

is **holomorphic** if the following **Cauchy-Riemann equations** hold

$$\frac{\partial u_j}{\partial x_i} = \frac{\partial v_j}{\partial y_i}, \quad \frac{\partial u_j}{\partial y_i} = -\frac{\partial v_j}{\partial x_i},$$

or equivalently, if each f_j admits an absolutely convergent power series expansion about each point of $\mathbb{C}^n$. It follows that every complex manifold of complex dimension n is a smooth real manifold of dimension $2n$.

We now give some simple examples of smooth manifolds, using Definition 1.1.4.

Examples 1.1.7. (1) **Euclidean space** $\mathbb{R}^n$. A smooth structure is given by an atlas consisting of only one chart $(\mathbb{R}^n, id)$. The maximal atlas generated by this atlas consists of all charts (U, ϕ), where U is an open set of $\mathbb{R}^n$ and ϕ is a diffeomorphism on it. This smooth structure on $\mathbb{R}^n$ is called the **standard smooth structure**.

A similar consideration shows that the complex n-space $\mathbb{C}^n$ is a smooth complex manifold of complex dimension n.

(2) **Vector space**. Any real vector space V of dimension n has a canonical smooth structure generated by the atlas consisting of all linear isomorphisms of V onto $\mathbb{R}^n$. Note that in this atlas any change of coordinates is a linear map and so indefinitely differentiable.

(3) **Manifold of matrices**. Let $\mathbb{K}$ denote the field $\mathbb{R}$ or $\mathbb{C}$, and $M(m, n, \mathbb{K})$ be the space of all $m \times n$ matrices with entries in $\mathbb{K}$. Taking the entries of matrices in lexicographic (or dictionary) order we may identify $M(m, n, \mathbb{K})$ with $\mathbb{K}^{mn}$ in the following way:

$$(a_{ij}) \leftrightarrow (a_{11}, a_{12}, \ldots, a_{1n}, a_{21}, a_{22}, \ldots, a_{2n}, \ldots, a_{m1}, a_{m2}, \ldots, a_{mn}).$$

This induces a smooth structure on $M(m, n, \mathbb{K})$, by Theorem 1.3.1 below. Thus $M(m, n, \mathbb{R})$ is a smooth manifold of dimension mn, and $M(m, n, \mathbb{C})$ is a smooth complex manifold of real dimension $2mn$.

Examples 1.1.8. (1) **Open subset**. An open set V of a smooth manifold M is itself a smooth manifold. The smooth structure is obtained by restrictions of coordinate charts. If $\mathcal{A}$ is a smooth atlas for M, then

$$\mathcal{A}_V = \{(U, \phi) \in \mathcal{A} \mid U \subset V\}$$

is a smooth atlas for V. Note that if (U, ϕ) and (V, ψ) are two compatible charts on M, then for any pair of open sets $U' \subset U$ and $V' \subset V$ the charts $(U', \phi|U')$ and $(V', \psi|V')$ are also compatible.

(2) **General linear group** $GL(n, \mathbb{K})$. If $n = m$, let us write the manifold of matrices $M(n, n, \mathbb{K})$ as $M(n, \mathbb{K})$. Then, the set $GL(n, \mathbb{K})$ of all non-singular matrices of order n forms an open subset of $M(n, \mathbb{K})$, since the determinant

function $\det : M(n, \mathbb{K}) \longrightarrow \mathbb{K}$ is continuous, being a polynomial map. Therefore $GL(n, \mathbb{K})$ is a smooth manifold.

$\Diamond$ **Exercise 1.2.** Suppose M is a topological manifold, and U, V are open sets of M such that $M = U \cup V$. Show that, if U and V have smooth structures which restricts to the same smooth structure on $U \cap V$, then there is a unique smooth structure on M which restrict to the smooth structure of U, and of V.

Example 1.1.9. Sphere S^n. This is the set of all unit vectors in $\mathbb{R}^{n+1}$

$$S^n = \{(x_1, \dots, x_{n+1}) \in R^{n+1} \mid x_1^2 + \cdots + x_{n+1}^2 = 1\}.$$

A smooth atlas is provided by two open sets U_+ and U_- obtained by deleting from S^n the north pole $P = (0, \dots, 0, 1)$ and the south pole $Q = (0, \dots, 0, -1)$ respectively, and the stereographic projections

$$\phi_+ : U_+ \longrightarrow \mathbb{R}^n, \text{ and } \phi_- : U_- \longrightarrow \mathbb{R}^n$$

from P and Q onto the equatorial plane $x_{n+1} = 0$, given by

$$\phi_\pm(x_1, \dots, x_{n+1}) = \left(\frac{x_1}{1 \mp x_{n+1}}, \dots, \frac{x_n}{1 \mp x_{n+1}} \right).$$

Clearly

$$\phi_\pm^{-1}(u_1, \dots, u_n) = \left(\frac{2u_1}{1 + \|u\|^2}, \dots, \frac{2u_n}{1 + \|u\|^2}, \mp \frac{1 - \|u\|^2}{1 + \|u\|^2} \right),$$

and therefore the change of coordinates

$$\phi_- \circ \phi_+^{-1} = \phi_+ \circ \phi_-^{-1} : \mathbb{R}^n - \{0\} \longrightarrow \mathbb{R}^n - \{0\}$$

is given by the smooth map $x \mapsto x / \parallel x \parallel^2$.

$\Diamond$ **Exercise 1.3.** Show that another smooth atlas of S^n is given by the $2n + 2$ coordinate charts (V_i^+, ψ_i^+), (V_i^-, ψ_i^-), $i = 1, \dots, n + 1$, where V_i^+ and V_i^- are the hemispheres

$$V_i^+ = \{x \in S^n : x_i > 0\}, \qquad V_i^- = \{x \in S^n : x_i < 0\}$$

and $\psi_i^+ : V_i^+ \longrightarrow \mathbb{R}^n$ and $\psi_i^- : V_i^- \longrightarrow \mathbb{R}^n$ are the projections onto the hyperplane $x_i = 0$

$$(x_1, \dots, x_{n+1}) \mapsto (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1}).$$

Show that these charts are C^∞ related to the charts (U_+, ϕ_+) and (U_-, ϕ_-) of Example 1.1.9.

Remark 1.1.10. It can be shown that S^2 (the Riemann sphere) has the structure of a complex analytic manifold. However, It is not known if such a structure exists on S^6.

Example 1.1.11. Real projective space $\mathbb{R}P^n$. This space is the quotient of S^n modulo the equivalence relation: $x \sim y$ if and only if $y = x$ or $y = -x, x, y \in S^n$. The equivalence classes are unordered pairs of antipodal points $\{x, -x\}$, which we denote by $[x]$. The topology on $\mathbb{R}P^n$ is the quotient topology induced by the canonical projection $\pi : S^n \longrightarrow \mathbb{R}P^n$ sending x to $[x]$ (the quotient topology on $\mathbb{R}P^n$ is obtained by specifying $U \subset \mathbb{R}P^n$ open if and only if $\pi^{-1}(U)$ is open in S^n). Relative to this topology, $\mathbb{R}P^n$ is a Hausdorff space and π is a continuous open map. If (U, ϕ) is a coordinate chart in S^n, where U does not contain any pair of antipodal points, then the restriction $\pi|U$ gives a homeomorphism $\pi_U : U \longrightarrow \pi(U)$, and $(\pi(U), \phi \circ \pi_U^{-1})$ is a coordinate chart in $\mathbb{R}P^n$. Any two of such charts $(\pi(U), \phi \circ \pi_U^{-1})$ and $(\pi(V), \psi \circ \pi_V^{-1})$ are C^∞ related, because restricting the maps ϕ, ψ, and π to the intersection $U \cap V$ we have

$$\psi \circ \pi^{-1} \circ (\phi \circ \pi^{-1})^{-1} = \psi \circ \phi^{-1}.$$

The resulting atlas can be extended to a maximal atlas.

$\Diamond$ **Exercises 1.4.** The projective space $\mathbb{R}P^n$ may also be considered as the quotient of $\mathbb{R}^{n+1} - \{0\}$ modulo the equivalence relation:

$$(x_0, \ldots, x_n) \sim (\lambda x_0, \ldots, \lambda x_n), \quad \lambda \in \mathbb{R} - \{0\}.$$

The equivalence classes are 1-dimensional subspaces or lines through the origin in $\mathbb{R}^{n+1}$. Let $\pi : \mathbb{R}^{n+1} - \{0\} \longrightarrow \mathbb{R}P^n$ be the canonical projection, and let $\mathbb{R}P^n$ be given the quotient topology so that π becomes a continuous open map.

(1) Show that this quotient space is homeomorphic to the quotient space of Example 1.1.11 obtained by identifying pairs of diametrically opposite points of S^n.

(2) For each i, $0 \le i \le n$, consider open subset U_i of $\mathbb{R}P^n$ given by

$$U_i = \{[x_0, \ldots, x_n] \mid x_i \neq 0\},$$

where $[x_0, \ldots, x_n] = \pi((x_0, \ldots, x_n))$. This is the set of all lines through the origin which intersect the hyperplane $x_i = 1$, and this is open in $\mathbb{R}P^n$ because

$$\pi^{-1}(U_i) = \mathbb{R}^{n+1} - \{\text{hyperplane } x_i = 0\}$$

is open in $\mathbb{R}^{n+1} - \{0\}$. Define $\phi_i : U_i \longrightarrow \mathbb{R}^n$ by

$$\phi_i([x_0, \ldots, x_n]) = \frac{1}{x_i}(x_0, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n).$$

Show that the family $\{(U_i, \phi_i)\}$ is a smooth atlas on $\mathbb{R}P^n$.

Hint. The inverse of ϕ_i is

$$\phi_i^{-1}(x_1, \ldots, x_n) = [x_1, \ldots, x_i, 1, x_{i+1}, \ldots, x_n].$$

So the change of coordinates between charts (U_i, ϕ_i) and (U_j, ϕ_j) is

$$\phi_j \circ \phi_i^{-1}(x_1, \ldots, x_n) = \frac{1}{x_{j+1}}(x_1, \ldots, x_j, x_{j+2}, \ldots, x_i, 1, x_{i+1}, \ldots, x_n),$$

assuming for convenience $j < i$.

We shall see in Exercise 1.16 (p.15) that the two smooth structures on $\mathbb{R}P^n$ are essentially the same, in a certain sense.

$\Diamond$ **Exercise 1.5. Complex projective space** $\mathbb{C}P^n$. This is the set of all 1-dimensional complex linear subspaces of $\mathbb{C}^{n+1}$ with the quotient topology obtained from the natural projection $\pi : \mathbb{C}^{n+1} - \{0\} \longrightarrow \mathbb{C}P^n$. Show that this can be given a smooth structure analogous to above construction for $\mathbb{R}P^n$

Example 1.1.12. Real Grassmann manifold. This is the set of all k-dimensional vector subspaces (or k-planes) of $\mathbb{R}^n$, and is denoted by $G_k(\mathbb{R}^n)$. Then $G_1(\mathbb{R}^n) = \mathbb{R}P^{n-1}$. If $M_k(n, k, \mathbb{R})$ denotes the space of all real $n \times k$ matrices of rank k, then a k-plane P may be represented by a matrix $A \in M_k(n, k, \mathbb{R})$ so that the column vectors of A generate P. Another matrix $B \in M_k(n, k, \mathbb{R})$ represents the same k-plane P if and only if there is a non-singular matrix $T \in GL(k, \mathbb{R})$ such that $B = AT$. Define an equivalence relation in $M_k(n, k, \mathbb{R})$ by $A \sim B$ if $B = AT$ for some $T \in GL(k, \mathbb{R})$. Then, if $M_k(n, k, \mathbb{R})/GL(k, \mathbb{R})$ denotes the quotient space of $M_k(n, k, \mathbb{R})$ modulo this equivalence relation, we have a bijection

$$f : M_k(n, k, \mathbb{R})/GL(k, \mathbb{R}) \longrightarrow G_k(\mathbb{R}^n).$$

We equip $G_k(\mathbb{R}^n)$ with the unique topology which makes f a homeomorphism. Thus we identify $G_k(\mathbb{R}^n)$ with $M_k(n, k, \mathbb{R})/GL(k, \mathbb{R})$ topologically. Let

$$\pi : M_k(n, k, \mathbb{R}) \longrightarrow G_k(\mathbb{R}^n)$$

denote the canonical projection which sends each matrix A to its equivalence class $[A]$. Then π is an open map.

Let α denote an increasing sequence of k integers $\alpha_1, \dots, \alpha_k$ such that

$$1 \leq \alpha_1 < \alpha_2 < \cdots < \alpha_k \leq n.$$

If A is an $n \times k$ matrix, let A_α denote the $k \times k$ submatrix of A composed of the rows $\alpha_1, \dots, \alpha_k$, and $A_{\alpha'}$ denote the $(n-k) \times k$ submatrix composed of the remaining rows. If $A \in M_k(n, k, \mathbb{R})$ represents a k-plane P such that A_α is non-singular, then the matrix $C = AA_\alpha^{-1}$ also represents P, and the submatrix C_α is the identity matrix I_k of order k. We will call the matrix C the α-representative of P. For a fixed sequence α, consider the subset U_α of $G_k(\mathbb{R}^n)$ defined by

$$U_\alpha = \{[A] \in G_k(\mathbb{R}^n) \mid A_\alpha \in GL(k, \mathbb{R})\}.$$

Then U_α is open in $G_k(\mathbb{R}^n)$, because $U_\alpha = \pi \circ g_\alpha^{-1}(\mathbb{R} - \{0\})$, where

$$g_\alpha : M_k(n, k; \mathbb{R}) \to \mathbb{R}$$

is the continuous function $A \mapsto \det A_\alpha$. Clearly the open sets U_α cover $G_k(\mathbb{R}^n)$. Define a map $\phi_\alpha : U_\alpha \longrightarrow \mathbb{R}^{k(n-k)}$ by $\phi_a([A]) = C_{\alpha'}$, where $C = AA_\alpha^{-1}$ and $C_{\alpha'}$ is considered as a point of $\mathbb{R}^{k(n-k)}$. Note that ϕ_α is well-defined, because $(AT)_\alpha = A_\alpha T$, $T \in GL(k, \mathbb{R})$, and that it is a homeomorphism. We shall show that the family $\{(U_\alpha, \phi_\alpha)\}$ is a smooth atlas for $G_k(\mathbb{R}^n)$.

Let $P \in U_\alpha \cap U_\beta$ and C be the α-representative of P such that $C_\alpha = I_k$. Since P belongs to U_β also, $C_\beta \in GL(k, \mathbb{R})$. Then the change of coordinates is given by $C_{\alpha'} \mapsto (CC_\beta^{-1})_{\beta'}$. This is a smooth map, since each entry of $(CC_\beta^{-1})_{\beta'}$ is a rational function of the entries of $C_{\alpha'}$.

It remains to show that $G_k(\mathbb{R}^n)$ is a second countable Hausdorff space. The second countability follows, because the open map $\pi : M_k(n, k, \mathbb{R}) \longrightarrow G_k(\mathbb{R}^n)$ sends a countable basis to a countable basis. Next, we have to show that two k-planes P_1 and P_2 admit disjoint neighbourhoods. This is certainly true if P_1 and P_2 belong to the same coordinate neighbourhood U_α, because $\phi_\alpha(U_\alpha)$ is Hausdorff. So suppose that $P_1 \in U_\alpha$ and $P_2 \in U_\beta$ but neither of them belongs to $U_\alpha \cap U_\beta$.

Let $h_{\alpha\beta} : G_k(U_\alpha) \longrightarrow \mathbb{R}$ be the map given by

$$h_{\alpha\beta}([A]) = \det(AA_\alpha^{-1})_\beta.$$

Then $h_{\alpha\beta}$ is continuous, and we have $U_\alpha \cap U_\beta = h_{\alpha\beta}^{-1}(\mathbb{R} - \{0\})$. Since $(AT)_\alpha = A_\alpha T$ for a non-singular T, it follows that $h_{\alpha\beta} = h_{\beta\alpha}^{-1}$. Now the assumptions about P_1 and P_2 imply that $h_{\alpha\beta}(P_1) = 0$ and $h_{\beta\alpha}(P_2) = 0$. Therefore if I is the open interval $(-1, 1)$, the sets $V_1 = h_{\alpha\beta}^{-1}(I)$ and $V_2 = h_{\beta\alpha}^{-1}(I)$ are open neighbourhoods of P_1 and P_2 respectively. They are disjoint, otherwise any k-plane P in $V_1 \cap V_2$ would belong to $U_\alpha \cap U_\beta$ and satisfy the inequalities $|h_{\alpha\beta}(P)| < 1$ and $|h_{\beta\alpha}(P)| < 1$ contradicting the fact that $h_{\alpha\beta}(P) = h_{\beta\alpha}(P)^{-1}$. Therefore $G_k(\mathbb{R}^n)$ is Hausdorff.

$\Diamond$ **Exercise 1.6. Complex Grassmann manifold**. Show that the complex Grassmann manifold $G_k(\mathbb{C}^n)$, which is the set of all k-planes in $\mathbb{C}^n$, is a smooth complex manifold of real dimension $2k(n-k)$.

Example 1.1.13. Product of manifolds. If M and N are smooth manifolds with smooth structures $\{(U_i, \phi_i)\}$ and $\{(V_r, \psi_r)\}$ respectively, then the Cartesian product $M \times N$ is a smooth manifold with atlas $\{(U_i \times V_r, \phi_i \times \psi_r)\}$. Any two such charts are smoothly compatible, because

$$(\phi_j \times \psi_s) \circ (\phi_i \times \psi_r)^{-1} = (\phi_j \times \psi_s) \circ (\phi_i^{-1} \times \psi_r^{-1}) = (\phi_j \circ \phi_i^{-1}) \times (\psi_s \circ \psi_r^{-1}),$$

which is a smooth map.

In particular, the n-torus $T^n = S^1 \times \cdots \times S^1$ (S^1 appearing n times) is a smooth manifold.

1.2. Smooth maps between manifolds

Each of Definitions 1.1.3 and 1.1.4 provides a notion of a smooth map between manifolds. The first definition of a smooth map in terms of sheaf of smooth functions is as follows.

Definition 1.2.1. If $(M, \mathcal{F}_M)$ and $(N, \mathcal{F}_N)$ are smooth manifolds, then a morphism $f : (M, \mathcal{F}_M) \longrightarrow (N, \mathcal{F}_N)$ is a **smooth map**. Explicitly, a continuous map $f : M \longrightarrow N$ is a smooth map if $g \in \mathcal{F}_N(U)$ implies $g \circ f \in \mathcal{F}_M(f^{-1}(U))$, for every open set U of N.

For the second definition of a smooth map in terms of coordinate charts, suppose $f : M \longrightarrow N$ is a map, $p \in M$, and (U, ϕ) and (V, ψ) are coordinate charts at p and $f(p)$ respectively. Then the map

$$\psi \circ f \circ \phi^{-1} : \phi(U \cap f^{-1}(V)) \longrightarrow \psi(V)$$

is called a **local representation** of f at p for the pair of coordinate systems (ϕ, ψ).

Definition 1.2.2. A map $f : M \longrightarrow N$ is **smooth**, if its local representation at every point $p \in M$ is a smooth map for some, and hence for all pairs of coordinate systems ϕ and ψ at p and at $f(p)$.

Observe that this definition is independent of the choice of coordinate systems. If f is smooth at p for a pair (ϕ, ψ), then it is smooth at p for every other compatible pair (ϕ_1, ψ_1). Because, the transition maps $\phi \circ \phi_1^{-1}$ and $\psi \circ \psi_1^{-1}$ are smooth, and so the composition

$$\psi_1 \circ f \circ \phi_1^{-1} = (\psi_1 \circ \psi^{-1}) \circ (\psi \circ f \circ \phi^{-1}) \circ (\phi \circ \phi_1^{-1})$$

is smooth.

Lemma 1.2.3. *Each of the definitions of smooth maps implies that the composition of smooth maps between manifolds is smooth.*

Proof. If $f : (M, \mathcal{F}_M) \longrightarrow (N, \mathcal{F}_N)$ and $g : (N, \mathcal{F}_N) \longrightarrow (R, \mathcal{F}_R)$ are smooth maps according to the first definition, then so is their composition $g \circ f$. For, if $\lambda \in \mathcal{F}_R(W)$, then $\lambda \circ g \in \mathcal{F}_N(g^{-1}(W))$, and so

$$\lambda \circ g \circ f \in \mathcal{F}_M(f^{-1}g^{-1}(W)) = \mathcal{F}_M((g \circ f)^{-1}(W)).$$

The assertion is also true for the second definition. Because for suitable coordinate charts (U, ϕ), (V, ψ), and (W, θ) in M, N, and R respectively, the map

$$\theta \circ (g \circ f) \circ \phi^{-1} = (\theta \circ g \circ \psi^{-1}) \circ (\psi \circ f \circ \phi^{-1})$$

is smooth, being the composition of smooth maps between open subsets of Euclidean spaces. □

Proposition 1.2.4. *The two definitions of a smooth map are equivalent.*

Proof. By Definition 1.1.3, for a coordinate chart (U, ϕ) in M, a function $\lambda : U \longrightarrow \mathbb{R}$ is smooth if and only if $\lambda \circ \phi^{-1} : \phi(U) \longrightarrow \mathbb{R}$ is smooth. It follows from this that both ϕ and ϕ^{-1} are smooth according to the first definition. Therefore if f is also smooth like these maps, then the composition $\psi \circ f \circ \phi^{-1}$ is smooth. Thus the second definition follows from the first.

Conversely, assume the second definition of smoothness, and suppose that $f : M \longrightarrow N$ is a map such that $\lambda \circ f$ is smooth whenever $\lambda : W \longrightarrow \mathbb{R}$ is a smooth function on an open neighbourhood W of $f(p)$ in N, $p \in M$. Take coordinate charts (U, ϕ) at p and (V, ψ) at $f(p)$. Then, each component $y_i = u_i \circ \psi$ of ψ is a smooth function on V, since its local representation for the pair of charts (V, ψ) and $(\mathbb{R}, id)$ is the smooth function u_i. Therefore each $y_i \circ f$ is a smooth function on $U \cap f^{-1}(V)$, by hypothesis. This means that $y_i \circ f$ has smooth local representation $u_i \circ \psi \circ f \circ \phi^{-1}$. But these are components of the local representation of f. So f is smooth as in the second definition. Thus the first definition follows from the second. $\square$

$\Diamond$ **Exercises 1.7.** Denote by $C^\infty(M)$ the set of all real-valued smooth functions on M. Then, by Definition 1.2.2, $f \in C^\infty(M)$ if for each $p \in M$ there is a coordinate chart (U, ϕ) with $p \in U$ such that $f \circ \phi^{-1} \in C^\infty(\phi(U))$.

(a) Show that the set $C^\infty(M)$ has the following properties.

(1) If $f : M \longrightarrow \mathbb{R}$ is a function such that each $p \in M$ has an open neighbourhood on which f coincides with a function in $C^\infty(M)$, then $f \in C^\infty(M)$.
(2) If $f_1, \ldots, f_k \in C^\infty(M)$, and $u \in C^\infty(\mathbb{R}^k)$, then $u(f_1, \ldots, f_k) \in C^\infty(M)$.

(b) Show that sums, products, and constant multiplies of smooth functions are also smooth, and therefore $C^\infty(M)$ is an algebra over $\mathbb{R}$.

$\Diamond$ **Exercises 1.8.** (a) Show that, for product of manifolds, the projections

$$M \times N \longrightarrow M \quad \text{and} \quad M \times N \longrightarrow N$$

are smooth maps.

(b) Show that if $f : M \longrightarrow M'$ and $g : N \longrightarrow N'$ are smooth maps, then the product map $f \times g : M \times N \longrightarrow M' \times N'$ is also smooth.

Definition 1.2.5. A map $f : M \longrightarrow N$ is called a **diffeomorphism** if f is a bijection and both f and f^{-1} are smooth maps.

For example, if (U, ϕ) is a coordinate chart on M, then $\phi : U \longrightarrow \mathbb{R}^n$ is a diffeomorphism onto its image, since its local representation in the pair of charts (U, ϕ) and $(\phi(U),\ \text{id})$ is the identity map.

$\Diamond$ **Exercises 1.9.** Show that (1) the real projective space $\mathbb{R}P^1$ is diffeomorphic to the circle S^1, and (2) the complex projective space $\mathbb{C}P^1 = G_1(\mathbb{C}^2)$ is diffeomorphic to the sphere S^2.

$\Diamond$ **Exercise 1.10.** Show that there is a natural diffeomorphism between the Grassmann manifolds $G_k(\mathbb{R}^n)$ and $G_{n-k}(\mathbb{R}^n)$.

Smooth maps are defined on open subsets of a manifold. The definition can be extended over arbitrary subsets of a manifold in the following way.

Definition 1.2.6. A map f from a subset S of a manifold M to a manifold N is **smooth** if it can be locally extended to a smooth map. Explicitly, f is smooth, if each point $p \in S$ admits an open neighbourhood U in M and a smooth map $F : U \longrightarrow N$ such that $F|S \cap U = f$.

The local extendability condition of f is equivalent to saying that all the partial derivatives of f exist and are continuous, by Whitney's extension theorem (Whitney [**58**]).

$\Diamond$ **Exercise 1.11.** Show that if $n < m$, and $\mathbb{R}^n$ is considered as the subset

$$\{(x_1, \ldots, x_m) \mid x_{n+1} = \cdots = x_m = 0\}$$

of the first n coordinates of $\mathbb{R}^m$, then the usual smooth maps on $\mathbb{R}^n$ and those obtained by using the above definition are the same.

A map f from a subset S of a manifold M to a subset K of a manifold N is a **diffeomorphism** if it is a bijection and both f and f^{-1} are smooth maps.

$\Diamond$ **Exercise 1.12.** Show that if S, R, K are subsets of manifolds M, N, P respectively, and $f : S \longrightarrow R$, $g : R \longrightarrow K$ are smooth maps, the the composition $g \circ f : S \longrightarrow K$ is smooth (in particular, if $M = N$ and $S \subset R$, then $g|S$ is smooth). Moreover, if f, g are diffeomorphisms, then so is $g \circ f$.

It follows that a subset S in a Euclidean space $\mathbb{R}^m$ is a smooth manifold of dimension n if it is locally diffeomorphic to $\mathbb{R}^n$, that is, if each point of S has an open neighbourhood in S (in the relative topology) which is diffeomorphic to an open subset of $\mathbb{R}^n$. Here is an example.

Example 1.2.7. Space of matrices of rank k. Let $M_k(m, n, \mathbb{R})$ be the space of all real $m \times n$ matrices of rank k, where $0 < k \leq \min(m, n)$, with the induced topology of $M(m, n, \mathbb{R})$. Then $M_k(m, n, \mathbb{R})$ is a smooth manifold of dimension $k(m + n - k)$. To see this, take an element $E_0 \in M_k(m, n, \mathbb{R})$. We may assume by permuting the rows and columns, if necessary, that E_0 is of the form

$$E_0 = \begin{pmatrix} A_0 & B_0 \\ C_0 & D_0 \end{pmatrix},$$

where A_0 is a non-singular $k \times k$ matrix. Then, we can find an $\epsilon > 0$ such that if A is a $k \times k$ matrix and if each entry of $A - A_0$ has absolute value less than ϵ, then A is non-singular. Let

$$U = \left\{ E = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \mid \text{absolute value of each entry of } A - A_0 < \epsilon \right\}.$$

A matrix E as above has the same rank as the matrix

$$\begin{pmatrix} I_k & 0 \\ X & I_{m-k} \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & B \\ XA + C & XB + D \end{pmatrix},$$

where I_k is the $k \times k$ identity matrix and X is any $(m - k) \times k$ matrix. Taking $X = -CA^{-1}$, we find that the rank of E is exactly k if and only if $D = CA^{-1}B$. Let V be the open set in the Euclidean space of dimension $mn - (m - k)(n - k) = k(m + n - k)$ consisting of matrices of the form

$$\begin{pmatrix} A & B \\ C & 0 \end{pmatrix},$$

where each entry of $A - A_0$ has absolute value less than ϵ. Then the map

$$\begin{pmatrix} A & B \\ C & CA^{-1}B \end{pmatrix} \longrightarrow \begin{pmatrix} A & B \\ C & 0 \end{pmatrix}$$

is a diffeomorphism of the neighbourhood $U \cap M_k(m, n, \mathbb{R})$ of E_0 onto V. Since E_0 is an arbitrary element of $M_k(m, n, \mathbb{R})$, $M_k(m, n, \mathbb{R})$ is a smooth manifold of dimension $k(m + n - k)$.

$\Diamond$ **Exercise 1.13.** Show that if M and N are smooth manifolds, and there is a diffeomorphism of M onto a subset S of N, then S is a smooth manifold.

$\Diamond$ **Exercise 1.14.** The graph of a map $f : M \longrightarrow N$ is the set

$$\Gamma(f) = \{(x, f(x)) \in M \times N \mid x \in M\}.$$

Show that if f is smooth, then the map $F : M \longrightarrow \Gamma(f)$ defined by $F(x) = (x, f(x))$ is a diffeomorphism. Conclude that $\Gamma(f)$ is a smooth manifold. In particular, the diagonal set $\triangle$ in $M \times M$, which is $\Gamma(\mathrm{Id}_M)$, is a smooth manifold.

1.3. Induced smooth structures

If $f : X \longrightarrow S$ is a bijection from a topological space X onto a set S, then f induces a unique topology on S so that f is a homeomorphism. The following theorem gives analogous result for smooth manifolds.

Theorem 1.3.1. *Let M be a smooth manifold with a smooth structure $\mathcal{A}$, X a topological space, and $f : M \longrightarrow X$ a homeomorphism. Then X can be endowed with a unique smooth structure $\mathcal{A}_f$ so that f becomes a diffeomorphism.*

This smooth structure $\mathcal{A}_f$ on X is said to be obtained by **transporting** $\mathcal{A}$ by f. Also $\mathcal{A}_f$ is called the structure induced from $\mathcal{A}$ by f.

Proof. If $\mathcal{A} = \{(U_i, \phi_i)\}$, then take $\mathcal{A}_f = \{(f(U_i), \phi_i \circ f^{-1})\}$. Then the transition maps for $\mathcal{A}$ and $\mathcal{A}_f$ are the same:

$$(\phi_i \circ f^{-1}) \circ (\phi_j \circ f^{-1})^{-1} = \phi_i \circ \phi_j^{-1},$$

and the local representation of f with respect to a pair of coordinate systems $(\phi_i, \phi_i \circ f^{-1})$ is the identity map. Therefore X is a manifold of the same dimension as M, and f is a diffeomorphism. $\square$

Example 1.3.2. Let M be the circle S^1 in the plane, and X be the space obtained from the interval $[-1, 1]$ by collapsing its end points $-1, 1$ (i.e. X is the quotient space of $[-1, 1]$ modulo the equivalence relation : $t \sim s$ if $t, s \in \{-1, 1\}$), and

$$\pi : [-1, 1] \longrightarrow X$$

be the canonical projection. The map $g : [-1, 1] \longrightarrow S^1$, defined by $g(t) = e^{i\pi t}$, is a continuous surjection, and passes onto the quotient to give a continuous bijection $\overline{g} : X \longrightarrow S^1$ so that $g = \overline{g} \circ \pi$. Since X is compact and S^1 is Hausdorff, $\overline{g}$ is a homeomorphism. Then transporting the differentiable structure of the circle S^1 onto X by means of $(\overline{g})^{-1}$, we may identify X with S^1 as a differentiable manifold, and write $S^1 = [-1,1]/\{-1,1\}$.

Theorem 1.3.3. *If $f : M \longrightarrow N$ is a diffeomorphism between manifolds, then the smooth structure on N induced by f as a homeomorphism is the same as the original smooth structure on N.*

PROOF. Suppose that the smooth structures on M and N are given respectively by the atlases $\mathcal{A} = \{(U_i, \phi_i)\}$, and $\mathcal{B} = \{(V_j, \psi_j)\}$. Then any chart $(f(U_i), \phi_i \circ f^{-1})$ of the induced structure $\mathcal{A}_f$ on N is compatible with any chart (V_j, ψ_j) of $\mathcal{B}$, where $f(U_i) \cap V_j \neq \emptyset$. Because, the transition map

$$\psi_j \circ (\phi_i \circ f^{-1})^{-1} : \phi_i \circ f^{-1}(f(U_i) \cap V_j) \longrightarrow \psi_j(f(U_i) \cap V_j)$$

is nothing but $\psi_j f \phi_i^{-1}$ which is the local representation of the diffeomorphism f in terms of the charts (U_i, ϕ_i) and (V_j, ψ_j) (note that

$$f^{-1}(f(U_i) \cap V_j) \subset U_i \quad \text{and} \quad f(U_i) \cap V_j \subset V_j).$$

$\square$

Definition 1.3.4. Two smooth structures $\mathcal{A}$ and $\mathcal{B}$ on the same topological space X are called **equivalent** if there is a homeomorphism $f : X \longrightarrow X$ such that $\mathcal{B} = \mathcal{A}_f$. Clearly, $\mathcal{A}$ and $\mathcal{B}$ are equivalent if and only if there is a diffeomorphism $f : (X, \mathcal{A}) \longrightarrow (X, \mathcal{B})$.

Note that this equivalence is weaker than compatibility conditions described earlier in connection with Definition 1.1.4. The structures $\mathcal{A}$ and $\mathcal{B}$ on X are compatible or identical if the identity map

$$\mathrm{id} \ : (X, \mathcal{A}) \longrightarrow (X, \mathcal{B})$$

is a diffeomorphism, otherwise $\mathcal{A}$ and $\mathcal{B}$ distinct. For example, the standard smooth structure $\mathcal{A}$ on $\mathbb{R}$ can be transported onto itself by means of the homeomorphism $f : \mathbb{R} \longrightarrow \mathbb{R}$ given by $f(x) = x$ if $x \geq 0$, and $f(x) = 2x$ if $x \leq 0$, obtaining a distinct but equivalent structure $\mathcal{A}_f$ on $\mathbb{R}$. The structures $\mathcal{A}$ and $\mathcal{A}_f$ are distinct, because f is not differentiable at $x = 0$.

$\Diamond$ **Exercise 1.15.** Show that the atlases $\{(\mathbb{R}, \phi_i)\}$, each consisting of a single chart $(\mathbb{R}, \phi_i)$, where $\phi_i(x) = x^{2i+1}$, $i = 0, 1, 2, \ldots$, generate distinct smooth structures on $\mathbb{R}$. Also show that all these structures are equivalent to the standard one.

$\Diamond$ **Exercise 1.16.** Show that the two differential structures on the projective space $\mathbb{R}P^n$ as defined in Example 1.1.11, and Exercise 1.4 (p.8) are equivalent.

We will show in Theorem 6.3.2 that any two smooth structures on the circle S^1 are equivalent. This is also true for $\mathbb{R}$. In 1942 T. Rado [**39**] proved that every topological manifold of dimension 2 admits a unique structure, and in 1952 E. Moise [**33**] proved the same fact for topological manifolds of dimension 3. In fact, by this time it was known that any topological manifold of dimensions ≤ 3 admits a smooth structure which is unique up to diffeomorphism. The analogous question in higher dimensions still remains largely unsolved.

The theory of differentiable structures on spheres S^n was originated with the sensational work of J. Milnor [**28**] who discovered that there are 28 smooth structures on the sphere S^7 whose manifold topologies are the standard topology of S^7, but no two of them are equivalent. Of these structures, 27 correspond to what are known as **exotic spheres**, or **Milnor's spheres**. These manifolds are certain S^3-bundles over S^4.

These spheres can also be realised as submanifolds $\mathbb{C}^5$ or $\mathbb{R}^{10}$ by the system of two equations

$$z_1^{5+8k} + z_2^3 + z_3^2 + z_4^2 + z_5^2 = 0, \quad |z_1|^2 + |z_2|^2 + |z_3|^2 + |z_4|^2 + |z_5|^2 = 1,$$

where z_i are complex variables and k takes values from 0 to 27. This ingenious construction is due to E. Brieskorn and F. Hirzebruch.

M. Kervaire and J. Milnor [**20**] also found many other spheres which do not have a unique smooth structure. For example, S^8 has two distinct smooth structures, and S^{11} has 992 distinct smooth structures. The number of smooth structures on S^n is finite for any $n \neq 4$.

A result of J. Stallings [**45**] says that if $n \neq 4$, then any smooth structure on $\mathbb{R}^n$ is equivalent to the standard one. The case $n = 4$ was settled by S. Donaldson who proved, using results of M. Freedman, that $\mathbb{R}^4$ possesses a smooth structure which is not equivalent to its standard structure. It is also known that the number of non-equivalent classes of smooth structures on $\mathbb{R}^4$ is uncountable. This was shown by R. Gompf and C. Taubes. Other remarkable consequences of Freedman's and Donaldson's work are that there are many examples of compact topological manifolds of dimension 4 with no differentiable structure, and there are many pairs of compact smooth manifolds of dimension 4 that are homeomorphic but not diffeomorphic. All these works may be found in Journal of Differential Geometry, Vol. 17, 18, 21, and 25.

1.4. Immersions and Submersions

Convention. From now on, by a manifold we shall always mean a smooth manifold, unless it is stated explicitly otherwise. Sometimes we call a manifold M of dimension n an n-manifold, if it be necessary to specify its dimension.

We recall from calculus the process of derivation which assigns to each differentiable map and each point of its domain a linear map.

Definition 1.4.1. Let $U \subset \mathbb{R}^n$ be an open set, and $a \in U$. Then a map $f : U \longrightarrow \mathbb{R}^m$ is **differentiable** at a if there is a linear map $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ such that

$$\lim_{u \to a} \frac{\|f(u) - f(a) - L(u-a)\|}{\|u-a\|} = 0.$$

The linear map L is unique. For, if $L' : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is another such linear map, then we have for $v \neq 0$

$$\frac{\|L(v) - L'(v)\|}{\|v\|} = \lim_{t \to 0} \frac{\|L(tv) - L'(tv)\|}{\|tv\|}$$

$$\leq \lim_{t \to 0} \frac{\|f(a+tv) - f(a) - L(tv)\|}{\|tv\|} + \lim_{t \to 0} \frac{\|f(a+tv) - f(a) - L'(tv)\|}{\|tv\|} = 0,$$

and so $L(v) = L'(v)$ for all $v \in \mathbb{R}^n$.

The linear map L is called the **derivative map** (or **total derivative**) of f at a, and is denoted by $df_a : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. Its value at $v \in \mathbb{R}^n$ is given by

$$df_a(v) = \lim_{t \to 0} \frac{f(a+tv) - f(a)}{t}.$$

For future reference, we list some well-known results.

Proposition 1.4.2. *The derivative map enjoys the following properties.*

(1) *If df_a exists, then f is continuous at a.*
(2) *If f is a constant map, then $df_a = 0$.*
(3) *If f is a linear map, then $df_a = f$.*
(4) *If $f, g : U \longrightarrow \mathbb{R}^m$ are differentiable at a, then $f+g$ is differentiable at a, and $d(f+g)_a = df_a + dg_a$.*
(5) *If $\lambda : U \longrightarrow \mathbb{R}$ and $f : U \longrightarrow \mathbb{R}^m$ are differentiable at a, then λf is differentiable at a, and $d(\lambda f)_a = \lambda(a) df_a + f(a) d\lambda_a$.*
(6) (Chain Rule). If $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ are open sets, and $f : U \longrightarrow V$, $g : V \longrightarrow \mathbb{R}^p$ are differentiable maps, then their composition $g \circ f$ is differentiable, and, for each $a \in U$,

$$d(g \circ f)_a = dg_{f(a)} \circ df_a.$$

Notice that if f is a diffeomorphism, then df_a is a linear isomorphism for all $a \in \mathbb{R}^n$.

If $m = 1$, and $(\alpha_1, \ldots, \alpha_n)$ is an orthonormal basis of $\mathbb{R}^n$ with coordinate functions $u_1, \ldots, u_n$ so that, for $p \in \mathbb{R}^n$, $u_i(p) = \langle p, \alpha_i \rangle$ is the i-th coordinate of p, then $df_a(\alpha_i)$ is the i-th partial derivative $\partial f / \partial u_i(a)$ of f at a. Setting $v = v_1\alpha_1 + \cdots + v_n\alpha_n$, we have

$$df_a(v) = v_1 \frac{\partial f}{\partial u_1}(a) + \cdots + v_n \frac{\partial f}{\partial u_n}(a),$$

by the above properties. In general, if $(\beta_1, \ldots, \beta_m)$ is an orthonormal basis of $\mathbb{R}^m$ so that

$$f(u) = \sum_{i=1}^{m} f_i(u)\beta_i,$$

where the components $f_i : U \longrightarrow \mathbb{R}$ are continuous and satisfy $f_i(u) = \langle f(u), \beta_i \rangle$, then df_a exists if and only if df_{ia} exists, and in that case

$$df_a(v) = \sum_{i=1}^{m} df_{ia}(v)\beta_i = \sum_{i=1}^{m} \left(v_1 \frac{\partial f_i}{\partial u_1}(a) + \cdots + v_n \frac{\partial f_i}{\partial u_n}(a) \right) \beta_i.$$

It follows that the matrix of the linear map df_a with respect to the bases α_i and β_j is the Jacobian matrix

$$Jf(a) = \left(\frac{\partial f_i}{\partial u_j}(a) \right).$$

Note that $f : U \longrightarrow \mathbb{R}^m$ is a C^1-map if and only if the map

$$df : U \longrightarrow L(\mathbb{R}^n, \mathbb{R}^m)$$

sending a to df_a, where $L(\mathbb{R}^n, \mathbb{R}^m)$ denotes the vector space of linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$, is continuous.

Let f be a smooth function from an open set V of an n-manifold M into $\mathbb{R}$. Then, for every chart (U, ϕ) on M with $U \cap V \neq \emptyset$, the function $f \circ \phi^{-1} : \phi(U \cap V) \longrightarrow \mathbb{R}$ is smooth. If $\phi = (x_1, \ldots, x_n)$, $x_i = u_i \circ \phi$, then the **partial derivative** of f with respect to x_i at $p \in U \cap V$, is defined by

$$\frac{\partial f}{\partial x_i}(p) = \frac{\partial (f \circ \phi^{-1})}{\partial u_i}(\phi(p)).$$

Let M and N be manifolds of dimension n and m respectively. If

$$f : M \longrightarrow N$$

is a smooth map, and $\phi = (x_1, \ldots, x_n)$ and $\psi = (y_1, \ldots, y_m)$ are coordinate systems in M and N respectively, then the functions $f_i = y_i \circ f$ of $x_1, \ldots, x_n$ are called the **components** of f. The **Jacobian matrix** of f relative to the pair of coordinate systems (ϕ, ψ) is defined to be the $m \times n$ matrix

$$Jf = \left(\frac{\partial f_i}{\partial x_j} \right).$$

Note that this is nothing but the Jacobian matrix Jg of the local representation $g = \psi \circ f \circ \phi^{-1}$. The **rank** of f at p is defined to be the rank of $Jf(p)$. The definition is independent of the local representation of f. This may be seen easily. Suppose that $g = \psi \circ f \circ \phi^{-1}$ and $g' = \psi' \circ f \circ \phi'^{-1}$ are two local representations of f at p for the pairs of coordinate charts (U, ϕ), (V, ψ) and (U', ϕ'), (V', ψ') respectively. We may suppose that $U = U'$ and $V = V'$, by replacing U, U' by $U \cap U'$ and V, V' by $V \cap V'$. Then $g' = (\psi' \circ \psi^{-1}) \circ g \circ (\phi \circ \phi'^{-1})$. This proves the assertion, since $\phi \circ \phi'^{-1}$ and $\psi' \circ \psi^{-1}$ are diffeomorphisms.

We will now prove some theorems which will provide the keys to understanding the local behaviour of a smooth map of maximum rank.

Theorem 1.4.3 (**Inverse Function Theorem).** *Let M and N be manifolds of the same dimension n, and $f : U \longrightarrow V$ be a smooth map, where U and V are open sets of M and N respectively. Then, if rank $f = n$ at a point $p \in U$, there exists an open neighbourhood W of p in U such that $f|W$ is a diffeomorphism onto an open neighbourhood of $f(p)$ in V.*

The converse is also true. This we shall see in Chapter 3 when we read the derivative map between tangent spaces.

Proof. The theorem is just the Inverse Function Theorem of Calculus when $M = N = \mathbb{R}^n$, and its proof follows trivially from this special case. By hypothesis, any local representation $g = \psi \circ f \circ \phi^{-1}$ of f has rank n at the point $\phi(p)$, and therefore there is an open neighbourhood W' of $\phi(p)$ on which g is a diffeomorphism. Then the restriction of f to $W = \phi^{-1}(W')$ is also a diffeomorphism. □

The next theorem generalises this result, when $\dim M \leq \dim N$.

Definition 1.4.4. Let M and N be manifolds of dimension n and m respectively. A smooth map $f : M \longrightarrow N$ is called an *immersion* at $p \in M$ if $n \leq m$ and rank $f = n$ at p. It is called a *submersion* at p if $n \geq m$ and rank $f = m$ at p. The map f is called an **immersion**, or a **submersion**, if it is so at each point of M.

Also, f is called an **embedding** if it an immersion, and a homeomorphism onto its image $f(M)$. If $n = m$, then a surjective embedding is a diffeomorphism.

.

Examples 1.4.5. (1) If $n \leq m$, the standard inclusion map $i : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ given by $(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_n, 0, \ldots, 0)$ is an embedding. It is called the **canonical embedding**.

(2) If $n \geq m$, the projection map $s : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ onto the first m coordinates given by $(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_m)$ is a submersion. It is called the **canonical submersion**.

The following examples show that an injective immersion may not be an embedding.

Example 1.4.6. The map $f : [0, 2\pi] \longrightarrow \mathbb{R}^2$ given by $f(t) = (\sin 2t, -\sin t)$ is an immersion. As t varies from 0 to 2π, the image point traces the lower half of the figure "8$''$ in the clockwise direction, and then traces the upper half in the anti-clockwise direction. (The Cartesian equation of the curve is $x^2 = 4y^2(1 - y^2)$.) It is not an embedding, because there is a crossing at the origin. The restriction $f|(0, 2\pi)$ is an injective immersion, but not an embedding, as it is not a homeomorphism onto its image (the ends are not joined). However, the restriction $f|(0, \pi)$ is a embedding, as the image is the lower half of the figure '8' without the origin.

Example 1.4.7. Consider the map $f : \mathbb{R} \longrightarrow S^1 \times S^1$ given by

$$f(t) = (e^{2\pi i \alpha t}, e^{2\pi i \beta t}),$$

where α/β is irrational, The map is an immersion, since df/dt is never zero. It is injective, since $f(t_1) = f(t_2)$ implies that both $\alpha(t_1 - t_2)$ and $\beta(t_1 - t_2)$ are integers, which is not possible unless $t_1 = t_2$. It is not hard to show that the image $f(\mathbb{R})$ is an everywhere dense curve winding around the torus $S^1 \times S^1$. Therefore f is far from being an embedding, because the image of an embedding cannot be dense (see Proposition 1.5.4).

Note that the fact that $\mathbb{R}$ is not compact plays an essential role in these examples. Indeed, we have the following simple result.

$\Diamond$ **Exercise 1.17.** Show that if M is a compact manifold, then any injective immersion $M \longrightarrow N$ is an embedding.

Definition 1.4.8. Two smooth maps $f : M \longrightarrow N$ and $f' : M' \longrightarrow N'$ are called equivalent up to diffeomorphism if there exist diffeomorphisms $\phi : M \longrightarrow M'$ and $\psi : N \longrightarrow N'$ such that $\psi \circ f = f' \circ \phi$.

We will show in the next two theorems that any immersion is locally equivalent to the canonical embedding i, and any submersion is locally equivalent to the canonical submersion s.

Theorem 1.4.9 (Local Immersion Theorem). *Let M and N be manifolds of dimension n and m respectively. If $f : M \longrightarrow N$ is an immersion at $p \in M$, then there is a local representation of f at p which is the canonical embedding i.*

Proof. Let $g = \psi \circ f \circ \phi^{-1}$ be a local representation of f at p for a pair of coordinate systems (ϕ, ψ). We may suppose without loss of generality that $\phi(p) = 0$ and $\psi(f(p)) = 0$, and that the matrix of g at 0 is of the form

$$Jg(0) = \begin{pmatrix} A \\ B \end{pmatrix},$$

where A is a non-singular $n \times n$ matrix (the last condition may be realised by permuting the coordinates in ψ, if necessary). By changing the coordinates in $\mathbb{R}^m$ by a linear transformation $\mathbb{R}^m \longrightarrow \mathbb{R}^m$ whose matrix is

$$\begin{pmatrix} A^{-1} & O \\ -BA^{-1} & I_{m-n} \end{pmatrix},$$

where I_{m-n} is the identity matrix of order $m-n$ and O is a null matrix, the matrix $Jg(0)$ may be given the following form

$$\begin{pmatrix} A^{-1} & O \\ -BA^{-1} & I_{m-n} \end{pmatrix} \cdot \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} I_n \\ O \end{pmatrix}.$$

Define a map $h : U \times \mathbb{R}^{m-n} \longrightarrow \mathbb{R}^m$, where U is the domain of g in $\mathbb{R}^n$, by

$$h(x, y) = g(x) + (0, y).$$

Then $g = h \circ i$, where $i : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is the canonical embedding $x \mapsto (x, 0)$, and the matrix $Jh(0)$ is I_m. By the inverse function theorem, h is a local diffeomorphism at $0 \in \mathbb{R}^m$, and we have

$$\psi \circ f \circ \phi^{-1} = g = h \circ i \Rightarrow (h^{-1} \circ \psi) \circ f \circ \phi^{-1} = i.$$

Thus the local representation of f at p for the pair of coordinate systems $(\phi, h^{-1} \circ \psi)$ is the canonical embedding i. $\square$

The following exercise points out that locally there is no distinction between immersion and embedding.

$\Diamond$ **Exercise 1.18.** Show that if $f : M \longrightarrow N$ is an immersion, then each point $p \in M$ has an open neighbourhood U such that $f|U$ is an embedding.

Theorem 1.4.10 (Local Submersion Theorem). *Let M and N be manifolds of dimension n and m respectively. If $f : M \longrightarrow N$ is a submersion at $p \in M$, then there is a local representation of f at p which is the canonical submersion s.*

Proof. As before, suppose that $g = \psi \circ f \circ \phi^{-1}$ be a local representation of f at p for a pair of coordinate systems (ϕ, ψ) such that $\phi(p) = 0$, $\psi(f(p)) = 0$, and that the Jacobian matrix of g at 0 is

$$Jg(0) = \begin{pmatrix} I_m & O \end{pmatrix},$$

after a linear change of coordinates in $\mathbb{R}^n$. Then, the map $h : U \longrightarrow \mathbb{R}^n$ given by $h(x) = (g(x), x_{m+1}, \ldots, x_n)$ has the Jacobian matrix I_n at $x = 0$, and we have $g = s \circ h$. Therefore $\psi \circ f \circ (h \circ \phi)^{-1}$ is the canonical submersion s. $\square$

$\Diamond$ **Exercises 1.19.** (a) Show that any submersion is an open map (i.e. maps an open set onto an open set).

(b) Show that if M is compact and N is connected, then any submersion $f : M \longrightarrow N$ is surjective.

(c) Show that there is no submersion of a compact manifold into a Euclidean space.

Proposition 1.4.11. *Let M, N, and P be manifolds, and $f : M \longrightarrow N$ be a surjective submersion. Then a map $g : N \longrightarrow P$ is smooth if and only if the composition $g \circ f : M \longrightarrow P$ is smooth.*

PROOF. If g is smooth, then $g \circ f$ is smooth by composition. To prove the converse, note that g is necessarily continuous, and, since the problem is local, we may suppose that f is the projection $(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_m)$ from $\mathbb{R}^n$ onto $\mathbb{R}^m$, where $n = \dim M, m = \dim N$, and $n \geq m$. Then, by hypothesis, the map $g \circ f : (x_1, \ldots, x_n) \mapsto g(x_1, \ldots, x_m)$ is smooth. Therefore the map $g : (x_1, \ldots, x_m) \mapsto g(x_1, \ldots, x_m)$ is smooth. This means that g is smooth on $f(M)$, and hence on N, since f is surjective. □

◊ **Exercise 1.20.** Show that if f and g are as in this proposition, then g is a submersion if and only if their composition $g \circ f$ is a submersion.

◊ **Exercises 1.21.** (a) Show that if $f : M \longrightarrow N$ is a surjective submersion, then for each $x \in M$ there exist an open neighbourhood U of $f(x)$ in N, and a smooth map $g : U \longrightarrow M$ such that $f \circ g$ is the identity map on U.

The map g is called a local section of f.

(b) Suppose that $f : M \longrightarrow N$ is a smooth map such that every point of M is in the image of a smooth local section of f. Show that f is a submersion.

◊ **Exercise 1.22.** If $f : M \longrightarrow N$ is a map and $y \in N$, then $f^{-1}(y)$ is called the fibre of f over y. Suppose that f is a surjective submersion. Show that if $g : M \longrightarrow P$ is a smooth map that is constant on the fibres of f, then there is a unique smooth map $h : N \longrightarrow P$ such that $h \circ f = g$.

◊ **Exercise 1.23.** Show that a smooth map $f : M \longrightarrow N$ is a diffeomorphism if and only if it is bijective and a submersion.

◊ **Exercise 1.24.** Let M, N, and P be manifolds, and $f : M \longrightarrow N$ be an immersion. Then show that a continuous map $g : P \longrightarrow M$ is smooth if and only if their composition $f \circ g : P \longrightarrow N$ is smooth.

◊ **Exercise 1.25.** Prove the implicit function theorem in the following form. If $f : U \longrightarrow \mathbb{R}$, U open in $\mathbb{R}^n$, is a smooth map with $f(p) = q$ and $\partial f / \partial u_i(p) \neq 0$ for some i, then there is a smooth function

$$u_i = g(u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_n)$$

whose graph in some open neighbourhood of p in U is the set of solutions of the equation $f(u) = q$.

1.5. Submanifolds

Definition 1.5.1. Let N be an m-manifold. Then a subset M of N is called a **submanifold** of dimension n if for each point $p \in M$ there is a coordinate chart (U, ϕ) at p in N such that ϕ maps $M \cap U$ homeomorphically onto an open subset of $\mathbb{R}^n \subset \mathbb{R}^m$, where $\mathbb{R}^n$ is considered as the subspace of the first n coordinates in $\mathbb{R}^m$

$$\mathbb{R}^n = \{(x_1, \ldots, x_m) \in \mathbb{R}^m \mid x_{n+1} = \cdots = x_m = 0\}.$$

Then the collection

$$\{(M \cap U, \phi|M \cap U) \mid (U, \phi) \text{ is a chart in } N, M \cap U \neq \emptyset\}$$

is a smooth atlas of M.

$\Diamond$ **Exercise 1.26.** Show that a submanifold M of a manifold N is a second countable Hausdorff space.

Lemma 1.5.2. *Let M and N be manifolds of dimension n and m respectively. If M is a submanifold of N, then for each point $p \in M$ there is an open neighbourhood U of p in N and a submersion $g : U \longrightarrow \mathbb{R}^{m-n}$ such that $g^{-1}(0) = M \cap U$.*

PROOF. By the above definition, there is a coordinate chart $\phi : U \longrightarrow \mathbb{R}^m$ about p in N such that if $\mathbb{R}^m = \mathbb{R}^n \times \mathbb{R}^{m-n}$, then $\phi^{-1}(\mathbb{R}^n \times \{0\}) = M \cap U$. Then $g = \pi \circ \phi$, where $\pi : \mathbb{R}^m \longrightarrow \mathbb{R}^{m-n}$ is the projection onto the second factor, is a submersion with $g^{-1}(0) = M \cap U$. $\square$

Proposition 1.5.3. *A subset A of an m-manifold N is a submanifold if and only if A is the image of a smooth embedding $f : M \longrightarrow N$, where M is an n-manifold and $n \leq m$.*

PROOF. If A is a submanifold of N, then it follows from the natural smooth structure on A derived from that of N that the inclusion of A in N is a smooth embedding. Conversely, suppose $f : M \longrightarrow N$ is a smooth embedding and $A = f(M)$. Then, by the local immersion theorem, for each $p \in M$ there exist a coordinate system $y_1, \ldots, y_m$ in an open neighbourhood V of $f(p)$ in N such that $A \cap V = \{q \in V | y_{n+1}(q) = \cdots = y_m(q) = 0\}$, and the restrictions of the remaining coordinate functions $y_1, \ldots, y_n$ to $A \cap V$ form a local chart on A at $f(p)$. Therefore A is a submanifold of N. $\square$

Proposition 1.5.4. *If M is an n-submanifold of an m-manifold N where $n < m$, then M is not a dense subset of N.*

PROOF. There is a coordinate chart (V, ψ) of N such that $U = M \cap V$ is non-empty, and $\psi(U) \subset \mathbb{R}^n \times \{0\}$. Then the non-empty open set

$$\psi^{-1}(\mathbb{R}^n \times (\mathbb{R}^{m-n} - \{0\}))$$

of N lies in V and does not intersect U. So M cannot be dense in N. $\square$

$\Diamond$ **Exercises 1.27.** Let M, N, and P denote manifolds, where M is a submanifold of N. Then show that

(1) if $f : N \longrightarrow P$ is a smooth map, then the restriction $f|M$ is also smooth; moreover, if f is an immersion, then $f|M$ is also an immersion.

(2) if M is a subset of N such that the inclusion $M \hookrightarrow N$ is an immersion, and $f : P \longrightarrow N$ is a smooth map with $f(P) \subset M$, then the map $f : P \longrightarrow M$ obtained by restricting the range of f may not be continuous. However, if

$$f : P \longrightarrow M$$

is continuous, then it is also smooth.

Definition 1.5.5. Let $f : M \longrightarrow N$ be a smooth map. Then a point $p \in M$ is called a **critical point** of f if f is not a submersion at p. Other points of M are called **regular points** of f. A point $q \in N$ is called a **critical value** of f if $f^{-1}(q)$ contains at least one critical point. Other points of N (including those for which $f^{-1}(q)$ is empty) are called **regular values** of f.

Theorem 1.5.6. *Let M and N be manifolds of dimension n and m respectively, where $n \geq m$. If $q \in N$ is a regular value of a smooth map $f : M \longrightarrow N$, then $f^{-1}(q)$ is a submanifold of M of dimension $n - m$.*

Proof. Since f is a submersion at a point $p \in f^{-1}(q)$, we can choose local coordinate systems about p and q such that $f(x_1, \ldots, x_n) = (x_1, \ldots, x_m)$, and q corresponds to $(0, \ldots, 0)$. Therefore, if U is the coordinate neighbourhood at p on which the functions $x_1, \ldots, x_n$ are defined, then $f^{-1}(q) \cap U$ is the set of points $(0, \ldots, 0, x_{m+1}, \ldots, x_n)$. Thus the functions $x_{m+1}, \ldots, x_n$ form a coordinate system on the relative open set $f^{-1}(q) \cap U$ of $f^{-1}(q)$. □

We may apply the theorem in the following situation. Let $m > n$, and N be an m-manifold. Let $f : N \longrightarrow \mathbb{R}^{m-n}$ be a smooth map. Then $M = f^{-1}(0)$ is the solution set of the system of equations

$$f_1(x_1, \ldots, x_n) = 0, \ldots, \ldots, f_{m-n}(x_1, \ldots, x_n) = 0,$$

where $f_i : N \longrightarrow \mathbb{R}$ are the components of f.

Proposition 1.5.7. *If f, N, and M are as above and rank $f = m - n$ at each point of N, then M is an n-dimensional submanifold of N.*

Proof. The proof follows immediately from the previous theorem. □

The converse is true locally.

Proposition 1.5.8. *Every n-submanifold M of an m-manifold N is locally definable as the set of common zeros of a set of functions*

$$f_1, \ldots, f_{m-n} : U \longrightarrow \mathbb{R}$$

such that

$$\mathit{rank} \left(\frac{\partial f_i}{\partial x_j} \right) = m - n,$$

where U is a coordinate neighbourhood in N of a point in M with coordinates $x_1, \ldots, x_m$.

Proof. The proof follows immediately from the local immersion theorem. If $p \in M$, then there exists local coordinate system $x_1, \ldots, x_m$ defined on a neighbourhood U of p in N such that $M \cap U$ is given by the equations

$$x_{n+1} = 0, \ldots, x_m = 0.$$

□

Proposition 1.5.7 may be illustrated by an example.

Example 1.5.9. Stiefel manifold. A k-frame in $\mathbb{R}^n$ is an ordered sequence of k orthonormal vectors $(v_1, \ldots, v_k)$, $v_i \in \mathbb{R}^n$. Such a k-frame may be represented by an $n \times k$ matrix A whose column vectors are $v_1, \ldots, v_k$. The orthonormality means that $A^t A = I_k$, where A^t denotes the transpose of A and I_k is the identity matrix of order k. The set of all k-frames in $\mathbb{R}^n$ is called a Stiefel manifold, and is denoted by $V_k(\mathbb{R}^n)$. This is a subspace of the space of $n \times k$ matrices $M(n, k, \mathbb{R})$, which we identify with $\mathbb{R}^{nk}$ by writing the entries of a matrix in some fixed order. Note that $V_n(\mathbb{R}^n)$ is the group of real orthogonal matrices $O(n)$.

Any orthogonal matrix $T \in O(n)$ sends a k-frame $v_1, \ldots, v_k$ to another k-frame $Tv_1, \ldots, Tv_k$. Moreover, given two k-frames $v_1, \ldots, v_k$ and $w_1, \ldots, w_k$, there is always an orthogonal matrix $T \in O(n)$ such that $w_i = Tv_i$, $1 \leq i \leq k$. These statements can be verified easily using elementary Linear Algebra.

We shall now show that the Stiefel manifold $V_k(\mathbb{R}^n)$ is a smooth manifold of dimension $nk - k(k+1)/2$. The entries of a matrix $(x_{ij}) \in V_k(\mathbb{R}^n)$ satisfy the following $r = k(k+1)/2$ equations

$$\langle v_i, v_j \rangle = \sum_{s=1}^{n} x_{si} \cdot x_{sj} = \delta_{ij},\ 1 \leq i \leq j \leq k,$$

where δ_{ij} is the Kronecker delta : $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$. Define a map $f : M(n, k, \mathbb{R}) \longrightarrow \mathbb{R}^r$ by $f(A) = A^t A - I_k$. Then $f^{-1}(0) = V_k(\mathbb{R}^n)$, and the components of f are

$$f_{ij} = \sum_{s=1}^{n} x_{si} x_{sj} - \delta_{ij},\ i, j = 1, \cdots, k,\ i \leq j.$$

Take an element $A_0 = (a_{ij}) \in V_k(\mathbb{R}^n)$, where $a_{ij} = \delta_{ij}$, $i = 1, \ldots, n$, $j = 1, \ldots, k$. Then some little computations show that the Jacobian matrix $Jf(A_0)$ of f at A_0 has rank r, because it has r columns of the form $\lambda_1 e_1, \ldots, \lambda_r e_r$, where e_i is the standard i-th unit vector in $\mathbb{R}^r$, and each λ_i is either 1 or 2. It follows from this that the rank of the Jacobian matrix $Jf(A)$ at any matrix $A \in V_k(\mathbb{R}^n)$ is also r. This may be seen in the following way. For A and A_0, we can find an orthogonal matrix $B \in O(n)$ such that $A = BA_0$. Then the left multiplication by B, $L_B : M(n, k, \mathbb{R}) \longrightarrow M(n, k, \mathbb{R})$, given by $L_B(C) = BC$, is a diffeomorphism with inverse $L_{B^{-1}}$ (note that L_B is smooth, because the entries of the matrix BC are polynomials in the entries of the matrix C), and we have $f \circ L_B = f$. Then, by chain rule

$$Jf(A_0) = J(f \circ L_B)(A_0) = Jf(A) \cdot J(L_B)(A_0).$$

Thus rank Jf(A) = rank Jf(A_0) = r, and f is a submersion. Therefore $V_k(\mathbb{R}^n)$ is a submanifold of $M(n, k, \mathbb{R})$ of codimension $k(k+1)/2$.

1.6. Further examples of manifolds

Definition 1.6.1. A **Lie group** is a smooth manifold G together with a group structure on it such that the multiplication operation $G \times G \longrightarrow G$ defined by $(g, h) \mapsto gh$, and the operation of taking inverses $G \longrightarrow G$ defined by $g \mapsto g^{-1}$ are smooth maps.

If G is a topological manifold and the group operations are continuous maps, then G is called a **topological group**.

The additive group of real numbers $(\mathbb{R}, +)$ is a commutative Lie group, so is the circle S^1 considered as the set of complex numbers of norm 1 with complex multiplication. The most important Lie group is the general linear group $GL(n, \mathbb{C})$ of $n \times n$ complex nonsingular matrices with matrix multiplication. We have already seen in Example 1.1.8(2) that $GL(n, \mathbb{C})$ is a manifold of dimension $2n^2$. Moreover, the matrix multiplication is a polynomial map of the entries of the matrices, and the matrix inversion is a rational map of the entries; so they are infinitely differentiable.

An abstract subgroup and a submanifold H of a Lie group G, which is a Lie group with its induced smooth structure, is called a **Lie subgroup**. Such a group arises as the image of a homomorphism between Lie groups which is also an embedding.

An important theorem about Lie subgroups says that if H is an abstract subgroup and a closed subset of a Lie group G, then H is a Lie subgroup of G (this is called the Cartan's criterion, see Chevalley [**5**], p. 135). The celebrated theorem of Peter-Weyl says that any compact Lie group can be embedded in $GL(n, \mathbb{C})$ as a Lie subgroup. The proofs are far beyond the scope of this book.

Examples 1.6.2. A **classical matrix group** is a Lie subgroup of $GL(n, \mathbb{C})$. Here are some examples.

(1) **Real general linear group** $GL(n, \mathbb{R})$ consisting of matrices A such that $A = \overline{A}$, where $\overline{A}$ is the complex conjugate of A; if $A = (a_{ij})$, then $\overline{A} = (\overline{a_{ij}})$.
(2) **Special linear group** $SL(n, \mathbb{C})$ consisting of matrices A such that $\det A = 1$.
(3) **Real special linear group** $SL(n, \mathbb{R}) = SL(n, \mathbb{C}) \cap GL(n, \mathbb{R})$.
(4) **Complex orthogonal group** $O(n, \mathbb{C})$ consisting of matrices A such that $A = A^*$, where A^* is the inverse transpose of A; $A^* = (A^{-1})^t$.
(5) **Unitary group** $U(n)$ consisting of matrices A such that $\overline{A} = A^*$.
(6) **Special unitary group** $SU(n) = SL(n, \mathbb{C}) \cap U(n)$.
(7) **Real orthogonal group** $O(n)$ consisting of matrices A such that
$$A = \overline{A} = A^*.$$
Therefore $O(n) = GL(n, \mathbb{R}) \cap U(n) = GL(n, \mathbb{R}) \cap O(n, \mathbb{C})$.
(8) **Real special orthogonal group** $SO(n) = SL(n, \mathbb{C}) \cap O(n)$.

All the subgroups are closed subsets of $GL(n,\mathbb{C})$, because each of the first five groups is defined by a closed condition, and each of the others is the intersection of two closed sets. Each of the last five groups is compact, because the absolute values of entries of matrices are bounded by 1, and therefore each is a compact subset of the Euclidean $2n^2$-space.

We shall prove that these are submanifolds of $GL(n,\mathbb{C})$. For this purpose we need to look at a basic construction which is the exponential map.

Lemma 1.6.3. *If A is a matrix in $M(n,\mathbb{C})$, then the series*

$$I_n + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots ,$$

where I_n denotes the identity matrix of order n, converges absolutely, and uniformly on any compact subset of $M(n,\mathbb{C})$.

Proof. If A has entries a_{ij}, let $a_{ij}^{(k)}$ denote the entries of the k-th power A^k, where k is an integer ≥ 0. We write $A^0 = I_n$, and $A^1 = A$. It follows by induction on k that if $|a_{ij}| \leq a$, then $|a_{ij}^{(k)}| \leq (na)^k$. Indeed, if the result is true for some $k > 1$, then

$$\left|a_{ij}^{(k+1)}\right| = \left|\sum_{r=1}^{n} a_{ir}^{(k)} a_{rj}\right| \leq n(na)^k a = (na)^{k+1}.$$

Therefore each of the n^2 series

$$\sum_{k=0}^{\infty} \frac{1}{k!} a_{ij}^{(k)}$$

converges absolutely, and uniformly on the set of all matrices A for which $|a_{ij}| \leq a$. This means that the series

$$I_n + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots$$

converges absolutely, and uniformly on any compact subset of the space of $M(n,\mathbb{C})$. $\square$

Writing $\exp A = \sum_{k=0}^{\infty}(1/k!)A^k$, we get a map $\exp : M(n,\mathbb{C}) \longrightarrow M(n,\mathbb{C})$, which sends A to $\exp A$. This is called the **matrix exponential map**.

Lemma 1.6.4. *The matrix exponential map* $\exp$ *is analytic, and its image is a subset of $GL(n,\mathbb{C})$. Also* $\exp$ *sends a neighbourhood of* 0 *in* $M(n,\mathbb{C})$ *diffeomorphically onto a neighbourhood of I_n in $GL(n,\mathbb{C})$.*

Proof. The map exp is analytic, by Hartog's theorem ([**16**], Theorem 2.2.8, p. 28), since the entries of $\exp A$ are analytic functions of the entries a_{ij} of A.

If $B \in GL(n,\mathbb{C})$, then $(BAB^{-1})^k = BA^kB^{-1}$, and therefore

$$\exp(BAB^{-1}) = B(\exp A)B^{-1}.$$

If $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A, then there is a matrix $B \in GL(n, \mathbb{C})$ such that BAB^{-1} is an upper triangular matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$. Therefore $B(\exp A)B^{-1}$ is also an upper triangular matrix with diagonal entries $\exp \lambda_1, \ldots, \exp \lambda_n$, and hence

$$\det(\exp A) = \det(B(\exp A)B^{-1}) = \exp \lambda_1 \cdot \cdots \cdot \exp \lambda_n = \exp(\text{trace } A).$$

Thus $\exp A$ is non-singular, and so exp maps $M(n, \mathbb{C})$ into the open subset $GL(n, \mathbb{C})$ of $M(n, \mathbb{C})$.

A simple computation shows that the Jacobian of exp at $\exp(0) = I_n$ is the $n^2 \times n^2$ identity matrix I_{n^2}. Note that the (i, j) entry of $\exp A$ is

$$\delta_{i,j} + a_{i,j} + \text{higher degree terms}.$$

Therefore, by the Inverse Function Theorem, exp maps a neighbourhood of 0 in $M(n, \mathbb{C})$ diffeomorphically onto a neighbourhood of I_n in $GL(n, \mathbb{C})$. $\square$

Lemma 1.6.5. *If $A, B \in M(n, \mathbb{C})$ commute, then*

$$\exp(\lambda A + \mu B) = \exp(\lambda A) \cdot \exp(\mu B), \ \textit{where } \lambda, \mu \in \mathbb{R}.$$

Proof. For fixed A and B, each side of the equality is analytic function of the real variables λ and μ. Since $AB = BA$, the Cauchy product of the series in the left hand side is equal to the series in the right hand side

$$\sum_{k=0}^{\infty} \frac{(\lambda A)^k}{k!} \cdot \sum_{r=0}^{\infty} \frac{(\mu B)^r}{r!} = \sum_{m=0}^{\infty} \sum_{p=0}^{m} \frac{(\lambda A)^{m-p}}{(m-p)!} \cdot \frac{(\mu B)^p}{p!} = \sum_{m=0}^{\infty} \frac{(\lambda A + \mu B)^m}{m!}.$$

Since the convergence is absolute, the two sides are equal, by an analog of Cauchy's theorem of convergence. $\square$

It follows that if A is a fixed matrix, then the map $t \longrightarrow \exp tA$ is a smooth homomorphism of the additive group of real numbers $(\mathbb{R}, +)$ into $GL(n, \mathbb{C})$. Moreover,

$$\exp(-A) = (\exp A)^{-1}, \quad \exp(A^t) = (\exp A)^t, \quad \exp(\overline{A}) = \overline{\exp A}.$$

The first one follows from the above homomorphism, and the next two from the definition of exp.

Now consider the vector spaces

$$S_r = \{A \in M(n, \mathbb{C}) | A = \overline{A}\}, \qquad S_0 = \{A \in M(n, \mathbb{C}) | \text{ trace } A = 0\},$$

$$S_{ss} = \{A \in M(n, \mathbb{C}) | A^t + A = 0\}, \qquad S_{sh} = \{A \in M(n, \mathbb{C}) | A^t + \overline{A} = 0\}.$$

Then S_r is the set of all real matrices, and S_{ss} (resp. S_{sh}) is the set of all skew-symmetric (resp. skew-Hermitian) matrices. We may identify S_r, S_0, S_{ss}, and S_{sh} with the Euclidean spaces so that

$$\dim S_r = n^2, \ \dim S_0 = 2n^2 - 2, \ \dim S_{ss} = n(n-1), \ \dim S_{sh} = n^2.$$

Also note that $S_r \cap S_{ss} = S_r \cap S_{sh} = S_r \cap S_{sh} \cap S_0$ is the vector space of all real skew-symmetric matrices, and we have

$$\dim(S_r \cap S_0) = n^2 - 1, \ \dim(S_r \cap S_{sh}) = n(n-1)/2, \ \dim(S_{sh} \cap S_0) = n^2 - 1,$$

Let U be a neighbourhood of 0 in $M(n, \mathbb{C})$ which is mapped diffeomorphically onto an open neighbourhood of I_n in $GL(n, \mathbb{C})$ by the map exp. Suppose that U is small enough so that $|\operatorname{trace} A| < 2\pi$, for all $A \in U$. Let

$$-U = \{-A | A \in U\}, \quad U^t = \{A^t | A \in U\}, \quad \overline{U} = \{\overline{A} | A \in U\}.$$

Set $V = U \cap (-U) \cap U^t \cap \overline{U}$.

Theorem 1.6.6. *In the following table for each matrix group G of the first row there corresponds a vector space $\mathfrak{g}$ of the second row so that G and $\mathfrak{g}$ appear in the same column. The third row gives the dimension of $\mathfrak{g}$.*

G	$GL(n,\mathbb{C})$	$O(n,\mathbb{C})$	$U(n)$	$SU(n)$	$O(n)$	$SO(n)$
$\mathfrak{g}$	$M(n,\mathbb{C})$	S_{ss}	S_{sh}	$S_{sh} \cap S_0$	$S_r \cap S_{ss}$	$S_r \cap S_{ss} \cap S_0$
dim	n^2	$n(n-1)$	n^2	$n^2 - 1$	$\frac{n(n-1)}{2}$	$\frac{n(n-1)}{2}$

Then exp *maps $\mathfrak{g} \cap V$ diffeomorphically onto an open neighbourhood of I_n in G. Therefore G is a manifold, and its dimension is as given in the third row.*

Proof. The proofs are consequences of the following facts, which have already been established.

(1) $A \in M(n, \mathbb{C}) \Rightarrow \exp A \in GL(n, \mathbb{C})$,

(2) $A \in S_{ss} \Rightarrow \exp A \in O(n, \mathbb{C})$,

(3) $A \in S_{sh} \Rightarrow \exp A \in U(n)$,

(4) $A \in S_{sh} \cap S_0 \Rightarrow \exp A \in SU(n)$.

(5) $A \in S_r \cap S_{sh} \Rightarrow \exp A \in O(n)$.

(6) $A \in S_r \cap S_{sh} \cap S_0 \Rightarrow \exp A \in SO(n)$.

Each of the reverse implications is true if $A \in V$ also. □

◊ **Exercise 1.28.** Show that exp maps $S_0 \cap V$ and $S_r \cap S_0 \cap V$ diffeomorphically onto an open neighbourhood of I_n in $SL(n, \mathbb{C})$ and $SL(n, \mathbb{R})$ respectively.

Remark 1.6.7. It is worth mentioning in passing that the vector space $\mathfrak{g}$ is the Lie algebra of the Lie group G.

A vector space V over $\mathbb{K}$ (which is either $\mathbb{R}$ or $\mathbb{C}$) is a **Lie algebra** if there is a binary operation $V \times V \longrightarrow V$ sending a pair of vectors (A, B) to a vector $[A, B]$ of V, called the **Lie bracket** of A and B. The bracket operation is required to satisfy the following properties:

(1) it is skew-symmetric: $[A, B] = -[B, A]$,

(2) it is bilinear over $\mathbb{K}$: $[r_1A_1 + r_2A_2, B] = r_1[A_1, B] + r_2[A_2, B]$,
(3) it satisfies the **Jacobi identity**

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0.$$

For example each of the vector spaces $\mathfrak{g}$ appearing in the second row of the above table is a Lie algebra with the bracket defined by $[A, B] = AB - BA$.

We shall see in Example 3.1.5 that the Lie algebra $\mathfrak{g}$ is the tangent space of the Lie group G at the unit element $e \in G$.

1.7. Quotient manifolds

Let M be a manifold, and ρ be an equivalence relation on M. Let M/ρ be the quotient set of equivalence classes of M modulo ρ. The canonical surjection $\pi : M \longrightarrow M/\rho$, which maps an element of M onto its equivalence class, induces the quotient topology on M/ρ. In general it is not true that there is a smooth structure on the quotient set M/ρ such that the canonical projection $\pi : M \longrightarrow M/\rho$ is a submersion.

Definition 1.7.1. If the quotient space M/ρ can be given the structure of a smooth manifold so that the projection $\pi : M \longrightarrow M/\rho$ becomes a submersion, then M/ρ is called a **quotient manifold** of M.

Note that this means that M/ρ must be a second countable Hausdorff space. In general, the quotient space M/ρ may not be Hausdorff even if M is so. However, M/ρ is second countable if M is so, because if $\{U_i\}$ is a countable basis of M, then $\{\pi(U_i)\}$ is a similar basis of M/ρ.

Note further that if the quotient space M/ρ admits a quotient manifold structure, then it must be unique. For, if $\mathcal{A}$ and $\mathcal{B}$ are two quotient manifold structures on M/ρ, then the corresponding canonical projections

$$\pi : M \longrightarrow (M/\rho, \mathcal{A}) \quad \text{and} \quad \pi : M \longrightarrow (M/\rho, \mathcal{B})$$

are the same, and are related by $\pi = \mathrm{Id}_{M/\rho} \circ \pi$, and so $\mathrm{Id}_{M/\rho}$ is a smooth map by Proposition 1.4.11, in fact, a diffeomorphism, and therefore $\mathcal{A} = \mathcal{B}$.

Example 1.7.2. A surjective submersion $\phi : M \longrightarrow N$ determines an equivalence relation ρ on M by $x\rho y$ if $\phi(x) = \phi(y)$. Then there is a bijection

$$\psi : M/\rho \longrightarrow N$$

given by $\pi(x) \mapsto \phi(x)$, where $\pi : M \longrightarrow M/\rho$ is the canonical projection. Transporting the differentiable structure of N onto M/ρ by ψ, M/ρ becomes a quotient manifold diffeomorphic to N.

Definition 1.7.3. A Lie group G is said to act on a manifold M as a transformation group if there is a smooth map $\eta : G \times M \longrightarrow M$ such that, writing $\eta(g, x)$ as gx, we have

(i) $ex = x$, where e is the unit in G, and $x \in M$,

(ii) $(g_1g_2)x = g_1(g_2x)$, where $g_1, g_2 \in G$ and $x \in M$.

The map η is called an **action** of G on M, and M is called a G-manifold. If η is continuous, then the action is a continuous action.

More precisely, η is called a left action of G on M. The analogous notion of right G-action on M, $M \times G \longrightarrow M$, is obvious. One has to replace (i) and (ii) by (i$'$) $xe = x$, and (ii$'$) $x(g_1g_2) = (xg_1)g_2$. A right action can be converted to a left action by defining gx to be xg^{-1}. Similarly, a left action can be converted to a right action. Thus any result about left action can be translated into a result about right action, and vice versa. We will prefer to use left actions.

If $\mathrm{Diff}(M)$ is the group of diffeomorphisms of M onto itself, where the group operation is the composition of maps, then the action η gives rise to a homomorphism

$$\eta' : G \longrightarrow \mathrm{Diff}(M).$$

The maps η and η' correspond to each other by means of the relation $\eta(g, x) = \eta'(g)(x)$. Note that if η is a right action, then η' will be an anti-homomorphism $(\eta'(g_1g_2) = \eta'(g_2)\eta'(g_1))$.

The group G is said to act **effectively** on M if the homomorphism η' is a monomorphism, in other words, if $gx = x$ for all $x \in M$ implies $g = e$. It is said to act **freely** (or **without fixed points**) on M if $gx = x$ for some $x \in M$ implies $g = e$.

The set $R = \{(x, gx) | x \in M, g \in G\}$ is an equivalence relation on M. The equivalence class of x is the set $Gx = \{gx | g \in G\}$; it is called the **orbit of** G **through** x. The set of equivalence classes M mod R is denoted by M/G. The quotient space M/G is called the **orbit space** of M. The group G is said to act **transitively** on M, if M/G consists of only one orbit, M itself. This means that for any two points $x, y \in M$ there is a $g \in G$ such that $y = gx$.

Examples 1.7.4. The natural action of $GL(n, \mathbb{R})$ on $\mathbb{R}^n$ is given by matrix multiplication $(A, x) \longrightarrow A \cdot x$, where x is considered as a column matrix. The action has exactly two orbits $\{0\}$ and $\mathbb{R}^n - \{0\}$.

The restriction of the natural action to $O(n) \times \mathbb{R}^n$ is an action of $O(n)$ on $\mathbb{R}^n$. Here the orbits are the origin, and spheres centred at the origin.

Further restriction $O(n) \times S^{n-1} \longrightarrow S^{n-1}$ gives a transitive action of $O(n)$ on S^{n-1}.

Definition 1.7.5. A group G acts **discontinuously** on M if given any point $x \in M$ and any sequence of distinct points $\{g_n\}$ in G, the sequence $\{g_nx\}$ does not converge to any point of M, or equivalently, each orbit Gx is a closed discrete subset of M.

A group G acts **properly discontinuously** on M if

(i) every $x \in M$ has an open neighbourhood U such that $gU \cap U = \emptyset$, for every $g \in G$, $g \neq e$,

(ii) for every $x, y \in M$ with $y \notin Gx$, there exist open neighbourhoods U and V of x and y respectively such that $gU \cap V = \emptyset$ for all $g \in G$.

Here gU denotes the subset $\{gx \mid x \in U\}$ of M.

The condition (i) implies that G acts freely on M. Because, if $x \in M$ and $gx = x$ for some $g \in G$, then, for any neighbourhood U of x, $x \in gU \cap U$ and so $gU \cap U \neq \emptyset$, which is not possible unless $g = e$.

The condition (ii) is equivalent to saying that any two points $x, y \in M$ which are not equivalent, have open neighbourhoods U and V respectively such that $gU \cap hV = \emptyset$ for all $g, h \in G$. Because, $gU \cap hV = (h^{-1}g)U \cap V$.

$\Diamond$ **Exercise 1.29.** Show that the following conditions (iii) and (iv) are equivalent.

(iii) For each compact subset K of M the set $gK \cap K$ is non-empty for only finitely many $g \in G$.

(iv) Every $x \in M$ has an open neighbourhood U such that the set $gU \cap U$ is non-empty for only finitely many $g \in G$

Show that the condition (i) implies the condition (iv), and that a free G-action with the condition (iv) implies the condition (i).

$\Diamond$ **Exercise 1.30.** Show that the condition (ii) is equivalent to saying that the orbit space M/G is a Hausdorff space.

Remark 1.7.6. The condition (iii) is the condition of a properly discontinuous action as considered by Thurston and others for the study of geometric manifolds (see [**40**]). A more general definition of a proper action may be given as follows.

An action of G on M is **proper** if and only if for every compact subset $K \subset M$, the set
$$G_K = \{g \in G \mid gK \cap K \neq \emptyset\}$$
is compact. Alternatively, the action is proper if and only if the map
$$E : G \times M \longrightarrow G \times M$$
given by $E(g, x) = (gx, x)$ is a proper map (i.e. the inverse image of any compact set is compact). To see this, suppose that E is proper. Then
$$\begin{aligned} G_K &= \{g \in G \mid \text{ there is an } x \in K \text{ such that } gx \in K\} \\ &= \{g \in G \mid \text{ there is an } x \in M \text{ such that } E(g, x) \in K \times K\} \\ &= \pi(E^{-1}(K \times K)), \end{aligned}$$
where $\pi : G \times M \longrightarrow G$ is the projection on the first factor. There G_K is compact. Conversely, suppose that G_K is compact for any compact set $K \subset M$. If $L \subset M \times M$ is compact, consider the compact set $K = \pi_1(L) \cup \pi_2(L) \subset M$, where $\pi_1, \pi_2 : M \times M \longrightarrow M$ are the projections on the first and the second factor. Then $E^{-1}(L)$ is a subset of
$$E^{-1}(K \times K) \subset \{(g, x) \in G \times M \mid gx \in K \text{ and } x \in K\} \subset G_K \times K.$$

Since $E^{-1}(L)$ is closed by continuity, it is a closed subset of the compact set

$$G_K \times K,$$

and therefore compact.

A **discrete group** is a topological group G all of whose points are open. Equivalently, G is discrete if and only if $\{e\}$ is open in G. It can be shown that if we consider M as a metric space, and if a discrete subgroup G of the group of smooth isometries of M acts freely and discontinuously on M, then G acts properly discontinuously on M. Recall that an isometry of M preserves distances, and is a diffeomorphism of M.

$\diamondsuit$ **Exercise 1.31.** Show that an action of a discrete group G on a manifold M is smooth if and only if for each $g \in G$, the map $x \mapsto gx$ is a smooth map $M \longrightarrow M$. For example, $\mathbb{Z}^n$ acts smoothly on $\mathbb{R}^n$ by translation

$$(m_1, \dots, m_n) \cdot (x_1, \dots, x_n) = (m_1 + x_1, \dots, m_n + x_n).$$

Proposition 1.7.7. *Any closed discrete subgroup H of a Lie group G acts properly discontinuously on G by left translations: $g \mapsto hg$.*

Proof. Since H is a discrete subgroup of G, there is an open neighbourhood U of e in G such that $H \cap U = \{e\}$, Since the map $\alpha : G \times G \longrightarrow G$, given by $(g_1, g_2) \mapsto g_1 g_2^{-1}$, is continuous, there exists an open neighbourhood V of e in G such that $VV^{-1} \subset U$. From this, we may conclude that only $h \in H$ such that $V \cap hV \neq \emptyset$ is $h = e$. Because, if $V \cap hV \neq \emptyset$, then there exist $v_1, v_2 \in V$ such that $v_1 = hv_2$, and so $h = v_1 v_2^{-1} \in VV^{-1} \subset U$, and therefore we must have $h = e$. This proves the condition (i).

Next, consider the coset space G/H with projection $\pi : G \longrightarrow G/H$. Since the map α is continuous, and $H \subset G$ is closed, the set

$$\alpha^{-1}(H) = \{(g_1, g_2) \in G \times G \mid g_1 = hg_2 \text{ for some } h \in H\}$$

is a closed subset of $G \times G$. Now, if Hg_1 and Hg_2 are distinct points of G/H, then $(g_1, g_2) \notin \alpha^{-1}(H)$, and therefore there exists open neighbourhoods W_1 of g_1 and W_2 of g_2 in G such that $(W_1 \times W_2) \cap \alpha^{-1}(H) = \emptyset$. Then $\pi(W_1)$ and $\pi(W_2)$ are disjoint open neighbourhoods of Hg_1 and Hg_2 respectively, and therefore $hW_1 \cap W_2$ must be an empty set for all $h \in H$. This proves the condition (ii). $\square$

Theorem 1.7.8. *If a group G acts properly discontinuously on a manifold M, then the orbit space M/G admits the structure of a quotient manifold of the same dimension as M so that the projection $\pi : M \longrightarrow M/G$ is a submersion.*

Proof. Each $x \in M$ has an open neighbourhood U on which the canonical surjection $\pi : M \longrightarrow M/G$ is a homeomorphism. To see this, take U as in condition (i) above. Then, if $x, y \in U$, we cannot have $y = gx$ for any $g \in G$, unless $g = e$. Therefore, the map $\pi : U \longrightarrow \pi(U) = V$ is injective, and hence a homeomorphism, because a continuous bijection is a homeomorphism if it is an open map. We may assume by shrinking U, if necessary, that U is the domain of a chart (U, ϕ) in M. We shall show that the pairs (V, ψ), where $V = \pi(U)$ and $\psi = \phi \circ (\pi|U)^{-1}$, form an atlas on M/G. If (V_1, ψ_1) and (V_2, ψ_2) are two such pairs obtained from charts (U_1, ϕ_1) and (U_2, ϕ_2) on M, then, on writing $\lambda_i = (\pi|U_i)^{-1}$, the transition map between them is given by $\psi_2 \circ \psi_1^{-1} = \phi_2 \circ \lambda_2 \circ \lambda_1^{-1} \circ \phi_1^{-1}$. We must therefore show that

$$\lambda_2 \circ \lambda_1^{-1} : \lambda_1(V_1 \cap V_2) \longrightarrow \lambda_2(V_1 \cap V_2)$$

is a smooth map.

For each $g \in G$, let η_g denote the diffeomorphism $\eta_g : M \longrightarrow M$ given by $\eta_g(p) = gp$. We shall show that each point $p \in \lambda_1(V_1 \cap V_2)$ has an open neighbourhood on which $\lambda_2 \circ \lambda_1^{-1}$ coincides with η_g. This will prove that $\lambda_2 \circ \lambda_1^{-1}$ is smooth. So take a point $p \in \lambda_1(V_1 \cap V_2)$, and let $q = \lambda_2(\lambda_1^{-1}(p))$. Then p and q are equivalent, and so there is a $g \in G$ such that $q = gp$. Then $U' = U_1 \cap g^{-1}U_2$ is an open neighbourhood of p, and, since $\pi(U_1) \cap \pi(g^{-1}U_2) = V_1 \cap V_2$, we have $U' \subset \lambda_1(V_1 \cap V_2)$. Now, for any $x \in U'$, $\pi(x) \in V_1 \cap V_2$, and therefore $\lambda_2(\lambda_1^{-1}(x)) = y$, where y is the unique point in U_2 such that $\pi(y) = \pi(x)$. But $gx \in U_2$ and $\pi(gx) = \pi(x) \in V_2$. Therefore $gx = y$. Thus $\lambda_2 \circ \lambda_1^{-1}$ and η_g agree on U', and so $\lambda_2 \circ \lambda_1^{-1}$ is smooth.

It remains to be shown that (1) M/G has a countable basis for its topology, and (2) M/G is Hausdorff. Since M has a countable basis for its open sets, and π is an open map, (1) follows, because if $\{U_\alpha\}$ is a countable basis for M. then $\{\pi(U_\alpha)\}$ is a countable basis for M/G. As for (2), take $y_1, y_2 \in M/G$ so that $y_1 = \pi(x_1)$ and $y_2 = \pi(x_2)$, $x_1, x_2 \in M$. By condition (ii), there exist open neighbourhoods U_1 and U_2 of x_1 and x_2 respectively such that $gU_1 \cap U_2 = \emptyset$ for all $g \in G$. Then $V_1 = \pi(U_1)$ and $V_2 = \pi(U_2)$ are open neighbourhoods of y_1 and y_2 respectively, and we must have $V_1 \cap V_2 = \emptyset$. Because, if $V_1 \cap V_2 \neq \emptyset$, then $\pi^{-1}(V_1 \cap V_2) = \pi^{-1}(V_1) \cap \pi^{-1}(V_2) \neq \emptyset$. But

$$\pi^{-1}(V_1) = \cup_{g_1 \in G}\ g_1 U_1, \qquad \pi^{-1}(V_2) = \cup_{g_2 \in G}\ g_2 U_2.$$

Therefore there exist $g_1, g_2 \in G$ such that $g_1 U_1 \cap g_2 U_2 \neq \emptyset$. Since this contradicts condition (ii), $V_1 \cap V_2 = \emptyset$. Thus M/G is a manifold. Finally, it follows trivially from above that π is a local diffeomorphism, and hence a submersion. □

Corollary 1.7.9. *If H is a discrete subgroup of a Lie group G, then the coset space G/H is a manifold, and the projection $\pi : G \longrightarrow G/H$ is a smooth map.*

Proof. This is a combination of Proposition 1.7.7 and Theorem 1.7.8. □

It may be remarked that a classical result says that for any closed subgroup H of a Lie group G, the coset space G/H is a manifold.

Example 1.7.10. Cylinder. Let the manifold M be the plane $\mathbb{R}^2$. The additive group of integers $\mathbb{Z}$ acts properly discontinuously on $\mathbb{R}^2$ by the action $\mathbb{Z}\times\mathbb{R}^2 \longrightarrow \mathbb{R}^2$ given by $(n,(x,y)) \mapsto (x+n,y)$. This group is generated by the transformation $(x,y)\mapsto(x+1,y)$. The quotient manifold $\mathbb{R}^2/\mathbb{Z}$ is the cylinder.

Example 1.7.11. Möbius band. Consider $\mathbb{Z}$ as generated by the single transformation $(x,y)\mapsto(x+1,-y)$ on $\mathbb{R}^2$. Then it acts on $\mathbb{R}^2$ properly discontinuously. In this case, the quotient manifold $\mathbb{R}^2/\mathbb{Z}$ is the Möbius band.

Example 1.7.12. Torus. The group $\mathbb{Z}\times\mathbb{Z}$ acts properly discontinuously on $\mathbb{R}^2$ by

$$((m,n),(x,y)) \mapsto (x+m, y+n).$$

The transformation group is generated by the two transformations

$$(x,y)\mapsto(x+1,y), \qquad (x,y)\mapsto(x,y+1).$$

The quotient manifold $\mathbb{R}^2/\mathbb{Z}\times\mathbb{Z}$ is the torus.

Example 1.7.13. Klein bottle. Consider $\mathbb{Z}\times\mathbb{Z}$ as a properly discontinuous transformation group of $\mathbb{R}^2$ generated by two transformations

$$(x,y)\mapsto(x+1,-y), \qquad (x,y)\mapsto(x,y+1).$$

In this case the quotient manifold $\mathbb{R}^2/\mathbb{Z}\times\mathbb{Z}$ is the Klein bottle.

The study of properly discontinuous transformation groups will be incomplete unless we solve the following exercises (see [**4**]).

$\Diamond$ **Exercises 1.32.** Let G be a transformation group acting properly discontinuously on a manifold M, and F be a subset of M consisting of just one element from each equivalence class modulo G. Then the closure $\overline{F}$ is called a **fundamental region** of the action.

(1) Show that the equivalence relation on M induced by G gives an equivalence relation ρ on $\overline{F}$ so that if two distinct points of $\overline{F}$ are equivalent, then they belong to the topological boundary $\overline{F} - \text{Int } F$ of F.

(2) Let A be a subset of M, and $R[A]$ denote the set $\{g\in G|\overline{F}\cap gA \neq \emptyset\}$.

Show that if $A\subseteq B$, then $R[A]\subseteq R[B]$.

(3) A fundamental region is called **normal** if each point $x\in M$ has an open neighbourhood U such that $R[U]$ is a finite set.

Show that if $\overline{F}$ is a normal fundamental region, then any $x\in M$ has an open neighbourhood U such that $R[U]=R[x]$. Moreover, if V is another open neighbourhood of x contained in U, then $R[V]=R[x]$ also.

(4) Let $\pi: M\longrightarrow M/G$ and $\eta:\overline{F}\longrightarrow \overline{F}/\rho$ be the canonical surjections. Show that the map $\lambda:\overline{F}/\rho\longrightarrow M/G$ defined by $\eta(x)\mapsto\pi(x)$, $x\in\overline{F}$ is a homeomorphism, provided $\overline{F}$ is normal.

Remark. The homeomorphism λ is actually a diffeomorphism. This may be seen by transporting the smooth structure of M/G to $\overline{F}/\rho$ by means of λ. We say that $\overline{F}/\rho$ is a geometric model of M/G.

(5) Show that the fundamental region $\overline{F}$ of each of the actions in Examples 1.7.10 and 1.7.11 is the infinite strip $\{(x,y) \in \mathbb{R}^2 \mid 0 \leq x \leq 1\}$, and of each of the actions in Examples 1.7.12 and 1.7.13 is the square $[0,1] \times [0,1]$, and they are normal. The respective quotient spaces $\overline{F}/\rho$ are obtained by identifying the points of the boundary of the strip or the square in the following way:
  (i) Cylinder. $(0,y) \equiv (1,y)$,
  (ii) Möbius band. $(0,y) \equiv (1,-y)$,
  (iii) Torus. $(x,0) \equiv (x,1)$ and $(0,y) \equiv (1,y)$,
  (iv) Klein bottle. $(x,0) \equiv (x,1)$ and $(0,y) \equiv (1,1-y)$.

Example 1.7.14. Lens spaces. Lens spaces $L(p,q)$, where p and q are relatively prime integers with $p > 0$, are quotient manifolds of S^3 modulo a properly discontinuous action of a group G defined in the following way. Identify $\mathbb{R}^4 \equiv \mathbb{C} \times \mathbb{C}$ so that

$$S^3 = \{(z_1, z_2) \in \mathbb{C} \times \mathbb{C} \mid |z_1|^2 + |z_2|^2 = 1\}.$$

Then G is the transformation group of S^3 generated by the transformation $(z_1, z_2) \mapsto (z_1\omega, z_2\omega^q)$, where ω is the pth root of unity $e^{2\pi i/p}$. Note that G is isomorphic to the group $\mathbb{Z}_p$, and that if $q \equiv q' \pmod{p}$, then $L(p,q) = L(p,q')$.

$\Diamond$ **Exercise 1.33.** Let G be any transformation group of M, and f be a diffeomorphism of M such that $f \notin G$. Then show that the set

$$G_1 = \{fgf^{-1} \mid g \in G\}$$

is a transformation group of M. Moreover, if M/G admits the structure of a quotient manifold, then M/G_1 admits a differentiable structure diffeomorphic to M/G.

$\Diamond$ **Exercise 1.34.** Show that the Lens spaces $L(p,q)$ and $L(p,q')$ are diffeomorphic if either $q + q' \equiv 0 \pmod{p}$ or $qq' \equiv 1 \pmod{p}$.

1.8. Manifolds with boundary and corner

We shall first extend the notion of manifolds so as to include manifolds with boundary. For example, the disk $D^n = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$ is a manifold with boundary which is the $(n-1)$-sphere

$$S^{n-1} = \{x \in \mathbb{R}^n \mid \|x\| = 1\}.$$

Let $\mathbb{R}^n_+$ and $\partial\mathbb{R}^n_+$ denote the subsets of $\mathbb{R}^n$ given by

$$\mathbb{R}^n_+ = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_1 \geq 0\}, \quad \partial\mathbb{R}^n_+ = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_1 = 0\}.$$

We call $\mathbb{R}^n_+$ the half space of $\mathbb{R}^n$, and $\partial\mathbb{R}^n_+$ the boundary of $\mathbb{R}^n_+$ (a more general definition says that a half space in $\mathbb{R}^n$ is an affine hyperplane, but we will not consider this). Note that we may identify $\partial\mathbb{R}^n_+$ with $\mathbb{R}^{n-1} \subset \mathbb{R}^n$.

Lemma 1.8.1. *Any linear isomorphism $\mathbb{R}^n \longrightarrow \mathbb{R}^n$, which maps $\partial\mathbb{R}^n_+$ onto itself, maps $\mathbb{R}^n_+$ onto itself.*

PROOF. The proof is obvious. Because, we may identify $\partial\mathbb{R}^n_+ \times \mathbb{R}$ with $\mathbb{R}^n$ by the linear isomorphism $\alpha(v_0, r) = v_0 + re_1$ (e_1 = unit vector along the first coordinate axis), so that $\mathbb{R}^n_+ = \alpha(\partial\mathbb{R}^n_+ \times \mathbb{R}_+)$. $\square$

If U is an open subset in $\mathbb{R}^n_+$, then its boundary ∂U is the subset $\partial U = U \cap \partial\mathbb{R}^n_+$, and its interior Int (U) is the subset Int $(U) = U - \partial U$. Thus Int (U) is open in $\mathbb{R}^n$, and ∂U is open in $\mathbb{R}^{n-1}$.

We may define smooth maps on open subsets of $\mathbb{R}^n_+$ by means of Definition 1.2.6. Thus a map $f : U \longrightarrow V$, where U is open in R^n_+ and V open in $\mathbb{R}^m_+$, is smooth if for each $x \in U$ there exist an open neighbourhood U_1 of x in $\mathbb{R}^n$, an open neighbourhood V_1 of $f(x)$ in $\mathbb{R}^m$, and a smooth map $f_1 : U_1 \longrightarrow V_1$ such that $f_1|U \cap U_1 = f|U \cap U_1$.

It may be mentioned that a result called Whitney's extension theorem says that a map f defined on the open set U of $\mathbb{R}^n_+$ given by $x_1 > 0$ extends to a smooth map g on $\mathbb{R}^n$ if and only if f and all its partial derivatives extend to continuous maps on $\mathbb{R}^n_+$. Whitney's proof which gives more general result than this one may be found in [**58**].

The notion of derivative of map also extends naturally. Consider a smooth map $f : U \longrightarrow \mathbb{R}^m$, where U is open in $\mathbb{R}^n_+$. Then, if $x \in$ Int (U), we already know what is df_x. If $x \in \partial U$, then, since f is smooth at x, f extends to a smooth map F in an open neighbourhood of x in $\mathbb{R}^n$. In this case, we define df_x to be the derivative map dF_x, which is a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$. The definition is independent of the choice of the extension F, that is, if F' is another local extension of f, then $dF'_x = dF_x$. To see this, note that if V and V' are the domains of F and F' respectively, and if $\{x_j\}$ is a sequence of points in $V \cap V' \cap$ Int (U) converging to x, then, since F and F' agree on $V \cap V' \cap$ Int (U), we have $dF_{x_j} = dF'_{x_j}$, as sequences in the vector space $L(\mathbb{R}^n, \mathbb{R}^m)$ of linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$. This implies, as $j \to \infty$, that $dF_x = dF'_x$, because the derivative maps $dF,\ dF' : V \cap V' \longrightarrow L(\mathbb{R}^n, \mathbb{R}^m)$ are continuous.

It follows that the definition of differentiability of $f : U \longrightarrow \mathbb{R}^m$ at a point $p \in U$ may be obtained from Definition 1.4.1, just by supposing U is an open subset of the half space $\mathbb{R}^n_+$ and keeping the other things the same. The derivative map $df_a : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ in the new situation will have the same properties (1)-(6) of Proposition 1.4.2.

$\Diamond$ **Exercises 1.35.** Show that

(1) if $f : U \longrightarrow \mathbb{R}^m$ is differentiable at $a \in U$, where U is open in $\mathbb{R}^n_+$, then

$$\begin{aligned} df_a(v) &= \lim_{t\to 0+} \frac{f(a+tv) - f(a)}{t} \text{ if } v \in \mathbb{R}^n_+ \\ &= \lim_{t\to 0-} \frac{f(a+tv) - f(a)}{t} \text{ if } -v \in \mathbb{R}^n_+. \end{aligned}$$

(2) if $f, g : U \longrightarrow \mathbb{R}^m$ are differentiable maps, where U is open in $\mathbb{R}^n$, such that f and g agree on $U \cap \mathbb{R}^n_+$, then $df_a = dg_a$ for $a \in U \cap \mathbb{R}^n_+$.

Lemma 1.8.2. *If $f : U \longrightarrow \mathbb{R}^m_+$ is differentiable, where U is open in $\mathbb{R}^n$, such that f maps $a \in U$ into $f(a) \in \partial\mathbb{R}^m_+$, then df_a maps $\mathbb{R}^n$ into $\partial\mathbb{R}^m_+$.*

Proof. Let $v \in \mathbb{R}^n$. Then

$$df_a(v) = \lim_{t\to 0} \frac{f(a+tv) - f(a)}{t}.$$

Therefore given an $\epsilon > 0$, there is a $\delta > 0$ such that if $a + tv \in U$, then

$$\left\| df_a(v) - \frac{f(a+tv) - f(a)}{t} \right\| < \epsilon$$

for all $t \in (-\delta, \delta)$, $t \neq 0$. For $t \in (-\delta, \delta)$, $t \neq 0$, let u_t denote the element

$$df_a(v) - \frac{f(a+tv) - f(a)}{t}.$$

Then

$$tu_t = t df_a(v) - f(a+tv) + f(a).$$

Let $F_1[v]$ denote the first coordinate of the vector v. Then, since $-f(a) \in \partial\mathbb{R}^m_+ \subset \mathbb{R}^m_+$, and $f(a+tv) \in \mathbb{R}^m_+$, we have

$$t \cdot F_1[df_a(v) - u_t] = F_1[f(a+tv) - f(a)] \geq 0.$$

Therefore, if $0 < t < \delta$, then

$$F_1[df_a(v)] \geq F_1[u_t] > -\epsilon,$$

and if $-\delta < t < 0$, then

$$F_1[df_a(v)] \leq F_1[u_t] < \epsilon.$$

Letting $\epsilon \to 0$, we get $F_1[df_a(v)] = 0$. Therefore $df_a(v) \in \partial\mathbb{R}^m_+$. ☐

Theorem 1.8.3. (**Invariance of interior and boundary**)
Let $f : U \longrightarrow V$ be a diffeomorphism, where U and V are open subsets of $\mathbb{R}^n_+$. Then

(a) $x \notin \partial U \Leftrightarrow f(x) \notin \partial V$,

(b) $f \mid \text{Int } (U)$, and $f \mid \partial U$ are diffeomorphisms.

Proof. The derivative map $df_a : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is an isomorphism for each $a \in U$, by Proposition 1.4.2. Therefore, by the preceding lemma, no interior point of U can be mapped onto a boundary point of V, and conversely. Thus f induces bijections $\text{Int } U \longrightarrow \text{Int } V$ and $\partial U \longrightarrow \partial V$. These are actually diffeomorphisms, because the restriction of f to any subset of U is always a smooth map. ☐

Definition 1.8.4. A second countable Hausdorff space M is called a smooth n**-manifold with boundary** if is satisfies all the conditions of a smooth manifold, with the exception that now we allow coordinate neighbourhoods to map onto open subsets in $\mathbb{R}^n_+$.

If $\phi : U \longrightarrow V \subset \mathbb{R}^n_+$ is such a coordinate chart, where U is open in M and V is open in $\mathbb{R}^n_+$, then a point of $\phi^{-1}(\partial V)$ is called a boundary point for the chart (U, ϕ). The definition does not depend on the chart. For, if (U, ϕ) and (V, ψ) are two coordinate charts around $x \in M$ with $\phi(x) \in \partial\mathbb{R}^n_+$ and $\psi(x) \in \text{Int } (\mathbb{R}^n_+)$, then the diffeomorphism $\psi \circ \phi^{-1}$ will map a boundary point of $\mathbb{R}^n_+$ onto an interior point of $\mathbb{R}^n_+$. This is not possible by the invariance of boundary as described in Theorem 1.8.3. The collection of all boundary points is the **boundary** of M, which is denoted by ∂M.

Theorem 1.8.5. *The boundary ∂M of an n-manifold M is a manifold of dimension $n-1$, and ∂M has no boundary.*

PROOF. We have already seen that if x is a boundary point with respect to one coordinate system, then it remains a boundary point relative to any other coordinate system. If $\phi : U \longrightarrow V \subset \mathbb{R}^n_+$ is a coordinate chart in M, then $\phi^{-1}(\partial V) = U \cap \partial M$ is an open set in ∂M, and $(U \cap \partial M, \lambda \circ \phi)$ is a coordinate chart for ∂M, where $\lambda : \partial R^n_+ \longrightarrow \mathbb{R}^{n-1}$ is a linear isomorphism. The collection of all such charts is a smooth atlas on ∂M. Thus the boundary ∂M is a manifold of dimension $n-1$. $\square$

The interior of M is the set $\text{Int } M = M - \partial M$. It is a manifold of the same dimension as M, and it has no boundary. A manifold is called a **closed manifold** if it is compact and has no boundary.

$\Diamond$ **Exercise 1.36.** Show that if $f : M \longrightarrow N$ is a diffeomorphism, then $f(\partial M) = \partial N$ and $f(\text{Int } M) = \text{Int } N$.

The notion of submanifold can also be extended.

Definition 1.8.6. An m-submanifold N of an n-manifold M with boundary satisfies the same conditions as when M is without boundary, except that, for every coordinate chart (U, ϕ), $\phi : U \longrightarrow \mathbb{R}^n_+$, $\phi^{-1}(\mathbb{R}^m_+) = U \cap N$, where $\mathbb{R}^m_+$ is the subspace of the first m coordinates in $\mathbb{R}^n_+$.

A map on a manifold with boundary is smooth, if it is locally extendable to a smooth map. The concepts of rank, immersion, submersion, embedding, and diffeomorphism remain exactly the same as before. However, there are two kinds of submanifolds N of M arising from two kinds of embeddings, namely, embeddings of a manifold into a manifold without boundary, or embeddings of a manifold into a manifold with boundary. Consider, for example, a closed interval I embedded in $\mathbb{R}^n_+$; I may lie entirely in $\text{Int } (\mathbb{R}^n_+)$, or I may have a boundary point in $\partial\mathbb{R}^n_+$. The two cases are essentially distinct, although Proposition 1.5.3 holds for each of them. For example, given two submanifolds of $\mathbb{R}^n_+$ of the first kind, there exists a diffeomorphism of $\mathbb{R}^n_+$ carrying one to the other, but there cannot exist a diffeomorphism of $\mathbb{R}^n_+$ carrying a submanifold of the first kind into one of the second kind (why?).

In general, there is no relation between ∂N and ∂M, when N is a submanifold of M. We define a special kind of submanifold N whose boundary is nicely placed in the ambient manifold M.

Definition 1.8.7. An m-submanifold N of an n-manifold M is a **neat submanifold** of M if N is a closed subset of M, and if in a neighbourhood of a point of N the pair (M, N) is locally like $(\mathbb{R}^n, \mathbb{R}^m)$ or $(\mathbb{R}^n_+, \mathbb{R}^m_+)$, that is, if

(a) each point $p \in N$ has a chart (U, ϕ) at p in M, where $\phi : U \longrightarrow \mathbb{R}^n_+$, such that $\phi^{-1}(\mathbb{R}^m_+) = U \cap N$,
(b) each point $p \in \partial N$ has a chart (U, ϕ) at p in M, where $\phi : U \longrightarrow \mathbb{R}^n_+$, such that $\phi^{-1}(\partial\mathbb{R}^m_+) = U \cap \partial N$,

The definition implies that N meets ∂M in the same way as $\mathbb{R}^m_+$ meets $\partial\mathbb{R}^n_+$. Indeed, $\partial\mathbb{R}^m_+ = \mathbb{R}^m_+ \cap \partial\mathbb{R}^n_+$ implies $\partial N = N \cap \partial M$. In particular, if $\partial N = \emptyset$, then N is disjoint from ∂M, and so N is a submanifold of Int M. Note that a curve with end points in a manifold with boundary is not a neat submanifold of M unless its end points lie in ∂M.

$\diamondsuit$ **Exercise 1.37.** Show that a closed subset A of an n-manifold M is a neat submanifold of dimension m if and only if at each point $p \in A$ there is a chart (U, ϕ) in M and a submersion $f : U \longrightarrow \mathbb{R}^{n-m}$ such that f is also a submersion on $U \cap \partial M$, and $f^{-1}(0) = U \cap A$.

$\diamondsuit$ **Exercise 1.38.** Extend Definition 1.5.5 of regular value of a smooth map

$$f : M \longrightarrow N$$

as follows. A point $q \in N$ is a regular value of f if (1) f is a submersion at every point $p \in f^{-1}(q)$, and (2) $f \mid \partial M$ is a submersion at every point $p \in f^{-1}(q) \cap \partial M$. If $p \in \text{Int } M$, then the condition (2) does not arise, and if $p \in \partial M$, then condition (1) is redundant, as it follows from the condition (2).

Show that if q is a regular value of f, then $f^{-1}(q)$ is a neat submanifold of M.

We next consider manifolds with corners. In the category of manifolds with boundary, the product does not exist. For example, the square $I \times I$, where I denotes the unit interval $0 \leq t \leq 1$, is not a manifold with boundary. The definition breaks down at the four vertices of the square. In order to obtain products in the category, we are forced to introduce manifolds with corners. Let $\mathbb{R}^n_{++}$ be the subspace of $\mathbb{R}^n$ consisting of points whose first two coordinates are non-negative

$$\mathbb{R}^n_{++} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n | x_1 \geq 0,\ x_2 \geq 0\}.$$

The boundary $\partial\mathbb{R}^n_{++}$ of $\mathbb{R}^n_{++}$ consists of points for which $x_1x_2 = 0$, and the interior Int $\mathbb{R}^n_{++}$ consists of points for which $x_1x_2 > 0$. The subset $\Lambda(\mathbb{R}^n_{++})$ of points for which $x_1 = x_2 = 0$ is called the corner of $\mathbb{R}^n_{++}$. We may identify $\Lambda(\mathbb{R}^n_{++})$ with the subspace $\mathbb{R}^{n-2}$ of $\mathbb{R}^n$.

If U is open in $\mathbb{R}^n_{++}$, we define

$$\text{Int }(U) = U \cap \text{Int }\mathbb{R}^n_{++},\ \partial U = U \cap \partial\mathbb{R}^n_{++},\ \Lambda(U) = U \cap \Lambda\mathbb{R}^n_{++}.$$

Then $U = \text{Int }(U) \cup \partial U$, $\text{Int }(U) \cap \partial U = \emptyset$.

We may define a smooth map $U \longrightarrow \mathbb{R}^m_{++}$, where U is open in $\mathbb{R}^n_{++}$, and its derivative $df_a : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, $a \in U$, as before. The derivative satisfies the properties (1)-(6) of Proposition 1.4.2. The following exercises are analogous to Lemma 1.8.2 and Theorem 1.8.3.

Let, for $i = 1, 2$, $\mathbb{R}^n_{i+} = \{x \in \mathbb{R}^n \mid x_i \geq 0\}$. Then

$$\mathbb{R}^n_{++} = \mathbb{R}^n_{1+} \cap \mathbb{R}^n_{2+}, \quad \partial\mathbb{R}^n_{++} \subseteq \partial\mathbb{R}^n_{1+} \cup \partial\mathbb{R}^n_{2+}, \quad \Lambda(\mathbb{R}^n_{++}) = \partial\mathbb{R}^n_{1+} \cap \partial\mathbb{R}^n_{2+}.$$

$\Diamond$ **Exercise 1.39.** Let $f : U \longrightarrow \mathbb{R}^m_{++}$ be a smooth map, where U is open in $\mathbb{R}^n_{++}$, and let $x \in U$. Show that

(i) if $f(a) \in \partial\mathbb{R}^n_{i+}$, then $df_a(\mathbb{R}^n) \subset \partial\mathbb{R}^n_{i+}$,

(ii) if $f(a) \in \Lambda(\mathbb{R}^n_{++})$, then $df_a(\mathbb{R}^n) \subset \Lambda(\mathbb{R}^n_{++})$.

$\Diamond$ **Exercise 1.40.** If $f : U \longrightarrow V$ is a diffeomorphism, where U, V are open subsets of $\mathbb{R}^n_{++}$, then show that

(i) $x \in \text{Int }(U) \Leftrightarrow f(x) \in \text{Int }(V)$,

(ii) $x \in \partial U \Leftrightarrow f(x) \in \partial V$,

(iii) $x \in \Lambda(U) \Leftrightarrow f(x) \in \Lambda(V)$.

Definition 1.8.8. A topological n-manifold M is a **manifold with corner** if it satisfies all the defining conditions of a smooth manifold, except that the coordinate neighbourhoods may now map onto open subsets of $\mathbb{R}^n_{++}$.

Points corresponding to $\partial\mathbb{R}^n_{++}$ constitute the boundary ∂M of M. Points corresponding to the corner $\Lambda(\mathbb{R}^n_{++})$ constitute the corner $\Lambda(M)$ of M. Clearly $\Lambda(M)$ is a smooth manifold of dimension $n-2$.

$\Diamond$ **Exercise 1.41.** Show that

(i) Int (M) is an open dense subset of M,

(ii) $\Lambda(M)$ is a manifold without boundary,

(iii) ∂M is a manifold without boundary, if $\Lambda(M) = \emptyset$.

For example, if M_1 and M_2 are manifolds with boundary, then the product $M_1 \times M_2$ is a manifold with corner, where

$$\partial(M_1 \times M_2) = \partial M_1 \times M_2 \cup M_1 \times \partial M_2, \quad \text{and} \quad \Lambda(M_1 \times M_2) = \partial M_1 \times \partial M_2.$$

Here the boundary $\partial(M_1 \times M_2)$ is not a manifold, but the corner $\Lambda(M)$ is.

Remark 1.8.9. Infinite dimensional manifolds Although it will not concern us, it is worth noting that all the notions developed so far will hold for more general kind of manifolds. These are manifolds whose coordinate charts are modelled on an infinite dimensional vector space, such as Hilbert space or Banach space (just as a finite dimensional manifold is modelled on $\mathbb{R}^n$). These are called infinite dimensional manifolds, in particular **Hilbert manifolds** or **Banach manifolds** if the vector space happens to be a Hilbert space or a Banach space. Banach manifolds arise (under certain conditions) as space of continuous (or C^r) maps between two finite dimensional manifolds. It may be noted that there are results of infinite dimensional manifolds which strongly contrast with the case of finite dimensional manifolds. For example, a paracompact manifold modelled on the Banach space ℓ^2, which is the vector space of sequences of real numbers $(a_1, a_2, \cdots)$ with $\sum_i a_i^2$ convergent, can be embedded in ℓ^2 (see Eells and Elworthy [**9**]). This is often unexpected in finite dimensional case : S^n cannot be embedded in $\mathbb{R}^n$. Again, a result of Kuiper [**22**] says that the infinite dimensional general linear group $GL(\ell^2)$ acting on ℓ^2 is a contractible space. On the other hand, the ordinary general linear group $GL(n, \mathbb{R})$ has non-trivial homotopy and homology groups.

CHAPTER 2

APPROXIMATION THEOREMS AND WHITNEY'S EMBEDDING

Perhaps the most important property of a manifold which opens up various developments of manifold theory is that a manifold can be embedded in a Euclidean space as a closed subspace. This is called Whitney's embedding theorem. Thus any manifold may be considered as a submanifold of a Euclidean space. Originally manifolds were introduced in this way, and Whitney's theorem reconciles this earlier concept of manifold with its modern abstract definition. We will prove the theorem by way of certain approximations. If we list the hierarchy of maps between manifolds as continuous, smooth, immersion, injective immersion, and embedding, then the approximation theorems say that each map in one of the above classes is approximable by a map in its immediate successor. Of course for a real analytic manifold, one can go one step further and prove that there is a real analytic embedding of a manifold in Euclidean space, but its proof lies much deeper and will not be considered here. Our approach to these approximation theorems will involve two important topological aspects of manifold, namely smooth partition of unity, and Sard's theorem. The first helps to construct smooth maps by piecing together local information, and the second leads to approximations.

In this chapter a manifold will always be a smooth manifold, and it may have boundary, unless it is stated explicitly otherwise.

2.1. Smooth partition unity

Recall that a covering $\mathcal{U}$ of a topological space X is called **locally finite** if each point of X has an open neighbourhood which intersects only finitely many members of $\mathcal{U}$. Another covering $\mathcal{V}$ of X is called a **refinement** of $\mathcal{U}$ if each member of $\mathcal{V}$ is contained in some member of $\mathcal{U}$. A Hausdorff space X is **paracompact** if every open covering of X admits an open locally finite refinement.

Theorem 2.1.1. *Every manifold M is paracompact.*

Proof. Since M is locally homeomorphic to either $\mathbb{R}^n$ or $\mathbb{R}^n_+$, it is locally compact (each of its points has a compact neighbourhood). Then each point $x \in M$ has an open neighbourhood V whose closure is compact. For, if U is

A. Mukherjee, *Differential Topology*, DOI 10.1007/978-3-319-19045-7_2

an arbitrary open neighbourhood of x, and K a compact neighbourhood of x, then $V = U \cap \text{Int } K \subset K$, and so $\overline{V}$ is compact, being a closed subset of a compact set. It follows then, since M is second countable, that M admits a countable basis $\{V_j\}$ such that each $\overline{V}_j$ is compact.

Then, there is an increasing sequence $K_1 \subset K_2 \subset \cdots \subset K_j \subset \cdots$ of compact subsets whose union is M such that $K_j \subset \text{Int } K_{j+1}$, for each j. Indeed, we may take $K_1 = \overline{V}_1$, and, assuming inductively that K_j has been defined, if m is the smallest integer $> j$ such that $K_j \subset V_1 \cup \cdots \cup V_m$, then we may take

$$K_{j+1} = \overline{V}_1 \cup \cdots \cup \overline{V}_m = \overline{V_1 \cup \cdots \cup V_m}.$$

Now let $\mathcal{U} = \{U_\alpha\}$ be an open covering of M. Choose a locally finite refinement $\mathcal{V}$ as follows. Let $K_{-1} = K_0 = \emptyset$, and, for each $j \geq 0$, consider open sets

$$(\text{Int } K_{j+2} - K_{j-1}) \cap U_\alpha, \quad U_\alpha \in \mathcal{U}.$$

These open sets cover the compact set $K_{j+1} - \text{Int } K_j$. Therefore, we can find a finite subcover $\mathcal{V}_j = \{V_1^j, \ldots, V_{\alpha(j)}^j\}$ ($\alpha(j)$ an integer). The collection $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2 \cup \cdots$ covers M, since the sets $K_{j+1} - \text{Int } K_j$ cover M. The covering $\mathcal{V}$ is a refinement of $\mathcal{U}$ (each V_i^j is contained in some U_α). It is locally finite, because if $x \in K_j$, then $\text{Int } K_{j+1}$ is a neighbourhood of x which intersects no member of $\mathcal{V}_k$ for $k > j + 1$. □

Remark 2.1.2. Actually we have constructed a locally finite refinement $\mathcal{V}$ which is countable.

Lemma 2.1.3. *Any open covering $\{U_\alpha\}$ of a manifold M has a locally finite countable refinement by coordinate neighbourhoods, each of which has compact closure.*

Proof. The proof follows the same line of arguments as that of the above theorem. One has only to choose the covering of each compact set $K_{j+1} - \text{Int } K_j$ suitably. Each point x of the open set

$$(\text{Int } K_{j+2} - K_{j-1}) \cap U_\alpha$$

has a coordinate neighbourhood V_x and a homeomorphism ϕ_x of V_x into $\mathbb{R}^n$ or $\mathbb{R}^n_+$ such that $V_x \subset (\text{Int } K_{j+2} - K_{j-1}) \cap U_\alpha$, and $\phi_x(V_x)$ contains a closed n-ball B^n with centre at $\phi_x(x)$, $n = \dim M$. Let $W_x = \phi_x^{-1}(\text{Int } B^n)$. With this choice, $\overline{W}_x$ will be compact. We may find a finite number of the W_x which cover $K_{j+1} - \text{Int } K_j$, and then proceed as before. □

A covering $\{V_i\}$ is called a **shrinking** of a covering $\{U_i\}$ if each $\overline{V}_i \subset U_i$.

Lemma 2.1.4 (Shrinking lemma). *Let $\mathcal{U} = \{U_i\}_{i \geq 1}$ be a countable locally finite open covering of M. Then there is another open covering $\{V_i\}$ of M such that $\overline{V}_i \subset U_i$ for every $i \geq 1$.*

Proof. We may assume that M is connected. Now write $\mathcal{U}_k = \cup_{i\geq k} U_i$. and construct the open sets V_i inductively in the following way.

The closed set $A_1 = U_1 - \mathcal{U}_2$ is contained in U_1, and so $M = A_1 \cup \mathcal{U}_2$. Since M is paracompact, it is normal, and therefore we may choose an open set V_1 such that $A_1 \subset V_1 \subset \overline{V}_1 \subset U_1$. Then $M = V_1 \cup \mathcal{U}_2$. Next, suppose that the open sets $V_1, \ldots, V_{k-1}$ have been chosen so that $\overline{V}_i \subset U_i$ for $i = 1, \ldots, k-1$, and $M = V_1 \cup \cdots \cup V_{k-1} \cup \mathcal{U}_k$. Then the closed set $A_k = U_k - (V_1 \cup \cdots \cup V_{k-1} \cup \mathcal{U}_{k+1})$ is contained in U_k, and we have

$$M = V_1 \cup \cdots \cup V_{k-1} \cup A_k \cup \mathcal{U}_{k+1}.$$

Choose an open set V_k such that $A_k \subset V_k \subset \overline{V}_k \subset U_k$. Then we have

$$M = V_1 \cup \cdots \cup V_k \cup \mathcal{U}_{k+1}.$$

To see that the collection $\{V_i\}$ is a covering of M, take any point $x \in M$. Then, since the covering $\mathcal{U}$ is locally finite, there is a largest m such that $x \notin U_k$ for $k \geq m$, that is, $x \notin \mathcal{U}_m$. Since $M = V_1 \cup \cdots \cup V_{m-1} \cup \mathcal{U}_m$, it follows that $x \in V_1 \cup \cdots \cup V_{m-1}$. This completes the proof. □

Theorem 2.1.5. *Every manifold is metrisable.*

Proof. Since every paracompact space is normal and a manifold is second countable, the proof follows trivially from Urysohn's metrisation theorem ([**15**], Theorem 2-46, p. 68). This theorem states that every second-countable normal space can be embedded topologically in infinite dimensional Hilbert coordinate space. This metric has nothing to do with the locally Euclidean structure on the manifold.

A natural metric on a manifold may be obtained from Smirnov's theorem ([**15**], Theorem 2-68, p. 81). This theorem states that a space is metrisable if and only if it is paracompact, and locally metrisable. Its proof uses paracompactness to pass from the local information to the global one.

Locally the topology of M is the same as the topology of $\mathbb{R}^n$, and therefore it is given by the standard metric in $\mathbb{R}^n$. If (U, ϕ) is a coordinate chart about a point $p \in M$ with coordinates $(x_1, \ldots, x_n)$ such that $\phi(U)$ is a convex set in $\mathbb{R}^n$, then we have a metric ρ on U such that for $x \in U$

$$\rho(x, p) = [(x_1 - x_1(p))^2 + \cdots + (x_n - x_n(p))^2]^{\frac{1}{2}},$$

or equivalently,

$$\rho(x, p) = \max\{|x_i - x_i(p)|\}.$$

□

We shall have occasions to use a bump function whose definition is as follows.

Definition 2.1.6. A **bump function** is a smooth function $\mathcal{B} : \mathbb{R} \longrightarrow \mathbb{R}$ such that

$$\mathcal{B}(x) = 0 \text{ if } x \leq 0,\ 0 < \mathcal{B}(x) < 1 \text{ if } 0 < x < 1,\ \mathcal{B}(x) = 1 \text{ if } x \geq 1.$$

To construct a bump function $\mathcal{B}$, first define

$$\begin{aligned} \psi(x) &= \exp\left(\frac{1}{x(x-1)}\right) \text{ if } 0 < x < 1 \\ &= 0 \text{ otherwise.} \end{aligned}$$

Then ψ is smooth, non-negative, and non-vanishing when $0 < x < 1$. Now define $\mathcal{B}$ by

$$\mathcal{B}(x) = \frac{\int_0^x \psi(t)dt}{\int_0^1 \psi(t)dt}.$$

Definition 2.1.7. The **support** of a function $f : M \longrightarrow \mathbb{R}$, denoted by $\mathrm{supp} f$, is the closure of the set of points of M where f is non-zero.

Lemma 2.1.8. *If $K \subset U \subset M$, where K is compact and U is open, then there is a smooth function $\mu : M \longrightarrow [0, \infty)$ such that $\mu(x) > 0$ if $x \in K$, and* $\mathrm{supp}\ \mu \subset U$.

PROOF. Define a smooth function $\alpha : \mathbb{R} \longrightarrow \mathbb{R}$ by $\alpha(x) = \mathcal{B}(1 - |x|)$. Then $\alpha(x) > 0$ if $|x| < 1$, and $\alpha(x) = 0$ if $|x| \geq 1$. Now construct for each $p \in K \subset U$ a smooth function $\mu_p : M \longrightarrow [0, \infty)$ such that $\mu_p(p) > 0$ and $\mathrm{supp}\ \mu_p \subset U$ in the following way. Choose local coordinates $(x_1, \ldots, x_n)$ about p with p corresponding to the origin such that $B_r = \{(x_1, \ldots, x_n) \mid |x_i| < r\} \subset U$, for a suitable $r > 0$. Define μ_p by

$$\begin{aligned} \mu_p(x) &= \alpha\Big(\frac{x_1}{r}\Big) \cdots \alpha\Big(\frac{x_n}{r}\Big) \text{ if } x \in B_r \\ &= 0 \text{ otherwise.} \end{aligned}$$

As p runs over K the open sets $\{x \in M \mid \mu_p(x) > 0\}$ cover K. By compactness, a finite number of them still cover K. Then the sum of the corresponding finite number of functions μ_p is the required function μ. $\square$

Definition 2.1.9. Let M be a manifold with an open covering $\mathcal{U} = \{U_i\}_{i \in A}$. Then a **smooth partition of unity** subordinate to $\mathcal{U}$ is a family of smooth functions $\{\lambda_i : M \longrightarrow \mathbb{R}\}_{i \in A}$ satisfying the following conditions.

(i) $\mathrm{supp}\lambda_i \subset U_i$ for all $i \in A$,
(ii) $0 \leq \lambda_i(x) \leq 1$ for all $x \in M$ and $i \in A$,
(iii) each $x \in M$ has a neighbourhood on which all but finitely many functions λ_i are identically zero,
(iv) $\sum_{i \in A} \lambda_i(x) = 1$ for all $x \in M$ (note that the sum is always finite by (iii)).

Lemma 2.1.10. *Let $\mathcal{U} = \{U_i\}_{i\in A}$ and $\mathcal{V} = \{V_j\}_{j\in B}$ be two open coverings of M such that $\mathcal{U}$ refines $\mathcal{V}$. Then, if $\mathcal{U}$ has a subordinate smooth partition of unity, so has $\mathcal{V}$.*

PROOF. Let $\{\lambda_i\}_{i\in A}$ be a smooth partition of unity subordinate to $\mathcal{U}$. Let $f : A \longrightarrow B$ be a map of the index sets so that $U_i \subset V_{f(i)}$, $i \in A$. Define $\mu_j : M \longrightarrow \mathbb{R}$ by

$$\mu_j(x) = \sum_{i\in f^{-1}(j)} \lambda_i(x).$$

It is easily checked that the conditions (i)–(iv) hold for the family $\{\mu_j\}$, when U_i are replaced by V_j. □

Remark 2.1.11. Some people call $\{\lambda_i\}$ a partition of unity subordinate to $\mathcal{V}$, In this case the condition (i) has to be replaced by the following condition:

"for every $i \in A$ there is a $j \in B$ such that supp $\lambda_i \subset V_j$".

Theorem 2.1.12. *Any manifold M with an open covering $\{U_i\}$ admits a smooth partition of unity subordinate to $\{U_i\}$.*

PROOF. We may assume that the given covering $\{U_i\}$ is countable and locally finite such that each of its members U_i is a coordinate neighbourhood with compact closure (Lemma 2.1.3). We may find another open covering $\{V_i\}$ of M such that $\overline{V}_i \subset U_i$ (Lemma 2.1.4). Now construct smooth functions $\mu_i : M \longrightarrow \mathbb{R}$ as described in Lemma 2.1.8 such that $\mu_i > 0$ on $\overline{V}_i$ and supp $\mu_i \subset U_i$. Then the function $\sum_i \mu_i$ is a well-defined positive smooth function, and the family of functions $\lambda_i = \mu_i / \sum_i \mu_i$ is the required smooth partition of unity. □

Lemma 2.1.13. *If $\{\lambda_i : U_i \longrightarrow \mathbb{R}\}$ is a smooth partition of unity on M, and $\{f_i : U_i \longrightarrow \mathbb{R}\}$ is a family of smooth functions, then the function $f : M \longrightarrow \mathbb{R}$ defined by $f(x) = \sum_i \lambda_i(x) f_i(x)$ is smooth.*

PROOF. Since the function $\lambda_i f_i$ is smooth on U_i and vanishes on a neighbourhood of $M - U_i$, it can be extended over M using the zero function on $M - U_i$. Therefore the sum $f = \sum_i \lambda_i f_i$ is smooth. □

Lemma 2.1.14. *For a map $f : U \longrightarrow \mathbb{R}^m$, where U is open in $\mathbb{R}^n_+$, the following conditions are equivalent.*

(a) *f is smooth, as defined in Definition* 1.2.6 *using local extendability condition,*

(b) *there is an open set V in $\mathbb{R}^n$ and a smooth map $F : V \longrightarrow \mathbb{R}^m$ such that $V \cap \mathbb{R}^n_+ = U$ and $F|U = f$.*

PROOF. The part (b) $\Rightarrow$ (a) is trivial. Next, assume (a). Then, for each $x \in U$, there is an open neighbourhood V_x of x in $\mathbb{R}^n$ and a smooth map

$F_x : V_x \longrightarrow \mathbb{R}^m$ such that $f = F_x$ on $U \cap V_x$. Let $W = \cup_{x \in U} V_x$. Then W is open in $\mathbb{R}^n$, and $U \subset W$.

The manifold W admits a partition of unity $\{\lambda_x\}$ subordinate to the covering $\{V_x\}$. Then $G = \sum_{x \in U} \lambda_x F_x$ is a smooth map from W to $\mathbb{R}^m$, by Lemma 2.1.13. On the other hand, there exists an open set V' of $\mathbb{R}^n$ such that $U = V' \cap \mathbb{R}^n_+$. Then taking $V = W \cap V'$ and $F = G|V$, we get the condition (b). □

Lemma 2.1.15. *If $f : M \longrightarrow \mathbb{R}$ is a positive continuous function, then there is a smooth function $g : M \longrightarrow \mathbb{R}$ such that*

$$0 < g(x) < f(x) \text{ for all } x \in M.$$

When M is compact, g may be taken to be a constant function.

Proof. As in the proof of Theorem 2.1.12, consider a locally finite covering $\{U_i\}$ of M, and another open covering $\{V_i\}$ such that $\overline{V}_i$ is compact and $\overline{V}_i \subset U_i$. Take a smooth partition of unity $\{\lambda_i\}$ such that $\lambda_i > 0$ on $\overline{V}_i$ and supp $\lambda_i \subset U_i$. Choose $\delta_i > 0$ smaller than the infimum of f on the compact set $\overline{V}_i$, and define $g : M \longrightarrow \mathbb{R}$ by $g(x) = \sum_i \delta_i \lambda_i(x)$. Then g is smooth. Since the sum $\sum_i \lambda_i(x)$ is finite and equal to 1, and the maximum of the corresponding δ_i is less than $f(x)$, we have $g(x) < f(x)$. Also $g(x) > 0$, since all δ_i are so. □

Lemma 2.1.16. *If K is a closed subset of a manifold M, and $f : K \longrightarrow \mathbb{R}$ is a smooth function, then f extends to a smooth function $F : M \longrightarrow \mathbb{R}$.*

Proof. In view of Definition 1.2.6, cover K by open sets U_i such that there exist smooth functions $g_i : U_i \longrightarrow \mathbb{R}$ with $g_i = f$ on $K \cap U_i$. The sets U_i and $M - K$ form an open covering of M. Let $\{\lambda_i\}$ be a smooth partition of unity subordinate to this covering. Then the smooth extension F of f is given by

$$F(x) = \begin{cases} \sum_i \lambda_i(x) g_i(x), & \text{if } x \notin M - K, \\ 0, & \text{otherwise.} \end{cases}$$

□

Lemma 2.1.17 (Smooth Urysohn's lemma). *If $K \subset U \subset M$, K closed, U open, then there is a smooth function $f : M \longrightarrow \mathbb{R}$ such that $0 \leq f \leq 1$, $f|K = 1$ and supp $f \subset U$.*

Proof. The open sets $U_1 = U$ and $U_2 = M - K$ form a covering of M. Let $\lambda_1 : U_1 \longrightarrow \mathbb{R}$ and $\lambda_2 : U_2 \longrightarrow \mathbb{R}$ be a smooth partition of unity subordinate to this covering. Then λ_1 extended over M by the zero function outside U_1 is a solution f of the problem. □

Theorem 2.1.18 (Whitney's weak embedding theorem). *If M is a compact n-manifold, then there is an embedding $f : M \longrightarrow \mathbb{R}^m$, where $m = r(n+1)$ for some integer $r > 0$.*

PROOF. Find a finite covering of M by coordinate charts (U_i, ϕ_i), $i = 1, \ldots, r$, and open sets V_i also covering M such that $\overline{V_i} \subset U_i$ for all i. By Lemma 2.1.17, there are C^∞ functions $\lambda_i : M \longrightarrow \mathbb{R}$ such that $\lambda_i | \overline{V_i} = 1$ and $\operatorname{supp} \lambda_i \subset U_i$. Let $\psi_i : M \longrightarrow \mathbb{R}^n$ be C^∞ maps given by

$$\begin{aligned} \psi_i(p) &= \lambda_i(p)\phi_i(p) \text{ if } p \in U_i, \\ &= 0 \text{ otherwise.} \end{aligned}$$

Define $f : M \longrightarrow \mathbb{R}^{r(n+1)}$ by

$$f(p) = (\psi_1(p), \ldots, \psi_r(p), \lambda_1(p), \ldots, \lambda_r(p)), p \in M.$$

Then the Jacobian matrix $J(f)$ has rank n at every point $p \in M$. Because, $J(\phi_i)$ has rank n, and if $p \in M$, then

$$(p \in V_i \text{ for some } i) \Rightarrow (\lambda_i(p) = 1) \Rightarrow (\phi_i = \psi_i \text{ on } \overline{V}_i) \Rightarrow (J(\phi_i) = J(\psi_i) \text{ at } p),$$

so $J(\psi_i)$, and hence $J(f)$, has rank n at p.

Also f is injective. Because, if $f(p) = f(q)$ then $\psi_i(p) = \psi_i(q)$ and $\lambda_i(p) = \lambda_i(q)$ for all i, and if $p \in V_j$ then $\lambda_j(p) = \lambda_j(q) = 1$, and so $\phi_j(p) = \phi_j(q)$, which implies $p = q$, ϕ_j being injective.

Thus f is an injective immersion. Since M is compact and $f(M)$ is Hausdorff, f is a homeomorphism onto its image, and hence it is an embedding. □

We remark that the theorem is unsatisfactory in that the value of m, which is the dimension of the Euclidean space, depends on the number r of coordinate neighbourhoods required to cover M, and therefore m may be much larger than we would like it. We will prove in Theorem 2.5.1 a stronger version of this theorem which removes the restriction of compactness and gives a much lower value for m, not depending on r.

◊ **Exercise 2.1.** Show that the projective space $\mathbb{R}P^n$ can be embedded in $\mathbb{R}^{(n+1)^2}$ by an embedding f where $f([x_0, \ldots, x_n])$ is a vector in $\mathbb{R}^{(n+1)^2}$ whose (i,j)-th coordinate in lexicographic order is $a_{ij} = x_i x_j / \sum_k x_k^2$, $i, j, k = 0, 1, \ldots, n$. The image of f consists of symmetric square matrices A of order $n+1$ such that $A \cdot A = A$ and trace of A is 1.

In a similar way, the complex projective space $\mathbb{C}P^n$ embeds in $\mathbb{C}^{(n+1)^2}$.

Proposition 2.1.19. *Any metric on a manifold M compatible with the topology of M can be turned into a complete metric giving the same topology of M.*

PROOF. The proof will be clear from the next two lemmas. □

Recall that a continuous map between topological spaces is **proper** if the inverse image of any compact set is compact.

Lemma 2.1.20. *If X is a metric space with metric ρ and $f : X \longrightarrow \mathbb{R}$ is a continuous proper map, then the map $\rho' : X \times X \longrightarrow \mathbb{R}$ given by*

$$\rho'(x, y) = \rho(x, y) + |f(x) - f(y)|, x, y \in X$$

is a complete metric on X which is compatible with the topology of X.

PROOF. Clearly ρ' is a metric. Let $\mathcal{T}$ and $\mathcal{T}'$ be the topologies on X induced by the metrics ρ and ρ' respectively. Then, since ρ' is continuous with respect to $\mathcal{T}$, we have $\mathcal{T}' \subset \mathcal{T}$. Conversely, take an open set U in $\mathcal{T}$ and a point x in U. Then we can find an $\epsilon > 0$ so that

$$B'(x, \epsilon) = \{y \in X \mid \rho'(x, y) < \epsilon\} \subset B(x, \epsilon) = \{y \in X \mid \rho(x, y) < \epsilon\} \subset U.$$

This means U is in $\mathcal{T}'$, and $\mathcal{T} = \mathcal{T}'$. Next, to see that ρ' is complete, take a Cauchy sequence $\{x_n\}$ in X with respect to the metric ρ'. Then there is a number $m > 0$ such that $\rho'(x_1, x_n) < m$ for all $n \geq 1$. Therefore

$$|f(x_1) - f(x_n)| < m$$

for all $n \geq 1$, and so

$$x_n \in f^{-1}([f(x_1) - m, f(x_1) + m]).$$

Since f is proper, the above set is compact, and the sequence $\{x_n\}$ converges to a limit in X. □

Lemma 2.1.21. *On a manifold M there always exists a proper smooth function $f : M \longrightarrow \mathbb{R}$.*

PROOF. Find an open covering of M by open sets with compact closure, and a smooth partition of unity $\{\lambda_i\}$ subordinate to a countable locally finite refinement of this covering. Since the refinement is countable, we may assume that the functions λ_i are indexed by integers $i > 0$. Then the function $f : M \longrightarrow [1, \infty)$ given by $f(x) = \sum_i i\lambda_i(x)$ is a well-defined, because all but a finite number of $\lambda_i(x)$ vanish. Now $f(x) \leq k$ implies at least one of the k functions $\lambda_1, \ldots, \lambda_k$ must not vanish at x (if all of them were zero, then $f(x)$ would be $\geq k + 1$). Therefore $f^{-1}([-k, k])$ is contained in the set

$$\cup_{i=1}^{k}\{x \in M : \lambda_i(x) \neq 0\}$$

which has compact closure. This implies that f is proper, because every compact subset of $\mathbb{R}$ is contained in some interval $[-k, k]$. □

We shall require some more facts about proper maps.

Recall that if X is a locally compact Hausdorff space, then its one-point compactification is a space $X^+ = X \cup \{\infty\}$ (∞ represents a point not in X) such that the topology of X^+ comprises all open sets in $X \subset X^+$ and all sets of the form $X^+ - K$, where K is a compact subset of X. This is the unique topology in X^+ which makes it a compact Hausdorff space with X as a subspace. If X is already compact, then ∞ is an isolated point in X^+.

Lemma 2.1.22. *Let $f : X \longrightarrow Y$ be a continuous map between locally compact Hausdorff spaces, and $f^+ : X^+ \longrightarrow Y^+$ be its extension obtained by setting $f^+(\infty) = \infty$. Then f is proper if and only if f^+ is continuous.*

Proof. Suppose that a continuous map f is proper, and U is an open set in Y^+. Then, if $U \subset Y$, $(f^+)^{-1}(U) = f^{-1}(U)$ is open, and if $U = Y^+ - K$ with $K \subset Y$ compact, then $(f^+)^{-1}(U) = X^+ - f^{-1}(K)$ is also open, since $f^{-1}(K)$ is compact and hence closed. Therefore f^+ is continuous.

Conversely, suppose f^+ is continuous, and $K \subset Y$ is a compact set. Then K is compact, and hence closed in Y^+. Then $(f^+)^{-1}(K)$ is closed in X^+, and hence compact and contained in X. Thus f is proper. $\square$

Corollary 2.1.23. *Let X be locally compact Hausdorff, and Y Hausdorff. Then a continuous injective proper map $f : X \longrightarrow Y$ is a homeomorphism onto its image.*

Proof. Since f is a continuous proper map from X onto $Z = f(X)$, it extends to a continuous map of the one-point compactifications

$$f^+ : X^+ \longrightarrow Z^+.$$

Since f^+ is a bijection from a compact space onto a Hausdorff space, it is a homeomorphism. Therefore f is a homeomorphism onto its image. $\square$

Corollary 2.1.24. *A proper injective immersion from a manifold M into a manifold N is an embedding.*

Proof. An injective immersion which is a homeomorphism onto its image is an embedding. $\square$

Corollary 2.1.25. *Any continuous proper map $f : X \longrightarrow Y$ between locally compact Hausdorff spaces is a closed map.*

Proof. Let C be any closed set in X, and $D = f(C)$. Then $C^+ = C \cup \{\infty\}$ is closed, and hence compact, in X^+. Since f^+ is continuous, $f^+(C^+)$ is compact, and hence closed, in Y^+. Therefore $f(C) = f^+(C^+) \cap Y$ is closed in Y. $\square$

2.2. Smooth approximations to continuous maps

Definition 2.2.1. Let M be a manifold, and N be a manifold with a metric ρ. Let $\delta : M \longrightarrow \mathbb{R}$ be a positive continuous function, and f, g be smooth maps from M to N. Then g is called a **δ-approximation** to f if

$$\rho(f(x), g(x)) < \delta(x),$$

for all $x \in M$.

The (fine) C^0**-topology** on the set $C^\infty(M, N)$ of smooth maps from M to N is a topology where the neighbourhood basis of $f \in C^\infty(M, N)$ comprises all sets of the form

$$B_0(f, \delta) = \{g \in C^\infty(M, N) \mid \rho(f(x), g(x)) < \delta(x)\}.$$

Thus g is a δ-approximation to f, if $g \in B_0(f, \delta)$.

The C^0 topology can be extended to the superset $C^0(M, N)$ of all continuous maps from M to N. This topology on $C^0(M, N)$ is larger than the compact open topology on $C^0(M, N)$. We shall prove this result in §8.2. We shall also see in there that the C^0 topology on $C^\infty(M, N)$ does not depend on the choice of the metric on N.

Lemma 2.2.2. *Let U be an open subset in $\mathbb{R}^n$ (or $\mathbb{R}^n_+$), and $f : U \longrightarrow \mathbb{R}$ be a continuous function such that f is smooth on an open set $V \subset U$. Let U' and V' be two other open sets in U such that $\overline{U}' \subset V'$, $\overline{V}' \subset U$, and $\overline{V}'$ is compact. Let $\delta : U \longrightarrow \mathbb{R}$ be a positive continuous function. Then there is a continuous function $g : U \longrightarrow \mathbb{R}$ such that g is smooth on $V \cup U'$, $g = f$ on $U - V'$, and $|g(x) - f(x)| < \delta(x)$ for all $x \in U$.*

The last condition means that g is a δ-approximation to f.

Proof. Let δ_0 be the positive minimum of the function δ on the compact set $\overline{V}'$. Then, by Weierstrass approximation theorem (see Dieudonné [**6**], (7.4.1), p. 139), there is a polynomial $p(x)$ such that

$$|p(x) - f(x)| < \delta_0 \text{ for } x \in \overline{V}'.$$

Let $h : U \longrightarrow \mathbb{R}$ be a smooth function such that $0 \le h \le 1$, $h = 1$ on $\overline{U}'$, and $h = 0$ on $U - V'$, as given by Lemma 2.1.17. Define $g : U \longrightarrow \mathbb{R}$ by

$$g(x) = h(x)p(x) + (1 - h(x))f(x), \ x \in U.$$

Then $g = f$ on $U - V'$, and $g = p$ on $\overline{U}'$. The last condition shows that g is smooth on U'. Also g is smooth on V, since p, h, f are smooth on it. Finally, on $\overline{V}'$ we have

$$|g(x) - f(x)| = |h(x)| \, |p(x) - f(x)| \ < \ \delta_0.$$

This completes the proof, since $g = f$ on $U - V'$. □

Theorem 2.2.3 (Smoothing theorem). *Let M and N be manifolds. Let K be a closed subset of M, and $f : M \longrightarrow N$ be a continuous map which is smooth on K. Then, there exist a positive continuous function $\delta : M \longrightarrow \mathbb{R}$, and a smooth map $g : M \longrightarrow N$ which agrees with f on K such that g is a δ-approximation to f.*

Remark 2.2.4. The possibility that $K = \emptyset$ is not ruled out.

Proof. For each $x \in M$, let A_x be a coordinate neighbourhood of x in M, and B_x be a coordinate neighbourhood of $f(x)$ in N such that $f(A_x) \subset B_x$. Let $C_x \subset A_x$ be the compact closure of a neighbourhood of x. We shall show that it is possible to choose a countable collection of such C so that their interiors cover M, and any C intersects only a finite number of the other C's of the collection.

For this purpose, we construct, as in the proof of Theorem 2.1.1, a sequence of compact sets $\{K_j\}$ covering M such that $K_j \subset \text{Int } K_{j+1}$. Then the compact sets $L_j = K_j - \text{Int } K_{j-1}$ also cover M, and $L_j \cap L_m = \emptyset$ if $m \neq j-1, j$, or $j+1$. For each $x \in L_j$, we choose coordinate neighbourhoods A_x, B_x, and a compact neighbourhood $C_x \subset A_x$ as above. By shrinking C_x, if necessary, we may suppose that it does not intersect L_m for $m \neq j-1, j$, or $j+1$. Choose a finite number of such C's whose interiors cover L_j, and doing this for each j construct a sequence of sets $\{C_n\}$ such that the Int C_n cover M and any member of the sequence intersects only a finite number of other members. Let $\{A_n\}$ and $\{B_n\}$ be the corresponding sequences of A_x's and B_x's respectively.

Define a sequence of closed sets S_k inductively as follows. Take S_0 as the given closed set K, and then take $S_k = S_{k-1} \cup C_k$, for $k \geq 1$. Then M is the union of the interiors of the sets S_k. We shall construct inductively a sequence of maps $f_k : M \longrightarrow N$, $k \geq 0$, such that

(1) $f_k(x) = f_r(x)$ for $x \in S_r$, if $r < k$,

(2) f_k is smooth on S_k,

(3) $\rho(f_k(x), f(x)) < \delta(x)$, $x \in M$,

(4) f_k maps C_r into B_r for all k and r.

(Here ρ is a metric on N, and δ is a given positive continuous function on M which we shall adjust for completing the inductive step.)

Define $f_0 = f$, and suppose f_r has been defined for $r \leq k$ satisfying these conditions. Let us write $F = f_k$. Then, since F is smooth on S_k, it is smooth on an open neighbourhood V of S_k. Let $D = C_{k+1} - V \cap C_{k+1}$. Then by (4), $f_k = F$ maps D into B_{k+1}. Choose an open set W in C_{k+1} such that $D \subset W$, and $F(W)$ is contained in B_{k+1}. Since $S_k \subset V$, $D \cap S_k = \emptyset$. Therefore we can find open sets U', V', and U with $\overline{V}'$ compact such that

$$D \subset U',\ \overline{U}' \subset V',\ \overline{V}' \subset U,\ U \subset W,\ U \cap S_k = \emptyset,$$

and U intersects only a finite number of the sets C_r. Since B_{k+1} is a coordinate neighbourhood in N, the map $F|U : U \longrightarrow B_{k+1}$ is given by its components $F^{(i)} : U \longrightarrow \mathbb{R}$, $i = 1, 2, \ldots, \dim N$. Then applying Lemma 2.2.2 to each component $F^{(i)}$, we get a map $F' : U \longrightarrow B_{k+1}$ such that $|{F'}^{(i)}(x) - F^{(i)}(x)| < \delta(x)$, $x \in U$, for each i, F' is smooth on $V \cup U'$, and $F' = F$ on $U - V'$. Define $f_{k+1} : M \longrightarrow N$ by

$$\begin{aligned} f_{k+1}(x) &= F(x) = f_k(x) \text{ if } x \notin U, \\ &= F'(x) \text{ if } x \in U. \end{aligned}$$

Then f_{k+1} satisfies (1), because if $x \in S_r \subset S_k$, $r \leq k$, then $x \notin U$ (as $U \cap S_k = \emptyset$), and so $f_{k+1}(x) = f_k(x) = f_r(x)$. Condition (2) holds, because $f_{k+1} = f_k$ is smooth on S_k, and F' is smooth on $V \cup U'$, which contains $(C_{k+1} \cap V) \cup D = C_{k+1}$. Condition (3) holds for f_k, and it will hold for f_{k+1} also, because

$$\rho(F'(x), F(x)) = \max\{|F'^{(i)}(x) - F^{(i)}(x)|\} < \delta(x),\ x \in U.$$

Condition (4) may be obtained by adjusting the size of δ; note that we need only to impose a finite number of restrictions on δ, since f_{k+1} differs from f_k only on V', and V' intersects only a finite number of the sets C_r.

Having constructed the sequence $\{f_k\}$, define $g : M \longrightarrow N$ by $g(x) = f_k(x)$ for $x \in S_k$. This gives g uniquely by (1), and g is smooth, since f_k is smooth on Int S_k and these sets cover M. Finally, $g(x) = f_0(x) = f(x)$ for $x \in S_0 = K$. □

2.3. Sard's theorem

Recall that the Lebesgue measure in $\mathbb{R}^n$ is given by a set function

$$\mu : \mathfrak{M} \longrightarrow [0, \infty]$$

satisfying certain axioms, where $\mathfrak{M}$ is a family of certain subsets of $\mathbb{R}^n$ that are called Lebesgue measurable sets. All open, closed, and compact subsets of $\mathbb{R}^n$ are Lebesgue measurable, so are all G_δ and F_σ subsets. We shall use the following properties of the Lebesgue measure: if S and T are Lebesgue measurable sets, then $\mu(S \cup T) \leq \mu(S) + \mu(T)$, and if $S \subset T$, then $\mu(S) \leq \mu(T)$.

An n-dimensional rectangle R in $\mathbb{R}^n$ is the Cartesian product of n intervals $I_1 \times \cdots \times I_n$; it is an n dimensional cube if all the intervals are of equal length. The Lebesgue measure of R is its volume $\mathrm{vol}(R)$ which is the product of the lengths of the n intervals. For an open set U in $\mathbb{R}^n$, $\mathrm{vol}(U) = \inf(\sum_i \mathrm{vol}(Q_i))$, where $\{Q_i\}$ is any sequence of n-dimensional cubes covering U.

A subset K of $\mathbb{R}^n$ has measure zero in $\mathbb{R}^n$ if for any $\epsilon > 0$, K can be covered by a countable collection of n-dimensional cubes such that the sum of their volumes is less than ϵ. This definition may also be given in terms of rectangles, or even n-dimensional balls. We may say that K has measure zero if and only if for any $\epsilon > 0$ there is an open set U such that $K \subset U$ and $\mathrm{vol}(U) < \epsilon$. A countable union of sets of measure zero has measure zero. For, if $K = K_1 \cup K_2 \cup \cdots$, and $K_i \subset U_i$ where U_i is open and $\mathrm{vol}(U_i) < \epsilon/2^i$, then $K \subset U = \cup U_i$ and $\mathrm{vol}(U) \leq \sum_i \mathrm{vol}(U_i) < \sum \epsilon/2^i = \epsilon$.

The following lemma shows that the condition of being a set of measure zero is invariant under smooth map.

Lemma 2.3.1. *If a subset A of $\mathbb{R}^n$ has measure zero in $\mathbb{R}^n$, and*

$$f : A \longrightarrow \mathbb{R}^m$$

is a smooth map, then $f(A)$ has measure zero in $\mathbb{R}^m$.

Proof. For each $p \in A$, f has a smooth extension on a neighbourhood of p in $\mathbb{R}^n$, which we still denote by f. By shrinking this neighbourhood, if necessary, we may suppose that f is smooth on a closed n-ball B centred at p. If $u_1, \ldots, u_n$ are the coordinate functions in $\mathbb{R}^n$, then the partial derivatives $\partial f_i / \partial u_j$ are bounded on the compact set B. Then by the fundamental theorem of calculus applied to each component f_i of f, together with the chain rule, we can find a constant c such that

$$\|f(x) - f(y)\| \leq c\|x - y\|$$

for all $x, y \in B$ (see Lemma 3.1.3 in Chapter 3). This is called the **Lipschitz estimate** for the smooth map f.

Now given an $\epsilon > 0$, take a countable covering $\{U_j\}$ of $A \cap B$ by open n-balls such that

$$\sum_j \mathrm{vol}(U_j) < \epsilon.$$

Then, by the Lipschitz estimate, $f(B \cap U_j)$ is contained in an n-ball V_j whose radius is not greater than c times the radius of U_j. It follows that $f(B \cap U_j)$ is contained in some of the balls of the collection $\{V_j\}$ whose total volume is not greater than $\sum_j \mathrm{vol}(V_j)$, which is less than $c^n \epsilon$. Since this can be made as small as we like, $f(A \cap B)$ has measure zero. Since $f(A)$ is a union of countably many such sets, it has also measure zero. □

Remark 2.3.2. The lemma may be false if f is only assumed to be continuous. For example, the subset $A = [0, 1]$ has measure zero in $\mathbb{R}^2$, but there exists a continuous map $f : A \longrightarrow \mathbb{R}^2$ whose image fills up the entire square $[0, 1] \times [0, 1]$, which is not a set of measure zero in $\mathbb{R}^2$.

This is the Hahn-Mazurkiewicz theorem which says that a topological space is a Peano space (i.e. a space which is compact, connected, locally connected,and metric) if and only if it is the image of the unit interval under a continuous map into a Hausdorff space (see [**15**], p. 129).

Theorem 2.3.3 (**Fubini**). *If K is a compact set in $\mathbb{R}^n$ such that each subset $K \cap (t \times \mathbb{R}^{n-1})$ has measure zero in the hyperplane $\mathbb{R}^{n-1}$, then K has measure zero in $\mathbb{R}^n$.*

Proof. This simple and elegant proof is due to Bredon [**3**]. We may assume that K is contained in the cube I^n, where I is the unit interval $[0, 1]$. Define a function $f : I \longrightarrow \mathbb{R}$ by

$$f(t) = \mu(K \cap ([0, t] \times I^{n-1})), \qquad t \in I,$$

where μ is the Lebesgue measure on $\mathbb{R}^n$. It is required to show that $f(1) = 0$. By hypothesis, given $\epsilon > 0$, there is an open set U in I^{n-1} such that

$$K \cap (t \times I^{n-1}) \subset t \times U \qquad \text{with } \mathrm{vol}(U) < \epsilon.$$

By compactness of K, there is a $h_0 > 0$ such that

$$K \cap ([t-h_0, t+h_0] \times I^{n-1}) \subset [t-h_0, t+h_0] \times U.$$

Then, for any h, $0 \leq h < h_0$,

$$K \cap ([0, t+h] \times I^{n-1}) \subset (K \cap ([0,t] \times I^{n-1})) \cup ([t, t+h] \times U)$$

can be covered by an open set of volume $< f(t) + \epsilon h$. Therefore

$$f(t+h) \leq f(t) + \epsilon h \qquad \text{for } 0 \leq h < h_0.$$

Similarly, we have

$$K \cap ([0,t] \times I^{n-1}) \subset (K \cap ([0, t-h] \times I^{n-1})) \cup ([t-h, t] \times U)$$

so that

$$f(t) \leq f(t-h) + \epsilon h \qquad \text{for } 0 \leq h < h_0.$$

Therefore

$$\left| \frac{f(t+h) - f(t)}{h} \right| \leq \epsilon, \qquad \text{for all } |h| < h_0.$$

Therefore f is differentiable at t and its derivative is zero. Since $f(0) = 0$, we have $f(1) = 0$ also. $\square$

Definition 2.3.4. A subset K of an n-manifold M is said to have **measure zero** if for each coordinate chart $\phi : U \longrightarrow \mathbb{R}^n$ (or $\mathbb{R}^n_+$) of M, the set $\phi(U \cap K)$ has measure zero in $\mathbb{R}^n$.

It is clear that if $K' \subset K \subset M$ and K has measure zero in M, then K' has also measure zero in M. It is also clear that if $\{K_n\}$ is a countable family of subsets of M such that each K_n has measure zero in M, then $\cup_n K_n$ has measure zero in M.

Let M and N be manifolds of dimension n and m respectively. Then, in view of Definition 1.5.5, a point $x \in M$ is a critical point of a smooth map $f : M \longrightarrow N$, and $f(x)$ is a critical value of f, provided the Jacobian matrix $Jf(x)$ has rank $< m$. A point $y \in N$ is a regular value of f, if it is not a critical value of f.

By convention, any point of N which is not in $f(M)$ is a regular value of f.

Thus, if $n < m$, then every point of M is a critical point of f, and if $n \geq m$ and $y \in f(M)$ is a regular value of f, then $Jf(x)$ has rank m at every point x of $f^{-1}(y)$.

Theorem 2.3.5 (Sard). *If $f : M \longrightarrow N$ is a smooth map of manifolds and C is the set of critical points of f in M, then $f(C)$ has measure zero.*

Remark 2.3.6. A more general version of Sard's theorem says that if

$$f : M \longrightarrow N$$

is a C^r map, where $\dim M = n$, $\dim N = m$, and $r > \max(0, n-m)$, then the set of critical values of f has measure zero. The smoothness condition is necessary, and a counter example (due to Whitney [**57**]) is available, if the inequality be refuted. Whitney constructed a C^1 function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ whose set of critical points C is homeomorphic to the open interval $(0,1) \subset \mathbb{R}$, and $f(C)$ is not a set of measure zero in $\mathbb{R}$. Here $r = \max(0, n-m)$. We will skip the proof of the general version, because Theorem 2.3.5 is adequate for our purpose.

Lemma 2.3.7. *If Sard's theorem is true for every smooth map*

$$f : U \longrightarrow \mathbb{R}^m,$$

where U is an open subset of $\mathbb{R}^n$, then it is also true for every smooth map $g : V \longrightarrow \mathbb{R}^m$, where V is an open subset of $\mathbb{R}^n_+$.

PROOF. Let $g : V \longrightarrow \mathbb{R}^m$, where V is an open subset of $\mathbb{R}^n_+$ be a smooth map, and C be the set of critical points of g. By Lemma 2.1.14(b), there is an open subset V' of $\mathbb{R}^n$ and a smooth map $g' : V' \longrightarrow \mathbb{R}^m$ such that $V = V' \cap \mathbb{R}^n_+$ and $g'|V = g$. If C' is the set of critical points of g', then $g'(C')$ is a set of measure zero in $\mathbb{R}^m$, by hypothesis. This implies $g(C)$ is a set of measure zero in $\mathbb{R}^m$, because $g(C) = g'(C) \subset g'(C')$. □

PROOF OF THEOREM 2.3.5. The proof is by induction on n which is the dimension of M. The starting point is $n = 0$ which is trivial. Therefore suppose that the theorem has been proved for all manifolds of dimensions $\leq n-1$. Next note using the Second Axiom of Countability that it suffices to consider only the special case when $f : U \longrightarrow \mathbb{R}^m$, U being an open set of $\mathbb{R}^n_+$, and C is the critical set of f in U. In view of Lemma 2.3.7, we may suppose that U is an open set in the interior of $\mathbb{R}^n_+$, or in $\mathbb{R}^n$.

Let D be the set of points in C where the Jacobian matrix $J(f)$ vanishes. We shall show in the next two Lemmas 2.3.8 and 2.3.9 that both $f(D)$ and $f(C-D)$ have measure zero in $\mathbb{R}^m$. This will complete the proof of the theorem. □

Lemma 2.3.8. *The set $f(D)$ has measure zero in $\mathbb{R}^m$.*

PROOF. Let $f_1 : U \longrightarrow \mathbb{R}$ be the first component of f. Then, if Jf vanishes at a point x, Jf_1 also vanishes at x, and if K is the set of points where Jf_1 vanishes (K is also the set of critical points of f_1), then $f(D) \subset f_1(K) \times \mathbb{R}^{m-1}$. Therefore if $f_1(K)$ has measure zero in $\mathbb{R}$, then $f_1(K) \times \mathbb{R}^{m-1}$, and hence $f(D)$, has measure zero in $\mathbb{R}^m$, because $\mathbb{R}^{m-1}$ has measure zero in $\mathbb{R}^m$. Hence it is sufficient to prove the lemma for the case $m = 1$.

Let D_i be the set of points of U at which all the partial derivatives of f of order $\leq i$ vanish. We have then a descending sequence of closed subsets of U :

$$D = D_1 \supset D_2 \supset \cdots \supset D_n \supset \cdots .$$

We shall show in the next two sublemmas 1 and 2 that each of the sets

$$f(D_i - D_{i+1}),\ 1 \leq i < n, \text{ and } f(D_n)$$

has measure zero. This will complete the proof of the lemma. □

Sublemma 1. *The set $f(D_i - D_{i+1})$, $1 \leq i < n$, has measure zero.*

Proof. It suffices to show that each point p of $D_i - D_{i+1}$, has a neighbourhood V in U such that $f(V \cap (D_i - D_{i+1}))$ has measure zero. This will prove that $f(D_i - D_{i+1})$ has measure zero, because $D_i - D_{i+1}$ can be covered by countably many of such neighbourhoods, by the Second Axiom of Countability.

If $p \notin D_{i+1}$, there is an i-th order derivative of f, say g, which vanishes on D_i, but its Jacobian matrix Jg is non-zero at p. Then some partial derivative of g say $\partial g/\partial x_1$, is non-zero at p. Define a map $h : U \longrightarrow \mathbb{R}^n$ by $h(x) = (g(x), x_2, \ldots, x_n)$. The Jacobian of h is non-singular at p, and so h maps a neighbourhood V of p diffeomorphically onto an open set W of $\mathbb{R}^n$, by the inverse function theorem. The critical set of $f : V \longrightarrow \mathbb{R}$ is $V \cap (D_i - D_{i+1})$, since $J(f)$ vanishes on D_i. Therefore, since h^{-1} is a diffeomorphism, the critical set of the composition

$$k = f \circ h^{-1} : W \longrightarrow \mathbb{R}$$

is $h(V \cap (D_i - D_{i+1}))$. But $h(V \cap (D_i - D_{i+1})) = (0 \times \mathbb{R}^{n-1})$, and this set is also the critical set of the restriction $k' = k|(0 \times \mathbb{R}^{n-1}) \cap W$. Therefore, by induction (Sard's theorem is true for $n-1$),

$$k'((0 \times \mathbb{R}^{n-1}) \cap W) = f \circ h^{-1}((0 \times \mathbb{R}^{n-1}) \cap W) = f(V \cap (D_i - D_{i+1}))$$

has measure zero. □

Sublemma 2. *The set $f(D_n)$ has measure zero.*

Proof. Again, it will be enough to show that $f(D_n \cap Q)$ has measure for any n-cube Q in U. Let r be the edge length of Q, and k be a positive integer. Subdivide Q into k^n subcubes of edge length r/k, and hence of diameter $r\sqrt{n}/k$. Let $p \in D_n \cap Q$, and Q_1 be one of the subcubes containing p. By Taylor's theorem of order n, if $p + h \in Q_1$, then

$$|f(p+h) - f(p)| \leq A \cdot ||h||^{n+1} \leq A \cdot (r\sqrt{n}/k)^{n+1},$$

where A is a constant independent of k obtained as a uniform estimate of partial derivatives of f of order $n+1$. Therefore $f(D_n \cap Q_1)$ is contained in an interval of length B/k^{n+1}, where B is a constant independent of k. Hence $f(D_n \cap Q)$ is contained in a union of intervals of total length $\leq B \cdot k^n/k^{n+1} = B/k$. Since $\lim_{k\to\infty} B/k = 0$, $f(D_n \cap Q)$ has measure zero. □

Lemma 2.3.9. *The set $f(C - D)$ has measure zero in $\mathbb{R}^m$.*

Proof. Let $p \notin D$. Then some first order partial derivative of some component of f, say $\partial f_1/\partial x_1$, fails to vanish at p. As in the proof of Sublemma 1, an application of the inverse function theorem asserts that the map $h : U \longrightarrow \mathbb{R}^n$, where $h(x) = (f_1(x), x_2, \ldots, x_n)$, sends a neighbourhood V of p diffeomorphically onto an open set W of $\mathbb{R}^n$. Then the set C_1 of critical points of $g = f \circ h^{-1} : W \longrightarrow \mathbb{R}^m$ is precisely $h(V \cap C)$, and $g(C_1) = f(V \cap C)$.

Now g carries each $(t, x_2, \dots, x_n) \in W$ into the hyperplane $t \times \mathbb{R}^{m-1} \subset \mathbb{R}^m$. Let $g_t : W \cap (t \times \mathbb{R}^{n-1}) \longrightarrow t \times \mathbb{R}^{m-1}$ be the restriction of g. Since the Jacobian $J(g)$ of g is of the form

$$\begin{pmatrix} 1 & 0 \\ * & J(g_t) \end{pmatrix},$$

a point in $t \times \mathbb{R}^{n-1}$ is a critical point of g_t if and only if it is a critical point of g. By inductive hypothesis, the set of critical values of g_t has measure zero in $t \times \mathbb{R}^{m-1}$. Now Fubini's theorem implies that the set of critical values of g, that is, the set $f(V \cap C)$, is of measure zero. □

Corollary 2.3.10. *If* $\dim M < \dim N$*, then a smooth map* $f : M \longrightarrow N$ *cannot be surjective.*

PROOF. The critical set of f is M. Therefore, if f is onto, then $N = f(M)$ will have measure zero, which is not possible. □

Corollary 2.3.11. *If* $f : M \longrightarrow N$ *is a smooth map with set of critical points* C*, then the set* $N - f(C)$ *is dense in* N*.*

PROOF. A set of measure zero cannot contain a non-empty open set. □

Corollary 2.3.12. *If* $f_i : M \longrightarrow N$ *is a countable family of smooth maps, then the set of common regular values of all* f_i *is dense in* N*.*

PROOF. Any countable union of sets of measure zero has measure zero. □

2.4. Approximations by immersions

Lemma 2.4.1. *Let* U *be an open set in* $\mathbb{R}^n$ *or* $\mathbb{R}^n_+$*, and* $f : U \longrightarrow \mathbb{R}^m$ *a smooth map, where* $m \geq 2n$*. Then, given any* $\epsilon > 0$*, there is an* $m \times n$ *matrix* $A = (a_{ij})$ *with* $|a_{ij}| \leq \epsilon$ *such the map* $g : U \longrightarrow \mathbb{R}^m$ *given by* $g(x) = f(x) + A \cdot x$ *(*x *written as* $n \times 1$ *column matrix) is an immersion.*

PROOF. The Jacobian matrix of f at $x \in U$ is $Jf(x)$, and that of g at x is $Jg(x) = Jf(x) + A$. The problem is to choose an A so that $Jg(x)$ has rank n at any x, or equivalently, to choose an A from the complement of the set

$$\{B - Jf(x) \mid B \in M(m, n), \text{ rank } B = k < n, x \in U\},$$

where $M(m, n)$ is the space of $m \times n$ matrices.

Let $M_k(m, n)$ denote the space of $m \times n$ matrices of rank k. For any $k < n$, define a map $F_k : M_k(m, n) \times U \longrightarrow M(m, n)$ by $F_k(B, x) = B - Jf(x)$. Then F_k is smooth, and the domain of F_k has dimension $k(m + n - k) + n$ (see Example 1.2.7). Now the function $\alpha(k) = k(m + n - k) + n$ is monotone increasing with k for $k < n$, since its derivative $\alpha'(k) = m + n - 2k > 0$, if $k < n < m$. Therefore

$$k(n + m - k) + n \leq (n - 1)(n + m - n + 1) + n = nm - (m - 2n) - 1 < nm,$$

if $m \geq 2n$. By Sard's theorem, Image F_k has measure zero in $M(m, n)$. Therefore it is possible to find an $A \in M(m, n)$ as close to the zero matrix as we please so that A does not lie in Image F_k for any $k < n$. $\square$

Theorem 2.4.2. *Let $f : M \longrightarrow \mathbb{R}^m$ be a smooth map,* $\dim M = n$ *and $m \geq 2n$. Let $\delta : M \longrightarrow \mathbb{R}$ be a positive continuous function. Then f can be δ-approximated by an immersion $g : M \longrightarrow \mathbb{R}^m$. Moreover, if f is an immersion on a closed subset $K \subset M$, then g can be chosen so that $g(x) = f(x)$ for $x \in K$.*

Proof. If $f|K$ is an immersion, it is an immersion on an open neighbourhood U of K in M, by the local immersion theorem (Theorem 1.4.9). The open covering $\{U, M - K\}$ has a countable locally finite refinement $\{U_i\}$ such that each $\overline{U}_i$ is compact, and each U_i is a coordinate neighbourhood of a chart (U_i, ϕ_i). Re-index the sets U_i by positive and negative integers so that $U_i \subset U$ if and only if $i \leq 0$. By applying the shrinking lemma (Lemma 2.1.4) twice, construct open sets V_i and W_i such that $\{W_i\}$ is a covering of M, and

$$\overline{W}_i \subset V_i, \quad \overline{V}_i \subset U_i.$$

Let $\epsilon_i = \min \delta(x)$ for $x \in \overline{U}_i$.

We shall construct a sequence of smooth maps $f_k : M \longrightarrow \mathbb{R}^m$, $k \geq 0$, such that

(1) $f_0 = f$,

(2) $f_k = f_{k-1}$ on $M - \overline{V}_k$,

(3) f_k has rank n on the set $S_k = \cup_{r \leq k} \overline{W}_r$,

(4) $\|f_k(x) - f_{k-1}(x)\| < \epsilon_k / 2^k$, $x \in M$.

The proof of the theorem will follow, once such a sequence is constructed. Indeed, since the covering $\{U_i\}$ is locally finite, and (2) holds, the f_k's become equal on any compact set when k is sufficiently large. Thus the sequence $\{f_k\}$ converges to a smooth map g. By (2), g agrees with f on K, because, by our indexing convention, $U_k \subset M - U$ if $k \geq 1$, and therefore $f_k = f_{k-1}$ on U for all $k \geq 1$. Also, g has rank n everywhere on M, by (3), and it is a δ-approximation of f by (4).

We now proceed to construct the sequence by induction. Take $f_0 = f$, and suppose f_{k-1} has been defined satisfying the conditions. Since Condition (2) determines f_k outside $\overline{V}_k$, the problem of defining f_k lies entirely within $\overline{V}_k$, and so we may transform the problem to Euclidean space.

For any $m \times n$ matrix A, consider the map $F_A : \phi_k(U_k) \longrightarrow \mathbb{R}^m$ given by

$$F_A(x) = f_{k-1}\phi_k^{-1}(x) + \alpha(x)A(x) \text{ (x is an $n \times 1$ matrix)},$$

where α is a smooth map $\mathbb{R}^n \longrightarrow [0, 1]$ with

$$\alpha|\phi_k(\overline{W}_k) = 1 \text{ and } \alpha(\mathbb{R}^n - \phi_k(V_k)) = 0,$$

obtained from smooth Urysohn's lemma (Lemma 2.1.17). The map F_A will be used in the construction of f_k. For this purpose, we need to choose A in three ways in order to achieve Conditions (3) and (4) for f_k.

Firstly, we require rank F_A to be n on the set $R = \phi_k(S_{k-1} \cap \overline{V}_k)$. Now the Jacobian of F_A is

$$JF_A(x) = J(f_{k-1}\phi_k^{-1})(x) + A(x) \cdot J\alpha(x) + \alpha(x) \cdot A(x),$$

where $J\alpha$ is $1 \times n$ matrix. This gives a continuous map from $R \times M(m,n)$ to $M(m,n)$ sending (x, A) onto $JF_A(x)$. It maps $R \times (0)$ into the open subset $M_n(m,n)$ of $M(m,n)$. So if A is sufficiently small, this map will carry $R \times A$ into $M_n(m,n)$.

Secondly, by Lemma 2.4.1, we may choose A arbitrarily small so that

$$f_{k-1}\phi_k^{-1}(x) + \alpha(x)A(x)$$

has rank n on $\phi_k(U_k)$.

Finally, we choose A small enough so that $\|A(x)\| < \epsilon_k/2^k$ for all $x \in \phi_k(U_k)$.

Let A satisfy all these requirements. Then define $f_k : M \longrightarrow \mathbb{R}^m$ by

$$\begin{aligned} f_k(y) &= f_{k-1}(y) + \alpha(\phi_k(y)) \cdot A(\phi_k(y)) \text{ if } y \in U_k, \\ &= f_{k-1}(y) \text{ if } y \in M - \overline{V}_k. \end{aligned}$$

Two parts of the definition agree on the overlap $U_k - \overline{V}_k$, so f_k is smooth. Condition (2) follows from the definition. By the first choice of A, f_k has rank n on S_{k-1}, and by the second choice of A, f_k has rank n on $\overline{W}_k$, and so Condition (3) holds. Finally, the third choice of A ensures Condition (4) for f_k. This completes the proof. □

Theorem 2.4.3. *If* $\dim M = n$*, and* $m > 2n$*, then any immersion*

$$f : M \longrightarrow \mathbb{R}^m$$

can be δ*-approximated by an injective immersion* $g : M \longrightarrow \mathbb{R}^m$ *for any positive continuous function* $\delta : M \longrightarrow \mathbb{R}$*. Moreover, if* f *is injective on an open neighbourhood* U *of a closed subset* K *of* M*, then we may choose* g *so that it agrees with* f *on* U*.*

Proof. For each $x \in M$, there is an open neighbourhood V_x such that $f|V_x$ is an embedding. The open covering $\{U \cap V_x, (M-K) \cap V_x\}_{x \in M}$ is a refinement of the covering $\{U, M - K\}$. Choose a partition of unity $\{\lambda_i\}$ subordinate to this refined covering, and re-index the λ_i by positive and negative integers so that $\text{supp}\lambda_i \subset U$ if and only if $i \leq 0$.

We shall construct $g : M \longrightarrow \mathbb{R}^m$ as the limit of an infinite series

$$f + v_1\lambda_1 + v_2\lambda_2 + \cdots + v_k\lambda_k + \cdots,$$

where $v_k \in \mathbb{R}^m$ will be chosen by induction. Suppose that $v_1, \ldots, v_k$ have been chosen so that

(1) the map $f_k = f + \sum_{i=1}^{k} v_i\lambda_i$ is an immersion,

(2) $\|f_r - f_{r-1}\| < \delta/2^r$ for $r \leq k$.

(The choice of v_1 will be clear from our arguments below, for which we must take $f_0 = f$.)

Since $\|f_{k+1} - f_k\| = \|v_{k+1}\lambda_{k+1}\| \leq \|v_{k+1}\|$, we may get (2) for $r = k+1$, simply by choosing $\|v_{k+1}\| \leq \delta/2^{k+1}$. The requirement that

$$f_{k+1} = f_k + v_{k+1}\lambda_{k+1}$$

be an immersion can also be met merely by taking $\|v_{k+1}\|$ sufficiently small by the arguments of Lemma 2.4.1. These choices of v_k show that the limit g is an immersion approximating f as required.

For the requirement that $g = f + \sum_{i=1}^{\infty} v_i\lambda_i$ be injective, we need to adjust the v_k still further. For this purpose, let W_{k+1} be the open set of $M \times M$ consisting of pairs (x_1, x_2) such that $\lambda_{k+1}(x_1) \neq \lambda_{k+1}(x_2)$. Let $\phi_{k+1} : W_{k+1} \longrightarrow \mathbb{R}^m$ be the map defined by

$$\phi_{k+1}(x_1, x_2) = \frac{f_k(x_2) - f_k(x_1)}{\lambda_{k+1}(x_1) - \lambda_{k+1}(x_2)}.$$

Since the map is smooth, and $\dim W_{k+1} = 2n < m$, the set $\phi_{k+1}(W_{k+1})$ has measure zero, by Sard's theorem. Therefore, it is possible to choose v_{k+1} arbitrarily small so that $v_{k+1} \notin \phi_{k+1}(W_{k+1})$. Suppose that v_{k+1} has been chosen in this way. Then

$$f_{k+1}(x_1) - f_{k+1}(x_2) = (f_k(x_1) - f_k(x_2)) + v_{k+1}(\lambda_{k+1}(x_1) - \lambda_{k+1}(x_2)).$$

Since $v_{k+1} \notin \phi_{k+1}(W_{k+1})$, it follows that $f_{k+1}(x_1) = f_{k+1}(x_2)$ if and only if $f_k(x_1) = f_k(x_2)$ and $\lambda_{k+1}(x_1) = \lambda_{k+1}(x_2)$.

Choosing all the v_k in this way, suppose $g(x_1) = g(x_2)$. Now, since $\{\lambda_i\}$ is a partition of unity, there is an r sufficient large such that $\lambda_k(x_1) = \lambda_k(x_2) = 0$ for $k > r$. Therefore

$$f_r(x_1) = g(x_1) = g(x_2) = f_r(x_2).$$

This implies by the above constructions of the v_k that $\lambda_r(x_1) = \lambda_r(x_2) = 0$, and $f_{r-1}(x_1) = f_{r-1}(x_2)$. Proceeding in this way in the direction of decreasing r, we arrive at $f(x_1) = f(x_2)$, and $\lambda_i(x_1) = \lambda_i(x_2) = 0$ for all $i > 0$. The last conditions imply that $x_1,\ x_2$ cannot belong to supp λ_i for all $i > 0$. But, by our indexing convention, supp $\lambda_i \subset M - U$ for all $i > 0$, Therefore $x_1,\ x_2$ must lie in U, on which f is injective. Therefore $x_1 = x_2$, showing that g is injective. □

2.5. Whitney's embedding theorem

We have now all the materials in hand to prove Whitney's theorem.

Theorem 2.5.1. *Any manifold M of dimension n can be embedded in $\mathbb{R}^{2n+1}$ as a closed subspace of $\mathbb{R}^{2n+1}$.*

PROOF. Consider the proper map $g = j \circ f : M \longrightarrow \mathbb{R}^{2n+1}$, where $f : M \longrightarrow \mathbb{R}$ is the proper function constructed in Lemma 2.1.21, and

$$j : \mathbb{R} \longrightarrow \mathbb{R}^{2n+1}$$

is the inclusion map. Recall that f is given in terms of a partition of unity $\{\lambda_i\}$ on M as $f(x) = \sum_{i=1}^{\infty} i \ \lambda_i(x)$. Then, taking $\delta(x) = 1/2$ and applying Theorems 2.4.2 and 2.4.3, we get an injective immersion $h : M \longrightarrow \mathbb{R}^{2n+1}$ such that

$$\|g(x) - h(x)\| < 1 \quad \text{for all } x \in M.$$

We will show that this h is also a proper map. This will imply that h is an embedding, by Corollary 2.1.24, and the proof will be complete.

It is sufficient to show that for each integer $k > 0$ the inverse image by h of the ball $B_k = \{u \in \mathbb{R}^{2n+1} | \|u\| \leq k\}$ is compact in M. By the above inequality, if $\|h(x)\| \leq k$, then

$$\|g(x)\| \leq \|g(x) - h(x)\| + \|h(x)\| < 1 + k,$$

which implies that $x \in \cup_{i=1}^{k+1} \mathrm{supp} \lambda_i$ (see the proof of Lemma 2.1.21). Thus $h^{-1}(B_k)$ is contained in a compact set, and so it is compact.

Finally, since h is proper, its image is a closed subset of $\mathbb{R}^{2n+1}$, by Corollary 2.1.25. $\square$

Remark 2.5.2. The above proof by the method of approximation cannot be improved so as to get an embedding of an n-manifold M into an Euclidean space of dimension lower than $2n + 1$. For example, the immersion of S^1 into $\mathbb{R}^2$ for which the image crosses itself in the form of figure 8 cannot be approximated to any sufficiently close injective immersion whatsoever. In [**60**] Whitney employed a different method to prove that every n-manifold M can be embedded into $\mathbb{R}^{2n}$. He made deeper analysis on some homological conditions on M for removing double points. This result is best possible in the sense that there are manifolds of dimension n which cannot be embedded in $\mathbb{R}^{2n-1}$. For example, if $n = 2^r$, there is no embedding of the real projective space $\mathbb{R}P^n$ into $\mathbb{R}^{2n-1}$ (see Husemoller [**18**], Theorem 10.3, p. 262).

2.6. Homotopy of smooth maps

We will now extend the notion of homotopy to the smooth category. Two smooth maps are called smoothly homotopic if one can be deformed to the other through smooth maps. Here is the precise definition.

Definition 2.6.1. Two smooth maps $f, g : M \longrightarrow N$ are **smoothly homotopic** if there is a smooth map $H : M \times \mathbb{R} \longrightarrow N$ such that $H(x, 0) = f(x)$ and $H(x, 1) = g(x)$.

Thus we have a family of smooth maps $H_t : M \longrightarrow N$ given by $H_t(x) = H(x, t), t \in \mathbb{R}$. The smooth map H is called a **smooth homotopy** between f

and g. If H is just a continuous map, then f and g are continuously homotopic, or simply homotopic.

The smooth homotopy is defined for all $t \in \mathbb{R}$, rather than on the interval $I = [0, 1]$, because we want to avoid a technical difficulty, namely, $M \times I$ is not a smooth manifold when M has boundary. In §7.5, we shall show that $M \times I$ can be given a unique smooth structure. Then we will have no problem in replacing $\mathbb{R}$ by I in the above definition.

The portion of $\mathbb{R}$ outside I does not play any important role. Given H as above, we can always find a smooth map $\overline{H} : M \times \mathbb{R} \longrightarrow N$ such that $\overline{H}(x, t) = f(x)$ if $t \leq 0$ and $\overline{H}(x, t) = g(x)$ if $t \geq 1$. Just define $\overline{H}(x, t) = H(x, \mathcal{B}(t))$, where $\mathcal{B}(t)$ is a bump function (Definition 2.1.6). The smooth map $\overline{H}$ is called the **normalised homotopy** corresponding to the homotopy H. Note that H and $\overline{H}$ are smoothly homotopic by $F : M \times \mathbb{R} \times \mathbb{R} \longrightarrow N$, where $F(x, s, t) = H(x, (1 - s) \cdot t + s \cdot \mathcal{B}(t))$.

Lemma 2.6.2. *Smooth homotopy is an equivalence relation.*

Proof. That the relation is reflexive and symmetric are obvious. To see that it is transitive, take smooth maps f, g, and h from M to N, and let H and F be normalised smooth homotopies between f and g and between g and h respectively. Define $K : M \times \mathbb{R} \longrightarrow N$ by

$$\begin{aligned} K(x, t) &= H(x, 3t) && \text{if } t \leq 1/2, \\ &= F(x, 3t - 2) && \text{if } t \geq 1/2. \end{aligned}$$

This is a smooth map, since H and F are smooth maps and $K(x, t) = g(x)$ for $1/3 \leq t \leq 2/3$ so that two parts of the definition match together smoothly. Clearly K is a normalised homotopy between f and h. $\square$

Lemma 2.6.3. *If two smooth maps $f, g : M \longrightarrow N$ are continuously homotopic, then they are smoothly homotopic.*

Proof. Let $H : M \times \mathbb{R} \longrightarrow N$ be a normalised continuous homotopy between f and g. Then H is smooth on the closed set $M \times J$, where $J = (-\infty, 0] \cup [1, \infty)$, since $H \mid M \times (-\infty, 0] = f$ and $H \mid M \times [1, \infty) = g$. By the smoothing theorem (Theorem 2.2.3), there is a positive continuous function δ on M such that H can be δ-approximated by a smooth map $F : M \times \mathbb{R} \longrightarrow N$ which agrees with H on $M \times J$. $\square$

$\Diamond$ **Exercise 2.2.** Show that if m is sufficiently large, then any smooth map

$$f : M \longrightarrow \mathbb{R}^m$$

is δ-approximable by an embedding $g : M \longrightarrow \mathbb{R}^m$ which is homotopic to f by a smooth homotopy $H_t : M \longrightarrow \mathbb{R}^m$ so that each H_t is a δ-approximation to f.

Hint. $H_t(x) = (1 - t)f(x) + tg(x)$.

Definition 2.6.4. Two embeddings $f, g : M \longrightarrow N$ are **isotopic** if there exists a smooth homotopy $H : M \times \mathbb{R} \longrightarrow N$ such that for each $t \in \mathbb{R}$, the map

$$H_t : M \longrightarrow N$$

is an embedding.

Remark 2.6.5. If $H_t : M \longrightarrow N$ is an isotopy, and

$$\alpha : M_1 \longrightarrow M, \quad \beta : N \longrightarrow N_1$$

are embeddings, then $\beta \circ H_t \circ \alpha : M_1 \longrightarrow N_1$ is an embedding.

Proposition 2.6.6. *Any two embeddings $f, g : M \longrightarrow \mathbb{R}^m$ are isotopic, provided m is sufficiently large (in fact, $m \geq 2n + 2$, where $n = \dim M$).*

Proof. Since $\mathbb{R}^m$ is contractible to a point, the embeddings f and g are continuously homotopic, and hence homotopic by a smooth homotopy $H : M \times \mathbb{R} \longrightarrow \mathbb{R}^m$. If m is sufficiently large, H may be deformed to an embedding $F : M \times \mathbb{R} \longrightarrow \mathbb{R}^m$ which agrees with H on $(-\infty, 0] \cup [1, \infty)$. This F serves as the required isotopy between f and g. $\square$

2.7. Stability of smooth maps

We now consider a different set of problems, whose elegant formulation and presentation are influenced by Guillemin and Pollack [**12**]. Suppose, for example, an embedding f is deformed slightly to a map g; then we would like to pose the question whether g also an embedding.

Definition 2.7.1. Let $\mathcal{C}$ be a class of smooth maps from M to N defined by a property. Then $\mathcal{C}$ is called a **stable class** with respect to the property if for any $f \in \mathcal{C}$ and any smooth homotopy $f_t : M \longrightarrow N$ of f, there is an $\epsilon > 0$ such that $f_t \in \mathcal{C}$ for all $t < \epsilon$.

Theorem 2.7.2. *Each of the following classes of smooth maps from M to N, where M is compact and $\partial M = \partial N = \emptyset$, is a stable class:*

(1) *local diffeomorphisms,*
(2) *immersions,*
(3) *submersions,*
(4) *embeddings,*
(5) *diffeomorphisms.*

To this list of classes of maps, we may add one more class, namely, the class of maps transversal to a given submanifold A of N. We will read about this class of maps in Chapter 6, and show that locally the transversality condition is the same as the submersion condition (3). Therefore this class will be stable.

Proof. We shall prove only (2) and (4). Because, (1) is a special case of (2) when $\dim M = \dim N$, and the proof of (3) is essentially identical with the proof of (2). The proof of (5) will follow from (4) and the fact that a local diffeomorphism maps open sets into open sets.

Proof of (2). Let f_t be a smooth homotopy of an immersion f_0. Then the problem is to find an $\epsilon > 0$ so that $d(f_t)_x$ is injective for all points

$$(x, t) \in M \times [0, \epsilon) \subset M \times I.$$

Since M is compact, any open neighbourhood of $M \times \{0\}$ in $M \times I$ contains $M \times [0, \epsilon)$ if ϵ is small enough. Therefore, it is sufficient only to show that each point $(x_0, 0) \in M \times \{0\}$ has an open neighbourhood U in $M \times I$ such that $d(f_t)_x$ is injective for $(x, t) \in U$. Since this assertion is local, it is enough to consider only the case when M is an open subset of $\mathbb{R}^n$, and N is an open subset of $\mathbb{R}^m$.

Since $d(f_0)_{x_0}$ is injective, the Jacobian matrix $Jf_0(x_0)$ of f_0 at x_0 has a minor $R(x_0, 0)$ of order n whose determinant is non-zero. The function

$$M \times I \longrightarrow \mathbb{R},$$

which sends (x, t) to the determinant of the minor $R(x, t)$ of the Jacobian matrix $Jf_t(x)$ (formed by the same rows and columns as $R(x_0, 0)$) is continuous, since each entry of $R(x, t)$ is continuous on $M \times I$, and the determinant function is continuous. Therefore there is an open neighbourhood U of $(x_0, 0)$ in $M \times I$ such that $R(x, t)$ is non-singular for all $(x, t) \in U$. This completes the proof of (2).

Proof of (4). As shown in the above proof, f_0 is an immersion implies that f_t is an immersion for small values of t. We shall show that if f_0 is injective, then so is f_t for sufficiently small t. This will complete the proof of (4), because any injective immersion on a compact manifold is an embedding.

Suppose our assertion is false. Take a sequence of real numbers $\{t_k\}$ which converges to zero. For each k, we can find a pair of distinct points (x_k, y_k) of M such that $f_{t_k}(x_k) = f_{t_k}(y_k)$. Since M is compact, each of the sequences $\{x_k\}$ and $\{y_k\}$ has convergent subsequences. Denoting them by the same notations, let $\lim x_k = x_0$ and $\lim y_k = y_0$. Then

$$f_0(x_0) = \lim f_{t_k}(x_k) = \lim f_{t_k}(y_k) = f_0(y_0).$$

This implies that $x_0 = y_0$, since f_0 is injective.

Define a smooth map $G : M \times I \longrightarrow N \times I$ by $G(x, t) = (f_t(x), t)$ A simple computation shows that the Jacobian matrix $JG(x_0, 0)$ is

$$\begin{pmatrix} \boxed{Jf_0(x_0)} & * \\ 0, \cdots, 0 & 1 \end{pmatrix},$$

which is non-singular, since $Jf_0(x_0)$ is so. Then, by the inverse function theorem (Theorem 1.4.3), G is injective in a neighbourhood of $(x_0, 0)$. But for large

k, both (x_k, t_k) and (y_k, t_k) belong to this neighbourhood, and so $x_k = y_k$, which is a contradiction. Therefore we may conclude that f_t is injective when t is sufficiently small. $\square$

$\Diamond$ **Exercise 2.3.** Show that Theorem 2.7.2 is false if M is not compact, by constructing counterexamples to all the classes in the following way:

Define $f_t : \mathbb{R} \longrightarrow \mathbb{R}$ by $f_t(s) = s\lambda(ts)$, where $\lambda : \mathbb{R} \longrightarrow \mathbb{R}$ is a smooth function with $\lambda(s) = 1$ if $|s| < 1$, and $\lambda(s) = 0$ if $|s| > 2$.

CHAPTER 3

LINEAR STRUCTURES ON MANIFOLDS

The title of the chapter may seem to be all embracing, as almost any thing of the manifold theory may come under this heading. In fact, there is hardly any structure on manifolds that is not influenced by linear algebra. However, we have selected among others two structures on manifolds, namely, the symplectic and contact structures. These structures play substantial roles in mathematical physics, see, for example, Guillemin and Sternberg [**13**]. Our aim here is to build up necessary ingredients with a view to encounter these structures at the end. Apart from this, the ingredients are the basic tools that are used in the workshop of manifold theory in the constructions of various other structures. Lying at the root of these ingredients is the concept of tangent space at a point which is the first order approximation of the manifold in a neighbourhood of that point. Through tangent spaces we are able to transport many ideas of linear analysis into manifolds.

The main theorem of the chapter is Darboux-Weinstein theorem on the local classification of the symplectic structures, and its analogue for contact structures. For these accounts we have consulted Guillemin and Sternberg [**13**], McDuff and Lafontaine [**27**].

3.1. Tangent spaces and derivative maps

As before for an open set U of a manifold M, $C^\infty(U)$ denotes the set of all smooth functions from U to $\mathbb{R}$. Let $p \in M$, and $\widetilde{C}^\infty(p)$ be the union of all $C^\infty(U)$ as U runs over all open neighbourhoods of p. This is an algebra over $\mathbb{R}$, because if $f \in C^\infty(U)$, and $g \in C^\infty(V)$, then $f+g, fg \in C^\infty(U \cap V)$, and $\lambda f \in C^\infty(U)$ for all $\lambda \in \mathbb{R}$. Two functions f and g as above are said to be **equivalent** (or have the same **germ** at p) if $f = g$ in a neighbourhood of p. The quotient set $C^\infty(p)$ of $\widetilde{C}^\infty(p)$ under this equivalence relation is also an algebra, called the **algebra of germs of smooth functions** at p.

In fact, $C^\infty(p)$ is the quotient algebra $\widetilde{C}^\infty(p)/\widetilde{C}^\infty_0(p)$, where $\widetilde{C}^\infty_0(p)$ is the ideal consisting of functions which vanish in a neighbourhood of p (neighbourhood depending on the function).

Definition 3.1.1. A **tangent vector** of M at a point $p \in M$ is the geometric name of what is called a derivation of the algebra $C^\infty(p)$ on $\mathbb{R}$. It is

A. Mukherjee, *Differential Topology*, DOI 10.1007/978-3-319-19045-7_3

a linear functional $X_p : C^\infty(p) \longrightarrow \mathbb{R}$ satisfying the **Leibniz formula**

$$X_p(fg) = f(p) \cdot X_p(g) + g(p) \cdot X_p(f), \quad f, g \in C^\infty(p).$$

The formula implies that if f is a constant function, then $X_p f = 0$ for all $p \in M$.

The set $\tau(M)_p$ of all tangent vectors of M at p is called the **tangent space** of M at p, or the space of derivations at p. It is a vector space over $\mathbb{R}$, where the vector space operations are defined by $(X_p + Y_p)(f) = X_p(f) + Y_p(f)$, and $(\lambda X_p)(f) = \lambda X_p(f)$ for $X_p, Y_p \in \tau(M)_p$, $f \in C^\infty(p)$, and $\lambda \in \mathbb{R}$.

The geometric picture behind the definition will be clear after we prove that the dimension of the vector space $\tau(M)_p$ is n, which is also equal to the dimension of M.

Proposition 3.1.2. *If $\phi = (x_1, \ldots, x_n)$ is a coordinate system in M at p, then the operators*

$$\left[\frac{\partial}{\partial x_i}\right]_p : C^\infty(p) \longrightarrow \mathbb{R}, \quad i = 1, \ldots, n,$$

defined by $f \mapsto (\partial f/\partial x_i)(p)$ are tangent vectors of M at p, and they form a basis of the vector space $\tau(M)_p$.

Here $(\partial f/\partial x_i)(p)$ is the partial derivative as defined in §1.4.

We first prove a lemma.

Lemma 3.1.3. *Let $a \in \mathbb{R}^n$ and $f \in C^\infty(a)$. Then there exist functions $g_1, \ldots, g_n \in C^\infty(a)$ and a neighbourhood U of a in $\mathbb{R}^n$ contained in the intersection of the domains of $f, g_1, \ldots, g_n$ such that $g_i(a) = (\partial f/\partial u_i)(a), 1 \leq i \leq n$, and*

$$f(u) = f(a) + \sum_{i=1}^{n} (u_i - u_i(a)) \cdot g_i(u), u \in U,$$

where $u = (u_1, \ldots, u_n)$, $u_i : \mathbb{R}^n \longrightarrow \mathbb{R}$, is the coordinate system in $\mathbb{R}^n$. [1]

Proof. Define

$$g_i(u) = \int_0^1 \frac{\partial f}{\partial u_i}(t(u-a)+a)dt.$$

---

[1]The lemma is not true for C^r manifolds where $r < \infty$, because the functions g_i may not be always C^r. In this case the space of derivations at p is infinite dimensional, and the tangent space is defined to be the space spanned by $[\partial/\partial x_i]_p$, see W.F. Newns and A.G. Walker, Tangent planes to a differentiable manifold, J. London Math. Soc. 31 (1956), 400-407.

This is C^∞ in a neighbourhood of a, and $g_i(a) = (\partial f/\partial u_i)(a)$. Therefore

$$\begin{aligned} f(u) - f(a) &= \int_0^1 \frac{d}{dt} f(t(u-a)+a)dt \\ &= \int_0^1 \left\{ \sum_{i=1}^n \frac{\partial f}{\partial u_i}(t(u-a)+a) \cdot (u_i - u_i(a)) \right\} dt \\ &= \sum_{i=1}^n g_i(u) \cdot (u_i - u_i(a)). \end{aligned}$$

□

♢ **Exercise 3.1.** Show that if $a \in \mathbb{R}^n$ and $f \in C^\infty(a)$, then there exist functions $\lambda_{ij} \in C^\infty(a)$, and a neighbourhood U of a on which f and λ_{ij} are defined such that $\lambda_{ij}(a) = (\partial^2 f/\partial u_i \partial u_j)(a)$, and if $u \in U$ then

$$f(u) = f(a) + \sum_{i=1}^n (u_i - u_i(a)) \cdot \frac{\partial f}{\partial u_i}(a) + \sum_{i,j=1}^n (u_i - u_i(a))(u_j - u_j(a)) \cdot \lambda_{ij}(u).$$

Proof of the Proposition 3.1.2. That the operators $[\partial/\partial x_i]_p$ are tangent vectors is immediate from the definition. Next, for any $f \in C^\infty(p)$, use the lemma to write $f \circ \phi^{-1}$ in a neighbourhood of $a = \phi(p)$ as

$$f \circ \phi^{-1}(u) = f \circ \phi^{-1}(a) + \sum_{i=1}^n (u_i - u_i(a)) \cdot g_i(u),$$

where g_i is a C^∞ function in a neighbourhood of a with

$$g_i(a) = (\partial(f \circ \phi^{-1})/\partial u_i)(a).$$

Transferring this relation to a neighbourhood of p in M, write

$$f(x) = f(p) + \sum_{i=1}^n (x_i - x_i(p)) \cdot h_i(x),$$

where $h_i = g_i \circ \phi \in C^\infty(p)$ with $h_i(p) = g_i(a) = (\partial f/\partial x_i)(p)$, and then apply a derivation $X_p \in \tau(M)_p$ to f using the Leibniz formula :

$$X_p(f) = \sum_{i=1}^n X_p(x_i) \cdot h_i(p) = \sum_{i=1}^n X_p(x_i) \cdot \left[\frac{\partial f}{\partial x_i}\right]_p$$

(recall that $X_p(c) = 0$ for any constant function c). Thus in terms of the coordinate system $(x_1, \dots, x_n)$, X_p takes the form

$$X_p = \sum_{i=1}^n X_p(x_i) \cdot \left[\frac{\partial}{\partial x_i}\right]_p .$$

Now $[\partial/\partial x_i]_p(x_j) = (\partial(u_j \circ \phi \circ \phi^{-1})/\partial u_i)(a) = \delta_{ij}$ (Kronecker delta). Therefore the vectors $\{[\partial/\partial x_i]_p\}$ are linearly independent, as may be seen by evaluating a linear combination of these vectors on x_j in turn. □

It follows that if U is an open neighbourhood of p, then $\tau(U)_p = \tau(M)_p$, because the definition of $\tau(M)_p$ uses only $C^\infty(p)$, and not the entire M. Also, the tangent space $\tau(M)_p$ is isomorphic to $\mathbb{R}^n$, where $[\partial/\partial x_i]_p$ corresponds to the i-th unit coordinate vector of $\mathbb{R}^n$, and therefore the tangent space $\tau(\mathbb{R}^n)_p$ can be identified with the set of all pairs (p, v), where $v \in \mathbb{R}^n$.

A **smooth curve** in M is a smooth map $\sigma : I \longrightarrow M$, where I is an open interval in $\mathbb{R}$. For each $t_0 \in I$, σ gives rise to a tangent vector $\dot{\sigma}(t_0) : C^\infty(p) \longrightarrow \mathbb{R}$ of M at $p = \sigma(t_0)$ defined by

$$\dot{\sigma}(t_0)(f) = \left[\frac{d}{dt} f(\sigma(t))\right]_{t=t_0},$$

which is the derivative of f along σ at p. The components of $\sigma(t)$ with respect to a local coordinate system $(x_1, \ldots, x_n)$ at p are the real-valued functions $\sigma_i(t) = x_i(\sigma(t))$, and their derivatives $\dot{\sigma}_i(t_0) = (d(\sigma_i(t))/dt)(t_0)$ are the components of the tangent vector $\dot{\sigma}(t_0)$ with respect to the basis $[\partial/\partial x_i]_p$. Because, if (U, ϕ) is the coordinate chart at p for which $\phi = (x_1, \ldots, x_n)$, $x_i = u_i \circ \phi$, then $\sigma_i(t) = x_i(\sigma(t)) = u_i \circ \phi(\sigma(t))$, and therefore, by chain rule

$$\begin{aligned}
\dot{\sigma}(t_0)(f) &= \frac{d}{dt}\Big[(f \circ \phi^{-1}) \circ (\phi \circ \sigma)\Big](t_0) \\
&= \sum_i \frac{\partial(f \circ \phi^{-1})}{\partial u_i}(\phi \circ \sigma(t_0)) \cdot \frac{\partial(u_i \circ \phi \circ \sigma)}{\partial t}(t_0) \\
&= \sum_i \frac{\partial f}{\partial x_i}(p) \cdot \frac{dx_i}{dt}(t_0).
\end{aligned}$$

We also say that $\dot{\sigma}(t_0)$ is the **tangent vector** or **velocity vector** of σ at $\sigma(t_0)$. In the case when $M = \mathbb{R}^n$, this vector may be viewed as a line segment from $\sigma(t_0)$ to $\sigma(t_0) + \dot{\sigma}(t_0)$. Conversely, any tangent vector to M at p is associated to a smooth curve in this way. For, if $\phi = (x_1, \ldots, x_n)$ is a coordinate system at p, then a vector $\sum_i v_i[\partial/\partial x_i]_p \in \tau(M)_p$ is clearly tangent at p to the curve

$$t \mapsto \phi^{-1}(x_1(p) + tv_1, \ldots, x_n(p) + tv_n).$$

We may therefore define a tangent vector to M at p alternatively as follows. Consider the set of all smooth curves $\sigma : I \longrightarrow M$, where I is an open interval containing 0, such that $\sigma(0) = p$. Define an equivalence relation in this set by taking two curves σ and τ to be equivalent if $\dot{\sigma}(0) = \dot{\tau}(0)$. Then a tangent vector to M at p is an equivalence class of curves.

$\Diamond$ **Exercise 3.2.** Check that the relation on the set of all smooth curves as defined above is indeed an equivalence relation.

Example 3.1.4. Let $\langle\ ,\ \rangle$ denote the standard inner product in $\mathbb{R}^{n+1}$. Then the n-sphere S^n in $\mathbb{R}^{n+1}$ is given by

$$S^n = \{v \in \mathbb{R}^{n+1} | \langle v, v\rangle = 1\}.$$

Consider a smooth curve $\sigma : I \longrightarrow \mathbb{R}^{n+1}$ so that $\sigma(s) \in S^n$ for all $s \in I$, and $\sigma(0) = p$. Then $\langle \sigma(s), \sigma(s) \rangle = 1$. Differentiating this relation with respect to s at $s = 0$, we get

$$\langle \dot\sigma(0), \sigma(0) \rangle + \langle \sigma(0), \dot\sigma(0) \rangle = 0, \text{ or } \langle \dot\sigma(0), \sigma(0) \rangle = 0.$$

Since $\dot\sigma(0)$ is a vector in $\tau(S^n)_p$, the above relation says that $\tau(S^n)_p$ is the hyperplane in $\mathbb{R}^{n+1}$ orthogonal to $\sigma(0) = p$.

Example 3.1.5. With reference to the table given in Theorem 1.6.6, the tangent space at the unit element e $(= I_n)$ of a Lie group G of the first row is the corresponding Lie algebra $\mathfrak{g}$ in the second row. This may be seen in the following way.

Each element $A \in \mathfrak{g}$ gives rise to a smooth curve $\sigma(s) = \exp sA$ in G such that $\sigma(0) = e(= I_n)$. The velocity vector of σ at $s = 0$ is

$$\lim_{s \to 0} \frac{\exp sA - I_n}{s} = \lim_{s \to 0} (A + \frac{sA^2}{2!} + \cdots) = A.$$

Thus $\mathfrak{g} \subset \tau(G)_e$. Since the vector space $\mathfrak{g}$ and the manifold G have the same dimension, we have $\tau(G)_e = \mathfrak{g}$.

The problem may also be solved without any reference to the exp map, as the following example shows.

Example 3.1.6. For the special linear group $SL(n, \mathbb{C})$, a smooth curve $\sigma(s)$ in $SL(n, \mathbb{C})$ with $\sigma(0) = I_n$ satisfies $\det \sigma(s) = 1$. Differentiation of this relation with respect to s at $s = 0$ gives trace $\dot\sigma(0) = 0$. Thus $\tau(SL(n, \mathbb{C}))_{I_n} \subseteq \mathfrak{sl}(n, \mathbb{C})$, where $\mathfrak{sl}(n, \mathbb{C})$ is the space of complex matrices of order n having trace equal to zero (this space was denoted earlier by S_0 in §1.6, p.25). Since the vector space $\mathfrak{sl}(n, \mathbb{C})$ has the same dimension as the manifold $SL(n, \mathbb{C})$, we have $\tau(SL(n, \mathbb{C}))_{I_n} = \mathfrak{sl}(n, \mathbb{C})$.

Similarly, the tangent space at I_n of the real special linear group $SL(n, \mathbb{R})$ is the vector space of real traceless matrices $\mathfrak{sl}(n, \mathbb{R})$.

Definition 3.1.7. If $f : M \longrightarrow N$ is a smooth map between manifolds, then the **derivative map** or **differential** of f at a point $p \in M$ is a linear map $df_p : \tau(M)_p \longrightarrow \tau(N)_{f(p)}$ defined by

$$df_p(X_p)(g) = X_p(g \circ f), \; X_p \in \tau(M)_p, \; g \in C^\infty(f(p)).$$

Taking X_p as the velocity vector $\dot\sigma(0)$ of a smooth curve σ in M at $\sigma(0) = p$ with parameter t, the definition may be given in the following alternative form:

$$df_p(\dot\sigma(0))(g) = \frac{d}{dt}(g \circ f(\sigma(t)))(0).$$

We may rephrase the previous definition of the velocity vector $\dot\sigma(0)$ as follows

$$\dot\sigma(0) = d\sigma_0\Big(\frac{d}{dt}\Big),$$

where $d\sigma_0 : \tau(I)_0 = \mathbb{R} \longrightarrow \tau(M)_p$ is the derivative map of σ at 0, and d/dt is the basis of $\mathbb{R}$. Because

$$d\sigma_0\Big(\frac{d}{dt}\Big)(g) = \frac{d}{dt}g(\sigma(t))(0) = \dot\sigma(0)(g).$$

Let (U, ϕ) with $\phi = (x_1, \ldots, x_n)$, $x_i = u_i \circ \phi$, be a coordinate chart at p, and (V, ψ) with $\psi = (y_1, \ldots, y_m)$, $y_j = v_j \circ \psi$, be a coordinate chart at $q = f(p)$, where u_i (resp. v_j) are the coordinate functions on $\mathbb{R}^n$ (resp. $\mathbb{R}^m$). Then

$$\begin{aligned} df_p\left(\left[\frac{\partial}{\partial x_i}\right]_p\right)(g) &= \frac{\partial}{\partial x_i}(g\circ f)(p) = \frac{\partial}{\partial u_i}(g\circ f\circ\phi^{-1})(\phi(p)) \\ &= \frac{\partial}{\partial u_i}(g\circ\psi^{-1}\circ\psi\circ f\circ\phi^{-1})(\phi(p)) = \frac{\partial}{\partial u_i}(\overline{g}\circ\overline{f})(\phi(p)), \end{aligned}$$

where $\overline{f} = \psi\circ f\circ\phi^{-1} : \phi(U) \longrightarrow \psi(V)$, and $\overline{g} = g\circ\psi^{-1} : \psi(V) \longrightarrow \mathbb{R}$ are smooth maps. By the chain rule, the last expression is equal to

$$\sum_{j=1}^m \frac{\partial \overline{f}_j}{\partial u_i}(\phi(p)) \cdot \frac{\partial \overline{g}}{\partial v_j}(\psi(q)) = \sum_{j=1}^m \frac{\partial f_j}{\partial x_i}(p) \cdot \frac{\partial g}{\partial y_j}(q).$$

Therefore

$$df_p\left(\left[\frac{\partial}{\partial x_i}\right]_p\right) = \frac{\partial f_1}{\partial x_i}(p)\left[\frac{\partial}{\partial y_1}\right]_{f(p)} + \cdots + \frac{\partial f_m}{\partial x_i}(p)\left[\frac{\partial}{\partial y_m}\right]_{f(p)}.$$

Therefore the i-th column vector of the matrix of the linear map df_p with respect to the bases $[\partial/\partial x_i]_p$ and $[\partial/\partial y_j]_{f(p)}$ of the tangent spaces $\tau(M)_p$ and $\tau(N)_{f(p)}$ is

$$\Big(\frac{\partial f_1}{\partial x_i}(p), \ldots, \frac{\partial f_m}{\partial x_i}(p)\Big).$$

Therefore the matrix of df_p is the Jacobian matrix of f at p, as defined in §(1.4)

$$(Jf)(p) = \left(\frac{\partial f_i}{\partial x_j}(p)\right).$$

Thus if we represent a tangent vector $X_p = \sum_i a_i\Big(\frac{\partial}{\partial x_i}\Big)_p$ by the $n\times 1$ matrix $A = (a_i)$, then the tangent vector $df_p(X_p)$ is represented by the $m\times 1$ matrix $(Jf)(p)\cdot A$. In particular, for the coordinate chart (U, ϕ),

$$d\phi_p\Big(\sum_i a_i\Big(\frac{\partial}{\partial x_i}\Big)_p\Big) = (a_1, \ldots, a_n),$$

If $f : M \longrightarrow N$ and $g : N \longrightarrow L$ are smooth maps of manifolds, then

$$d(g\circ f)_p = dg_{f(p)}\circ df_p, \quad p\in M.$$

For. if $X_p \in \tau(M)_p$ and $h \in C^\infty(g(f(p)))$, then

$$d(g\circ f)_p(X_p)(h) = X_p(h\circ g\circ f) = df_p(X_p)(h\circ g) = dg_{f(p)}(df_p(X_p))(h).$$

In terms of local coordinates this computation exhibits the chain rule and the multiplicative behaviour of Jacobian matrices.

$\diamondsuit$ **Exercise 3.3.** Given a smooth map $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, define an affine map $g : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ by $g(x) = f(p) + df_p(x - p)$. Show that g is the best affine approximation of f in a neighbourhood of p in the sense that it is the only affine map with the property

$$\lim_{x \to p} \frac{f(x) - g(x)}{\|x - p\|} = 0.$$

$\diamondsuit$ **Exercises 3.4.** Prove the following assertions.

(1) If V is a vector space and $p \in V$, then $\tau(V)_p = V$, and if $f : V \longrightarrow W$ is a linear map between vector spaces, then $df_p = f$.
(2) If M is a submanifold of a manifold N, and $i : M \longrightarrow N$ is the inclusion map, then $di_p : \tau(M)_p \longrightarrow \tau(N)_{i(p)}$, $p \in M$, is also the inclusion map.
    Therefore $\tau(M)_p$ is a subspace of $\tau(N)_p$. Moreover, if P is another manifold, and $f : N \longrightarrow P$ is a smooth map, then $d(f|M)_p = df_p|\tau(M)_p$.

$\diamondsuit$ **Exercises 3.5.** Let M, N be manifolds, and $(p, q) \in M \times N$. Prove the following assertions.

(1) $\tau(M \times N)_{(p,q)} = \tau(M)_p \oplus \tau(N)_q$.
(2) If $\pi : M \times N \longrightarrow M$ is the projection $(p, q) \longrightarrow p$, then $d\pi_{(p,q)}$ is also the projection $(v, w) \longrightarrow v$.
(3) If, for a fixed point $q \in N$, $j : M \longrightarrow M \times N$ is the map $p \mapsto (p, q)$, then dj_p is the map $v \longrightarrow (v, 0)$.
(4) If $f : M \longrightarrow M \times M$ is the map $p \longrightarrow (p, p)$, then df_p is the map $v \longrightarrow (v, v)$.
(5) If Δ is the diagonal set of $M \times M$, then $\tau(\Delta)_{(p,p)}$ is the diagonal set of $\tau(M)_p \times \tau(M)_p$.
(6) If $f : M \longrightarrow N$ is a smooth map, and $F : M \longrightarrow M \times N$ is the smooth map $p \longrightarrow (p, f(p))$, then $dF_p(v) = (v, df_p(v))$.
(7) If $f : M \longrightarrow N$ is a smooth map, then the tangent space of the graph of f (which is Image F in (6)) at a point $(p, f(p))$ is the graph of $df_p : \tau(M)_p \longrightarrow \tau(N)_{f(p)}$.

$\diamondsuit$ **Exercise 3.6.** For an n-manifold M and $x \in M$, let $C_x(M)$ be the set of all pairs (c, v), where $c = (U, \phi)$ is a coordinate chart at x, and $v \in \mathbb{R}^n$. Define an equivalence relation in $C_x(M)$ by $(c, v) \sim (c', v')$ if $d(\phi' \circ \phi^{-1})_{\phi(x)}(v) = v'$. Show that

(1) the quotient set $C_x(M)/\sim$ may be identified with $\tau(M)_x$,
(2) if c is a chart at x, then the map $\lambda_c : \mathbb{R}^n \longrightarrow \tau(M)_x$, where $\lambda_c(v) = [(c, v)]$ (the equivalence class of (c, v)), is a linear isomorphism,
(3) if c and c' are charts at x, then $\lambda_{c'}^{-1} \circ \lambda_c = d(\phi' \circ \phi^{-1})_{\phi(x)}$,

(4) if $f : M \longrightarrow N$ is a smooth map, then, for any pair of charts $c = (U, \phi)$ at x and $c' = (U', \phi')$ at $f(x)$ with $f(U) \subset U'$,

$$df_x = \lambda_{c'} \circ d(\phi' \circ f \circ \phi^{-1})_{\phi(x)} \circ \lambda_c^{-1}.$$

In view of Definition 1.4.4, a smooth map $f : M \longrightarrow N$ is an immersion if $\dim M \leq \dim N$ and df_p is an injective map for every $p \in M$. A smooth map $f : M \longrightarrow N$ is a submersion if $\dim M \geq \dim N$ and df_p is a surjective map for every $p \in M$. For a diffeomorphism $f : M \longrightarrow N$, df_p is an isomorphism with inverse $[df_p]^{-1} = [df^{-1}]_p$. We may state the inverse function theorem (Theorem 1.4.3) in a slightly better way as: df_p is an isomorphism if and only if f induces a diffeomorphism of a neighbourhood of p in M onto a neighbourhood of $f(p)$ in N.

Definition 3.1.8. The **cotangent space** $\tau(M)_p^*$ of M at $p \in M$ is the dual of the vector space $\tau(M)_p$. It consists of all linear functions $\tau(M)_p \longrightarrow \mathbb{R}$.

Remark 3.1.9. If U is a coordinate neighbourhood in M at $p \in M$ with coordinates $(x_1, \ldots, x_n)$, then, for a smooth function $f : U \longrightarrow \mathbb{R}$, the differential

$$df_p : \tau(M)_p \longrightarrow \tau(\mathbb{R})_{f(p)} \simeq \mathbb{R}$$

is an element of $\tau(M)_p^*$, where $df_p([\partial/\partial x_i]_p) = (\partial f/\partial x_i)(p)$, identifying $\tau(\mathbb{R})_{f(p)}$ with $\mathbb{R}$ by means of the natural correspondence $\lambda[d/dt]_{f(p)} \longrightarrow \lambda$. In particular, $[dx_i]_p([\partial/\partial x_j]_p) = \delta_{ij}$. Therefore the differentials $[dx_1]_p, \ldots, [dx_n]_p$ form the dual basis of $\tau(M)_p^*$ corresponding to the basis $[\partial/\partial x_1]_p, \ldots, [\partial/\partial x_n]_p$ of $\tau(M)_p$. The derivative map $[dx_j]_p$ maps a tangent vector $v \in \tau(M)_p$ onto its j-th coordinate v_j with respect to the basis $\{[\partial/\partial x_i]_p\}$, and we have $df_p = \sum_{i=1}^n (\partial f/\partial x_i)(p)[dx_i]_p$.

3.2. Vector Fields and Flows

Definition 3.2.1. The **tangent bundle** $\tau(M)$ of M is the disjoint union of all tangent spaces $\tau(M)_p$ as p runs over M.

This is the set of all ordered pairs (p, v) such that $v \in \tau(M)_p$. The map $\pi : \tau(M) \longrightarrow M$, given by $(p, v) \mapsto p$, is called the **projection map** of the tangent bundle. The following theorem shows we can pull back the differential structure on M by π to obtain a unique differential structure on $\tau(M)$.

Theorem 3.2.2. *If M is a manifold of dimension n, then its tangent bundle $\tau(M)$ is a manifold of dimension $2n$.*

Proof. Each chart (U, ϕ) of M determines a map

$$\tau_\phi : \pi^{-1}(U) \longrightarrow \phi(U) \times \mathbb{R}^n \subset \mathbb{R}^n \times \mathbb{R}^n$$

given by $\tau_\phi(p,v) = (\phi(p), d\phi_p(v))$. Clearly, τ_ϕ is a bijection with inverse τ_ϕ^{-1} given by $\tau_\phi^{-1}(a,w) = (p, d\phi_p^{-1}(w))$ where $p = \phi^{-1}(a)$. For two compatible charts (U,ϕ) and (V,ψ) of M, the map

$$\tau_\psi \circ \tau_\phi^{-1} : \phi(U \cap V) \times \mathbb{R}^n \longrightarrow \psi(U \cap V) \times \mathbb{R}^n$$

is given by

$$\begin{aligned} \tau_\psi \circ \tau_\phi^{-1}(a,w) &= \tau_\psi(p, d\phi_p^{-1}(w)) \\ &= (\psi(p), d\psi_p \circ d\phi_p^{-1}(w)) \\ &= (\psi \circ \phi^{-1}(a), d\psi_p \circ d\phi_p^{-1}(w)), \end{aligned}$$

where $p = \phi^{-1}(a)$. Therefore $\tau_\psi \circ \tau_\phi^{-1}$ is a homeomorphism. It follows that $\tau(M)$ has a unique topology which makes each τ_ϕ a homeomorphism. Moreover, since $\tau_\psi \circ \tau_\phi^{-1}$ is a diffeomorphism, the family of charts $\{(\pi^{-1}(U), \tau_\phi)\}$ constitute a smooth atlas on $\tau(M)$. Thus $\tau(M)$ is a smooth manifold. □

◊ **Exercise 3.7.** Complete the proof of the above theorem by showing that $\tau(M)$ is second countable and Hausdorff. Also show that the projection $\pi : \tau(M) \longrightarrow M$ is a smooth map.

◊ **Exercise 3.8.** Show that a smooth map $f : M \longrightarrow N$ between manifolds induces a smooth map $df : \tau(M) \longrightarrow \tau(N)$ which is defined by $df(p,v) = (f(p), df_p(v))$.

Definition 3.2.3. A **vector field** X on M is a map $X : M \longrightarrow \tau(M)$ such that the value of X at $p \in M$ is a tangent vector $X_p \in \tau(M)_p$.

For any $f \in C^\infty(U)$, a vector field X defines a function $Xf : U \longrightarrow \mathbb{R}$ by $(Xf)(p) = X_p(f)$. A vector field X is called a **smooth vector field** if, for every $p \in M$, $f \in C^\infty(p)$ implies $Xf \in C^\infty(p)$ also.

Thus a smooth vector field X may be considered as a map

$$X : C^\infty(M) \longrightarrow C^\infty(M)$$

given by $f \mapsto Xf$. We have

(i) $X(\lambda f + \mu g) = \lambda X f + \mu Y f$,

(ii) $X(fg) = f(Xg) + (Xf)g$,

for $f, g \in C^\infty(M)$, and $\lambda, \mu \in \mathbb{R}$.

◊ **Exercise 3.9.** Show that a smooth vector field X on M is completely determined by its action on smooth functions on M satisfying the above properties (i) and (ii).

◊ **Exercise 3.10.** Show that if f is a constant function, then $Xf = 0$.

The set of all smooth vector fields on M is denoted by $\mathfrak{X}(M)$. This is a module over the ring $C^\infty(M)$, where the module operations are given by

$$(X+Y)f = Xf + Yf, \text{ and } (fX)g = f(Xg),$$

for $X, Y \in \mathfrak{X}(M)$ and $f, g \in C^\infty(M)$.

If (U, ϕ) is a coordinate chart in M with $\phi = (x_1, \ldots, x_n)$, then for each $i = 1, \ldots, n$, the assignment $p \mapsto [\partial/\partial x_i]_p$ is a smooth vector field $\partial/\partial x_i$ on U. The tangent vectors $([\partial/\partial x_i]_p)$ are linearly independent at each point $p \in U$. Therefore, if X is a vector field on U, then X may be written as

$$X = \sum_{i=1}^{n} X x_i \cdot \frac{\partial}{\partial x_i}.$$

The functions Xx_i are called the **components** of X. We shall write $Xx_i = X_i$.

$\Diamond$ **Exercise 3.11.** Show that a vector field X is smooth if and only if its components X_i are smooth for every coordinate system ϕ.

Definition 3.2.4. The **Lie bracket** $[X, Y]$ of vector fields $X = \sum_i X_i \partial/\partial x_i$ and $Y = \sum_i Y_i \partial/\partial x_i$ in local coordinates $(x_1, \ldots, x_n)$ on M is defined as the vector field

$$\sum_{i,j} \left(X_i \frac{\partial Y_j}{\partial x_i} - Y_i \frac{\partial X_j}{\partial x_i} \right) \frac{\partial}{\partial x_j}.$$

Lemma 3.2.5. (1) *We have for a smooth function $f : M \longrightarrow \mathbb{R}$*

$$[X, Y]f = X(Yf) - Y(Xf),$$

(2) $[X, Y]$ *is bilinear over* $\mathbb{R}$ *in* X *and* Y,

(3) $[X, Y] = -[Y, X]$, *and so* $[X, X] = 0$,

(4) *For vector fields X, Y, Z, the Jacobi identity holds,*

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

Proof. (1) In local coordinates with $X = \sum_i X_i \partial/\partial x_i$, $Y = \sum_i Y_i \partial/\partial x_i$, we have

$$[X, Y]f = \sum_{i,j} X_i \frac{\partial Y_j}{\partial x_i} \frac{\partial f}{\partial x_j} - \sum_{i,j} Y_i \frac{\partial X_j}{\partial x_i} \frac{\partial f}{\partial x_j} = X(Yf) - Y(Xf).$$

The other results (2)–(4) may be seen easily. $\square$

In view of Remark 1.6.7, the set of smooth vector fields $\mathfrak{X}(M)$ equipped with Lie brackets is a Lie algebra.

Lemma 3.2.6. *If X is a smooth vector field on an open neighbourhood U in M, and $p \in U$, then there is an open neighbourhood V of p in U, and a smooth vector field $\widehat{X}$ on M which agrees with X on V.*

Proof. Let K be a closed neighbourhood of p in U, and let V be the interior of K. Then, by Smooth Urysohn's Lemma (2.1.17), there is a smooth function $\phi : M \longrightarrow \mathbb{R}$ with support in U such that $\phi = 1$ on K. Then define a vector field $\widehat{X}$ on M by

$$\begin{aligned} \widehat{X}(q) &= \phi(q)X(q) \text{ if } q \in U, \\ &= 0 \text{ if } q \notin U. \end{aligned}$$

Clearly this is the required vector field. □

Definition 3.2.7. If X is a vector field on M, then an **integral curve** (or **solution curve**) of X at $p \in M$ is a smooth curve $\sigma : (-a, a) \longrightarrow M$ such that $\sigma(0) = p$ and $\dot{\sigma}(t) = X_{\sigma(t)}$.

Definition 3.2.8. Let U be an open set in M, and $\epsilon > 0$. Then a **local flow** is a smooth map $\phi : U \times (-\epsilon, \epsilon) \longrightarrow M$ such that, if $\phi_t : U \longrightarrow M$ is the smooth map sending $x \mapsto \phi(x, t)$, then

(a) $\phi_0(x) = x$ for all $x \in U$,

(b) $\phi_{s+t}(x) = \phi_s(\phi_t(x))$ for all $s, t \in (-\epsilon, \epsilon)$ and $x \in U$ for which all the points $(x, t), (x, s + t), (\phi_t(x), s)$ are in $U \times (-\epsilon, \epsilon)$.

The map ϕ is called a **flow** or **dynamical system** on M if the above definition holds when $U \times (-\epsilon, \epsilon)$ is replaced by $M \times \mathbb{R}$. In this case, each ϕ_t is a diffeomorphism with inverse $(\phi_t)^{-1} = \phi_{-t}$. The family $\{\phi_t\}$ therefore forms an abelian group. For this reason, ϕ is also called a **one-parameter group of diffeomorphisms**. Another way of saying this is that ϕ is an action of the Lie group $(\mathbb{R}, +)$ on M.

A flow is called **maximal** if it is not a proper restriction of any flow.

A flow ϕ on M defines a smooth vector field X on M in the following way. If $p \in M$, then X_p is the tangent vector to the curve $t \mapsto \phi_t(p)$ at $t = 0$. Note that X is smooth, because ϕ is smooth with respect to p and t simultaneously. The vector field X constructed in this way is called the **infinitesimal generator** of ϕ, and we say that X generates the flow ϕ. Thus the infinitesimal generator of a flow ϕ is the vector field for which the curves $t \mapsto \phi_t(p)$ are integral curves.

Sometimes the set $\{\phi_t(p)\}_{t\in\mathbb{R}}$ is called the **orbit** through p, noting that the action of ϕ is an action of the Lie group $(\mathbb{R}, +)$ on M.

Lemma 3.2.9. *Let $h : M \longrightarrow N$ be a diffeomorphism, and $\phi : M \times \mathbb{R} \longrightarrow M$ be a flow on M with infinitesimal generator X. Then the vector field $Y = dh(X)$ on N (defined by $Y_q = dh_p(X_p)$, $q \in N$ and $p = h^{-1}(q)$), is the infinitesimal generator of the flow $\psi : N \times \mathbb{R} \longrightarrow N$ given by $\psi = h \circ \phi \circ (h^{-1} \times \mathrm{id})$.*

Proof. Since $q \in N$ and $p = h^{-1}(q)$, we have

$$Y_q = dh_p(X_p) = dh_p\Big(\frac{d}{dt}\phi_t(p)\Big|_{t=0}\Big) = \frac{d}{dt}\Big(h \circ \phi_t(p)\Big)(0).$$

□

In Lemma 3.2.11 below, we shall show that any vector field X on M generates a local flow at each point of M. In the proof of this lemma, and also in some other places, we will have occasions to use the following existence and uniqueness theorem for a system of ordinary differential equations.

Theorem 3.2.10 (Picard's Theorem). *Let U be an open set in $\mathbb{R}^n$ and K a compact set in U. Let $f : U \longrightarrow \mathbb{R}^n$ be a smooth map. Then given a system of differential equations*

$$\frac{du}{dt} = f(u),$$

there is a unique smooth map $\lambda : K \times (-\epsilon, \epsilon) \longrightarrow U$, for some $\epsilon > 0$, such that for each $u_0 \in K$, the function $u : (-\epsilon, \epsilon) \longrightarrow U$ given by $u(t) = \lambda(u_0, t)$ satisfies the differential equations with initial condition $u(0) = u_0$.

It can be shown (see Hurewicz [**17**], p. 28) that the integral curves $u(t)$ exist and are unique, since its components are given by given by

$$u_i(t) = u_i(0) + \int_0^t f_i(u(s))ds, \quad i = 1, \dots, n,$$

where $u_i(0)$ is a constant and s is a dummy variable. The proof requires a local Lipschitz estimate for f, but this holds automatically for a smooth map, as we have already seen in the proof of Lemma 2.3.1.

Lemma 3.2.11. *Given a vector field X on a manifold M without boundary, and a point $p \in M$, there exist an open neighbourhood U of p in M, an $\epsilon > 0$, and a unique local flow $\phi : U \times (-\epsilon, \epsilon) \longrightarrow M$ whose infinitesimal generator is $X|U$.*

Proof. Let (V, h) be a coordinate chart at p with $h(V) = V' \subset \mathbb{R}^n$. Let U be another open set of M with $\overline{U}$ compact and contained in V. Define a vector field Y on V' by $Y = dh(X|V)$. We may write $Y = \sum_i f_i(\partial/\partial u_i)$, where f_i are smooth functions $V' \longrightarrow \mathbb{R}$, and u_i are coordinate functions in $\mathbb{R}^n$. Consider, for each $q = (q_1, \dots, q_n) \in h(\overline{U})$, the system of differential equations

$$\frac{du_i}{dt} = f_i(u)$$

with initial values $u_i(0) = q_i$. Then Theorem 3.2.10 gives a unique smooth map $\lambda : h(\overline{U}) \times (-\epsilon, \epsilon) \longrightarrow V'$, and hence a unique integral curve $\eta_q(t) = \lambda(q, t)$ of Y satisfying $\eta_q(0) = q$. Then, by Lemma 3.2.9, $\sigma_p(t) = h^{-1} \circ \lambda \circ (h \times \mathrm{id})(p, t)$ is an integral curve of X satisfying $\sigma_p(0) = p$, where $p = h^{-1}(q)$.

Define for each $s \in (-\epsilon, \epsilon)$ a smooth map $\phi_s : U \longrightarrow M$ by $\phi_s(p) = \sigma_p(s)$, where σ_p is the unique integral curve of X at p. Then we have

$$\phi_0 = id, \text{ and } \phi_s \circ \phi_t = \phi_{s+t}, \qquad s, t, s+t \in (-\epsilon, \epsilon).$$

The second relation follows from uniqueness, because both

$$\phi_s \circ \phi_t(p) = \sigma_{\sigma_p(t)}(s), \text{ and } \phi_{s+t}(p) = \sigma_p(s+t)$$

are integral curves, and they agree at $s = 0$. This completes the proof. □

◊ **Exercises 3.12.** Let $\mathcal{F}$ be the family of all local flows on M, and $\mathcal{V}$ be the family of vector fields on open subsets of M.

(1) Show that the map $\mathcal{F} \longrightarrow \mathcal{V}$ which takes a local flow ϕ to its infinitesimal generator X_ϕ is surjective, and if ϕ_1 and ϕ_2 are two local flows on M with domain D_1 and D_2 respectively such that $X_{\phi_1} = X_{\phi_2}$ on $D_1 \cap D_2$, then $\phi_1 = \phi_2$ on $D_1 \cap D_2$.

(2) Let $\mathcal{F}_X$ be the family of all local flows on M generated by a vector field $X \in \mathcal{V}$. Define a partial ordering in $\mathcal{F}_X$ by $\phi_1 < \phi_2$ if $D_1 \subset D_2$, where ϕ_1, ϕ_2, D_1, and D_2 are as in (1). Show that $\mathcal{F}_X$ has a unique maximal element. This is the **maximal local flow** generated by X.

Remark 3.2.12. The above lemma is meant for a manifold M which has no boundary. If M has boundary, then the proof remains valid at a non-boundary point p. But, at a boundary point p, the proof may break down, because the domain of the integral curve σ_p of the vector field X may not be an open interval. Explicitly, if $p \in \partial M$ and the tangent vector X_p points into (resp. out of) M (that is, if, in terms of a coordinate chart at p, $X_p = \sum_i \lambda_i \partial/\partial x_i$ has the first coordinate $\lambda_1 > 0$ (resp. $\lambda_1 < 0$)), then the domain of σ_p will be a half-open interval $[0, \epsilon)$ (resp. $(-\epsilon, 0]$) for some $\epsilon > 0$. This may also be the case even if $X_p \in \tau(\partial M)_p$ (that is, if $\lambda_1 = 0$). However, if X is tangent to ∂M, i.e. if $X_p \in \tau(\partial M)_p$ for all $p \in \partial M$, then our arguments will remain all right.

For a manifold M with boundary, the difficulty may be remedied by embedding M in a manifold N without boundary and extending the vector field X on M to a smooth vector field Y on N. Note that Y may be obtained gluing together local smooth extensions of X by means of a smooth partition of unity.

Definition 3.2.13. The **support** of a vector field X on M is the closure of the set of points of M where X does not vanish.

Theorem 3.2.14. *Let X be a smooth vector field with compact support on a manifold M without boundary (that is, X vanishes outside of a compact subset of M), then X generates a unique flow ϕ on M.*

Proof. Let K be a compact subset of M such that $X = 0$ outside of K. Then using Lemma 3.2.11, we can find for each $p \in K$ an open neighbourhood U_p of p in M, an $\epsilon_p > 0$, and a unique local flow $\phi^{(p)} : U_p \times (-\epsilon_p, \epsilon_p) \longrightarrow M$ whose infinitesimal generator is $X|U_p$. Choose a finite subcover of the compact set K by $U_{p_1}, \ldots, U_{p_k}$, and let $\epsilon = \min_{1 \le i \le k} \epsilon_{p_i}$. Then define $\phi_t : M \longrightarrow M$ for $|t| < \epsilon$ by

$$\phi_t(p) = \begin{cases} \phi^{(p_i)}(p, t) & \text{if } p \in U_{p_i}, \\ p & \text{if } p \in M - \cup_i U_{p_i}. \end{cases}$$

Note that if ϕ_t is defined, then $\phi_{t/n}$ is also defined for any integer $n > 1$, and that ϕ_t is the same as the n times iteration of $\phi_{t/n}$. We may use these properties to define ϕ_t for $|t| \ge \epsilon$ in the following way. Choose an integer n

so that $|t/n| < \epsilon$, and define $\phi_t = (\phi_{t/n})^n$. This is well-defined, smooth, and satisfies $\phi_s \circ \phi_t = \phi_{s+t}$. $\square$

Remark 3.2.15. The condition that X has compact support cannot be omitted. For example, the standard vector field d/dt on the open interval $(0,1)$ does not generate a flow on the interval. We do not have the map $\phi : (0,1) \times \mathbb{R} \longrightarrow (0,1)$ given by $\phi(x,t) = x + t$.

$\diamondsuit$ **Exercises 3.13.** (i) Show that on a compact manifold M without boundary there is a one-one correspondence between vector fields and flows.

(ii) Let ϕ be a flow generated by a vector field X on M, and $p \in M$. Show that $X_p = 0$ if and only if $\phi_t(p) = p$ for all t.

3.3. Exterior algebra

The vector spaces that we consider in this section are finite dimensional and over the field of real numbers $\mathbb{R}$. The letters V, W will denote vector spaces. The vector space of linear maps $V \longrightarrow W$ is denoted by $L(V,W)$. Then $L(V,\mathbb{R})$ is the dual V^* of V. The vector space of multilinear maps

$$V \times \cdots \times V \times V^* \times \cdots \times V^* \longrightarrow \mathbb{R},$$

where V appears k times, and V^* appears r times, is denoted by $L_k^r(V)$. The multilinearity means linearity in each coordinate when the other coordinates are held fixed.

An element of $L_k^r(V)$ is called a **mixed tensor of degree** (k,r) or, a tensor of covariant degree k, and contravariant degree r. A tensor is pure if $r = 0$ or $k = 0$. The elements of $L_k^0(V)$ and $L_0^r(V)$ are respectively called **covariant k-tensors** and **contravariant r-tensors**. Note that $L_1^0(V) = V^*$, and $L_0^1(V) = V$, . We shall take $L_0^0 = \mathbb{R}$.

In this section we shall be concerned only with pure covariant tensors, and therefore by a k-tensor we shall always mean a covariant k-tensor.

Definition 3.3.1. The **tensor product** of a k-tensor ω and a s-tensor η is a $(k+s)$-tensor, denoted by $\omega \otimes \eta$. It is defined by

$$(\omega \otimes \eta)(v_1, \dots, v_{k+s}) = \omega(v_1, \dots, v_k) \cdot \eta(v_{k+1}, \dots, v_{k+s}),\ v_i \in V.$$

This gives a bilinear map

$$\otimes : L_k^0(V) \times L_s^0(V) \longrightarrow L_{k+s}^0(V).$$

It may be seen easily that the product is not commutative. But it is associative, and it distributes over sum.

Proposition 3.3.2. *If $(\lambda_1, \dots, \lambda_n)$ is a basis of V^*, then the k-tensors*

$$\lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k}, \quad 1 \leq i_1, \dots, i_k \leq n,$$

constitute a basis over $\mathbb{R}$ of the vector space $L_k^0(V)$, and therefore

$$\dim(L_k^0(V)) = n^k.$$

Proof. Let $(u_1, \ldots, u_n)$ be the basis of V which is dual to the basis $(\lambda_1, \ldots, \lambda_n)$ of V^*, so that $\lambda_i(u_j) = \delta_{i,j}$ (Kronecker delta). Let I denote a sequence of k indices $1 \le i_1, \ldots, i_k \le n$ (repetitions allowed), and let u_I and λ_I be defined by

$$u_I = (u_{i_1}, \ldots, u_{i_k}), \text{ and } \lambda_I = \lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k}.$$

Then we have for a pair of index sequences I and J, each consisting of k indices,

$$\lambda_I(u_J) = 1 \text{ if } I = J, \text{ and } \lambda_I(u_J) = 0 \text{ if } I \neq J.$$

This follows because

$$\lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k}(u_{j_1}, \ldots, u_{j_k}) = \delta_{i_1 j_1} \cdots \delta_{i_k j_k}.$$

It also follows from multilinearity that any k-tensor ω can be written as a linear combination over $\mathbb{R}$ of the λ_I

$$\omega = \sum_I a_I \lambda_I,$$

where $a_I = \omega(u_I) \in \mathbb{R}$. Also two k-tensors ω and η are equal if and only if $\omega(u_I) = \eta(u_I)$ for every index sequence I. This implies that the λ_I are linearly independent. $\square$

A tensor ω in $L_k^0(V)$ is called **symmetric** if the value $\omega(v_1, \ldots, v_k)$, remains the same after any permutation of the variables $v_1, \ldots, v_k$, and ω is called **alternating** or **skew symmetric**, if the sign of $\omega(v_1, \ldots, v_k)$ is changed whenever any two of the variables $v_1, \ldots, v_k$ are interchanged. The collection of all alternating tensors form a linear subspace of $L_k^0(V)$, which we denote by $\Lambda^k(V)$. By convention all 1-tensors are alternating, and therefore $\Lambda^1(V) = V^*$. We set $\Lambda^0(V) = \mathbb{R}$.

Example 3.3.3. The map $\det : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \longrightarrow \mathbb{R}$ (n copies of $\mathbb{R}^n$), sending $(v_1, \ldots, v_n) \mapsto \det(v_1, \ldots, v_n)$, where $\det(v_1, \ldots, v_n)$ is the determinant of the matrix whose column vectors are the components of $v_1, \ldots, v_n$ with respect to the standard basis of $\mathbb{R}^n$, is an alternating n-tensor in $\Lambda^n(\mathbb{R}^n)$.

Definition 3.3.4. If $\phi_1, \phi_2, \ldots, \phi_k$ are elements in V^*, then a **monomial** $\phi_1 \cdot \phi_2 \cdots \phi_k$ of degree k is an alternating k-tensor defined by

$$\phi_1 \cdot \phi_2 \cdots \phi_k(v_1, v_2, \ldots, v_k) = \frac{1}{k!} \det(\phi_i(v_j)), \quad v_i \in V,$$

where on the right hand side we have the determinant of the $k \times k$ matrix $(\phi_i(v_j))$.

Proposition 3.3.5. *Let $(\lambda_1, \ldots, \lambda_n)$ be a basis of V^*. Then the monomials*

$$\lambda_{i_1} \cdots \lambda_{i_k}, \qquad 1 \le i_1 < \cdots < i_k \le n$$

constitute a basis over $\mathbb{R}$ of the vector space $\Lambda^k(V)$, and therefore

$$\dim \Lambda^k(V) = \binom{n}{k}.$$

PROOF. The proof is similar to that of Proposition 3.3.2, the only difference is that this time the index sequences should be increasing. The basis of $\Lambda^k(V)$ is given by the λ_I, where I is an increasing index sequence $1 \leq i_1 < \cdots < i_k \leq n$. $\square$

Let S_k denote the symmetric group of all permutations of the numbers from 1 to k. If $\sigma \in S_k$, set sgn $\sigma = +1$ or -1 depending on whether σ is representable as a product of an even or odd number of transpositions. This gives a homomorphism of S_k into the multiplicative group of two elements ± 1. Define an action (or representation) $S_k \times L_k^0(V) \longrightarrow L_k^0(V)$ of S_k on $L_k^0(V)$ by $(\sigma, \omega) \mapsto \omega^\sigma$, where ω^σ is a k-tensor defined by

$$\omega^\sigma(v_1, \ldots, v_k) = \omega(v_{\sigma(1)}, \ldots, v_{\sigma(k)}),\ v_i \in V.$$

Define an operator Alt : $\mathrm{L}_\mathrm{k}^0(\mathrm{V}) \longrightarrow \mathrm{L}_\mathrm{k}^0(\mathrm{V})$ by

$$\mathrm{Alt}(\omega) = \frac{1}{\mathrm{k}!} \sum_{\sigma \in \mathrm{S_k}} \mathrm{sgn}\ \sigma \cdot \omega^\sigma.$$

◊ **Exercise 3.14.** Show that the monomial of Definition 3.3.4 may be written as

$$\phi_1 \cdot \phi_2 \cdots \phi_k = \mathrm{Alt}(\phi_1 \otimes \cdots \otimes \phi_\mathrm{k}).$$

Proposition 3.3.6. *The operator* Alt *has the following properties.*

(1) Alt *is a linear map.*

(2) $(\mathrm{Alt}(\omega))^\sigma = \mathrm{sgn}\ \sigma \cdot \mathrm{Alt}(\omega)$, $\sigma \in S_k$.

(3) *The image of* Alt *is* $\Lambda^k(V)$.

(4) *The restriction of* Alt *to* $\Lambda^k(V)$ *is the identity map.*

(5) $\mathrm{Alt}^2 = \mathrm{Alt}$ (*that is,* Alt *is a projection operator*).

PROOF. The proofs consist of routine verifications. We shall prove only the last three statements.

(3) We have $\mathrm{sgn}^2 = 1$, and sgn $\sigma =$ sgn $\sigma\tau \cdot$ sgn τ. Also, as σ runs through S_k, so does $\sigma\tau$. Therefore

$$\begin{aligned}
\mathrm{Alt}(\omega)(v_{\tau(1)}, \ldots, v_{\tau(k)}) &= \frac{1}{k!} \sum_\sigma \mathrm{sgn}\ \sigma \cdot \omega(v_{\sigma\tau(1)}, \ldots, v_{\sigma\tau(k)}) \\
&= \frac{1}{k!}\ \mathrm{sgn}\ \tau \sum_\sigma \mathrm{sgn}\ \sigma\tau \cdot \omega(v_{\sigma\tau(1)}, \ldots, v_{\sigma\tau(k)}) \\
&= \mathrm{sgn}\ \tau \cdot \mathrm{Alt}(\omega)(v_1, \ldots, v_k).
\end{aligned}$$

Therefore $\mathrm{Alt}(\omega)$ is alternating.

(4) If ω is alternating, then sgn $\sigma \cdot \omega^\sigma = \omega$. Since S_k has $k!$ elements, summing over all $\sigma \in S_k$ we get $\mathrm{Alt}(\omega) = \omega$.

(5)

$$\begin{aligned}\mathrm{Alt}(\mathrm{Alt}(\omega)) &= \mathrm{Alt}\Big[\frac{1}{\mathrm{k}!}\sum_{\sigma}\mathrm{sgn}\ \sigma\cdot\omega^{\sigma}\Big]\\ &= \frac{1}{k!}\sum_{\sigma}\mathrm{sgn}\ \sigma\cdot\mathrm{Alt}(\omega^{\sigma}) = \frac{1}{\mathrm{k}!}\sum_{\sigma}(\mathrm{sgn}\ \sigma)^2\cdot\mathrm{Alt}(\omega) = \mathrm{Alt}(\omega)).\end{aligned}$$

Therefore $\mathrm{Alt}^2 = \mathrm{Alt}$ □

Definition 3.3.7. If $\omega \in \Lambda^k(V)$ and $\eta \in \Lambda^r(V)$ then their **wedge product** or **exterior product** $\omega \wedge \eta$ is defined by

$$\omega \wedge \eta = \frac{(k+r)!}{k!r!}\mathrm{Alt}(\omega \otimes \eta).$$

This gives a map $\Lambda^k(V) \otimes \Lambda^r(V) \longrightarrow \Lambda^{k+r}(V)$.

We shall denote the normalising factor $(k+r)!/k!r!$ by the letter R.

Lemma 3.3.8. *The wedge product $\wedge$ is bilinear.*

Proof. The wedge product is the composition of a bilinear map and a linear map

$$\Lambda^k(V) \times \Lambda^r(V) \longrightarrow L^0_{k+r}(V) \longrightarrow \Lambda^{k+r}(V),$$

where the first one is $\otimes$, and the second one is $R{\cdot}\mathrm{Alt}$. Therefore $\wedge$ is bilinear. □

Lemma 3.3.9. *If for a tensor ω, $\mathrm{Alt}(\omega) = 0$, then $\omega \wedge \eta = 0 = \eta \wedge \omega$, for any tensor η, and therefore $\mathrm{Alt}(\omega \otimes \eta) = \mathrm{Alt}(\eta \otimes \omega) = 0$ for any η.*

Proof. The symmetric group S_{k+r} of permutations of $\{1, 2, \dots, k+r\}$ has a subgroup G consisting of permutations which leave the last r integers $k+1, \dots, k+r$ fixed. The group G is isomorphic to the symmetric group S_k by the correspondence which sends $\sigma \in G$ to the permutation σ' obtained by restricting σ to $\{1, \dots, k\}$. Clearly, we have $\mathrm{sgn}\ \sigma = \mathrm{sgn}\ \sigma'$, and $(\omega \otimes \eta)^{\sigma} = \omega^{\sigma'} \otimes \eta$. Therefore

$$\sum_{\tau\in G}\mathrm{sgn}\ \tau\cdot(\omega\otimes\eta)^{\tau} = \Big[\sum_{\tau'\in \mathrm{S_k}}\mathrm{sgn}\ \tau'\cdot\omega^{\tau'}\Big]\otimes\eta = \mathrm{Alt}(\omega)\otimes\eta = 0.$$

Now the group S_{k+r} is a disjoint union of right cosets $G\cdot\sigma = \{\tau\sigma \mid \tau \in G\}$, and for each coset

$$\sum_{\tau\in G}\mathrm{sgn}\ \tau\sigma\cdot(\omega\otimes\eta)^{\tau\sigma} = \mathrm{sgn}\ \sigma\Big[\sum_{\tau\in\mathrm{G}}\mathrm{sgn}\ \tau\cdot(\omega\otimes\eta)^{\tau}\Big]^{\sigma} = 0.$$

Since $\omega \wedge \eta = R \cdot \mathrm{Alt}(\omega \otimes \eta)$ is the sum of these partial sums over the right cosets of G, we have $\omega \wedge \eta = 0$. Similarly $\eta \wedge \omega = 0$. □

Proposition 3.3.10. *The wedge product $\wedge$ is associative,*

$$(\omega \wedge \eta) \wedge \theta = \omega \wedge (\eta \wedge \theta).$$

Proof. Let $\omega \in \Lambda^k(V)$, $\eta \in \Lambda^r(V)$, and $\theta \in \Lambda^s(V)$. We shall show that each of $(\omega \wedge \eta) \wedge \theta$ and $\omega \wedge (\eta \wedge \theta)$ is equal to $S \cdot \mathrm{Alt}(\omega \otimes \eta \otimes \theta)$, where S is the integer

$$\frac{(k+r+s)!}{k!\ r!\ s!}.$$

Let

$$T = \frac{(k+r+s)!}{(k+r)!\ s!}, \quad \text{and as before} \quad R = \frac{(k+r)!}{k!\ r!}.$$

Then $S = TR$.

By definition $(\omega \wedge \eta) \wedge \theta = T \cdot \mathrm{Alt}((\omega \wedge \eta) \otimes \theta)$ Therefore, by linearity of Alt,

$$(\omega \wedge \eta) \wedge \theta - S \cdot \mathrm{Alt}(\omega \otimes \eta \otimes \theta) = \mathrm{T} \cdot \mathrm{Alt}([\omega \wedge \eta - \mathrm{R}(\omega \otimes \eta)] \otimes \theta).$$

Now

$$\mathrm{Alt}(\omega \wedge \eta - \mathrm{R}(\omega \otimes \eta)) = \mathrm{Alt}(\omega \wedge \eta) - \mathrm{R} \cdot \mathrm{Alt}(\omega \otimes \eta) = \omega \wedge \eta - \omega \wedge \eta = 0.$$

So by the above lemma, $\mathrm{Alt}([\omega\wedge\eta-\mathrm{R}(\omega\otimes\eta)]\otimes\theta) = 0$, and we have $(\omega\wedge\eta)\wedge\theta = S \cdot \mathrm{Alt}(\omega \otimes \eta \otimes \theta)$. The proof of the other part is similar. $\square$

Proposition 3.3.11. *The wedge product $\wedge$ is anti-commutative. This means that if $\omega \in \Lambda^k(V)$ and $\eta \in \Lambda^r(V)$, then*

$$\omega \wedge \eta = (-1)^{kr} \eta \wedge \omega.$$

Proof. It is enough to show that

$$\mathrm{Alt}(\omega \otimes \eta) = (-1)^{\mathrm{kr}} \mathrm{Alt}(\eta \otimes \omega).$$

To see this note that $\mathrm{Alt}(\omega \otimes \eta)(v_1, \ldots, v_{k+r})$

$$\begin{aligned} &= \frac{1}{(k+r)!} \sum_\sigma \mathrm{sgn}\ \sigma \cdot \omega(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) \cdot \eta(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+r)}) \\ &= \frac{1}{(k+r)!} \sum_\sigma \mathrm{sgn}\ \sigma \cdot \eta(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+r)}) \cdot \omega(v_{\sigma(1)}, \ldots, v_{\sigma(k)}). \end{aligned}$$

Now, if τ is the permutation

$$\{1, \ldots, k, k+1, \ldots, k+r\} \longrightarrow \{k+1, \ldots, k+r, 1, \ldots, k\},$$

then $\mathrm{sgn}\ \tau = (-1)^{\mathrm{kr}}$. Then, since $\mathrm{sgn}\ \sigma = \mathrm{sgn}\ \sigma\tau \cdot \mathrm{sgn}\ \tau$, and

$$\sigma\tau(j) = \sigma(k+j)\ (1 \le j \le r), \quad \sigma\tau(r+j) = \sigma(j)\ (1 \le j \le k),$$

we have

$$\begin{aligned} \mathrm{Alt}(\omega \otimes \eta)(v_1, \ldots, v_{k+r}) = &\frac{1}{(k+r)!} \textstyle\sum_\sigma \mathrm{sgn}\ \sigma\tau \cdot \mathrm{sgn}\ \tau \\ &\cdot (\eta(v_{\sigma\tau(1)}, \ldots, v_{\sigma\tau(r)}) \cdot \omega(v_{\sigma\tau(r+1)}, \ldots, v_{\sigma\tau(r+k)})) \\ = &\ \mathrm{sgn}\ \tau \cdot \mathrm{Alt}(\eta \otimes \omega)(v_1, \ldots, v_{k+r}). \end{aligned}$$

This completes the proof. $\square$

The vector spaces $\Lambda^k(V)$ are defined for $1 \leq k \leq n$, where $\dim V = n$. Also, we have $\Lambda^0(V) = \mathbb{R}$, which may be interpreted as the vector space of constant functions $V \longrightarrow \mathbb{R}$. We define the wedge product $\Lambda^0(V) \wedge \Lambda^k(V) \longrightarrow \Lambda^k(V)$ simply as the multiplication of a vector by a scalar: $f \wedge \omega = f \cdot \omega$. Note that if the number of indices in I is greater than n, then at least one of the indices in I must repeat, and thus $\lambda_I = 0$. This means that $\Lambda^k(V) = 0$ for $k > n$. The direct sum of these $n+1$ vector spaces is denoted by $\Lambda(V)$

$$\Lambda(V) = \Lambda^0(V) \oplus \Lambda^1(V) \oplus \cdots \oplus \Lambda^n(V).$$

This with the wedge product is a graded algebra, called the **exterior algebra** or **Grassmann algebra** over V. It is associative and anti-commutative, and has identity element $1 \in \Lambda^0(V)$.

Definition 3.3.12. A linear map $f : V \longrightarrow W$ induces a map

$$f^* : \Lambda(W) \longrightarrow \Lambda(V)$$

between the exterior algebras, where, for each $k \geq 0$, $f^* : \Lambda^k(W) \longrightarrow \Lambda^k(V)$ is defined in the following way. If $\omega \in \Lambda^k(W)$, then $f^*\omega \in \Lambda^k(V)$ is given by

$$(f^*\omega)(v_1, \ldots, v_k) = \omega(fv_1, \ldots, fv_k)$$

for all vectors $v_1, \ldots, v_k \in V$.

Note that for $k = 1$, $f^* : W^* \longrightarrow V^*$ is just the transpose map which sends $\phi \in W^*$ to $\phi \circ f \in V^*$.

It can be easily checked that f^* is linear and that

$$f^*(\omega_1 \wedge \omega_2) = f^*\omega_1 \wedge f^*\omega_2.$$

Moreover, if $f : V_1 \longrightarrow V_2$ and $g : V_2 \longrightarrow V_3$ are linear maps between vector spaces, then $(g \circ f)^* = f^* \circ g^*$. In particular, if f is an isomorphism, so is f^*.

Theorem 3.3.13. *Let* $\dim V = n$, *and* $f : V \longrightarrow V$ *be a linear isomorphism. Then* $f^* : \Lambda^n(V) \longrightarrow \Lambda^n(V)$ *is given by* $f^*\omega = (\det f) \cdot \omega$, *for every* $\omega \in \Lambda^n(V)$. *In particular, if* $\omega_1, \ldots, \omega_n \in \Lambda^1(V)$, *then*

$$f^*\omega_1 \wedge \cdots \wedge f^*\omega_n = (\det f) \cdot \omega_1 \wedge \cdots \wedge \omega_n.$$

Proof. Since $\dim(\Lambda^n(V)) = 1$, the linear map $f^* : \Lambda^n(V) \longrightarrow \Lambda^n(V)$ is multiplication by some scalar $\lambda \in \mathbb{R}$, so that $f^*(\omega) = \lambda\omega$ for all $\omega \in \Lambda^n(V)$. Now det is an element of $\Lambda^n(\mathbb{R}^n)$. Therefore, choosing a linear isomorphism $g : V \longrightarrow \mathbb{R}^n$, we find that $g^*(\det) \in \Lambda^n(V)$, and therefore $f^*g^*(\det) = \lambda g^*(\det)$. This gives

$$(g^*)^{-1}f^*g^*(\det) = \lambda(g^*)^{-1}g^*(\det) = \lambda(gg^{-1})^*(\det) = \lambda(\det),$$

or $(gfg^{-1})^*(\det) = \lambda(\det)$. Evaluating both sides of this equation on the standard basis $e_1, \ldots, e_n$ of $\mathbb{R}^n$, we get $\lambda = det(gfg^{-1}) = \det f$. Note that it follows from the definition of the alternating n-tensor det in Example 3.3.3 that for any linear map $h : \mathbb{R}^n \longrightarrow \mathbb{R}^n$

$$\det(he_1, \ldots, he_n) = \det h.$$

$\square$

3.4. Differential forms

In this section we shall consider exterior algebra associated to the tangent space at each point of a manifold.

Definition 3.4.1. Let M be a manifold with or without boundary. Then the **k-th exterior power bundle** $\Lambda^k(M)$ of M is the disjoint union of the vector spaces $\Lambda^k(\tau(M)_p)$ as p varies over M. The map $\pi : \Lambda^k(M) \longrightarrow M$, which carries the vector space $\Lambda^k(\tau(M)_p)$ onto the point p, is called the **projection map**.

Lemma 3.4.2. *If M is a manifold of dimension n, then the exterior power bundle $\Lambda^k(M)$ is a manifold of dimension $n + \binom{n}{k}$.*

PROOF. The proof is parallel to that of Theorem 3.2.2. A coordinate chart $\phi : U \longrightarrow \mathbb{R}^n$ in M produces isomorphisms

$$d\phi_p : \tau(M)_p \longrightarrow \mathbb{R}^n \text{ and } \lambda_p = (d\phi_p)^* : \Lambda^k(\mathbb{R}^n) \longrightarrow \Lambda^k(\tau(M)_p).$$

Then a coordinate chart for $\Lambda^k(M)$ is given by $\psi : \Lambda^k(U) \longrightarrow \phi(U) \times \Lambda^k(\mathbb{R}^n)$, where $\Lambda^k(U) = \pi^{-1}(U)$ and $\psi(p, \omega) = (\phi(p), \lambda_p^{-1}(\omega))$. It is easy to verify that these coordinate charts do actually define a smooth atlas for $\Lambda^k(M)$. $\square$

Definition 3.4.3. A **differential k-form** on M is a smooth map

$$\omega : M \longrightarrow \Lambda^k(M)$$

such that $\pi \circ \omega = \mathrm{id}$. Thus ω assigns to each point $p \in M$ an alternating k-tensor $\omega_p \in \Lambda^k(\tau(M)_p)$. The integer k is called the *degree* of ω, and we write $\deg \omega = k$.

$\Diamond$ **Exercise 3.15.** Show that the map ω is smooth if and only if for every set of smooth vector fields $X_1, \ldots, X_k$, the function $\omega(X_1, \ldots, X_k)$ defined on the intersection of the domains of $X_1, \ldots, X_k$ by

$$\omega(X_1, \ldots, X_k)(p) = \omega_p(X_{1p}, \ldots, X_{kp}),$$

is smooth, where X_{ip} is the value of X_i at p.

Since we are only interested in smooth forms, a k-form will implicitly mean a differentiable (or smooth) k-form. A 0-form is just a real-valued smooth function on M.

Let $\Omega^k(M)$ denote the space of k-forms on M. The **sum** of two k-forms ω_1 and ω_2 is a k-form $\omega_1 + \omega_2$, where

$$(\omega_1 + \omega_2)(p) = \omega_1(p) + \omega_2(p).$$

The product of a function $f \in C^\infty(M)$ and a k-form $\omega \in \Omega^k(M)$ is the k-form $f\omega$, where

$$(f\omega)(p) = f(p) \cdot \omega(p).$$

The **wedge product** of a k-form ω and an s-form η is the $(k+s)$-form $\omega \wedge \eta$ given by $(\omega \wedge \eta)(p) = \omega(p) \wedge \eta(p)$. Then the **anti-commutative law**

$$\omega \wedge \eta = (-1)^{ks} \eta \wedge \omega$$

follows from the anti-commutativity at each point.

The wedge product makes the space $\Omega(M) = \Omega^0(M) \oplus \cdots \oplus \Omega^n(M)$, $n = \dim M$, a graded algebra, which is called the exterior algebra of differential forms on M.

◊ **Exercise 3.16.** The space $\Lambda(M) = \Lambda^0(M) \oplus \Lambda^1(M) \oplus \cdots \oplus \Lambda^n(M)$, where $n = \dim M$, is the disjoint union of the vector spaces $\oplus_{k=0}^{n} \Lambda^k(\tau(M)_p)$, where p varies over M. Show that $\Lambda(M)$ is a manifold.

The manifold $\Lambda^1(M)$ is the union of cotangent spaces $\tau(M)_p^*$ as p varies over M. It is called the **cotangent bundle** of M, and is also denoted by $\tau(M)^*$. As described in Remark 3.1.9, any $f \in C^\infty(M)$ gives rise to a smooth 1-form df whose value at p is df_p. In a local coordinate system $(x_1, \ldots, x_n)$, we have

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \cdot dx_i,$$

and in terms of the same coordinate system, local coordinates in $\Lambda^k(M)$ are

$$x_1, \ldots, x_n, \text{ and } \binom{n}{k} \text{ monomials } dx_{i_1} \cdots dx_{i_k},$$

where $1 \leq i_1 < \cdots < i_k \leq n$, and a k-form ω has representation

$$\omega = \sum a_{i_1 \cdots i_k} dx_{i_1} \cdots dx_{i_k},$$

where $a_{i_1 \cdots i_k}$ are real-valued smooth functions on the coordinate neighbourhood whose value at p is

$$\omega_p \left(\left[\frac{\partial}{\partial x_{i_1}} \right]_p, \ldots, \left[\frac{\partial}{\partial x_{i_k}} \right]_p \right).$$

Definition 3.4.4. Given a smooth map $f : M \longrightarrow N$ and a k-form $\omega \in \Omega^k(N)$, the **pull-back** of ω by f is a k-form $f^*\omega$ on M given by

$$(f^*\omega)(p) = (df_p)^* \omega(f(p)),$$

where $(df_p)^*$ is the transpose of the differential $df_p : \tau(M)_p \longrightarrow \tau(N)_{f(p)}$.

If ω is a 0-form, that is, a smooth function on N, then $f^*\omega$ is the smooth function $\omega \circ f$ on M.

◊ **Exercise 3.17.** Show that

(i) f^* is a linear map $\Omega^k(N) \longrightarrow \Omega^k(M)$,

(ii) $f^*(\omega \wedge \eta) = (f^*\omega) \wedge (f^*\eta)$,

(iii) $\mathrm{id}^* = \mathrm{id}$, and $(g \circ f)^*\omega = f^* g^* \omega$, where f and g are composable smooth maps.

Explicit expression of the pullback $f^*\omega$ of a k-form ω on N by a smooth map $f : M \longrightarrow N$ in terms of local coordinates $(x_1, \ldots, x_n)$ at $p \in M$ and $(y_1, \ldots, y_m)$ at $f(p) \in N$ may be found out easily. First let us look at the action of f^* on the basic 1-forms dy_i. Since $(df_p)^* : \tau(N)^*_{f(p)} \longrightarrow \tau(M)^*_p$ is the transpose of the differential $df_p : \tau(M)_p \longrightarrow \tau(N)_{f(p)}$, the matrix of $([df]_p)^*$ is the transpose of the Jacobian matrix $(\partial f_i/\partial x_j)$ of f at p. Consequently,

$$f^* dy_i = \sum_j \frac{\partial f_i}{\partial x_j} dx_j = df_i.$$

Therefore, if $\omega = \sum_I a_I dy_I$, then, by the properties of f^* listed in the above exercise, we have

$$f^*\omega = \sum_I f^*(a_I) df_I,$$

where $f^*(a_I) = a_I \circ f$ and $df_I = df_{i_1} \wedge \cdots \wedge df_{i_k}$ for the index sequence $I = \{i_1, \ldots, i_k\}$. Also note that, by the above expression for f^*dy_i and Theorem 3.3.13, we have

$$f^*(dy_1 \wedge \cdots \wedge dy_n) = df_1 \wedge \cdots \wedge df_n = \det\left(\frac{\partial f_i}{\partial x_j}\right) dx_1 \wedge \cdots \wedge dx_n.$$

3.5. Derivations of algebra of differential forms

In this section we shall consider three types of derivations of the exterior algebra $\Omega(M)$. These are exterior, interior, and Lie derivatives. Although they are commonly called derivatives, the first two are actually anti-derivatives. Here is the precise definition.

Definition 3.5.1. A linear operator $D : \Omega(M) \longrightarrow \Omega(M)$ is called a **derivation** or **anti-derivation** according to whether $D(\omega \wedge \eta)$ is equal to

$$D(\omega) \wedge \eta + \omega \wedge D(\eta),$$

or

$$D(\omega) \wedge \eta + (-1)^{\deg \omega}\, \omega \wedge D(\eta).$$

Let us first consider the exterior derivative.

Definition 3.5.2. If U is an open set in $\mathbb{R}^n$ or $\mathbb{R}^n_+$, then the **exterior derivative operator** on $\Omega(U)$ is a sequence of linear maps

$$d : \Omega^k(U) \longrightarrow \Omega^{k+1}(U),\ k \geq 0,$$

defined in the following way.

(a) If f is a function in $\Omega^0(U)$, then df is the differential of f

$$df = \sum_{i=1}^n (\partial f_i/\partial x_i) dx_i,$$

(b) If $\omega = \sum f_I dx_I$ is a k-form in $\Omega^k(U)$, then $d\omega = \sum\ df_I \wedge dx_I$.

Example 3.5.3. In $\mathbb{R}^3$ there is an identification between 0-forms or smooth functions and 3-forms

$$f \leftrightarrow f dxdydz.$$

The identification between vector fields and 1-forms is given by

$$X = (f_1, f_2, f_3) \leftrightarrow f_1 dx + f_2 dy + f_3 dz;$$

1-forms also correspond bijectively with 2-forms

$$f_1 dx + f_2 dy + f_3 dz \leftrightarrow f_1 dydz + f_2 dzdx + f_3 dxdy.$$

The exterior derivative of a function corresponds to the gradient vector field

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz \leftrightarrow \nabla f = \Big(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\Big).$$

The exterior derivative of an 1-form corresponds to the curl

$$\begin{gathered} d(f_1 dx + f_2 dy + f_3 dz) \leftrightarrow \nabla \times f \quad \text{(vector product)} \\ = \Big(\frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z}\Big) dydz + \Big(\frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial x}\Big) dzdx + \Big(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}\Big) dxdy. \end{gathered}$$

The exterior derivative of a 2-form corresponds to divergence

$$d(f_1 dydz + f_2 dzdx + f_3 dxdy) \leftrightarrow \nabla \cdot f = \Big(\frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z}\Big).$$

Proposition 3.5.4 (Anti-derivative condition). *If ω and η are forms on U, then*

$$d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^{\deg \omega} \omega \wedge d\eta.$$

PROOF. It is sufficient to prove this for monomials $\omega = a_I dx_I$ and $\eta = b_J dx_J$.

$$\begin{aligned} d(\omega \wedge \eta) &= d(a_I b_J dx_I \wedge dx_J) \\ &= d(a_I b_J) \wedge dx_I \wedge dx_J \\ &= [(da_I) b_J + a_I (db_J)] \wedge dx_I \wedge dx_J \\ &= da_I \wedge dx_I \wedge b_J dx_J + (-1)^{\deg \omega} a_I dx_I \wedge db_J \wedge dx_J \\ &= d\omega \wedge \eta + (-1)^{\deg \omega} \omega \wedge d\eta. \end{aligned}$$

□

Proposition 3.5.5 (Cocycle condition). $d(d\omega) = 0$ *for any form ω on U.*

PROOF. If f is a function, then

$$d(df) = d\Big(\sum_i \frac{\partial f}{\partial x_i} dx_i\Big) = \sum_{i,j} \frac{\partial^2 f}{\partial x_j \partial x_i} dx_j \wedge dx_i = 0.$$

The sum is zero, because the mixed partial derivatives remains the same if i and j are interchanged.

If ω is a monomial $\omega = a_I dx_I$, then we have by Proposition 3.5.4

$$d(d\omega) = d(da_I \wedge dx_I) = d(da_I) \wedge dx_I - da_I \wedge d(dx_I).$$

This is zero, because $d(df) = 0$. □

Proposition 3.5.6 (Uniqueness). *If $D : \Omega^k(U) \longrightarrow \Omega^{k+1}(U)$, $k \geq 0$, is a sequence of linear maps satisfying anti-derivative and cocycle conditions such that $Df = df$ for functions $f \in \Omega^0(U)$, then $D = d$.*

Proof. For any smooth function f, we have

$$D(df) = D(Df) = 0,$$

and so by the anti-derivative condition

$$\begin{aligned} D(dx_I) &= D(dx_{i_1} \wedge \cdots \wedge dx_{i_k}) \\ &= \textstyle\sum_r \pm\, dx_{i_1} \wedge \cdots \wedge D(dx_{i_r}) \wedge \cdots \wedge dx_{i_k} = 0. \end{aligned}$$

Therefore, for $\omega = a_I dx_I$,

$$\begin{aligned} D\omega &= Da_I \wedge dx_I + a_I D(dx_I) \\ &= Da_I \wedge dx_I = da_I \wedge dx_I = d\omega. \end{aligned}$$

□

Proposition 3.5.7. *If $\phi : M \longrightarrow N$ is a smooth map, then*

$$d(\phi^*\omega) = \phi^*(d\omega).$$

Proof. Since ϕ^* is a homomorphism and d is an anti-derivative, the problem can be reduced to considering individual factors of a monomial, namely, to functions and differentials of functions. Now, for a smooth function f on N, we have

$$\phi^* df = df \circ d\phi = d(f \circ \phi) = d(\phi^* f), \text{ and}$$
$$d(\phi^* df) = d(d(f \circ \phi)) = 0 = \phi^*(0) = \phi^*(d(df)).$$

□

Definition 3.5.8. If ω is a k-form on a manifold M with boundary, then the exterior derivative $d\omega$ is defined in the following way. For each coordinate chart (U, ϕ) in M, define $d\omega$ on U by

$$d\omega|U = \phi^* d[(\phi^{-1})^*\omega].$$

Then $d\omega$ is well-defined, because the local definitions agree on the overlaps of the coordinate neighbourhoods. To see this, take another coordinate chart (V, ψ) in M with $U \cap V \neq \emptyset$, and consider the diffeomorphism $\eta = \phi \circ \psi^{-1}$. Then

$$\eta^* d[(\phi^{-1})^*\omega] = d[\eta^*(\phi^{-1})^*\omega] = d[(\psi^{-1})^*\omega].$$

Therefore

$$d\omega|V = \psi^* d[(\psi^{-1})^*\omega] = \psi^*\eta^* d[(\phi^{-1})^*\omega] = \phi^* d[(\phi^{-1})^*\omega] = d\omega|U.$$

◊ **Exercise 3.18.** Show that if M is a manifold with boundary, then the exterior derivative operator $d : \Omega(M) \longrightarrow \Omega(M)$ has all the properties of the exterior derivative on $\mathbb{R}^n$ or $\mathbb{R}^n_+$

We shall now define the interior derivative.

Definition 3.5.9. Let X be a vector field on M. Then the **interior derivative** along X (also called **interior multiplication** or **contraction** with X) is a linear map

$$i(X) : \Omega(M) \longrightarrow \Omega(M)$$

such that

(a) $i(X)[\Omega^0(M)] = 0$,

(b) for $k \geq 1$, $i(X) : \Omega^k(M) \longrightarrow \Omega^{k-1}(M)$ is defined by

$$(i(X)\omega)(X_1, \ldots, X_{k-1}) = \omega(X, X_1, \cdots, X_{k-1}),$$

for all vector fields $X_1, \ldots, X_{k-1}$ on M.

Lemma 3.5.10. (a) $i(X) \circ i(Y) = -i(Y) \circ i(X)$, *and hence* $i(X) \circ i(X) = 0$,

(b) *if* $\omega \in \Omega^1(M)$, *then* $i(X)\omega = \omega(X)$,

(c) $i(X)$ *is an anti-derivation:*

$$i(X)(\omega \wedge \eta) = (i(X)\omega) \wedge \eta + (-1)^{\deg \omega} \omega \wedge (i(X)\eta),$$

(d) *the operation* $i(X)$ *is uniquely defined.*

PROOF. The first three parts are immediate from the definition. Part (d) may be seen easily from the proof of Proposition 3.5.6. □

◊ **Exercise 3.19.** Prove the following general formula for 1-forms $\omega_1, \ldots, \omega_k$,

$$i(X)(\omega_1 \wedge \cdots \wedge \omega_k) = \sum_{i=1}^{k} (-1)^{i-1} \omega_i(X) \omega_1 \wedge \cdots \wedge \widehat{\omega}_i \wedge \cdots \wedge \omega_k,$$

where $\widehat{\omega}_i$ indicates that ω_i is omitted.

◊ **Exercise 3.20.** Let $\phi : M \longrightarrow N$ be a smooth map. Let X be a smooth vector field on M, and Y a smooth vector field on N such that $d\phi_p(X_p) = Y_{\phi(p)}$ for all $p \in M$. Let ω be a k-form on N. Then prove the following commutation law:

$$i(X)(\phi^* \omega) = \phi^*(i(Y)\omega).$$

Definition 3.5.11. Let X be a smooth vector field on M, and ω denote a smooth k-form, $k \geq 0$, on M. Then the **Lie derivative** of ω with respect to X, denoted by $\mathcal{L}_X(\omega)$, is a smooth k-form whose value at $x \in M$ is given in the following way. Let U be an open neighbourhood of x, and $\phi : U \times (-\epsilon, \epsilon) \longrightarrow M$ be the local flow whose infinitesimal generator is $X|U$. Then

$$\mathcal{L}_X(\omega)_x = \lim_{t \to 0} \frac{1}{t} [\phi_t^* \omega(x) - \omega(x)] = \frac{d}{dt}(\phi_t^* \omega)(0),$$

where $\phi_t^*\omega(x)(X_1, \cdots, X_k) = \omega(\phi(x,t))(d\phi_t(X_1), \cdots, d\phi_t(X_k))$.

The definition is independent of U, since ϕ_t is unique.

The definition also holds when ω is replaced by a vector field, or in general a tensor field.

Theorem 3.5.12. *The Lie derivation $\mathcal{L}_X$ has the following properties.*

(1) $\mathcal{L}_X : \Omega(M) \longrightarrow \Omega(M)$ is $\mathbb{R}$-linear (but not $C^\infty(M)$-linear)

$$\mathcal{L}_X(a\omega + b\eta) = a\mathcal{L}(\omega) + b\mathcal{L}(\eta),$$

where a and b are real numbers.

(2) If $f \in C^\infty(M)$, then $\mathcal{L}_{fX}\omega = df \wedge i(X)\omega + f\mathcal{L}_X\omega$.

(3) If f is a 0-form or function, then $\mathcal{L}_X(f) = df(X) = X(f)$.

(4) $(\mathcal{L}_X(df))(Y) = d(Xf)(Y)$.

(5) $\mathcal{L}_X(Y) = [X, Y]$.

(6) The Lie derivative commutes with the exterior derivative:

$$\mathcal{L}_X d = d\mathcal{L}_X.$$

(7) If $\omega \in \Omega^k(M)$ and $\eta \in \Omega^r(M)$, then

$$\mathcal{L}_X(\omega \wedge \eta) = (\mathcal{L}_X\omega) \wedge \eta + \omega \wedge (\mathcal{L}_X\eta).$$

(8) If T and S are covariant tensor fields, then

$$\mathcal{L}_X(T \otimes S) = (\mathcal{L}_X T) \otimes S + T \otimes (\mathcal{L}_X S).$$

(9) (**Cartan formula**) $\mathcal{L}_X\omega = i(X)d\omega + d(i(X)\omega)$.

Proof. The proofs consist of checking the statements against the definition. For example, the properties (3), (4), (5) , and (9) may be obtained as follows. Others may be proved similarly.

(3)

$$\begin{aligned}
(\mathcal{L}_X f)(x) &= \left(\lim_{t\to 0} \frac{\phi_t^* f - f}{t}\right)(x) \\
&= \lim_{t\to 0} \frac{f(\phi_t(x)) - f(x)}{t} \\
&= \lim_{t\to 0} \frac{f(\phi(x,t)) - f(x)}{t} \\
&= (\phi_x)'_0(f) = X_x(f) = Xf(x).
\end{aligned}$$

(4)

$$\begin{aligned}
(\mathcal{L}_X df)(Y) &= \lim_{t\to 0} \frac{\phi_t^* df - df}{t}(Y) \\
&= \lim_{t\to 0} \frac{d(\phi_t^* f) - df}{t}(Y) = d(\mathcal{L}_X f)(Y) = d(Xf)(Y).
\end{aligned}$$

(5) First note that

$$\begin{aligned}
\mathcal{L}_X(Yf) &= \lim_{t\to 0}\frac{\phi_t^*(Yf)-Yf}{t} = \lim_{t\to 0}\frac{(\phi_t^*Y)(\phi_t^*f)-Yf}{t}\\
&= \lim_{t\to 0}\left((\phi_t^*Y)\frac{\phi_t^*f-f}{t}+\frac{\phi_t^*Y-Y}{t}f\right)\\
&= Y(\mathcal{L}_Xf)+(\mathcal{L}_XY)f.
\end{aligned}$$

But $\mathcal{L}_Xf = Xf$, by (3). So $X(Yf) = \mathcal{L}_X(Yf) = Y(Xf) + (\mathcal{L}_XY)f$, and hence $(\mathcal{L}_XY)f = X(Yf) - Y(Xf) = [X,Y]f$.

(9) The formula is true when the form ω is a function f or a differential df. In fact, $i(X)df + d(i(X)df) = i(X)df + 0 = \mathcal{L}_Xf$, by (3), and $i(X)ddf + d(Xf) = 0 + df(X) = d(Xf) = \mathcal{L}_Xdf$, by (4).

The formula is also true for a general differential form ω. This will follow from the above results by a straightforward induction. □

◊ **Exercise 3.21.** Show that

$$\mathcal{L}_X\ i(Y) - i(Y)\mathcal{L}_X = i([X,Y]).$$

In particular, the Lie derivative $\mathcal{L}_X$ commutes with the interior derivative $i(X)$.

3.6. Darboux-Weinstein theorems

This section is devoted to reducing certain differential forms to canonical forms. We first consider the symplectic forms.

Let us recall some facts about bilinear form on a vector space V.

If $(e_1,\dots,e_n)$ is a basis of V, then any bilinear form $\omega : V\times V \longrightarrow \mathbb{R}$ gives rise to a matrix $A = (a_{ij})$, where $a_{ij} = \omega(e_i,e_j)$. This matrix determines the bilinear form uniquely. For, if $u = \sum u_ie_i$, and $v = \sum v_je_j$, then

$$\omega(u,v) = \sum a_{ij}u_iv_j.$$

The kernel of ω is given by

$$\operatorname{Ker}\omega = \{u\in V \mid \omega(u,v) = 0 \text{ for all } v\in V\}.$$

The bilinear form ω is called **non-degenerate** if $\operatorname{Ker}\omega = 0$, i.e. if the map

$$\overline{\omega} : V \longrightarrow V^*,$$

given by $\overline{\omega}(u)(v) = \omega(u,v)$, is a linear isomorphism. A subspace W of V is called non-degenerate, if, for every $u\in W$, $\omega(u,v) = 0$ for all $v\in V$ implies $u = 0$. We say that ω is null, if $\omega(u,v) = 0$ for all $u,v\in V$

The bilinear form ω is alternating if its matrix A is skew-symmetric.

Lemma 3.6.1. *Let V be a vector space of dimension n, and ω an alternating bilinear form on V. Then there is a basis $(v_1,\dots,v_n)$ of V^* and an integer $k\le n/2$ depending on ω such that*

$$\omega = v_1\wedge v_2 + v_3\wedge v_4 + \cdots + v_{2k-1}\wedge v_{2k}.$$

PROOF. The complement W of $\operatorname{Ker}\omega$ is non-degenerate. Also, W cannot be one-dimensional, otherwise ω will be null. Therefore W contains two non-zero vectors u and v such that $\omega(u,v) \neq 0$, dividing by some constant we may assume that $\omega(u,v) = 1$. If P is the plane generated by u and v, then the matrix of $\omega|P$ with respect to the basis (u,v) is

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

We have then $W = P \oplus P^{\perp}$, where $P^{\perp}$ is the complement of P in W, and it is non-degenerate. Repeating the arguments a finite number of times, we get a basis of V with respect to which the matrix of ω is

$$\begin{pmatrix} J & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & J & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & J & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix},$$

where J is the 2×2 matrix given above, and it occurs k times, where $k \leq n/2$. Now choosing a dual basis, we may get the required expression for ω. $\square$

The number $2k$ is called the **rank** of ω; it is also the dimension of Image $(\overline{\omega})$. If ω is non-degenerate, then $\overline{\omega}$ is an isomorphism, and $\dim V = 2k$.

A **symplectic vector space** is a pair (V,ω) consisting of a finite dimensional vector space V, and an alternating non-degenerate bilinear form ω on V. Thus the dimension of a symplectic vector space is necessarily even.

Two symplectic vector spaces (V_1,ω_1) and (V_2,ω_2) are isomorphic if there is a linear isomorphism $f : V_1 \longrightarrow V_2$ such that $f^*\omega_2 = \omega_1$. It will follow from the next theorem that any two symplectic vector spaces of the same dimension are always isomorphic.

Theorem 3.6.2. *If (V,ω) is a symplectic vector space, then there is a basis*

$$(e_1,\ldots,e_n,f_1,\ldots,f_n)$$

of V such that for all $1 \leq i,j \leq n$

$$\omega(e_i,e_j) = \omega(f_i,f_j) = 0, \quad \omega(e_i,f_j) = \delta_{ij}.$$

If $(u_1,\ldots,u_n,v_1,\ldots,v_n)$ is the basis of V^* dual to the above basis of V, then

$$\omega = u_1 \wedge v_1 + u_2 \wedge v_2 + \cdots + u_n \wedge v_n.$$

The matrix of $\overline{\omega} : V \longrightarrow V^*$ with respect to the above bases of V and V^* is

$$J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}.$$

Proof. The proof is an easy consequence of the above lemma. □

We now turn to manifolds.

Definition 3.6.3. A 2-form ω on a manifold M is called **closed** if $d\omega = 0$. It is called **non-degenerate** if each pair $(\tau(M)_x, \omega(x))$ is a symplectic vector space, or equivalently each $\overline{\omega}(x) : \tau(M)_x \longrightarrow \tau(M)^*_x$ is a linear isomorphism for all $x \in M$.

A non-degenerate closed 2-form ω on M is called a **symplectic form** or **symplectic structure** on M, and the pair (M, ω) is called a **symplectic manifold**.

Thus a non-degenerate 2-form ω on M sets up a bijection between vector fields X on M and 1-forms θ on M, where X and θ are related by

$$\theta = i(X)\omega.$$

This means that for any vector field Y on M, we have $\theta(Y) = \omega(X, Y) = [i(X)\omega](Y)$.

Example 3.6.4. In $\mathbb{R}^{2n}$ with coordinates $(x_1, \dots, x_n, y_1, \dots, y_n)$, the alternating bilinear form

$$\omega_0 = \sum_{i=1}^{n} dx_i \wedge dy_i$$

is a symplectic form on $\mathbb{R}^{2n}$, by Theorem 3.6.2. Also we have

$$(\omega_0)^n = \omega_0 \wedge \cdots \wedge \omega_0 = c \cdot dx_1 \wedge \cdots \wedge dx_n \wedge dy_1 \wedge \cdots \wedge dy_n,$$

where c is the constant $n!\ (-1)^{n(n-1)/2}$. Therefore the n-fold wedge product of ω_0 is never zero.

The form ω_0 is called the **canonical symplectic form** on $\mathbb{R}^{2n}$.

Two symplectic manifolds (M, ω) and (M', ω') are called **equivalent** if there is a diffeomorphism $\phi : (M, \omega) \longrightarrow (M', \omega')$ such that $\phi^*\omega' = \omega$. Such a diffeomorphism is called a **symplectic diffeomorphism**. In general, it is not known whether two given symplectic manifolds are equivalent. However, Darboux's theorem asserts that locally any two symplectic manifolds are equivalent.

Theorem 3.6.5 (Darboux). *At each point x of a symplectic manifold (M, ω) without boundary there is a coordinate chart $\phi : U \longrightarrow \mathbb{R}^{2n}$ such that $\phi(x) = 0$ and $\phi^*\omega_0 = \omega \mid U$, where ω_0 is the canonical symplectic form on $\mathbb{R}^{2n}$.*

The theorem was first proved by Darboux using induction on the dimension (see Sternberg [**47**]). Here we shall present a generalised version of the theorem which is due to Weinstein [**53**]. The proof is based on Moser's argument on the isotopy of symplectic forms.

Theorem 3.6.6 (**Darboux-Weinstein**). *Let M be a submanifold of a manifold N without boundary, and let ω_0 and ω_1 be two symplectic forms on N such that $\omega_0|M = \omega_1|M$. Then there is a diffeomorphism $\phi : U \longrightarrow N$, where U is a neighbourhood of M in N, such that $\phi|M = \mathrm{Id}$, and $\phi^*\omega_1 = \omega_0$ on U.*

The proof uses a result of advanced calculus of forms which we prove first.

Lemma 3.6.7. *Let M be a submanifold of a manifold N. Let $\{\phi_t\}$ be a one parameter family of smooth maps from M to N. Let X_t be a smooth family of vector fields $X_t : M \longrightarrow \tau(N)$ so that $X_t(x)$, $x \in M$, is the tangent vector to the curve $t \mapsto \phi_t(x)$ at $\phi_t(x)$:*

$$\frac{d}{dt}\phi_t = X_t \circ \phi_t, \;\; \phi_0 = \mathrm{Id}.$$

Then, for a one parameter family of differential forms $\{\omega_t\}$ on N, we have

$$\frac{d}{dt}\phi_t^*\omega_t = \phi_t^*\frac{d\omega_t}{dt} + \phi_t^*(i(X_t)d\omega_t) + d\phi_t^*(i(X_t)\omega_t).$$

Note that the right hand side is equal to

$$\phi_t^*\frac{d\omega_t}{dt} + \phi_t^*\mathcal{L}_{X_t}\omega_t,$$

by Cartan's formula (Theorem 3.5.12(9).

Proof. The proof is due to Guillemin and Sternberg [**13**]. First we consider a special case of the lemma when $M = N = \mathbb{R} \times Z$ (Z is a manifold), and for ϕ_t we take the map $\psi_t : \mathbb{R} \times Z \longrightarrow \mathbb{R} \times Z$ is given by

$$\psi_t(s, x) = (s + t, x),$$

and for ω_t we take any differential form θ_t on $\mathbb{R} \times Z$.

Now, let, for $s \in \mathbb{R}$, $i_s : Z \longrightarrow \mathbb{R} \times Z$ be the injection $p \mapsto (s, p)$, and $\pi : \mathbb{R} \times Z \longrightarrow Z$ be the projection onto Z. Then, since $\pi \circ i_s = \mathrm{Id}$, we have a direct sum decomposition

$$\Omega^2(\mathbb{R} \times Z) = \text{Image } \pi^* \oplus \operatorname{Ker} i_s^*,$$

where Image π^* can be identified with $\Lambda^2(Z)$, and $\operatorname{Ker} i_s^*$ consists of all forms of the type $ds \wedge \alpha$ with α a 1-form on Z. Thus any 2-form θ_t on $\mathbb{R} \times Z$ can be written uniquely as $\theta_t = (ds \wedge \alpha) + \beta$, where α and β are forms on Z. In terms of local coordinates $s, x_1, \ldots, x_n$ in $\mathbb{R} \times Z$, $n = \dim Z$, the forms α and β look like

$$\alpha = \sum_i a_i(s, x, t)dx_i, \text{ and } \beta = \sum_{i,j} b_{ij}(s, x, t)dx_i \wedge dx_j,$$

where x is the multi-variable $(x_1, \ldots, x_n)$. We may therefore suppose without loss of generality that

$$\theta_t = ds \wedge a(s, x, t)d\xi + b(s, x, t)d\eta,$$

where $a(s,x,t)$ and $b(s,x,t)$ are functions of s,x,t, $d\xi$ is a dx_i, and $d\eta$ is one of $dx_i \wedge dx_j$. Then

$$\psi_t^*\theta_t = ds \wedge a(s+t,x,t)d\xi + b(s+t,x,t)d\eta,$$

$$\begin{aligned}\tfrac{d\psi_t^*\theta_t}{dt} = \quad & ds \wedge \frac{\partial a}{\partial s}(s+t,x,t)d\xi + \frac{\partial b}{\partial s}(s+t,x,t)d\eta \\ & + ds \wedge \frac{\partial a}{\partial t}(s+t,x,t)d\xi + \frac{\partial b}{\partial t}(s+t,x,t)d\eta,\end{aligned}$$

$$\frac{d\theta_t}{dt} = ds \wedge \frac{\partial a}{\partial t}(s,x,t)d\xi + \frac{\partial b}{\partial t}(s,x,t)d\eta,$$

$$\psi_t^*\Big(\frac{d\theta_t}{dt}\Big) = ds \wedge \frac{\partial a}{\partial t}(s+t,x,t)d\xi + \frac{\partial b}{\partial t}(s+t,x,t)d\eta.$$

$$\text{(1)} \qquad \frac{d\psi_t^*\theta_t}{dt} - \psi_t^*\Big(\frac{d\theta_t}{dt}\Big) = ds \wedge \frac{\partial a}{\partial s}(s+t,x,t)d\xi + \frac{\partial b}{\partial s}(s+t,x,t)d\eta.$$

Also the tangent vector field along the curve $\psi_t(s,x)$ is $\dfrac{\partial}{\partial s}$, and

$$i\Big(\frac{\partial}{\partial s}\Big)\theta_t = a(s,x,t)d\xi.$$

Therefore

$$\psi_t^*\Big(i\Big(\frac{\partial}{\partial s}\Big)\theta_t\Big) = a(s+t,x,t)d\xi,$$

and so

$$\text{(2)} \qquad d\psi_t^*\Big(i\Big(\frac{\partial}{\partial s}\Big)\theta_t\Big) = \frac{\partial a}{\partial s}(s+t,x,t)ds \wedge d\xi + d_x a(s+t,x,t)d\xi,$$

where the symbol $d_x f$ denotes the the part of the differential of a function f on $\mathbb{R} \times Z$, which does not involve ds so that $df = (\partial f/\partial s)ds + d_x f$. Similarly, we have

$$d\theta_t = -ds \wedge d_x a(s,x,t)d\xi + \frac{\partial b}{\partial s}(s,x,t)ds \wedge d\eta + d_x b(s,x,t)d\eta,$$

by Proposition 3.5.4, and Definition 3.5.2. Therefore

$$i\Big(\frac{\partial}{\partial s}\Big)d\theta_t = -d_x a(s,x,t)d\xi + \frac{\partial b}{\partial s}(s,x,t)d\eta,$$

and so

$$\text{(3)} \qquad \psi_t^*\Big(i\Big(\frac{\partial}{\partial s}\Big)d\theta_t\Big) = -d_x a(s+t,x,t)d\xi + \frac{\partial b}{\partial s}(s+t,x,t)d\eta.$$

Subtracting the sum of the left hand sides of (2) and (3) from that of (1) we get zero. This proves

$$\text{(4)} \qquad \frac{d\psi_t^*\theta_t}{dt} = \psi_t^*\Big(\frac{d\theta_t}{dt}\Big) + d\psi_t^*\Big(i\Big(\frac{\partial}{\partial s}\Big)\theta_t\Big) + \psi_t^*\Big(i\Big(\frac{\partial}{\partial s}\Big)d\theta_t\Big),$$

which is the required equation in this special case.

For the general case, consider the map $\phi : \mathbb{R} \times M \longrightarrow N$ given by $\phi(s,x) = \phi_s(x)$ Then the curves $s \to \phi_s(x)$ in N are the images under ϕ of the lines parallel to $\mathbb{R}$ in $\mathbb{R} \times M$ through x. In other words

$$d\phi\Big(\frac{\partial}{\partial s}\Big)_{(t,x)} = X_t(x).$$

Let $\lambda : M \longrightarrow \mathbb{R} \times M$ be given by $\lambda(x) = (0,x)$. Then we can write ϕ_t as $\phi \circ \psi_t \circ \lambda$, where $\psi_t : \mathbb{R} \times M \longrightarrow \mathbb{R} \times M$ is as in the first part. Then

$$\phi_t^* \omega_t = \lambda^* \psi_t^* \phi^* \omega_t, \tag{5}$$

where ω_t is a form on N. This gives, since λ is independent of t,

$$\frac{d}{dt}\phi_t^* \omega_t = \lambda^* \frac{d}{dt}\psi_t^*(\phi^* \omega_t). \tag{6}$$

Then at a point $(t,x) \in \mathbb{R} \times M$, we have by a commutation law for interior derivative (Exercise 3.20, p. 93),

$$i\Big(\frac{\partial}{\partial s}\Big)\phi^* \omega_t = \phi^*\Big(i\Big(d\phi\frac{\partial}{\partial s}\Big)\omega_t\Big) = \phi^* i(X_t)\omega_t. \tag{7}$$

Now writing the formula (4) for $\theta_t = \phi^* \omega_t$, and applying λ^* on both sides, we get

$$\begin{aligned}\lambda^* \frac{d}{dt}\psi_t^*(\phi^*\omega_t) &= \lambda^* \psi_t^* \Big(\frac{d(\phi^*\omega_t)}{dt}\Big) \\ &\quad + \lambda^* d\psi_t^*\Big(i\Big(\frac{\partial}{\partial s}\Big)\phi^*\omega_t\Big) + \lambda^* \psi_t^*\Big(i\Big(\frac{\partial}{\partial s}\Big)d(\phi^*\omega_t)\Big).\end{aligned}$$

This is the required formula, in view of (5), (6), (7), and the commutation law for exterior derivative (Proposition 3.5.7). □

The following theorem will be used in the proof of Darboux-Weinstein theorem.

Theorem 3.6.8. *Let M be a submanifold of a manifold N without boundary. Then there is an open neighbourhood U of M in N (called a tubular neighbourhood) and a strong deformation retraction $\psi_t : U \longrightarrow U$ of U onto M.*

The proof is a consequence of the tubular neighbourhood theorem. We shall defer the proof until Chapter 7.

Proof of the Darboux-Weinstein theorem. Suppose that W is a tubular neighbourhood of M in N with strong deformation retraction $\psi_t : W \longrightarrow W$ so that $\psi_0 : W \longrightarrow M$, $\psi_1 = \mathrm{Id}$, and $\psi_t(x) = x$ for all $x \in M$. Then $\dot{\psi}_t : W \longrightarrow \tau(W)$ is a vector field whose value $\dot{\psi}_t(x)$ at $x \in W$ is the tangent vector to the curve $t \to \psi_t(x)$ at $\psi_t(x)$. Also if α is a k-form on W, then $\psi_t^*[i(\dot{\psi}_t)\alpha]$, and hence its integral

$$\int_0^1 \psi_t^*[i(\dot{\psi}_t)\alpha]dt,$$

is a $(k-1)$-form on W. (This is the integration with respect to the variable $t \in [0,1]$. For example, if $\psi_t^*[i(\dot{\psi}_t)\alpha] = a(x,t)\, dx_{i_1} \cdots dx_{i_{k-1}}$, then $I(\alpha) = \big(\int_0^1 a(x,t)dt\big)\, dx_{i_1} \cdots dx_{i_{k-1}}$. If a does contain t, then the integral is 0.) This defines a linear operator

$$I : \Omega^k(W) \longrightarrow \Omega^{k-1}(W)$$

by $I(\alpha) = \int_0^1 \psi_t^*[i(\dot{\psi}_t)\alpha]dt$.

Since $d\alpha/dt = 0$ (α is independent of t), we have by Lemma 3.6.7

$$\frac{d}{dt}[\psi_t^*\alpha] = \psi_t^*[i(\dot{\psi}_t)d\alpha] + d[\psi_t^*(i(\dot{\psi}_t)\alpha].$$

Integrating

$$\psi_1^*\alpha - \psi_0^*\alpha = \int_0^1 \psi_t^*[i(\dot{\psi}_t)d\alpha]dt + d\int_0^1 \psi_t^*[i(\dot{\psi}_t)\alpha]dt,$$

or,

$$\alpha - \psi_0^*\alpha = d(I(\alpha)) + I(d\alpha).$$

Now take $\alpha = \omega_1 - \omega_0$. Then, since $\alpha|M = 0$, $\psi_0^*\alpha = 0$. Therefore

$$\alpha = d(I(\alpha)) + I(d\alpha) = d(I(\alpha)),$$

since $d\alpha = 0$ (α being closed).

Thus we find a 1-form $\beta = I(\alpha)$ on W such that $\beta|M = 0$, and $\omega_1 = \omega_0 + d\beta$ on W. Since the restriction of the 2-form $\sigma_t = \omega_0 + t(d\beta)$ to M is a symplectic form (β being zero on M), for all $t \in \mathbb{R}$, there exists a smaller neighbourhood V of M, $V \subset W$, on which σ_t is still symplectic for all t. This means that

$$\overline{\sigma}_t(x) : \tau(N)_x \longrightarrow \tau(N)_x^*$$

is an isomorphism for all $x \in V$. Therefore we find a family of vector fields X_t on V such that $i(X_t)\sigma_t = -\beta$. By Cartan's formula

$$\mathcal{L}_{X_t}\sigma_t = i(X_t)d\sigma_t + d(i(X_t)\sigma_t).$$

Since $d\sigma_t = 0$, and $d(i(X_t)\sigma_t) = -d\beta = \omega_0 - \omega_1$, we have

$$\mathcal{L}_{X_t}\sigma_t + \frac{d\sigma_t}{dt} = 0, \tag{8}$$

by the definition of σ_t. Since $X_t = 0$ on M, there is a smaller neighbourhood U of M, $U \subset V$, on which X_t can be integrated to a time dependent embedding $\phi_t : U \longrightarrow N$ with $\phi_t(x) = x$ for $x \in M$. Then, by Lemma 3.6.7 and the equation (8), we have

$$\frac{d}{dt}(\phi_t^*\sigma_t) = \phi_t^*(\mathcal{L}_{X_t}\sigma_t + \frac{d\sigma_t}{dt}) = 0.$$

This means that $\phi_t^*\sigma_t = \phi_0^*\sigma_0 = \sigma_0$ on U, that is, $\phi_1^*\omega_1 = \omega_0$ on U. □

Having done these, we now turn to contact structures on manifolds. A contact structure is an odd dimension analogue of symplectic structure.

Definition 3.6.9. A 1-from ω on a manifold M of dimension $2n+1$ is called a **contact form** if, for each $x \in M$, the alternating bilinear map $d\omega(x)$ on $\tau(M)_x$ is non-degenerate on the kernel of $\omega(x) : \tau(M)_x \longrightarrow \mathbb{R}$. The union ξ of all hyperplanes $\xi(x) = \operatorname{Ker}\omega(x)$ as x varies over M is called a **hyperplane field** or **hyperplane distribution** on M. This assigns to each $x \in M$ the hyperplane $\xi(x) \subset \tau(M)_x$.

The non-degeneracy condition may be expressed by saying that $\omega \wedge (d\omega)^n$ is non-zero, because, by our previous discussion on symplectic form, we have

$$(d\omega \text{ is non-degenerate on } \xi) \Leftrightarrow ((d\omega)^n \neq 0 \text{ on } \xi) \Leftrightarrow \omega \wedge (d\omega)^n \neq 0.$$

Thus we may say that a contact form on M is a 1-form ω on M such that $\omega \wedge (d\omega)^n$ is nowhere vanishing on M.

Two contact forms ω and ω' on M are called **equivalent** if $\omega' = f \cdot \omega$, where f is a non-zero function on M. An equivalence class of contact forms is called a **contact structure** on M. Note that any two contact forms in the same contact structure determine the same hyperplane field. A pair (M, ω), where ω is any contact form representing a contact structure on M, is called a **contact manifold**.

Two contact manifolds (M, ω) and (M', ω') are called **diffeomorphic** if there is a diffeomorphism $\phi : M \longrightarrow M'$ such that $\phi^*\omega' = f \cdot \omega$, where f is a non-zero function on M.

Example 3.6.10. In $\mathbb{R}^{2n+1}$ with coordinates

$$(x_1, \ldots, x_n, y_1, \ldots, y_n, z)$$

the 1-form $\omega_0 = dz - \sum_i y_i dx_i$ is a contact form. By simple computations we have

$$d\omega_0 = \sum_i dx_i \wedge dy_i,$$

$$(d\omega_0)^n = c \cdot dx_1 \wedge \cdots \wedge dx_n \wedge dy_1 \wedge \cdots \wedge dy_n,$$

$$\omega_0 \wedge (d\omega_0)^n = c \cdot dx_1 \wedge \cdots \wedge dx_n \wedge dy_1 \wedge \cdots \wedge dy_n \wedge dz,$$

where c is the constant appearing in Example 3.6.4. Therefore $\omega_0 \wedge (\omega_0)^n$ is nowhere zero.

The form ω_0 is called the **canonical contact form** on $\mathbb{R}^{2n+1}$.

We now introduce the **Reeb vector field** on a contact manifold (M, ω). Note that on a vector space V with an alternating bilinear form α, the α-orthogonal complement of a subspace W of V is the subspace

$$W^\perp = \{v \in V \mid \alpha(v, w) = 0 \text{ for all } w \in W\}.$$

Then, we have $\dim(W^\perp) = \dim(V) - \dim(\overline{\alpha}(W))$. If α is non-degenerate on W, that is, $\overline{\alpha}|W$ is an isomorphism, then $\dim(\overline{\alpha}(W)) = \dim(W)$, and so

$$\dim(V) = \dim(W) + \dim(W^\perp).$$

Applying this to the hyperplane field $\xi = \operatorname{Ker}\omega$ of the contact form ω on M, we find that, for each $x \in M$,

$$\tau(M)_x = \xi_x \oplus \operatorname{Ker} d\omega(x),$$

where $\operatorname{Ker} d\omega(x) = \{u \in \tau(M)_x \mid d\omega(u,v) = 0 \text{ for all } v \in \tau(M)_x\}$ (this is actually $\operatorname{Ker} \overline{d\omega}(x)$). Therefore, by the one-one correspondence between vector fields and 1-forms induced by a non-degenerate 2-form, there is a unique vector field Y on M, called the Reeb vector field, determined by the conditions

$$i(Y)d\omega = 0, \text{ and } \omega(Y) = 1.$$

The first condition says that the value $Y(x)$ belongs to the one dimensional subspace $\operatorname{Ker} d\omega(x)$, for every $x \in M$. The second condition normalises Y. Note that, in view of Cartan's formula, the conditions imply that

$$\mathcal{L}_Y\omega = di(Y)\omega + i(Y)d\omega = 0.$$

Lemma 3.6.11. *Let M be a contact manifold with contact form ω and corresponding Reeb vector field Y. Then for any 1-form α on M such that $\alpha(Y) = 0$, there is a unique vector field X on M which takes values in* $\operatorname{Ker}\omega$ *such that $\alpha = i(X)d\omega$.*

Proof. Since $d\omega$ is non-degenerate on $\xi = \operatorname{Ker}\omega$, it determine a unique vector field X corresponding to the 1-form $\alpha|\xi$ on ξ so that $\alpha = i(X)d\omega$. Then $\alpha(Y) = (i(X)d\omega)(Y) = d\omega(X,Y) = 0$. Note that the uniqueness of X follows from the non-degeneracy of $d\omega$ on ξ. $\square$

We are now in a position to present a version of Darboux's theorem for contact forms. It is required to show that given a family of contact forms ω_t on M, there is a family of diffeomorphisms ϕ_t of M such that

$$\phi_t^*\omega_t = f_t\omega_0, \tag{9}$$

for some nowhere vanishing family of functions f_t on M. As before we need to find a family of vector fields X_t so that

$$\frac{d}{dt}\phi_t = X_t \circ \phi_t, \qquad \phi_0 = \mathrm{Id},$$

This means that we want to determine ϕ_t as the flow of X_t. Now differentiating (9) with respect to t, and using Lemma 3.6.7, we get

$$\phi_t^*\Big(\frac{d}{dt}\omega_t + \mathcal{L}_{X_t}\omega_t\Big) = \frac{df_t}{dt}\cdot\frac{1}{f_t}\phi_t^*\omega_t = g_t\phi_t^*\omega_t, \qquad g_t = \frac{1}{f_t}\frac{df_t}{dt}.$$

Applying $(\phi_t^*)^{-1}$ on both the sides of this,

$$\frac{d}{dt}\omega_t + \mathcal{L}_{X_t}\omega_t = h_t\omega_t, \qquad h_t = (\phi_t^*)^{-1}\circ g_t \circ \phi_t^*. \tag{10}$$

Therefore our problem reduces to finding both X_t and h_t so that (10) holds. To solve Equation (10), suppose that Y_t is the Reeb vector field for the given contact form ω_t, and take

$$h_t = i(Y_t)\frac{d\omega_t}{dt}.$$

Let α be the 1-form

$$\alpha = h_t\omega_t - \frac{d\omega_t}{dt}.$$

Then $\alpha(Y_t) = 0$, because $\omega_t(Y_t) = 1$, and $(d\omega_t/dt)(Y_t) = i(Y_t)(d\omega_t/dt) = h(t)$ by Lemma 3.5.10(b). Therefore, by Lemma 3.6.11, there is a unique vector field X_t taking values in $\operatorname{Ker}\omega_t$ such that $\alpha = i(X_t)d\omega_t$, or

$$\frac{d\omega_t}{dt} = h_t\omega_t - i(X_t)d\omega_t.$$

Since $d(i(X_t)\omega_t) = 0$, the last equation gives (10) in view of Cartan's formula. We have therefore proved the following theorem.

Theorem 3.6.12 (Darboux). *A contact manifold of dimension* $2n+1$ *is locally diffeomorphic to* $\mathbb{R}^{2n+1}$ *with the canonical contact structure.*

$\Diamond$ **Exercise 3.22.** Let ω be a differential 1-form on an n-manifold M such that $d\omega$ has rank k everywhere. Show that

(a) if $\omega \neq 0$ and $\omega \wedge (d\omega)^k = 0$ everywhere, then there is a coordinate system

$$x_1, \ldots, x_{n-k}, y_1, \ldots, y_k$$

about any point such that

$$\omega = x_1 dy_1 + \cdots + x_k dy_k.$$

(b) if $\omega \wedge (d\omega)^k \neq 0$ everywhere, then there is a coordinate system

$$x_1, \ldots, x_{n-k}, y_1, \ldots, y_k, \quad 2k+1 \leq n,$$

about any point such that

$$\omega = x_1 dy_1 + \cdots + x_k dy_k + dx_{k+1}.$$

CHAPTER 4

RIEMANNIAN MANIFOLDS

The metric on a manifold M that we considered so far comes from Smirnov's theorem (see Theorem 2.1.5), and also from the fact that M is embeddable in some Euclidean space. Of these, the second metric is more important for us, because the first metric has nothing to do with smooth structure, it may be obtained for any nice topological manifold. In this chapter we shall obtain another metric on M which gives the same topology of M as a manifold. This metric comes from a Riemannian structure on M, and is defined intrinsically (that is, without embedding M in Euclidean space). A Riemannian structure g on M is a field of non-singular, symmetric, second-order covariant tensors, and is one of the simplest geometric objects on M. The other topics that we shall consider in this chapter are connection, geodesic, and exponential map. The final goal is to prove the Hopf-Rinow theorem.

4.1. Riemannian Metric

Definition 4.1.1. A **Riemannian metric** g on a manifold M is a smooth positive definite symmetric 2-tensor field on M. This assigns to each point $p \in M$ a positive definite symmetric bilinear form or inner product on the tangent space $\tau(M)_p$

$$g_p : \tau(M)_p \times \tau(M)_p \longrightarrow \mathbb{R}.$$

Recall that positive definiteness means $g_p(v,v) > 0$ for all non-zero $v \in \tau(M)_p$, and $g_p(u,u) = 0$. A **Riemannian manifold** is a manifold with a Riemannian metric on it.

The length of a tangent vector $v \in \tau(M)_p$ is then defined in the usual way as

$$\|v\| = g_p(v,v)^{1/2}.$$

In terms of local coordinate system $(x_1, \dots, x_n)$ in M with basic vector fields $\delta_i = \partial/\partial x_i$, the local representation of g is given by

$$g = \sum_{i,j=1}^{n} g_{ij} dx_i dx_j,$$

where $g_{ij} = g(\delta_i, \delta_j)$ are real-valued functions on the coordinate neighbourhood U of the system. If $v, w \in \tau(M)_p$, and $v = \sum_i v_i \delta_i, w = \sum_j w_j \delta_j$, then

A. Mukherjee, *Differential Topology*, DOI 10.1007/978-3-319-19045-7_4

$g_p(v,w) = \sum_{i,j} g_{ij}(p) v_i w_j$. As in the case of 2-forms, g is smooth if and only the functions g_{ij} are smooth, or equivalently, if for every pair of smooth vector fields X, Y on U, the function $g(X,Y)$ is smooth on U. Also g is $\mathbb{R}$-bilinear, as well as $C^\infty(M)$-bilinear:

$$g(fX, Y) = fg(X,Y) = g(X, fY), \quad \text{for } f \in C^\infty(M).$$

Definition 4.1.1 is equivalent to the following. A Riemannian metric g on M is a positive definite symmetric $C^\infty(M)$-bilinear map

$$g : \mathfrak{X}(M) \times \mathfrak{X}(M) \longrightarrow C^\infty(M),$$

where $\mathfrak{X}(M)$ denotes the algebra (over the ring $C^\infty(M)$) of all smooth vector fields on M.

Example 4.1.2. The Euclidean space $\mathbb{R}^n$ with coordinates

$$(u_1, \dots, u_n)$$

has a natural Riemannian metric

$$g = \sum_{i,j=1}^{n} \delta_{ij} du_i du_j = \sum_{i=1}^{n} (du_i)^2.$$

Theorem 4.1.3. *Any manifold M admits a Riemannian metric.*

PROOF. Choose an open covering of M by coordinate neighbourhoods $\{U_i\}$. Let $x_{i1}, \dots, x_{in}$ be local coordinates in U_i. Using these coordinates define a metric g_i on U_i by

$$g_i = (dx_{i1})^2 + \cdots + (dx_{in})^2.$$

Let $\{\lambda_i\}$ be a smooth partition of unity subordinate to the covering $\{U_i\}$. Then

$$g = \sum_i \lambda_i g_i$$

is a well defined Riemannian metric on M. □

The necessity of the condition of paracompactness on M in the above theorem may be seen from the following negative result. The Alexandroff line or the long line is a smooth connected manifold of dimension one, but not a manifold in our sense (it is not second countable), see [**44**], Appendix A. It is known that this manifold cannot be given a Riemannian structure (Kneser [**21**]).

The importance of Riemannian metric is that it turns the tangent space at each point into an inner product space, which enables us to define angle between curves (that is, angle between tangent vectors of the curves at the point of intersection), and length of curves.

We shall use the terms 'path', 'curve', and 'parametrised curve' synonymously. Note that a curve $\sigma : [a,b] \longrightarrow M$ on a closed interval $[a,b]$ is smooth if it can be extended to a smooth map on an open interval containing $[a,b]$.

If $\sigma : [a,b] \longrightarrow M$ is a smooth curve in a Riemannian manifold M, then its **length** $\ell(\sigma)$ is defined by

$$\ell(\sigma) = \int_a^b \|\dot\sigma(t)\| dt = \int_a^b \sqrt{g(\dot\sigma(t), \dot\sigma(t))} dt, \quad \dot\sigma = \frac{d\sigma}{dt}.$$

A **reparametrisation** of σ is a curve $\sigma \circ \phi : [c,d] \longrightarrow M$, where $\phi : [c,d] \longrightarrow [a,b]$ is a diffeomorphism with positive derivative everywhere (i.e. an orientation preserving diffeomorphism[1]). Then $\sigma(t)$, $t \in [a,b]$, and $\sigma(\phi(u))$, $u \in [c,d]$, trace the same curve in M in the same direction. Also $\ell(\sigma) = \ell(\sigma \circ \phi)$, by the change of variable formula for integrals: if $t = \phi(u)$, $u \in [c,d]$ is new parameter, then $d\sigma/du = (d\sigma/dt) \cdot (dt/du)$, and

$$\begin{aligned} \int_c^d \left[g\Big(\frac{d\sigma}{du}, \frac{d\sigma}{du}\Big)\right]^{\frac{1}{2}} du &= \int_a^b \left[g\Big(\frac{d\sigma}{dt}, \frac{d\sigma}{dt}\Big) \cdot \Big(\frac{dt}{du}\Big)^2 \right]^{\frac{1}{2}} \frac{du}{dt}\, dt \\ &= \int_a^b \left[g\Big(\frac{d\sigma}{dt}, \frac{d\sigma}{dt}\Big)\right]^{\frac{1}{2}} dt. \end{aligned}$$

Thus the length of a curve is an invariant under reparametrisation. Any curve σ can be re parametrised with its arc length s as parameter, where

$$s = \phi(t) = \int_a^t \|\dot\sigma(t)\| dt$$

is a strictly increasing function $[a,b] \longrightarrow [0, \ell(\sigma)]$. The reparametrisation $\sigma \circ \phi^{-1} : [0, \ell(\sigma)] \longrightarrow M$ has tangent vectors of unit length at all points, as may be seen by the chain rule, and therefore $\ell(\sigma \circ (\phi^{-1} \mid [0,s])) = s$ for all $s \in [0, \ell(\sigma)]$.

A continuous map $\sigma : [a,b] \longrightarrow M$ is called a **piecewise smooth curve** if there is a partition $a = t_0 < t_1 < \cdots < t_k = b$ of $[a,b]$ such that $\sigma \mid [t_{i-1}, t_i]$ is smooth for $i = 1, \ldots, k$ (the left- and right-hand derivatives of σ at points $t_1, \ldots, t_{k-1}$ may be different). Then the length $\ell(\sigma)$ is given by

$$\ell(\sigma) = \sum_i \ell(\sigma | [t_{i-1}, t_i]),$$

that is, by $\int_a^b \|\dot\sigma(t)\| dt$ as in the smooth case.

◊ **Exercise 4.1.** For a piecewise smooth curve $\sigma : [a,b] \longrightarrow M$, define a continuous non-decreasing function $\phi : [a,b] \longrightarrow \mathbb{R}$ by

$$\phi(u) = \int_a^u \|\dot\sigma(t)\| dt.$$

(a) Show that ϕ is smooth at every point u where $\dot\sigma(u)$ exists and is non-zero.

(b) Show that if ℓ denotes the length of σ, then the map $\sigma \circ \phi^{-1} : [0, \ell] \longrightarrow M$ is well-defined and continuous, even if ϕ^{-1} may not be a function. Moreover, $\sigma \circ \phi^{-1}$ is smooth at a point $\phi(u)$ such that $\dot\sigma(u) \neq 0$.

[1] The concept of orientation is discussed in §5.5.

The next lemma concerns connectedness of manifolds. First note that for a manifold pathwise connectedness and connectedness are the same. A pathwise connected space is necessarily connected, because any two points lie in a connected subset of the space (continuous image of an interval being connected). Conversely, in a locally Euclidean space, connectedness implies pathwise connectedness, by an argument that we shall describe in the next lemma.

Two points p and q in M are called piecewise smoothly connected if there is a piecewise smooth curve in M whose image contains p and q.

Lemma 4.1.4. *For a manifold M 'joining by continuous curve' is equivalent to 'joining by piecewise smooth curve'.*

Proof. Clearly joining by piecewise smooth curve is an equivalence relation. An equivalence class is an open set in M, because if (U, ϕ) is a coordinate chart about a point $p \in M$ with $\phi(U) = V$ a convex set in an Euclidean space, then any point of U can be joined to p by a curve corresponding to a straight line in V. Again, an equivalence class is a closed set in M, because it is the complement of the union of other equivalence classes. It follows that a subset of M is open and closed if and only if it is a union of equivalence classes. This completes the proof. □

Remark 4.1.5. The lemma also holds if 'joining by piecewise smooth curve' is replaced by 'joining by smooth curve'. Because 'joining by smooth curve' is also an equivalence relation. This may be seen in the following way.

The relation is obviously reflexive and symmetric. To see transitivity, suppose that σ and τ are two smooth curves with images containing the pairs of points (x, y) and (y, z) respectively. Suppose without loss of generality that

$$\sigma(-1) = x,\ \sigma(0) = y = \tau(0),\ \tau(1) = z.$$

Let (U, ϕ) be a coordinate chart about y such that $\phi(U) = V$ is convex. Since σ and τ are continuous, there exists an ϵ with $0 < \epsilon < 1$ such that $|t| < \epsilon$ implies $\sigma(t),\ \tau(t) \in U$. Let $\lambda : \mathbb{R} \longrightarrow \mathbb{R}$ be a smooth function which is 0 near $t = -\epsilon$ and 1 near $t = \epsilon$. For example, we may take

$$\lambda(t) = \mathcal{B}\left(\frac{t}{\epsilon} + \frac{1}{2}\right),$$

where $\mathcal{B}$ is a bump function (Definition 2.1.6). Define a curve ω by

$$\begin{aligned} \omega(t) &= \sigma(t) \text{ if } t < -\epsilon, \\ &= (1 - \lambda(t))\sigma(t) + \lambda(t)\tau(t) \text{ if } -\epsilon < t < \epsilon, \\ &= \tau(t) \text{ if } t > \epsilon, \end{aligned}$$

where the second line is a convex combination in V. Then ω is smooth and its image contains x and z.

The equivalence classes are components of M, and M is connected if it has only one component.

If M is connected, a Riemannian metric g on M induces a metric

$$d : M \times M \longrightarrow \mathbb{R}$$

so that (M, d) becomes a metric space. The metric d is defined in the following way. Since M is connected, every pair of points $p_1, p_2 \in M$ can be joined by a piecewise smooth curve σ. Then $d(p_1, p_2)$ is defined by

$$d(p_1, p_2) = \inf \ell(\sigma),$$

where the infimum is taken over all piecewise smooth curves σ from p_1 to p_2. Clearly d is a pseudo-metric, that is, it is symmetric, and satisfies the triangle inequality and the condition that $d(p_1, p_2) = 0$ whenever $p_1 = p_2$. The triangle inequality follows, because if σ_1 and σ_2 are any two piecewise smooth curves from p_1 to p_2 and p_2 to p_3 respectively, then $d(p_1, p_3) \leq \ell(\sigma_1) + \ell(\sigma_2)$.

Theorem 4.1.6. *Let g be a Riemannian metric on a connected manifold M, and d be the induced pseudo-metric on M. Then d is a metric.*

We first prove an easy lemma.

Lemma 4.1.7. *Let σ be a piecewise smooth curve in an Euclidean space $\mathbb{R}^n$ from the origin to a point on the sphere of radius r centred at the origin. Then $\ell(\sigma) \geq r$, where $\ell(\sigma)$ is the length of σ with respect to the standard metric on $\mathbb{R}^n$. The equality holds if σ is a straight line segment.*

Proof. We have $\ell(\sigma) = \int_0^1 \|\dot\sigma(t)\| dt \geq \| \int_0^1 \dot\sigma(t) dt\| = \|\sigma(1) - \sigma(0)\| = r$. $\square$

We shall find a similar result for piecewise smooth curves in a Riemannian manifold in the proof of Lemma 4.4.7 below.

Proof of the theorem. We have only to show that $d(p_1, p_2) = 0$ implies that $p_1 = p_2$. Suppose that $p_1 \neq p_2$. Let $(x_1, \ldots, x_n)$ be a coordinate system in a coordinate neighbourhood U about p_1. Let B be an open ball with respect to these coordinates with centre at p_1 such that $\overline{B} \subset U$ (B is obtained in the following way: if the coordinate chart is (U, ϕ) with $\phi(p_1) = 0$, the origin in $\mathbb{R}^n$, and B_0 is an open ball in $\mathbb{R}^n$ with centre at 0 such that $\overline{B_0} \subset \phi(U)$, then take $B = \phi^{-1}(B_0)$). Suppose that B is such that $p_2 \notin B$. Define a function $f : \tau(U) = U \times \mathbb{R}^n \longrightarrow \mathbb{R}$ by

$$f(p, v_1, \ldots, v_n) = \Big[\sum_{i,j} g_{ij}(p) v_i v_j\Big]^{\frac{1}{2}} = \Big\| \sum_i v_i \frac{\partial}{\partial x_i}(p) \Big\|,$$

where $\|\ \|$ is the norm defined by the Riemannian metric g on U. Then $f|U \times S^{n-1}$ is positive and continuous, and therefore we can find a $k > 0$ such that

$$\frac{1}{k} \leq f|\overline{B} \times S^{n-1} \leq k.$$

Let $\| \ \|'$ be the Euclidean norm in $U \times \mathbb{R}^n$. Then since

$$\Big\| \sum_i v_i \frac{\partial}{\partial x_i}(p) \Big\|' = \Big[\sum_i v_i^2 \Big]^{\frac{1}{2}}$$

is 1 on $\overline{B} \times S^{n-1}$, we have

$$\frac{1}{k} \Big\| \sum_i v_i \frac{\partial}{\partial x_i}(p) \Big\|' \leq \Big\| \sum_i v_i \frac{\partial}{\partial x_i}(p) \Big\| \leq k \Big\| \sum_i v_i \frac{\partial}{\partial xi}(p) \Big\|' \tag{1}$$

on $\overline{B} \times S^{n-1}$, and hence on $\overline{B} \times \mathbb{R}^n$ since we may replace v_i by λv_i, $\lambda \in \mathbb{R}$, in these inequalities (the expressions in (1) being homogeneous in the v_i's).

Now, let σ be a piecewise smooth curve from p_1 to p_2, and σ_1 be the portion of σ within $\overline{B}$ from p_1 to the first point of intersection of σ and the boundary of B. Let r be the radius of B.

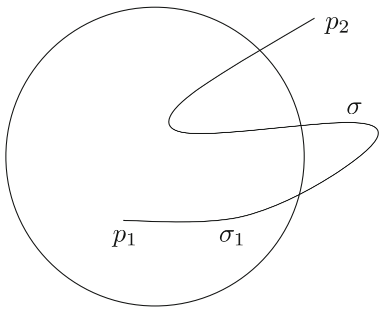


FIGURE 4.1

We have then

$$d(p_1, p_2) = \inf \|\sigma\| \geq \inf \|\sigma_1\| \geq \frac{1}{k} \inf \|\sigma_1\|' \geq \frac{1}{k} r > 0,$$

by the first inequality of (1) and the above lemma. Thus $p_1 \neq p_2 \Rightarrow d(p_1, p_2) \neq 0$. This shows that d is a metric on M. $\square$

Theorem 4.1.8. *The topology on a connected Riemannian manifold M defined by the metric d is equivalent to the topology of M as a manifold. Therefore d is a continuous function on $M \times M$.*

PROOF. The manifold topology has a basis consisting of coordinate neighbourhoods. Let p be an arbitrary point of M, and (U, ϕ) a coordinate chart at p. Then any open neighbourhood of p in U has two topologies, namely, the metric topology induced by the metric d, and the manifold topology induced by the Euclidean metric d' via the homeomorphism ϕ. It is enough to show that these topologies are the same, or explicitly, to show that there is an open neighbourhood B of p with $\overline{B} \subset U$ and a number $k > 0$ such that

$$\frac{1}{k} d' \leq d \leq k\ d', \text{ on } \ B \times B. \tag{2}$$

This will complete the proof of the theorem.

Let B be an open ball with respect to the coordinates in U with centre at p such that $\overline{B} \subset U$. Then, as shown in (1) of the previous theorem, there is a number $k > 0$ such that

$$\frac{1}{k}\|v\|' \leq \|v\| \leq k\|v\|'$$

for any vector $v \in \tau(M)_q$ with $q \in \overline{B}$, where $\|v\|'$ and $\|v\|$ are the norms with respect to the metrics d' and d respectively. This means that for any piecewise smooth curve σ lying entirely in B, we have

$$\frac{1}{k}\|\sigma\|' \leq \|\sigma\| \leq k\|\sigma\|'.$$

Therefore we need only to be concerned about curves that do not lie entirely in B. We shall show by reducing the size of B that for any piecewise smooth curve τ having end points in B but goes outside of B, there is a piecewise smooth curve σ with the same end points lying entirely in B such that $\|\tau\| \geq \|\sigma\|$. This will enable us to compute d on $B \times B$ by restricting ourselves only to curves lying entirely in B. This will establish (2), and complete the proof of the theorem.

For this purpose, suppose that r is the radius of the ball B, and r_1 is a number such that $r_1 \leq r/(2k^2+1)$. Let B_1 be a concentric open ball of radius r_1 such that $\overline{B_1} \subset B$. Let p_1 and p_2 be points in B_1, and τ be a piecewise smooth curve in U from p_1 to p_2. Let τ_1 be the portion of τ in B from p_1 to the first point of intersection of τ with the boundary of B, and σ be the straight line segment from p_1 to p_2 in B_1.

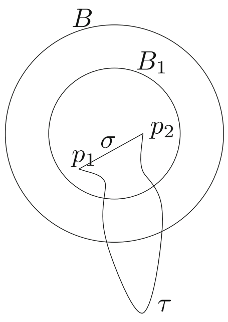


FIGURE 4.2

We have then

$$\|\tau\| \geq \|\tau_1\| \geq \frac{1}{k}\|\tau_1\|' \geq \frac{1}{k}(r - r_1) \geq 2kr_1 \geq k\|\sigma\|' \geq \|\sigma\|,$$

and we get what we wanted to show. □

4.2. Geodesics on a Manifold

A geodesic on a manifold is a smooth curve that gives minimum arc length between points. Therefore it is analogous to a straight line in Euclidean space. However, there is a difference, whereas two points in $\mathbb{R}^n$ can be joined by a unique line segment, there may be many geodesics between two points on a general manifold.

Given two points p and q on a manifold M, the problem is to find, out of all smooth curves joining p and q, those which give minimum arc length. This problem is too general to give a solution. For example, although the lengths of curves between p and q certainly have a greatest lower bound, it is by no means clear that there exists a curve of this length. However, it is possible to formulate a more restricted problem in terms of certain differential equations whose solution curves may be regarded as stationary or critical points of certain functional on the space of smooth curves joining p and q. It is customary to call any curve defined by these differential equations a geodesic, whether it is a curve of minimum arc length or not.

In proper mathematical language, geodesics are solutions of a one dimensional variational problem of the calculus of variation. To describe this, consider the space $\mathcal{C}$ of smooth curves $\sigma : [a, b] \longrightarrow M$ which satisfy the boundary conditions $\sigma(a) = p$ and $\sigma(b) = q$. The **Lagrangian** of the variational problem for $\mathcal{C}$ is a smooth function $G : \tau(M) \longrightarrow \mathbb{R}$. Associated to the Lagrangian G is the functional $\mathcal{G} : \mathcal{C} \longrightarrow \mathbb{R}$ given by

$$\mathcal{G}(\sigma) = \int_a^b G(\sigma(t), \dot{\sigma}(t))dt, \ \ \sigma \in \mathcal{C}.$$

The functional $\mathcal{G}$ is called the **action integral**. It is said to attain an extremal value at a smooth curve $\sigma_0 \in \mathcal{C}$ if either $\mathcal{G}(\sigma_0) \leq \mathcal{G}(\sigma)$ for all $\sigma \in \mathcal{C}$, or $\mathcal{G}(\sigma_0) \geq \mathcal{G}(\sigma)$ for all $\sigma \in \mathcal{C}$. In calculus of variation we are interested in finding a method which determines curves σ for which the action integral $\mathcal{G}(\sigma)$ is least. This is called the **principle of least action**.

Example 4.2.1. In classical mechanics of a particle in $\mathbb{R}^3$, kinetic energy (KE) is the energy of motion, while potential energy (PE) is the energy determined by position. If a particle of mass m is moving with velocity v, then its $\mathrm{KE} = mv^2/2$, and its PE is a smooth function $U : \mathbb{R}^3 \longrightarrow \mathbb{R}$. The Lagrangian of the system $\tau(\mathbb{R}^3) \longrightarrow \mathbb{R}$ is given by $(p, v) \mapsto mv^2/2 - U(p)$. The motion of the particle takes place along a curve which minimises the action integral associated to this Lagrangian.

The following classical theorem gives necessary conditions for a curve σ_0 to give an extremal value of the action integral $\mathcal{G}(\sigma)$ for an arbitrary Lagrangian G.

Theorem 4.2.2. *If a curve $\sigma_0 \in \mathcal{C}$ gives an extremal value of the action integral $\mathcal{G}(\sigma)$, then the following partial differential equations, called the* **Euler-Lagrange equations**, *are satisfied along σ_0.*

$$\frac{d}{dt}\left(\frac{\partial G}{\partial v_i}\right) - \frac{\partial G}{\partial x_i} = 0, \quad i = 1, \ldots, n,$$

where the variables x_i and v_i in $G(x_1, \ldots, x_n, v_1, \ldots, v_n)$ are local coordinates in $\tau(M)$, and they are treated as independent (we may consider $v = (v_1, \ldots, v_n) \in \mathbb{R}^n = \tau(\mathbb{R}^n)_x$ as a tangent vector at a point $x = x(\sigma(t)) = (x_1(\sigma(t)), \ldots, x_n(\sigma(t)))$ of $\mathbb{R}^n$ representing $\dot{x}$).

Proof. It suffices to prove the theorem when $M = \mathbb{R}^n$. Let $\mathcal{D}$ denote the space of smooth curves $\lambda : [a, b] \longrightarrow \mathbb{R}^n$ such that $\lambda(a) = \lambda(b) = 0$, Consider a curve $\omega \in \mathcal{C}$ given by $\omega = \sigma_0 + \epsilon\lambda$, where ϵ is a small real variable so that ω is close to σ_0. Then $\mathcal{G}(\sigma_0 + \epsilon\lambda)$ is a real valued function of the real variable ϵ, and it will take an extremal value at $\epsilon = 0$ if

$$\frac{d}{d\epsilon}\mathcal{G}(\sigma_0 + \epsilon\lambda)\Big|_{\epsilon=0} = 0. \tag{3}$$

Now differentiating under the sign of integration and using the chain rule, we have

$$\frac{d}{d\epsilon}\mathcal{G}(\sigma_0 + \epsilon\lambda)\Big|_{\epsilon=0} = \int_a^b \sum_{j=1}^n \left[\frac{\partial G}{\partial x_j}\lambda_j(t) + \frac{\partial G}{\partial v_j}\dot{\lambda}_j(t)\right] dt. \tag{4}$$

Again, integrating by parts,

$$\begin{aligned} \int_a^b \frac{\partial G}{\partial v_j}\dot{\lambda}_j dt &= \left[\frac{\partial G}{\partial v_j}\lambda_j\right]_a^b - \int_a^b \frac{d}{dt}\left(\frac{\partial G}{\partial v_j}\right)\lambda_j dt \\ &= -\int_a^b \frac{d}{dt}\left(\frac{\partial G}{\partial v_j}\right)\lambda_j dt, \end{aligned} \tag{5}$$

since λ_j vanishes at $t = a$ and $t = b$. Therefore, substituting (5) in (4), we get from (3)

$$\int_a^b \sum_{j=1}^n \left[\frac{\partial G}{\partial x_j} - \frac{d}{dt}\left(\frac{\partial G}{\partial v_j}\right)\right]\lambda_j dt = 0 \tag{6}$$

for all curves $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathcal{D}$. From this it will follow that

$$\frac{\partial G}{\partial x_j} - \frac{d}{dt}\left(\frac{\partial G}{\partial v_j}\right) = 0 \text{ for each } j = 1, \ldots, n. \tag{7}$$

For example, the jth equation in (7) is obtained in the following way. Choose a curve $\lambda \in \mathcal{D}$ by taking $\lambda_i = 0$ if $i \neq j$, and λ_j as a smooth function $f : [a, b] \longrightarrow \mathbb{R}$ such that $f(a) = f(b) = 0$. If we call such a function f admissible, then (6) gives

$$\int_a^b f(t)g(t)dt = 0, \text{ for all admissible functions } f,$$

where $g(t) = \dfrac{\partial G}{\partial x_j} - \dfrac{d}{dt}\left(\dfrac{\partial G}{\partial v_j}\right)$ is a continuous function of t on $[a,b]$. We have to show that $g \equiv 0$ on $[a,b]$. Suppose, if possible, that $g(c) \neq 0$, say $g(c) > 0$, for some $c \in (a,b)$. Then $g > 0$ in a neighbourhood (α, β) of c in (a,b), where $a < \alpha < c < \beta < b$. Therefore, if we take an admissible function f defined by

$$f(t) = \begin{cases} (t-\alpha)^2(\beta - t)^2 & \text{if } \alpha \le t \le \beta, \\ 0 & \text{otherwise,} \end{cases}$$

then we would have

$$\int_a^b f(t)g(t)dt = \int_\beta^\alpha f(t)g(t)dt > 0,$$

since $f > 0$ and $g > 0$ in $[\alpha, \beta]$. Therefore the assumption that $g(c) > 0$ is false. Similarly, $g(c)$ cannot be negative. This proves the theorem. □

In a manifold M with Riemannian metric g, there are two naturally associated Lagrangian functions L and E given by

$$L(p,v) = g_p(p,v)^{1/2}, \quad E(p,v) = (b-a)g_p(p,v),$$

where $v \in \tau(M)_p$, and $a < b$ are real numbers. The action integral corresponding to L is

$$\mathcal{L}(\sigma) = \int_a^b g_p(\dot\sigma(t), \dot\sigma(t))^{1/2} dt, \quad \sigma \in \mathcal{C},$$

which is the length of σ from a to b. The action integral corresponding to E is

$$\mathcal{E}(\sigma) = (b-a)\int_a^b g_p(\dot\sigma(t), \dot\sigma(t)) dt, \quad \sigma \in \mathcal{C},$$

and this is the energy density of σ from a to b. Note that in the last expression, we may omit the factor $(b-a)$, because if $\mathcal{E}$ attains an extremal value at σ, then so does $\mathcal{E}/(b-a)$.

Definition 4.2.3. A smooth curve in $\mathcal{C}$ is called a **geodesic** if it is an extremal value of the energy functional $\mathcal{E}$.

The definition of geodesic involves the energy functional $\mathcal{E}$, rather than the length functional $\mathcal{L}$. The reason behind this is that the length $\|v\|$ of a tangent vector v is not a differentiable function of v; the difficulty arises at vectors of zero length. Another difficulty with $\mathcal{L}(\sigma)$ is that it is invariant under reparametrisation; therefore we cannot expect to find a unique solution curve to the extremal problem. On the other hand, the energy functional $\mathcal{E}(\sigma)$ is a smooth function of the tangent vector, the integrand being the square of length of the tangent vector. Moreover, $\mathcal{E}(\sigma)$ does depend on the parametrisation. The following lemma shows that functionals $\mathcal{E}$ and $\mathcal{L}$ have the same critical points in a certain sense.

Lemma 4.2.4. *A geodesic is also an extremal value of the length functional $\mathcal{L}$ provided it is represented by a parameter proportional to the arc length parameter*

Proof. The Schwarz inequality for continuous functions says that

$$\left(\int_a^b fg\;dt\right)^2 \le \left(\int_a^b f^2 dt\right)\left(\int_a^b g^2 dt\right),$$

with equality if and only if f and g are linearly dependent over $\mathbb{R}$. Applying this we get

$$[\mathcal{L}(\sigma)]^2 = \Big[\int_a^b g_p(\dot\sigma(t),\dot\sigma(t))^{1/2}dt\Big]^2$$

$$\le \int_a^b dt\cdot\int_a^b g_p(\dot\sigma(t),\dot\sigma(t))dt = (b-a)\int_a^b g_p(\dot\sigma(t),\dot\sigma(t))dt = \mathcal{E}(\sigma)$$

with equality if and only if $g_p(\dot\sigma(t),\dot\sigma(t))$ is constant, that is, if and only if ds/dt is constant, because the arc length function along σ is

$$s = s(t) = \int_a^t g_p(\dot\sigma(t),\dot\sigma(t))^{1/2}dt,$$

and so $ds/dt = g_p(\dot\sigma(t),\dot\sigma(t))^{1/2}$. This proves the lemma. $\square$

In a local coordinate system $x = (x_1,\dots,x_n)$, where $x_i(t) = x_i(\sigma(t))$ on σ, we may write

$$\frac{1}{b-a}E(\sigma) = \frac{1}{b-a}E(x,\dot x) = \sum_{i,j} g_{ij}(x)\dot x_i\dot x_j. \tag{8}$$

Proposition 4.2.5. *The following differential equations are satisfied along any geodesic*

$$\frac{d^2x_k}{dt^2} + \sum_{i,j}\Gamma^k_{ij}\frac{dx_i}{dt}\frac{dx_j}{dt} = 0. \tag{9}$$

In fact, these are the Euler- Lagrange equations in the local coordinate system for the Lagrangian $E_1 = E/(b-a)$. The coefficients Γ^k_{ij} are called the **Christoffel symbols** associated to g; they will be defined in the proof of the proposition.

Proof. We have from (8)

$$\frac{\partial E_1}{\partial x_i} = \sum_{j,k}\frac{\partial g_{jk}}{\partial x_i}\dot x_j\dot x_k, \text{ and}$$

$$\begin{aligned}
\frac{d}{dt}\Big(\frac{\partial E_1}{\partial \dot x_i}\Big) &= 2\frac{d}{dt}\Big(\sum_j g_{ij}\dot x_j\Big)\\
&= 2\sum_j\Big(\sum_k\frac{\partial g_{ij}}{\partial x_k}\dot x_k\dot x_j + g_{ij}\ddot x_j\Big)\\
&= \sum_{k,j}\frac{\partial g_{ij}}{\partial x_k}\dot x_k\dot x_j + \sum_{j,k}\frac{\partial g_{ik}}{\partial x_j}\dot x_j\dot x_k + 2\sum_j g_{ij}\ddot x_j,
\end{aligned}$$

where the first two terms of the last line are equal, by interchanging j and k. Then the Euler-Lagrange equations

$$\frac{d}{dt}\Big(\frac{\partial E_1}{\partial \dot{x}_i}\Big) - \frac{\partial E_1}{\partial x_i} = 0$$

become after division by 2

$$\sum_j g_{ij}\ddot{x}_j + \frac{1}{2}\Big(\sum_{j,k} \frac{\partial g_{ij}}{\partial x_k} + \frac{\partial g_{ik}}{\partial x_j} - \frac{\partial g_{jk}}{\partial x_i}\Big)\dot{x}_j\dot{x}_k = 0. \tag{10}$$

Let (g^{ij}) denote the inverse of the matrix (g_{ij}). Then, multiplying both sides of (10) by $g^{\ell i}$ and summing over i, we get

$$\ddot{x}_\ell + \frac{1}{2}\sum_i g^{\ell i}\sum_{j,k}\left(\frac{\partial g_{ij}}{\partial x_k} + \frac{\partial g_{ik}}{\partial x_j} - \frac{\partial g_{jk}}{\partial x_i}\right)\dot{x}_j\dot{x}_k = 0.$$

Taking

$$\Gamma^\ell_{jk} = \frac{1}{2}\sum_i g^{\ell i}\left(\frac{\partial g_{ij}}{\partial x_k} + \frac{\partial g_{ik}}{\partial x_j} - \frac{\partial g_{jk}}{\partial x_i}\right),$$

we get (9) after changing $j,\ k,\ \ell$ to $i,\ j,\ k$. This completes the proof. $\square$

4.3. Riemannian connection and geodesics

The Christoffel symbols Γ^k_{ij} are the components of the unique Riemannian connection on M. We now describe what is a connection, and an alternative definition of geodesic in terms of connection.

As before, let $\mathfrak{X}(M)$ denote the module over $C^\infty(M)$ of all smooth vector fields on M.

Definition 4.3.1. A **connection** on M is a map

$$\nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) \longrightarrow \mathfrak{X}(M),$$

such that, for $X, Y, Z \in \mathfrak{X}(M)$, and $f \in C^\infty(M)$, we have after writing $\nabla(X,Y) = \nabla_X(Y)$

(a) $\nabla_X(Y+Z) = \nabla_X(Y) + \nabla_X(Z)$,

(b) $\nabla_{(X+Y)}(Z) = \nabla_X(Z) + \nabla_Y(Z)$,

(c) $\nabla_X(fY) = (Xf)Y + f\nabla_X(Y)$,

(d) $\nabla_{(fX)}(Y) = f\nabla_X(Y)$.

These properties imply that $\nabla(X,Y)$ is $C^\infty(M)$-linear in X, and $\mathbb{R}$-linear in Y. At a point $p \in M$, the $\mathbb{R}$-bilinear map

$$\nabla_p : \tau(M)_p \times \mathfrak{X}(M) \longrightarrow \tau(M)_p,$$

given by $\nabla_p(X_p, Y) = (\nabla_X(Y))_p$, is called the **covariant derivative** of Y in the direction X_p.

Let U be a coordinate neighbourhood in M with local coordinates $(x_1, \ldots, x_n)$. Let δ_i denote the basic vector field $\partial/\partial x_i$ on U, $i = 1, \ldots, n$. Then a connection ∇ on U defines n^3 real functions Γ_{ij}^k, $i, j, k = 1, \ldots, n$, on U such that

$$\nabla_{\delta_i}(\delta_j) = \sum_k \Gamma_{ij}^k \delta_k.$$

The coefficients are denoted by the same notations as the Christoffel symbols, because ultimately, as we shall see, they will be identical with the Christoffel symbols.

Then, if $X = \sum_i X_i \delta_i$ and $Y = \sum_j Y_j \delta_j$ are vector fields on U, we have the above properties

$$\nabla_X(Y) = \sum_k \left(\sum_i X_i \frac{\partial Y_k}{\partial x_i} + \sum_{i,j} X_i Y_j \Gamma_{ij}^k \right) \delta_k. \tag{11}$$

Conversely, given the functions Γ_{ij}^k, and vector fields X, Y, we may define $\nabla_X(Y)$ by the formula (11), which may easily be seen to satisfy the conditions (a) - (d). Thus the functions Γ_{ij}^k completely determine the connection ∇ on U. These are called the components of ∇. A connection is called **smooth** if its components are so.

Recall that if $\sigma : I \longrightarrow M$ is a smooth curve in M on an open interval I, then a vector field V along σ is a map $I \longrightarrow \ni\!\in(M)$ which assigns to each $t \in I$ a vector $V_t \in \tau(M)_{\sigma(t)}$. A function f along σ is just a function $f : I \longrightarrow \mathbb{R}$. For example, the velocity vector field $\dot\sigma(t)$ is a vector field along σ, and the speed $\|\dot\sigma(t)\|$ is a function along σ.

Definition 4.3.2. If ∇ is a smooth connection on M, and V is a smooth vector field along σ, then the **covariant derivative of V along σ** is a vector field $\frac{DV}{dt}$ along σ defined by

$$\frac{DV}{dt} = \nabla_{\dot\sigma(t)}(Y),$$

where Y is a vector field on M which induces V, that is, $V_t = Y_{\sigma(t)}, t \in I$.

$\lozenge$ **Exercise 4.2.** Show that the operation $V \longrightarrow \frac{DV}{dt}$ satisfies the properties

(a) $\frac{D(V+W)}{dt} = \frac{DV}{dt} + \frac{DW}{dt}$,

(b) $\frac{D(fV)}{dt} = \frac{df}{dt} V + f \frac{DV}{dt}$,

where V and W are vector fields along σ, and f is a smooth function $I \longrightarrow \mathbb{R}$.

In a coordinate neighbourhood U with coordinates $(x_1, \ldots, x_n)$, a curve $\sigma : I \longrightarrow U$ has components $x_i(t) = x_i(\sigma(t))$ and a vector field V on U has a unique expression $V = \sum_j V_j \delta_j$ along σ, where V_j are real functions along σ.

Then, we have by (11)

$$\frac{DV}{dt} = \sum_k \left(\frac{dV_k}{dt} + \sum_{i,j} \frac{dx_i}{dt} \Gamma^k_{ij} V_j \right) \delta_k. \tag{12}$$

Definition 4.3.3. A vector field V along σ is called **parallel along** σ if $\frac{DV}{dt} = 0$.

The following theorem gives the existence of parallel fields along σ.

Theorem 4.3.4. *Given a smooth curve $\sigma : [a,b] \longrightarrow M$ and a tangent vector $V_0 \in \tau(M)_{\sigma(a)}$, there is a unique parallel vector field V along σ which extends V_0.*

The vector V_t is said to be obtained from V_0 by parallel translation along σ. For each $t \in [a,b]$, the correspondence $V_0 \mapsto V_t$ gives a linear map

$$P_t : \tau(M)_{\sigma(a)} \longrightarrow \tau(M)_{\sigma(t)}$$

called **parallel translation**.

Proof. Consider the system of differential equations obtained from (12)

$$\frac{dV_k}{dt} + \sum_{i,j} \frac{dx_i}{dt} \Gamma^k_{ij} V_j = 0, \quad k = 1, \ldots, n.$$

By the existence and uniqueness theorem of ordinary differential equations, there exist solutions $V_k(t)$ which are determined uniquely by the initial values $V_k(0)$. Note that the solutions can be defined for all $t \in [a,b]$, since the equations are linear. □

Definition 4.3.5. A connection ∇ on a Riemannian manifold M is said to be **compatible with the Riemannian metric** g (or **a Riemannian connection**) if the parallel translation P_t along any curve σ preserves the inner product

$$g_{\sigma(a)}(v,w) = g_{\sigma(t)}(P_t(v), P_t(w)), \quad v, w \in \tau(M)_{\sigma(t)},$$

that is, the inner product is constant along σ. In this case the parallel translation P_t will be a linear isomorphism, and it will preserve length of a vector and angle between vectors.

We denote the inner product of two tangent vectors v and w by $\langle v, w \rangle$.

Lemma 4.3.6. *Let the connection ∇ be compatible with the metric g. Then*

(a) *for vector fields V and W along σ,*

$$\frac{d}{dt}\langle V, W\rangle = \left\langle \frac{DV}{dt}, W \right\rangle + \left\langle V, \frac{DW}{dt} \right\rangle,$$

(b) *for vector fields V, W on M,and any tangent vector $X_p \in \tau(M)_p$,*

$$X_p\langle V, W\rangle = \langle \nabla_{X_p}(V), W_p\rangle + \langle V_p, \nabla_{X_p}(W)\rangle.$$

Proof. (a) Let $P_1, \dots, P_n$ be parallel vector fields along σ which are orthonormal at every point of σ. These fields are obtained by choosing a set of orthonormal tangent vectors at a point of σ and extending them to parallel fields along σ. Then the vector fields V and W are represented by $V = \sum V_i P_i$, and $W = \sum W_j P_j$, where $V_i = \langle V, P_i \rangle$ and $W_j = \langle W, P_j \rangle$, so that $\langle V, W \rangle = \sum V_i W_i$, and

$$\frac{DV}{dt} = \sum \frac{dV_i}{dt} P_i, \qquad \frac{DW}{dt} = \sum \frac{dW_j}{dt} P_j.$$

(The last expressions follow from Exercise 4.2, p.117.) Therefore

$$\frac{d}{dt}\Big\langle V, W \Big\rangle = \sum \Big(\frac{dV_i}{dt} W_i + V_i \frac{dW_i}{dt}\Big) = \Big\langle \frac{DV}{dt}, W \Big\rangle + \Big\langle V, \frac{DW}{dt} \Big\rangle.$$

(b) This part is obtained from (a) by choosing a curve σ so that $\dot\sigma(0) = X_p$. □

Definition 4.3.7. If X and Y are vector fields on M, let $[X, Y]$ denote their **Lie bracket**. This is a vector field defined by

$$[X, Y]\lambda = X(Y\lambda) - Y(X\lambda),$$

where $\lambda \in C^\infty(M)$.

◊ **Exercises 4.3.** Prove the following results:

(1) If $X, Y, Z \in \mathfrak{X}(M)$, $\lambda, \mu \in C^\infty(M)$, then

(a) $[X, Y] = -[Y, X]$, $[X, X] = 0$,

(b) $[X + Y, Z] = [X, Z] + [Y.Z]$,

(c) $[\lambda X, \mu Y] = \lambda\mu[X, Y] + \lambda(X\mu)Y - \mu(Y\lambda)X$,

(d) $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$.

The expression in (d) is called the **Jacobi identity**.

(2) If (U, ϕ) is a coordinate chart in M, $X, Y \in \mathfrak{X}(M)$, and $\delta_1, \dots, \delta_n$ are the basic vector fields on U, then $[\delta_i, \delta_j] = 0$ on U. Also, if $X = \sum_i \alpha_i \delta_i$ and $Y = \sum_j \beta_j \delta_j$, then

$$[X, Y] = \sum_{i,j} \left(\alpha_i \frac{\partial \beta_j}{\partial x_i} - \beta_i \frac{\partial \alpha_j}{\partial x_i} \right) \delta_j \text{ on } U.$$

Definition 4.3.8. A connection ∇ on M is called **symmetric** if

$$\nabla_X(Y) - \nabla_Y(X) = [X, Y],$$

for all $X, Y \in \mathfrak{X}(M)$

Lemma 4.3.9. *A connection is symmetric if and only if* $\Gamma^k_{ij} = \Gamma^k_{ji}$.

PROOF. Applying the definition to $X = \delta_i$ and $Y = \delta_j$, we get $\Gamma_{ij}^k = \Gamma_{ji}^k$, since $[\delta_i, \delta_j] = 0$. Conversely, if $\Gamma_{ij}^k = \Gamma_{ji}^k$, it can be verified easily using (11) that ∇ is symmetric on any coordinate neighbourhood. □

Theorem 4.3.10 (Fundamental Theorem of Riemannian Geometry). *A Riemannian manifold M has a unique symmetric connection which is compatible with the Riemannian metric. (This connection is the Riemannian connection on M.)*

PROOF. Uniqueness. Let M admit a symmetric connection compatible with the Riemannian metric g. Then applying Lemma 4.3.6(b) to the vector fields $\delta_i, \delta_j, \delta_k$, we get

$$\delta_i\langle\delta_j, \delta_k\rangle = \langle\nabla_{\delta_i}(\delta_j), \delta_k\rangle + \langle\delta_j, \nabla_{\delta_i}(\delta_k)\rangle.$$

Since $\langle\delta_j, \delta_k\rangle = g_{jk}$, this gives

$$\delta_i g_{jk} = \langle\nabla_{\delta_i}(\delta_j), \delta_k\rangle + \langle\delta_j, \nabla_{\delta_i}(\delta_k)\rangle.$$

By permutations of i, j, k, we have

$$\delta_j g_{ik} = \langle\nabla_{\delta_j}(\delta_i), \delta_k\rangle + \langle\delta_i, \nabla_{\delta_j}(\delta_k)\rangle.$$

$$\delta_k g_{ij} = \langle\nabla_{\delta_k}(\delta_i), \delta_j\rangle + \langle\delta_i, \nabla_{\delta_k}(\delta_j)\rangle.$$

These give

$$\langle\nabla_{\delta_i}(\delta_j), \delta_k\rangle = \frac{1}{2}(\delta_i g_{jk} + \delta_j g_{ik} - \delta_k g_{ij}).$$

This is called the **first Christoffel identity**. Since the left hand side of this identity is $\sum_\ell \Gamma_{ij}^\ell \; g_{lk}$, multiplication by the inverse $(g^{k\ell})$ of the matrix $(g_{\ell k})$ gives

$$\Gamma_{ij}^\ell = \frac{1}{2}\sum_k \Big(\delta_i g_{jk} + \delta_j g_{ik} - \delta_k g_{ij}\Big) g^{k\ell}. \tag{13}$$

This is called the **second Christoffel identity**. It expresses the function Γ_{ij}^k in terms of the metric tensor g, and hence shows that the connection ∇ is uniquely determined by g.

Existence. One obtains a connection on M by defining its components Γ_{ij}^l by the formula (13). It is easily checked that the resulting connection is symmetric and compatible with the metric. □

We are now in a position to present the second definition of geodesic.

Definition 4.3.11. A smooth curve $\sigma : I \longrightarrow M$ is called a **geodesic** if its tangent vector field $\dot\sigma(t)$ is parallel along σ.

In terms of the local coordinates $(x_1, \dots, x_n)$. The equation $\frac{D\dot\sigma(t)}{dt} = 0$ for a geodesic takes the form

$$\frac{d^2 x_k}{dt^2} + \sum_{i,j} \Gamma_{ij}^k \frac{dx_i}{dt}\frac{dx_j}{dt} = 0.$$

This is nothing but the Euler-Lagrange equations we encountered in (9). Therefore the two definitions of geodesic are the same.

It follows then that by parallel translation, a tangent vector to a geodesic $\sigma : [a, b] \longrightarrow M$ goes to a tangent vector to the geodesic

$$P_t(\dot\sigma(a)) = \dot\sigma(t).$$

Lemma 4.3.12. *If $\sigma(t)$ is a geodesic, then the length of the velocity vector field $\dot\sigma(t)$ is independent of t.*

Proof. By Lemma 4.3.6(a),

$$\frac{d}{dt}\left\|\dot\sigma(t)\right\|^2 = \frac{d}{dt}\left\langle \dot\sigma(t), \dot\sigma(t)\right\rangle = 2\left\langle \frac{D\dot\sigma(t)}{dt}, \dot\sigma(t)\right\rangle = 0,$$

since

$$\frac{D\dot\sigma(t)}{dt} = 0.$$

□

We shall close this section with a consequence of symmetric connection on M. This will be used in the proof of Gauss Lemma in the next section.

A **parametrised surface** in M is a smooth map $f : U \longrightarrow M$, where U is an open set in $\mathbb{R}^2$ with coordinates (r, t). For a fixed $(r_0, t_0) \in U$, the curve $r \mapsto f(r, t_0)$ (resp. $t \mapsto f(r_0, t)$) is called the r-parametric (resp. t-parametric) curve passing through (r_0, t_0). The tangent vector field along the r-parametric (resp. t-parametric) curve is given by $df(\partial/\partial r)$ (resp. $df(\partial/\partial t)$), and is denoted by $\partial f/\partial r$ (resp. $\partial f/\partial t$). A smooth vector field V along a parametrised surface f is a smooth assignment $(r, t) \mapsto V_{(r,t)} \in \tau(M)_{f(r,t)}$. Then V restricts to a smooth vector field along r-parametric (resp. t-parametric) curve. Its covariant derivative is denoted by $DV/\partial r$ (resp. $DV/\partial t$).

Lemma 4.3.13. *If $f(r, t)$ is a parametrised surface in a Riemannian manifold M with a symmetric connection, then*

$$\frac{D}{\partial r}\Big(\frac{\partial f}{\partial t}\Big) = \frac{D}{\partial t}\Big(\frac{\partial f}{\partial r}\Big).$$

Proof. In terms of local coordinates

$$\frac{\partial f}{\partial r} = \sum_i \frac{\partial f_i}{\partial r}\delta_i, \text{ and } \frac{\partial f}{\partial t} = \sum_i \frac{\partial f_i}{\partial t}\delta_i.$$

Then by (12)

$$\frac{D}{\partial r}\Big(\frac{\partial f}{\partial t}\Big) = \sum_k \left(\frac{\partial^2 f_k}{\partial r \partial t} + \sum_{i,j} \Gamma^k_{ij}\frac{\partial f_i}{\partial r}\frac{\partial f_j}{\partial t}\right)\delta_k,$$

$$\frac{D}{\partial t}\Big(\frac{\partial f}{\partial r}\Big) = \sum_k \left(\frac{\partial^2 f_k}{\partial t \partial r} + \sum_{i,j} \Gamma^k_{ij}\frac{\partial f_i}{\partial t}\frac{\partial f_j}{\partial r}\right)\delta_k.$$

The last two expressions are equal for a smooth map f, since $\Gamma^k_{ij} = \Gamma^k_{ji}$. □

4.4. Exponential maps

We shall use the following notation:

$$\begin{aligned} B(0_p, r) &= \{v \in \tau(M)_p \mid \|v\| = g_p(0_p, v) < r\}. \\ B(p, r) &= \{q \in M \mid d(p, q) < r\}, \end{aligned}$$

where 0_p denotes the zero vector in $\tau(M)_p$.

Proposition 4.4.1. *For any point $p \in M$ there is a relatively compact open neighbourhood U of p in M and a number $\epsilon > 0$ such that for each point $q \in U$ and each tangent vector $v \in \tau(M)_q$ of length $\|v\| < \epsilon$ there is a unique geodesic $\sigma(t)$ satisfying $\sigma(0) = q$ and $\dot\sigma(0) = v$. The geodesic $\sigma(t)$ is defined for $|t| < 2$, it lies in M and depends smoothly on q, v, and t.*

Proof. In a local coordinate system $x = (x_1, \ldots, x_n)$ on a coordinate neighbourhood V of p, where $p = (p_1, \ldots, p_n)$, $p_i = x_i(p)$, a tangent vector $v \in \tau(M)_q$ at a point $q \in V$ can be written as $v = \sum_i v_i \partial_i$, $v_i \in \mathbb{R}$. Consider the following system of equations obtained from the geodesic equations (9).

$$\begin{aligned} \frac{dx_i}{dt} &= y_i, \\ \frac{dy_i}{dt} &= -\sum_{j,k} \Gamma^i_{jk}(x_1, \ldots, x_n) y_j y_k, \end{aligned}$$

with initial conditions $(x_1, \ldots, x_n, y_1, \ldots, y_n) = (p_1, \ldots, p_n, v_1, \ldots, v_n)$. By the existence and uniqueness theorem of differential equations (Theorem 3.2.10) adapted on the neighbourhood V, for some $\epsilon_1 > 0$, some relatively compact neighbourhood U of p in V, and some open set B in $\tau(M)$ consisting of vectors of length $< \epsilon_2$ with $\pi(B) = U$, where $\pi : \tau(M) \longrightarrow M$ is the projection, there is a unique solution $x = f(q, v, t)$, where $f : B \times (-\epsilon_1, \epsilon_1) \longrightarrow \tau(M)$ is smooth, satisfying the above initial conditions and depending smoothly on q, v, t. Thus for each $(q, v) \in B$, the curve $\xi(t) = \pi f(q, v, t)$ gives a unique geodesic $(-\epsilon_1, \epsilon_1) \longrightarrow M$ satisfying the conditions $\xi(0) = q$ and $\dot\xi(0) = v$.

To obtain the geodesic σ of the proposition, let ϵ be a positive number $< \epsilon_1\epsilon_2/2$. Then, if $q \in U$, v is a tangent vector in $\tau(M)_q$ of length $\|v\| < \epsilon$, and $|t| < 2$, we have

$$|\epsilon_1 t/2| < \epsilon_1 \text{ and } \|2v/\epsilon_1\| < \epsilon_2.$$

So $f(q, 2v/\epsilon_1, \epsilon_1 t/2)$ is defined, and the geodesic $\eta(s) = f(q, 2v/\epsilon_1, s)$, where $|s| < \epsilon_1$, has the property that $\eta(0) = q$ and $\dot\eta(0) = 2v/\epsilon_1$. Introduce a change of parameter $(-\epsilon_1, \epsilon_1) \longrightarrow (-2, 2)$ by $t = 2s/\epsilon_1$ and define $\sigma : (-2, 2) \longrightarrow M$ by $\sigma(t) = \eta(t\epsilon_1/2)$. Then σ is also a geodesic (note that if $t \mapsto \eta(t)$ is a geodesic, then so is $t \mapsto \eta(ct)$, where c is a constant), and it satisfies the conditions $\sigma(0) = \eta(0) = q$ and $\dot\sigma(0) = (\epsilon_1/2)\dot\eta(0) = v$. This completes the proof. $\square$

Let p be a point in M, and W_p denote the open ball $B(0_p, \epsilon)$ in $\tau(M)_p$. Then for each $v \in W_p$ there is a unique geodesic $\sigma_v(t)$ with $\sigma_v(0) = p$ and $\dot{\sigma}_v(0) = v$, where $\sigma_v(t)$ is defined for $|t| \leq 1$.

Definition 4.4.2. The **exponential map** at p is the map

$$\exp_p : W_p \longrightarrow M$$

defined by $\exp_p(v) = \sigma_v(1)$.

Then $\exp_p$ is a smooth map, because it is a restriction of the smooth map πf of the last proposition.

The name 'exponential map' is derived from the fact that when M is the general linear group $GL(n, \mathbb{R})$, it coincides with the matrix exponential map

$$A \longrightarrow e^A = \sum_k \frac{A^k}{k!}$$

from the vector space of $n \times n$ matrices $M(n, \mathbb{R})$ into $GL(n, \mathbb{R}$ (see §1.6). It can be shown easily that $\sigma(t) = e^{tA}$ is the geodesic in $GL(n, \mathbb{R}$ with tangent A at $t = 0$.

Let W be the subset of $\tau(M)$ such that if $(p, v) \in W$, then $\exp_p v$ is defined. Then W is an open set, and we define

$$\exp : W \longrightarrow M$$

by $\exp(p, v) = \exp_p(v)$. Note that exp is also smooth by the same reason as before.

Lemma 4.4.3. *For $p \in M$ and $v \in W_p$, the unique geodesic σ_v with $\sigma_v(0) = p$ and $\dot{\sigma}_v(0) = v$ is given by the relation*

$$\sigma_v(t) = \exp_p(tv).$$

PROOF. Since $v \in W_p$, $tv \in W_p$ also for any t with $|t| < 1$. Let σ_{tv} be the unique geodesic with direction tv at p. Let I be an open interval which is the domain of the geodesic σ_v. For each non-zero $t \in I$, let $f_t : \mathbb{R} \longrightarrow \mathbb{R}$ be the diffeomorphism $s \mapsto ts$. Then $J = f_t^{-1}(I) = \{s \in \mathbb{R} \mid ts \in \mathbb{R}\}$ is an open interval. Define $\sigma : J \longrightarrow M$ by $\sigma(s) = \sigma_v(ts)$. Then σ is also a geodesic satisfying $\sigma(0) = p$ and $\dot{\sigma}(0) = tv$. By the uniqueness of geodesic, $\sigma_{tv}(s) = \sigma(s) = \sigma_v(ts)$ for all s for which $ts \in I$. Then, taking $s = 1$ we get $\sigma_v(t) = \sigma_{tv}(1) = \exp_p(tv)$. □

Proposition 4.4.4. *The exponential map $\exp_p$ is a diffeomorphism of an open neighbourhood of 0_p in $\tau(M)_p$ onto an open neighbourhood of p in M.*

PROOF. In view of the inverse function theorem, it is sufficient to prove that the map $(d\exp_p)_{0_p} : \tau(\tau(M)_p))_{0_p} \longrightarrow \tau(M)_p$ is non-singular. Now, every element of $\tau(\tau(M)_p)_{0_p}$ is of the form $\dot{\lambda}(0)$, where $\lambda(t) = tv$ for some $v \in \tau(M)_p$. Then, by Lemma 4.4.3, we have

$$(d\exp_p)_{0_p}(\dot{\lambda}(0)) = \frac{d}{dt}(\exp_p \circ \lambda)(0) = \dot{\sigma}_v(0) = v,$$

where σ_v is the unique geodesic with $\sigma_v(0) = p$ and $\dot{\sigma}_v(0) = v$. Therefore the matrix of $(d\exp_p)_{0_p}$ is the identity matrix, and it is non-singular. □

Suppose that δ is a positive number $< \epsilon$ such that $\exp_p$ is a diffeomorphism on $B(0_p, \delta)$. Let r be any number with $0 < r < \delta$, and $S(0_p, r)$ be the boundary of the open ball $B(0_p, r)$. Then $S(0_p, r)$ is a submanifold of codimension one (or a hypersurface) in $\tau(M)_p$, being the solution set of the equation $g_p(v, v) = r^2$. Therefore $\exp_p(S(0_p, r))$ is a submanifold of M. Note That $S(0_p, sr) = sS(0_p, r)$ for $s > 0$.

Lemma 4.4.5 (Gauss Lemma). *If r is as above, then the geodesics emanating from $p \in M$ meet the images by $\exp_p$ of the spheres $S(0_p, r)$ orthogonally.*

Proof. Let $t \longrightarrow v(t)$ be a curve in $\tau(M)_p$ with $\|v(t)\| = 1$. Then our problem is to show that the curves $t \longrightarrow \exp_p(r_0 v(t))$ in $B(0_p, \delta)$, $0 < r_0 < \delta$, are orthogonal to the radial geodesics $r \longrightarrow \exp_p(rv(t_0))$, $0 < r < \delta$. In other words, if $f(r, t)$ is the parametric surface $f(r, t) = \exp_p(rv(t))$, then the problem is to show that the parametric curves are orthogonal, that is,

$$\left\langle \frac{\partial f}{\partial r}, \frac{\partial f}{\partial t} \right\rangle = 0,$$

for all (r, t).

We have by Lemma 4.3.6(a),

$$\frac{\partial}{\partial r}\left\langle \frac{\partial f}{\partial r}, \frac{\partial f}{\partial t} \right\rangle = \left\langle \frac{D}{\partial r}\left(\frac{\partial f}{\partial r}\right), \frac{\partial f}{\partial t} \right\rangle + \left\langle \frac{\partial f}{\partial r}, \frac{D}{\partial r}\left(\frac{\partial f}{\partial t}\right) \right\rangle. \tag{14}$$

The first term on the right is zero, because the curve $r \longrightarrow f(r, t)$ is a geodesic. The second term is also zero, because

$$\left\langle \frac{\partial f}{\partial r}, \frac{D}{\partial r}\left(\frac{\partial f}{\partial t}\right) \right\rangle = \left\langle \frac{\partial f}{\partial r}, \frac{D}{\partial t}\left(\frac{\partial f}{\partial r}\right) \right\rangle = \frac{1}{2}\frac{\partial}{\partial t}\left\langle \frac{\partial f}{\partial r}, \frac{\partial f}{\partial r} \right\rangle = 0.$$

The first equality follows from Lemma 4.3.13. The second equality follows from Lemma 4.3.6(a) (taking $V = W = \partial f/\partial r$). The third equality follows, because the length of a tangent vector to a geodesic is constant.

Therefore $\langle \partial f/\partial r, \partial f/\partial t\rangle$ is independent of r. But for $r = 0$, we have $f(0, t) = \exp_p(0) = p$, and therefore $\partial f/\partial t = 0$ at $(0, t)$. It follows finally that $\langle \partial f/\partial r, \partial f/\partial t\rangle$ vanishes identically, concluding the proof. □

Corollary 4.4.6. *If $p \in M$, and r be chosen as in Gauss Lemma, then $\exp_p(B(0_p, s)) = B(p, s)$ for any $s \leq r$*

Lemma 4.4.7. *Let $B(0_p, \delta)$ be an open ball in $\tau(M)_p$ on which $\exp_p$ is a diffeomorphism, and $U = \exp_p(B(0_p, \delta))$. For any $q \in U$, let σ be the unique geodesic joining p and q, and λ be any piecewise smooth curve which joins p and q. Then $\ell(\lambda) \geq \ell(\sigma)$ with equality occurring only if λ is a piecewise smooth reparametrisation of σ.*

PROOF. Let v_1 be a vector in $\tau(M)_p$ with $\|v_1\| = 1$ and r_1 be a positive number $< \delta$ so that $q = \exp_p(r_1 v_1)$, and $r_1 v_1$ belongs to the closed ball $\overline{B}(0_p, r_1)$. Using polar coordinates in $\tau(M)_p$, we may write any piecewise smooth curve $\lambda : [0,1] \longrightarrow \overline{B}(0_p, r_1)$ joining p and q as $\lambda(t) = \exp_p(r(t)v(t))$, where $r(0) = 0$, $r(1) = r_1$, $v(1) = v_1$, and $\|v(t)\| = 1$. The vector $r(t)v(t)$ is obtained by $\exp_p^{-1}$ followed by a projection, so the functions $r(t)$ and $v(t)$ are piecewise smooth on $[0,1]$. We may write $\lambda(t) = f(r(t), t)$, where $f(r,t) = \exp_p(rv(t))$. Then

$$\dot{\lambda}(t) = \frac{d\lambda}{dt} = \frac{\partial f}{\partial r}\dot{r}(t) + \frac{\partial f}{\partial t}.$$

By Lemma 4.4.5, $\partial f/\partial r$ and $\partial f/\partial t$ are orthogonal. Since $\partial f/\partial r$ is tangent to a geodesic with initial direction $v(t)$, $\|\partial f/\partial r\| = \|v(t)\| = 1$. Therefore

$$\|\dot{\lambda}(t)\|^2 = |\dot{r}(t)|^2 + \left\|\frac{\partial f}{\partial t}\right\|^2 \geq \|\dot{r}(t)\|^2,$$

and equality occurs only if $\partial f/\partial t = 0$, or equivalently $dv/dt = 0$, or $v = v_1$. It follows that

$$\ell(\lambda) = \int_0^1 \|\dot{\lambda}(t)\| dt \geq \int_0^1 \|\dot{r}(t)\| dt = |r(1) - r(0)| = r_1$$

with equality occurring only if $v = v_1$ and $dr/dt > 0$. But then $r(t) = \exp_p(r(t)v_1)$ which is a reparametrisation of the geodesic

$$\sigma(\theta) = \exp_p(\theta r_1 v_1)$$

by the change of parameter $\theta = r(t)/r_1$

In the above arguments, we have assumed that $\lambda([0,1]) \subset \overline{B}(0_p, r)$. If this is not the case, we proceed in the same way, considering the first point where λ meets the boundary of $\exp_p(B(0_p, r_1))$.

Since the length of the geodesic joining p and q is $\|r_1 v_1\| = r_1$, we have $\ell(\lambda) \geq \ell(\sigma)$. □

Corollary 4.4.8. *If p and q are any two points in M, and λ is a piecewise smooth curve from p to q such that $\ell(\lambda) = d(p,q)$, then λ is a geodesic. In particular, λ is smooth.*

PROOF. Since λ minimises arc length from p to q, it minimises arc length locally. Therefore λ is locally a piecewise smooth reparametrisation of a geodesic by the above lemma. This is enough to conclude that λ is a geodesic. □

By the zero-section in the tangent bundle $\tau(M)$, we shall mean the subset of all zeros $0_p \in \tau(M)_p$. Recall that W_p is a neighbourhood of 0_p in $\tau(M)_p$ on which $\exp_p$ is a diffeomorphism. Let W be the union of all W_p as p runs over M, then W is an open neighbourhood of the zero-section in $\tau(M)$ (it is in fact a small disk bundle (see Chapter 5)). We also have a map

$$\text{Exp } : W \longrightarrow M \times M.$$

given by Exp $(p, v) = (p, \exp_p(v))$. The map Exp is also called exponential map. It is smooth, since $\pi_1 \circ$Exp$= \pi_1$ and $\pi_2 \circ$Exp $= exp$ are so, $\pi_i : M \times M \longrightarrow M$, $i = 1, 2$ being the projection maps.

Lemma 4.4.9. *The map Exp is a local diffeomorphism of an open neighbourhood of the zero-section in $\tau(M)$ onto an open neighbourhood of the diagonal in $M \times M$.*

Proof. There is a natural isomorphism between the tangent space $\tau(W)_{(p,0_p)}$ and $\tau(M)_p \times \tau(\tau(M)_p)_{0_p}$, while the tangent space to $M \times M$ at $\text{Exp}(p, 0_p) = (p, p)$ is naturally isomorphic to $\tau(M)_p \times \tau(M)_p$. In terms of these isomorphisms $(d\text{Exp})_{(p,0_p)}$ is the identity on the first factor and $(d \exp_p)_{0_P}$ on the second factor. Therefore the matrix of $(d\text{Exp})_{(p,0_p)}$ is the identity matrix. Now the inverse function theorem gives diffeomorphisms of open neighbourhoods of points $(p, 0_p)$ in $\tau(M)$ onto open neighbourhoods of (p, p) in $M \times M$. $\square$

Lemma 4.4.10. *If K is a compact subset of M, there is an $\epsilon > 0$ such that, for every $p \in K$, $\exp_p$ is defined on $B(0_p, \epsilon)$ and maps it diffeomorphically onto $B(p, \epsilon)$.*

Proof. By Lemma 4.4.9, for each $p \in M$, there is a neighbourhoods W_p of $(p, 0)$ in $\tau(M)$ which is mapped by Exp diffeomorphically onto a neighbourhood of (p, p) in $M \times M$. We may suppose that W_p comprises of points $(q, v) \in \tau(M)$, where q belongs to an open neighbourhood U_p of p in M and $\|v\| < \epsilon_p$. We map also suppose that U_p is sufficiently small so that $U_p \times U_p \subset$ Exp (W_p). Then, any two points of U_p can be joined by a unique geodesic of length less than ϵ_p. Therefore, by Corollary 4.4.6, for any $q \in U_p$, $\exp_q$ maps $B(0_q, \epsilon_p)$ diffeomorphically onto $B(q, \epsilon_p)$. By compactness, a compact set K is covered by a finite number of U_p, to each of which there corresponds an ϵ_p. Then, taking $\epsilon = \min \epsilon_p$ gives the desired result. $\square$

4.5. Hopf-Rinow theorem

Definition 4.5.1. A manifold is *geodesically complete at a point $p \in M$*, if the map $\exp_p$ can be defined on the entire tangent space $\tau(M)_p$. If M is geodesically complete at every point $p \in M$, then M is called **geodesically complete**.

Lemma 4.5.2. *If a Riemannian manifold M is complete as a metric space, then it is geodesically complete.*

PROOF. Suppose that M is complete, but not geodesically complete. Let $\sigma_v(t)$ be a geodesic in M which is defined only if t is in a finite open interval I. Let b be a boundary point of I, say the right hand end point. Let $\{t_n\}$ be a Cauchy sequence in I converging to b. Then $\{\sigma_v(t_n)\}$ is a Cauchy sequence in M, and hence it converges to a point $p \in M$. The point p does not depend on the choice of the sequence $\{t_n\}$. By Proposition 4.4.1, there is a relatively compact open neighbourhood U of p such that any geodesic with domain I can be extended to a bigger open interval J. This contradiction shows that M must be geodesically complete. $\square$

Let $p \in M$. For each integer $r > 0$ consider the sets

$$\begin{aligned} B_r &= \{x \in M \,|\, d(p,x) \le r\}, \\ C_r &= \{x \in B_r \,|\, p \text{ and } x \text{ can be joined by a geodesic of length } d(p,x)\}. \end{aligned}$$

Lemma 4.5.3. *If M is geodesically complete at p, then*

(a) *C_r is compact.*
(b) *If $C_r = B_r$ for some r, and $d(p,x) > r$ for some $x \in M$, then there is a point $y \in M$ such that $d(p,y) = r$, and $d(p,x) = d(x,y) + r$.*
(c) *If $C_r = B_r$, then there is an $\epsilon > 0$ such that $C_{r+\epsilon} = B_{r+\epsilon}$.*
(d) *$C_r = B_r$ if and only if $C_s = B_s$ for all $s < r$.*
(e) *$C_r = B_r$ for all r.*

PROOF. (a) Since M is geodesically complete at $p \in M$, we can define a function $f : \tau(M)_p \longrightarrow \mathbb{R}$ by

$$f(v) = d(p, \exp_p v) - \|v\|.$$

Then, for any $r > 0$, the set D_r given by

$$D_r = f^{-1}(0) \cap \overline{B}(0_p, r) = \{v \in \tau(M)_p | d(p, \exp_p v) = \|v\| \le r\},$$

where $\overline{B}(0_p, r)$ is the closed ball in $\tau(M)_p$, that is, a closed and bounded subset of $\tau(M)_p$, and hence it is compact by Heine-Borel theorem. The compactness of C_r follows from the facts that $C_r = \exp_p(D_r)$ and $\exp_p$ is continuous.

(b) By the definition of the metric d, for each integer $n > 0$ there is a smooth curve σ_n from p to x such that $l(\sigma_n) < d(p,x) + 1/n$. Let p_n be the last point on σ_n in B_r so that $d(p, p_n) = r$. Since $C_r = B_r$ is compact, a subsequence of $\{p_n\}$ converges to a point $y \in C_r$, we may assume that the sequence $\{p_n\}$ itself converges to y. Then

$$d(p,x) = \lim_{n\to\infty} d(p, p_n) = r,$$

and

$$d(p,x) \le d(p,y) + d(y,x) = r + d(x,y).$$

On the other hand,

$$\begin{aligned} l(\sigma_n) &= l(\text{ part of } \sigma_n \text{ from } p \text{ to } p_n) + l(\text{ part of } \sigma_n \text{ from } p_n \text{ to } x) \\ &\geq r + d(p_n, x). \end{aligned}$$

Therefore,

$$\begin{aligned} d(p,x) &> l(\sigma_n) - 1/n \\ &\geq r + d(p_n, x) - 1/n, \end{aligned}$$

which has the limit $r + d(x,y)$, and therefore

$$d(p,x) \geq r + d(x,y).$$

So the equality holds, and the proof is complete.

(c) By Lemma 4.4.10, there exists $\epsilon > 0$ such that, for each $q \in B_r$, $\exp_p$ maps $B(0_p, 2\epsilon)$ diffeomorphically onto $B(p, 2\epsilon)$, since $B_r = C_r$ is compact. We shall show that $B_{r+\epsilon} \subseteq C_{r+\epsilon}$. This will complete the proof, since $B_{r+\epsilon} \supseteq C_{r+\epsilon}$ always.

For any $x \in B_{r+\epsilon}$, we may suppose that $d(p,x) > r$, because if $d(p,x) \leq r$, then we will have $x \in B_r = C_r \subset C_{r+\epsilon}$. Then (b) gives a $y \in B_r$ such that $d(p,y) = r$ and $d(p,x) = r + d(x,y)$. Therefore $d(x,y) < \epsilon$, and so there is a geodesic σ_1 from x to y with $\ell(\sigma_1) = d(x,y)$. As $y \in B_r = C_r$, p and y can be joined by a geodesic σ_2 with $\ell(\sigma_2) = d(p,y) = r$. Then $\sigma_1 + \sigma_2$ is a piecewise smooth curve from p to x with $\ell(\sigma_1+\sigma_2) = \ell(\sigma_1)+\ell(\sigma_2) = d(x,y)+r = d(p,x)$. By Corollary 4.4.8, $\sigma_1 + \sigma_2$ is a geodesic, and therefore $x \in C_{r+\epsilon}$.

(d) The 'only if' part follows trivially. Because, if $C_r = B_r$ and $s < r$, then $C_s = B_s \cap C_r = B_s \cap B_r = B_s$. For the 'if' part, it suffices only to show that $B_r \subseteq C_r$, because the other inclusion is always true. Take $x \in B_r$. Now if $x \in B_s$ also, for some $s < r$, we have by the hypothesis that $x \in C_s \subset C_r$. So suppose that $d(p,x) = r$. Take a sequence $r_k = r - 1/k$, $k = 1, 2, \ldots$. Then, since $d(p,x) = r > r_k$, we have by (b) that for each k there is an element x_k such that $d(p, x_k) = r_k$ and $d(p,x) = r_k + d(x, x_k)$. Then $x_k \in B_{r_k} = C_{r_k} \subset C_r$, and $\lim_{k\to\infty} d(x, x_k) = 0$. Therefore $x \in C_r$ by the compactness of C_r. This proves $C_r = B_r$.

(e) Let $I = \{r \in \mathbb{R} \mid C_r = B_r\}$. Then, using the results (a) to (c), one can show easily that I is closed and equal to the interval $[0, \infty)$. Therefore $C_r = B_r$ for all r. $\square$

Theorem 4.5.4 (Hopf-Rinow). *For a Riemannian manifold M the following assertions* (i) *to* (iv) *are equivalent, and each of them implies the assertion* (v)*.*

(i) *M is complete.*
(ii) *All bounded closed subsets of M are compact.*
(iii) *For some point $p \in M$, all geodesics from p are infinitely extendable.*
(iv) *All geodesics are infinitely extendable.*

(v) *Any two points p and q of M can be joined by a geodesic whose length is $d(p, q)$.*

PROOF. The result (ii) $\Rightarrow$ (i) is a classical result of topology, and the result (i) $\Rightarrow$ (iv) has already been proved in Lemma 4.5.2. That (iv) $\Rightarrow$ (iii) is trivial. Therefore we have only to prove that (iii) $\Rightarrow$ (ii) and (iii) $\Rightarrow$ (v).

Suppose that (iii) is true for a particular point $p \in M$. Then any bounded subset of M is in a B_r, and therefore, by Lemma 4.5.3(a) and (e), any closed bounded subset of M is compact, which is (ii). Finally, since (iii) and (iv) are equivalent, we may allow p to be any point of M, and find a B_r which contains q. By Lemma 4.5.3(e), $B_r = C_r$, and so we get (v). $\square$

4.6. Totally geodesic submanifolds

Let S be a connected submanifold of M. Then a Riemannian structure on M induces a Riemannian structure on S so that if d_M and d_S are the distances in M and S, then $d_M(p, q) \leq d_S(p, q)$ for $p, q \in S$. In order to distinguish between geodesics in M and S, we will call them M-geodesics and S-geodesics.

Lemma 4.6.1. *If a curve σ in S is an M-geodesic, then it is an S-geodesic.*

PROOF. Let $p \in S$, and $B(0_p, \delta)$ be an open ball in $\tau(M)_p$ which is mapped diffeomorphically by $\exp_p$ on to an open neighbourhood U of p in M. Let $q \in U$, and σ be an M-geodesic in U joining p and q. Then, by Lemma 4.4.7,

$$\ell(\sigma) = d_M(p, q) \leq d_S(p, q) \leq \ell(\sigma).$$

Therefore $\ell(\sigma) = d_S(p, q)$, and σ is a curve of shortest length in S, that is, σ is an S-geodesic. $\square$

Definition 4.6.2. Let S be a connected submanifold of a Riemannian manifold M, and $p \in S$. Then S is called *geodesic at p* if every M-geodesic which is tangent to S at p is a curve in S, and hence an S-geodesic.

The submanifold S is called **totally geodesic** if it is geodesic at each of its points.

Lemma 4.6.3. *If a connected submanifold S of M is geodesic at $p \in S$, then any S-geodesic through p is also an M-geodesic.*

PROOF. Let σ be an S-geodesic through p, and τ be the maximal M-geodesic tangent to S at p. Then by the above lemma, τ is an S-geodesic through p, and so $\sigma \subset \tau$. Therefore σ is an M-geodesic. $\square$

$\lozenge$ **Exercise 4.4.** Show that if M is a complete Riemannian manifold, and S is a totally geodesic submanifold of M, then S is also complete.

$\lozenge$ **Exercise 4.5.** Let S be a totally geodesic submanifold of M with inclusion map $i : S \longrightarrow M$. Show that each point $p \in S$ has an open neighbourhood U in S on which i is an isometry.

Theorem 4.6.4. *Let M be a manifold and S be a totally geodesic submanifold of M. Then M-parallel translation along curves in S transports tangent vectors to S to tangent vectors to S*

PROOF. Let $\dim M = n$, $\dim S = s$, and $p \in S$. Then there is coordinate neighbourhood U of p in M with coordinate functions $x_1, \dots, x_n$ such that the set

$$V = \{q \in U \mid x_i(q) = 0 \ for \ s < i \leq n\}$$

is a coordinate neighbourhood in S with coordinates functions $x_1, \dots, x_s$ restricted to V.

Let $\sigma(t)$ be a curve in V with $\sigma(0) = p$, and $X(t)$ be a vector field tangent to M along σ such that $X(0) \in \tau(S)_p$, and $X(t)$ is M-parallel along σ. Then, if $X(t) = \sum_{i=1}^n X_i(t)\delta_i$, $\delta_i = \partial/\partial x_i$, its components satisfy the equations

$$\frac{dX_k}{dt} + \sum_{i,j} \frac{dx_i}{dt} \cdot \Gamma_{ij}^k X_j = 0, \quad 1 \leq i, j, k \leq n. \tag{15}$$

$$X_k(0) = 0, \quad x_k(t) = 0 \text{ for } s < k \leq n.$$

Let $\beta(t)$ be an M-geodesic tangent to S at p, where t is the arc length parameter measured from p. Then, if t is sufficiently small, $x_k(t) = x_k(\beta(t))$ satisfy

$$\frac{d^2 x_k}{dt^2} + \sum_{i,j} \Gamma_{i,j}^k \frac{dx_i}{dt} \frac{dx_j}{dt} = 0, \ 1 \leq i, j, k \leq n, \tag{16}$$

$$\frac{dx_k}{dt}(0) = 0, \ s < k \leq n.$$

Throughout the remaining of the proof the indices i, j, k, and a, b will vary in the following way: $1 \leq i, j, k \leq n$, and $s < a, b \leq n$.

Since S is totally geodesic, the above M-geodesic β is also S-geodesic. Since, $\beta(t) \in V$ for small t, we have $x_a(t) = 0$. Since any S-geodesic is also M-geodesic, the equations (16) give

$$\Gamma_{ij}^a(q) = 0 \text{ for } \ q \in V. \tag{17}$$

For the curve σ above, we have $(dx_a/dt)(t) = 0$. Therefore, in view of (17) we get

$$\frac{dX_a}{dt} + \sum_{i,b} \Gamma_{jb}^a \frac{dx_j}{dt} X_b(t) = 0 \ \text{ on } \sigma. \tag{18}$$

Since $X(0) \in \tau(S)_p$, then $X_b(0) = 0$. Therefore we must have $X_b(t) = 0$, by the uniqueness theorem for the system of linear differential equations (18). Therefore $x(t)$ is tangent to S.

Now, let $\gamma : I \longrightarrow S$ be any curve, and $Y(t)$ be an M-parallel vector field along γ such that $Y(0) \in \tau(S)_{\gamma(0)}$. The set of all $t \in I$ such that $Y(t) \in$

$\tau(S)_{\gamma(t)}$ is closed in I. The set is also open by the above argument. Therefore $Y(t) \in \tau(S)_{\gamma(t)}$ for all $t \in I$. This completes the proof. $\square$

$\Diamond$ **Exercise 4.6.** Show that if S is complete, then the converse of the above theorem is also true

$\Diamond$ **Exercise 4.7.** Let M be a Riemannian manifold, and S a connected complete submanifold of M. Show that S is totally geodesic if and only if M-parallel translation of tangent vectors to S along curves in S coincides with S-parallel translation.

CHAPTER 5

VECTOR BUNDLES ON MANIFOLDS

5.1. Vector bundles

The theory of vector bundles provides a very elegant and concise language to describe many phenomena in manifolds. In its simplest situation, a vector bundle over a manifold M is a manifold E which is the disjoint union of a family of vector spaces $\{E_x\}$, indexed by $x \in M$. For example, a product $E = M \times \mathbb{R}^n$ is a vector bundle so that if $\pi : E \longrightarrow M$ is the natural projection map and $x \in M$, then $E_x = \pi^{-1}(x)$. This is called a product bundle over M. In general, a vector bundle of dimension n over M is obtained by gluing together a family of product bundles $\{U_j \times \mathbb{R}^n\}$ over members of an open covering $\{U_j\}$ of M by means of an action of the group of $GL(n, \mathbb{R})$ on $\mathbb{R}^n$ in a nice way. Here is the precise definition.

Definition 5.1.1. A **vector bundle** of dimension n is a triple (E, M, π) consisting of a pair of manifolds E and M connected by a smooth surjective map $\pi : E \longrightarrow M$ satisfying the following conditions.

VB1. For each $x \in M$, the inverse image $E_x = \pi^{-1}(x)$ is an n-dimensional vector space over $\mathbb{R}$,

VB2. For each $x \in M$, there is an open neighbourhood U of x and a diffeomorphism $\phi : U \times \mathbb{R}^n \longrightarrow \pi^{-1}(U)$ such that

(i) the following diagram commutes

$$\begin{tikzcd} U \times \mathbb{R}^n \arrow[rr, "\phi"] \arrow[dr, "p_1"'] & & \pi^{-1}(U) \arrow[dl, "\pi"] \\ & U & \end{tikzcd}$$

where p_1 is the projection onto the first factor,

(ii) for each $y \in U$, the map $\phi_y : \mathbb{R}^n \longrightarrow \pi^{-1}(y)$, defined by $\phi_y(v) = \phi(y, v)$, is a linear isomorphism.

Note that locally π is the composition of a diffeomorphism ϕ^{-1} followed by a submersion p_1, therefore π is a submersion.

A. Mukherjee, *Differential Topology*, DOI 10.1007/978-3-319-19045-7_5

The vector bundle is also denoted by the map $\pi : E \longrightarrow M$ (and sometimes E itself is called the vector bundle, by an abuse of language). The manifold E is called the **total space** of the bundle, the manifold M its **base space**, and the map π its **projection**. The inverse image $E_x = \pi^{-1}(x), x \in M$, is called the **fibre over** x. The condition VB2 is called the **local triviality**; the pair (U, ϕ) is called a **vector bundle chart**, or VB-chart, and U is called a **trivialising open set**. A collection $\{(U_i, \phi_i)\}$ of VB-charts is called a **vector bundle atlas** or VB-atlas.

Notice that the dimension of a vector bundle E is actually the dimension of its fibre, and is not the dimension of E as a manifold. The function $M \longrightarrow \mathbb{R}$ given by $x \mapsto \dim E_x$ is a locally constant function, and therefore it is a constant on each component of M. If the function is constant on the whole of M, then the common value $\dim E_x$, for all $x \in M$, is the dimension of the vector bundle E.

If $\pi : E \longrightarrow M$ is a vector bundle and S is a submanifold of M, then $E|S$ will denote the vector bundle $\pi : \pi^{-1}(S) \longrightarrow S$.

Example 5.1.2. If V is a finite dimensional vector space over $\mathbb{R}$, then the projection $\pi : M \times V \longrightarrow M$ is a vector bundle. This is called a **product bundle**.

Example 5.1.3. The tangent bundle $\tau(S^n)$ of the n-sphere S^n is a vector bundle $\pi : E \longrightarrow S^n$ whose total space E is

$$E = \{(x, v) \in \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} \mid \|x\| = 1, \langle x, v\rangle = 0\},$$

and the projection π is $\pi(x, v) = x$. To see the local triviality, consider the open covering $\{U_i\}$ of S^n given by

$$U_i = \{x = (x_1, \ldots, x_{n+1}) \in \mathbb{R}^{n+1} \mid \|x\| = 1, x_i \neq 0\},\ 1 \leq i \leq n+1,$$

(U_i is not connected, it has two components corresponding to $x_i > 0$ and $x_i < 0$). Define $\phi_i : U_i \times \mathbb{R}^n \longrightarrow \pi^{-1}(U_i)$ by

$$\phi_i(x, v) = (x, f_i(v) - \langle x, f_i(v)\rangle \cdot x),$$

where $f_i : \mathbb{R}^n \longrightarrow \mathbb{R}^{n+1}$ is given by

$$f_i(v_1, \ldots, v_n) = (v_1, \ldots, v_{i-1}, 0, v_i, \ldots, v_n).$$

Then ϕ_i satisfies $\pi \circ \phi_i = p_1$, and is linear on the fibres. Since $x_i \neq 0$, x lies outside $\operatorname{Image} f_i$, and so $v \neq 0$ implies $f_i(v) - \langle x, f_i(v)\rangle \cdot x \neq 0$. Thus ϕ_i is injective (and hence isomorphism) on each fibre. It may be seen easily that ϕ_i is a diffeomorphism.

Example 5.1.4. If $\pi_1 : E_1 \longrightarrow M$ and $\pi_2 : E_2 \longrightarrow M$ are vector bundles over M, let

$$E_1 \oplus E_2 = \{(v_1, v_2) \in E_1 \times E_2 \mid \pi_1(v_1) = \pi_2(v_2)\}.$$

Then $\pi : E_1 \oplus E_2 \longrightarrow M$ given by $\pi(v_1, v_2) = \pi_1(v_1) = \pi_2(v_2)$ is a vector bundle, whose fibre over $x \in M$ is the direct sum $\pi_1^{-1}(x) \oplus \pi_2^{-1}(x)$. We defer the proof until §5.2.

The bundle $E_1 \oplus E_2$ is call the **Whitney sum** of E_1 and E_2.

Example 5.1.5. The tangent bundle $\tau(M)$ of a manifold M is a vector bundle with dimension equal to $\dim M$. The charts constructed in Theorem 3.2.2 to show that $\tau(M)$ is a manifold are actually VB-charts.

♢ **Exercise 5.1.** Show that the cotangent bundle $\tau(M)^*$ is a vector bundle over M.

Definition 5.1.6. If (E, M, π) and (E', M', π') are vector bundles, then a **morphism** or **bundle map** $(f, g) : (E, M, \pi) \longrightarrow (E', M', \pi')$ consists of a pair of smooth maps $f : E \longrightarrow E'$ and $g : M \longrightarrow M'$ such that

(i) the following diagram commutes

$$\begin{tikzcd} E \arrow[r, "f"] \arrow[d, "\pi"'] & E' \arrow[d, "\pi'"] \\ M \arrow[r, "g"] & M' \end{tikzcd}$$

(ii) the restriction of f to each fibre

$$f_x = f|E_x : E_x \longrightarrow E'_{g(x)}, \qquad x \in M,$$

is a linear map.

A pair $(\mathrm{Id}_E, \mathrm{Id}_M)$ is the identity morphism. The composition of two morphisms (f, g) and (f', g') is defined to be $(f' \circ f, g' \circ g)$, and this is again a morphism.

A morphism $(f, \mathrm{Id}_M) : (E, M, \pi) \longrightarrow (E', M, \pi')$ over the same base space M is called a **homomorphism**, and sometimes it is denoted simply by $f : E \longrightarrow E'$. It is called a **monomorphism, epimorphism**, or **isomorphism** according to whether each f_x is a monomorphism, epimorphism, or isomorphism. It is called a **bundle equivalence** if f is an isomorphism and a diffeomorphism. In this case, each $f|E_x$ is a linear isomorphism, and its inverse is the restriction to E'_x of f^{-1}, and we write $E \approx E'$.

Lemma 5.1.7. *A homomorphism $f : (E, M, \pi) \longrightarrow (E', M, \pi')$ is a bundle equivalence if and only if f is an isomorphism.*

Proof. It suffices to show that a smooth map $f : E \longrightarrow E'$ over M which maps each fibre isomorphically onto a fibre, is a diffeomorphism. Define a map $g : E' \longrightarrow E$ by $g(\alpha') = f_x^{-1}(\alpha')$, where $\alpha' \in E'$, and $\pi'(\alpha') = x$. Then f will be a diffeomorphism, if we show that g is a smooth map.

Let $\phi : U \times \mathbb{R}^n \longrightarrow \pi^{-1}(U)$ and $\phi' : U \times \mathbb{R}^n \longrightarrow (\pi')^{-1}(U)$ be VB-charts for E and E' respectively corresponding to a common trivialising open set U in M. Then $(\phi')^{-1} \circ f \circ \phi$ is of the form $(x, v) \mapsto (x, h(x)v)$, where $h : U \longrightarrow L(\mathbb{R}^n, \mathbb{R}^n)$. Thus the map $f : \pi^{-1}(U) \longrightarrow (\pi')^{-1}(U)$ over U is smooth if and only if h is smooth. Also f is an isomorphism on each fibre if and only if Image $h \subset GL(n, \mathbb{R})$. Since the map $GL(n, \mathbb{R}) \longrightarrow GL(n, \mathbb{R})$ given by $\lambda \mapsto \lambda^{-1}$ is smooth, the map $h^{-1} : U \longrightarrow GL(n, \mathbb{R})$, given by $h^{-1}(x) = h(x)^{-1}$, is also smooth, if h is so. Using these facts, we find

$$f|\pi^{-1}(U) \text{ smooth } \Rightarrow (\phi')^{-1} \circ f \circ \phi \text{ smooth } \Rightarrow h \text{ smooth } \Rightarrow$$
$$h^{-1} \text{ smooth } \Rightarrow \phi^{-1} \circ g \circ \phi' \text{ smooth } \Rightarrow g|(\pi')^{-1}(U) \text{ smooth }.$$

Since this is true for any common trivialising open set U, g is a smooth map. □

A vector bundle is called **trivial** if it is equivalent to a product bundle.

Example 5.1.8. The normal bundle $\nu(S^n)$ of S^n of the n-sphere S^n, $n \geq 1$, is the bundle $\pi : E \longrightarrow S^n$, where

$$E = \{(x, v) \in \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} \mid \|x\| = 1, v = \lambda x, \lambda \in \mathbb{R}\},$$

and π is $\pi(x, v) = x$. Define $\phi : E \longrightarrow S^n \times \mathbb{R}$ and $\psi : S^n \times \mathbb{R} \longrightarrow E$ by

$$\phi(x, v) = (x, \langle x, v\rangle), \text{ and } \psi(x, \lambda) = (x, \lambda x).$$

Then ϕ and ψ are bundle homomorphisms inverse to one another, and so $\nu(S^n)$ is a trivial bundle.

Example 5.1.9. Let θ_k be the product bundle $S^n \times \mathbb{R}^k \longrightarrow S^n$. Then we have a bundle equivalence

$$\tau(S^n) \oplus \theta_1 \approx \theta_{n+1}.$$

The bundle isomorphism $f : \tau(S^n) \oplus \theta_1 \longrightarrow S^n \times \mathbb{R}^{n+1}$ is given by $f(x, (v, \lambda)) = (x, \lambda x + v)$, where $v \in \tau(S^n)_x$ and $\lambda \in \mathbb{R}$. Its inverse $g : S^n \times \mathbb{R}^{n+1} \longrightarrow \tau(S^n) \oplus \theta_1$ is given by $g(x, w) = \big(x, (w - \langle x, w\rangle \cdot x, \langle x, w\rangle)\big)$, where $w \in \mathbb{R}^{n+1}$.

Implicit in the above definition of vector bundle is an important role of the general linear group $GL(n, \mathbb{R})$, which appears in a transition law (the equation (1) below) between VB-charts. This may be described in the following way. The definition of a vector bundle $\pi : E \longrightarrow M$ guarantees the existence of a trivialising covering $\{U_i\}_{i\in\Lambda}$ of M (Λ = index set) and the VB-charts $\phi_i : U_i \times \mathbb{R}^n \longrightarrow \pi^{-1}(U_i)$ satisfying VB2. Then, for any $i, j \in \Lambda$ with $U_i \cap U_j \neq \emptyset$, the diffeomorphism

$$\phi_j^{-1} \circ \phi_i : (U_i \cap U_j) \times \mathbb{R}^n \longrightarrow (U_i \cap U_j) \times \mathbb{R}^n$$

must be of the form

$$\phi_j^{-1} \circ \phi_i(x, v) = (x, g_{ji}(x)v) \tag{1}$$

for a unique smooth map $g_{ji} : U_i \cap U_j \longrightarrow GL(n, \mathbb{R})$. In fact, if

$$\phi_{ix} : \mathbb{R}^n \longrightarrow \pi^{-1}(x)$$

is the map $\phi_{ix}(v) = \phi_i(x, v)$, then

$$g_{ji}(x) = \phi_{jx}^{-1} \circ \phi_{ix}$$

is a linear isomorphism $\mathbb{R}^n \longrightarrow \mathbb{R}^n$. The family of maps $\{g_{ji}\}$ is called a **cocycle**. They satisfy the following condition, called the **cocycle condition**,

$$g_{kj}(x) \cdot g_{ji}(x) = g_{ki}(x), \qquad x \in U_i \cap U_j \cap U_k. \tag{2}$$

Putting $i = j = k$, and then $k = i$ in this condition, we get

$$g_{ii}(x) = \text{ id, for all } x \in U_i, \text{ and } g_{ij}(x) = (g_{ji}(x))^{-1} \text{ for all } x \in U_i \cap U_j.$$

A more general definition of a vector bundle replaces the group $GL(n, \mathbb{R})$ by a subgroup G of it, and includes the transition law (1) as a third condition VB3, where the g_{ji} are now maps from $U_i \cap U_j$ into G satisfying the cocycle condition (2). It is then called a vector bundle with structure group G (cf.Definition 5.1.19 below). If G is the trivial group $\{1\}$, then it follows from VB2 that E is the product bundle $M \times \mathbb{R}^n$.

Theorem 5.1.10. (i) *Given an open covering $\{U_i\}_{i \in \Lambda}$ of M and smooth maps*

$$g_{ji} : U_i \cap U_j \longrightarrow G, \qquad G \subseteq GL(n, \mathbb{R}),$$

defined for all $i, j \in \Lambda$ for which $U_i \cap U_j \neq \emptyset$, and satisfying the cocycle condition (2)*, there is a vector bundle $\overline{\pi} : \overline{E} \longrightarrow M$ having its associated cocycle as $\{g_{ji}\}$.*

(ii) *If $\pi : E \longrightarrow M$ is a vector bundle with associated cocycle $\{g_{ji}\}$, and $\overline{\pi} : \overline{E} \longrightarrow M$ is the vector bundle constructed using the g_{ji} as in* (i) *above, then the two bundles E and $\overline{E}$ are equivalent.*

Proof. (i) Consider the set of all triples $(x, v, i) \in M \times \mathbb{R}^n \times \Lambda$ such that $x \in U_i$. Define an equivalence relation $\sim$ on this set by $(x, v, i) \sim (y, w, j)$ if and only if $x = y \in U_i \cap U_j$ and $w = g_{ji}(x)v$. That this is an equivalence relation follows from the cocycle condition. Let $\overline{E}$ be the resulting quotient space. Define $\overline{\pi} : \overline{E} \longrightarrow M$ by $\overline{\pi}([x, v, i]) = x$, where $[x, v, i]$ is the equivalence class of (x, v, i). Let $\overline{U}_i = \overline{\pi}^{-1}(U_i)$. Define $\overline{\phi}_i : U_i \times \mathbb{R}^n \longrightarrow \overline{U}_i$ by $\overline{\phi}_i(x, v) = [x, v, i]$. It may be seen without difficulty that there is a unique smooth structure on $\overline{E}$ with respect to which each $\overline{\phi}_i$ is a diffeomorphism. It follows that $\overline{\pi} : \overline{E} \longrightarrow M$ is a smooth vector bundle M. The cocycle $\{\overline{g}_{ji}\}$ of $\overline{E}$ is $\{g_{ji}\}$, because

$$\begin{aligned} g_{ji}(x)v = w \quad &\Leftrightarrow \quad (x, w, j) \sim (x, v, i) \\ &\Leftrightarrow \quad \overline{\phi}_j(x, w) = \overline{\phi}_i(x, v) \ \Leftrightarrow \overline{g}_{ji}(x)v = w. \end{aligned}$$

(ii) Define a map $f : E \longrightarrow \overline{E}$ as follows. If $\alpha \in E$ with $\pi(\alpha) = x \in U_i$ and $\phi_i^{-1}(\alpha) = (x, v)$, then set $f(\alpha) = \overline{\phi}_i(x, v)$. If $x \in U_j$ also and $\phi_j^{-1}(\alpha) = (x, w)$, then $w = g_{ji}(x)v$ by (1), and so $\overline{\phi}_i(x, v) = \overline{\phi}_j(x, w)$. Therefore f is well defined. Since $f|\pi^{-1}(U_i)$ is the composition $\overline{\phi}_i \circ \phi_i^{-1}$ of two diffeomorphisms, f is a diffeomorphism. □

Lemma 5.1.11. *If $\{(U_i, \phi_i)\}$ and $\{(U_i, \phi_i')\}$ are two VB-atlases of a vector bundle E over M defined for the same trivialising open covering $\{U_i\}$ of M and* *with cocycles* $\{g_{ji}\}$ *and* $\{g'_{ji}\}$ *then there exist smooth maps* $\lambda_i : U_i \longrightarrow GL(n, \mathbb{R})$ *for each* i *such that*

$$g'_{ji}(x) = \lambda_j(x) \circ g_{ji}(x) \circ \lambda_i(x)^{-1} \quad \text{for } x \in U_i \cap U_j.$$

PROOF. At each point $x \in U_i$, ϕ_{ix} and ϕ'_{ix} differ by a linear isomorphism of $\mathbb{R}^n$, These determine a smooth map $\lambda_i : U_i \longrightarrow GL(n, \mathbb{R})$ such that $\phi_{ix} = \phi'_{ix} \circ \lambda_i(x)$. Therefore

$$g'_{ji}(x) = \phi'^{-1}_{jx} \circ \phi'_{ix} = \lambda_j(x) \circ \phi_{jx}{}^{-1} \circ \phi_{ix} \circ \lambda_i(x)^{-1} = \lambda_j(x) \circ g_{ji}(x) \circ \lambda_i(x)^{-1}.$$

□

Definition 5.1.12. Two cocycles $\{g_{ji}\}$ and $\{g'_{ji}\}$ related as in the above lemma are called **equivalent**.

Theorem 5.1.13. *Two vector bundles* E *and* E' *over* M *are isomorphic if and only if their cocycles* $\{g_{ji}\}$ *and* $\{g'_{ji}\}$ *relative to the same trivialising covering* $\{U_i\}$ *of* M *are equivalent.*

PROOF. Suppose that $f : E \longrightarrow E'$ is an isomorphism. Then the relation $\phi_j^{-1} \circ \phi_i(x, v) = (x, g_{ji}(x)v)$ implies $(f\phi_j)^{-1} \circ (f\phi_i)(x, v) = (x, g_{ji}(x)v)$, and therefore $\{(U_i, f\phi_i)\}$ is a VB-atlas for E' with cocycle $\{g_{ji}\}$. Then the cocycles $\{g_{ji}\}$ and $\{g'_{ji}\}$ are equivalent, by Lemma 5.1.11.

Conversely, suppose that $g'_{ji}(x) = \lambda_j(x) \circ g_{ji}(x) \circ \lambda_(x)^{-1}$ for $x \in U_i \cap U_j$ and smooth maps $\lambda_j : U_j \longrightarrow GL(n, \mathbb{R})$. Define smooth maps

$$f_j : U_j \times \mathbb{R}^n \longrightarrow U_j \times \mathbb{R}^n$$

by $f_j(x, v) = (x, \lambda_j(x)v)$, and then define $f : E \longrightarrow E'$ by setting $f = \phi'_j \circ f_j \circ \phi_j^{-1}$ on $\pi^{-1}(U_j)$. Then the definitions of f agree on $U_i \cap U_j$, because $f\phi_j = \phi'_j f_j$ implies that $f\phi_i = \phi'_i f_i$ on $U_i \cap U_j$. To see this, take $(x, v) \in (U_i \cap U_j) \times \mathbb{R}^n$. Then

$$\begin{aligned} f\phi_i(x, v) &= f\phi_j(x, g_{ji}(x)v) = \phi'_j f_j(x, g_{ji}(x)v) \\ &= \phi'_j(x, \lambda_j(x) \circ g_{ji}(x)v) = \phi'_j(x, g'_{ji}(x) \circ \lambda_i(x)v) \\ &= \phi'_i(x, \lambda_i(x)v) = \phi'_i f_i(x, v). \end{aligned}$$

Therefore f is well-defined. Also f is an isomorphism, since it is so locally. □

◊ **Exercise 5.2.** Define a more general equivalence relation between cocycles

$$g_{ji} : U_i \cap U_j \longrightarrow G, \ i, j \in \Lambda \ \text{ and } \ g'_{\beta\alpha} : U'_\alpha \cap U'_\beta \longrightarrow G, \ \alpha, \beta \in \Lambda'.$$

of two vector bundles having the same base, fibre, and structure group in the following way:

The cocycles $\{g_{ji}\}$ and $\{g'_{\beta\alpha}\}$ are equivalent if and only if there exist smooth maps $\overline{g}_{\alpha i} : U_i \cap U'_\alpha \longrightarrow G$, $i \in \Lambda$, $\alpha \in \Lambda'$, such that

$$\begin{aligned} \overline{g}_{\alpha i} &= \overline{g}_{\alpha j} \cdot g_{ji} \ \text{ on } \ U_i \cap U_j \cap U'_\alpha, \\ \overline{g}_{\beta j} &= g'_{\beta\alpha} \cdot \overline{g}_{\alpha j} \ \text{ on } \ U_j \cap U'_\alpha \cap U'_\beta. \end{aligned}$$

The previous definition of equivalence (when the trivialising coverings are the same) may be obtained from this by taking $\lambda_i = (\overline{g}_{ii})^{-1}$.

Show that two vector bundles are isomorphic if and only if their cocycles are equivalent in the general sense.

Definition 5.1.14. A **section** of a vector bundle $\pi : E \longrightarrow M$ is a smooth map $s : M \longrightarrow E$ such that $\pi \circ s = \mathrm{Id}_M$.

For example, a section of the tangent bundle $\tau(M)$ is a vector field on M; a section of the cotangent bundle $\tau^*(M)$ is a differential 1-form on M.

A set of sections $(s_1, \ldots, s_n)$ over an open set U of M is called a **frame** on U if for every $x \in U$, the vectors $s_1(x), \cdots, s_n(x)$ form a basis of the fibre E_x.

$\Diamond$ **Exercise 5.3.** Show that a vector bundle $\pi : E \longrightarrow M$ is a trivial bundle if and only if it admits a frame on M.

Remarks 5.1.15. (1) Let us notice at once that the local triviality can be described quite explicitly as follows. For each point $x \in M$ there is an open neighbourhood U of x and a frame $(s_1, \ldots, s_n)$ on U such that the map $\phi : U \times \mathbb{R}^n \longrightarrow \pi^{-1}(U)$ given by

$$\phi(y, (v_1, \ldots, v_n)) = \sum_{i=1}^{n} v_i s_i(y)$$

is a diffeomorphism.

(2) Every section of a product bundle $M \times \mathbb{R}^n$ has the form $x \mapsto (x, f(x))$, where $f : M \longrightarrow \mathbb{R}^n$ is a smooth map. In fact, there is a bijection from the set of sections of $M \times \mathbb{R}^n$ to the set of smooth maps $M \longrightarrow \mathbb{R}^n$, where a section s corresponds to the map $p_2 \circ s$, p_2 being the projection onto the second factor.

(3) For the equivalent bundle $\overline{E}$ of Theorem 5.1.10(ii), a section s may be written as $s(x) = [x, s_i(x), i]$, where $s_i : U_i \longrightarrow \mathbb{R}^n$ is a smooth map, called the local representative of s on U_i. The local representatives are related by the formula $s_j = g_{ji} \circ s_i$ on $U_i \cap U_j$.

Example 5.1.16. The tangent bundle $\tau(G)$ of a Lie group G is trivial. In particular, $\tau(S^3)$ is trivial, since S^3 is a Lie group (unit quaternions).

To see this, let $R_g : G \longrightarrow G$ denote the right translation by $g \in G$ defined by $R_g(h) = h \cdot g$. This is a diffeomorphism with inverse $R_{g^{-1}}$. Therefore the differential $dR_g : \tau(G)_e \longrightarrow \tau(G)_g$, $e =$ identity element of G, is a linear isomorphism. Let $\dim G = n$, and $\{v_1, \ldots, v_n\}$ be a basis of $\tau(G)_e$. Define vector fields

$$X_i : G \longrightarrow \tau(G),\ i = 1, \cdots, n,$$

by $X_i(g) = dR_g(v_i)$. These are smooth vector fields, since the multiplication in G is smooth. One checks immediately that the map

$$f : G \times \mathbb{R}^n \longrightarrow \tau(G)$$

defined by $f(g, \lambda_1, \ldots, \lambda_n) = (g, \sum_i \lambda_i X_i(g))$ is a vector bundle isomorphism.

The section, which maps each $x \in M$ onto the zero vector of the vector space E_x, is called the **zero-section**. It is an embedding of M into E.

Definition 5.1.17. A **metric** on a vector bundle $E \longrightarrow M$ is a smooth map which assigns to each $x \in M$ a positive definite symmetric bilinear form or inner product on the fibre E_x

$$\langle \ , \ \rangle_x : E_x \times E_x \longrightarrow \mathbb{R}.$$

We shall show in Example 5.2.4(5) that the collection of all metrics on E form a vector bundle $S^2(E)$ over M, whose fibre over $x \in M$ is the vector space of all inner products on the fibre E_x. Then a metric on E will be a smooth section of the vector bundle $S^2(E)$.

A metric on the tangent bundle $\tau(M)$ is a Riemannian metric on M, which we have discussed already in Chapter 4.

The standard metric $\langle \ , \ \rangle$ on $\mathbb{R}^n$ defines a metric on the product bundle $M \times \mathbb{R}^n$ by

$$\langle (x, v), (x, w) \rangle_x = \langle v, w \rangle.$$

Theorem 5.1.18. *Any vector bundle E over M admits a metric.*

Proof. The proof is similar to that of Theorem 4.1.3. One has to take metrics on the trivial bundles $E|U_i$, for some trivialising open covering $\{U_i\}$ of M, and then splice them together using a smooth partition of unity. □

At this point it is well worth considering for better perspective more general objects called fibre bundles. A fibre bundle is like a vector bundle, except that the role of $\mathbb{R}^n$ is replaced by a manifold F. Let G be a Lie group acting effectively on F (see Definition 1.7.3). Then G may be identified with a subgroup of the group Diff(F) of diffeomorphisms of F.

Definition 5.1.19. A **fibre bundle** over M with fibre F and structure group G is a submersion $\pi : E \longrightarrow M$ possessing a fibre bundle atlas $\{(U_i, \phi_i)\}$, $i \in \Lambda$, which consists of an open covering $\{U_i\}$, $i \in \Lambda$, of M (called a trivialising covering), and diffeomorphisms

$$\phi_i : U_i \times F \longrightarrow \pi^{-1}(U_i)$$

such that $\pi \circ \phi_i = p_1$, and such that the transition maps

$$\phi_{ji} : (U_i \cap U_j) \times F \longrightarrow (U_i \cap U_j) \times F,$$

defined by $\phi_{ji} = \phi_j^{-1} \circ \phi_i$, satisfy $\phi_{ji}(x, y) = (x, g_{ji}(x, y))$, where

$$g_{ji} : U_i \cap U_j \longrightarrow G$$

are smooth maps.

The collection of maps $\{g_{ji}\}$ is called a cocycle. They satisfy the cocycle condition (2).

This becomes the definition of a vector bundle in the special case when $G \subseteq GL(n, \mathbb{R})$ and $F = \mathbb{R}^n$. All our previous definitions and results up to Theorem 5.1.13 apply to fibre bundles, provided we replace $GL(n, \mathbb{R})$ by G and $\mathbb{R}^n$ by F, and 'linear isomorphism' by 'diffeomorphism', wherever they occur.

Definition 5.1.20. A fibre bundle $\pi : P \longrightarrow M$ is called a **principal G-bundle** if its fibre is the same as its structure group G, where G acts on itself by left translation.

$\diamondsuit$ **Exercises 5.4.** Let $\pi : P \longrightarrow M$ be a principal G-bundle with a trivialising covering $\{U_j\}$ of M.

(i) Show that there is a unique free right action of G on P, which is simply-transitive on each fibre $\pi^{-1}(x)$, $x \in M$, such that for each fibre bundle chart $\phi_i : U_i \times G \longrightarrow \pi^{-1}(U_i)$ we have $\phi_i(x, g) \cdot h = \phi_i(x, gh)$ for $x \in U_i$ and $g, h \in G$, and such that the orbits αG, $\alpha \in P$, coincide with the fibre $\pi^{-1}(x)$, where $x = \pi(\alpha)$.

(ii) Define $f_j : \pi^{-1}(U_j) \longrightarrow G$ by $f_j(\alpha) = \phi_{jx}^{-1}(\alpha)$ where $x = \pi(\alpha)$. Show that

(1) $f_j(\alpha \cdot g) = f_j(\alpha) \cdot g$,

(2) $f_j \circ \phi_j(\alpha, g) = g$,

(3) $\phi_j(\pi(\alpha), f_j(\alpha)) = \alpha$,

(4) $g_{ji}(\pi(\alpha)) \cdot f_i(\alpha)) = f_j(\alpha),\ \pi(\alpha) \in U_i \cap U_j$.

Hint (i). First define a right action of G on $U_i \times G$ by $(x, g) \cdot h = (x, gh)$. Then use it to define a right action of G on $\pi^{-1}(U_i)$ by $\phi_i(x, g) \cdot h = \phi_i(x, gh)$. The definition is independent of the choice of charts, because

$$(\phi_j^{-1} \circ \phi_i(x, g)) \cdot h = \phi_j^{-1} \circ \phi_i((x, g) \cdot h).$$

Let $\pi : E \longrightarrow M$ be a fibre bundle with fibre F, structure group G, trivialising covering $\{U_j\}$, and cocycle $g_{ji} : U_i \cap U_j \longrightarrow G$. Then we can construct as in Theorem 5.1.10(i) a principal G-bundle $P \longrightarrow M$ using the same M, the same $\{U_j\}$, the same $\{g_{ji}\}$, but replacing $\mathbb{R}^n$ by G and allowing G to act on itself by left translation (note that here G need not be a subgroup of $GL(n, \mathbb{R})$). This principal bundle is called the **associated principal bundle** of the fibre bundle $\pi : E \longrightarrow M$. The bundles have the same cocycle. It follows that two fibre bundles having the same fibre and group are equivalent if and only if their associated principal bundles are equivalent (cf. Definition 5.1.12).

The following example shows that the reverse process is true,

Example 5.1.21. (**Associated bundle**) Let $\pi : P \longrightarrow M$ be a principal G-bundle with trivialising covering $\{U_j\}$, and cocycle $g_{ji} : U_i \cap U_j \longrightarrow G$. Let G act effectively on the left of a manifold F. Then there is a fibre bundle $\pi_F : P[F] \longrightarrow M$ with fibre F, group G, and cocycle $\{g_{ji}\}$, called the fibre bundle associated to the principal bundle P. The construction is given below.

Define a right action of G on $P \times G$ by $(\alpha, y) \cdot g = (\alpha g, g^{-1}y)$ for $g \in G$, $\alpha \in P$, and $y \in F$. Let $P[F] = P \times G/G$ be the quotient space. Denote the equivalence class of (α, y) by $[\alpha, y]$. Then $[\alpha g, y] = [\alpha, gy]$. Define $\pi_F : P[F] \longrightarrow M$ by $\pi_F[\alpha, y] = \pi(\alpha)$. This is well-defined and continuous.

If $\{\phi_j : U_j \times G \longrightarrow \pi^{-1}(U_j)\}$ is an atlas for P, define an atlas

$$\{\psi_j : U_j \times F \longrightarrow \pi_F^{-1}(U_j)\}$$

by $\psi_j(x, y) = [\phi_j(x, e), y]$, $x \in U_j$, $y \in F$, e unit in G. Then ψ_j is a homeomorphism with inverse $\psi_j^{-1} : \pi_F^{-1}(U_j) \longrightarrow U_j \times F$ given by

$$\psi_j^{-1}[\alpha, y] = (\pi(\alpha), f_j(\alpha) \cdot y),$$

where $f_j : \pi^{-1}(U_j) \longrightarrow G$ is the smooth map defined in Exercise 5.4 (ii), p. 141. We can endow a smooth structure on $\pi_F^{-1}(U_j)$ so that ψ_j becomes a diffeomorphism. Also $\pi_F \circ \psi_j = p_1$. Thus $P[F]$ is a fibre bundle over M with fibre F and structure group G. Its cocycle is $\{g_{ji}\}$, because

$$\begin{aligned} \psi_{jx}^{-1} \circ \psi_{ix}(y) &= \psi_{jx}^{-1}[\phi_i(x, 1), y] \\ &= f_j(\phi_i(x, 1)) \cdot y = g_{ji}(x) \cdot f_i\phi_i(x, 1) \cdot y = g_{ji}(x) \cdot y, \end{aligned}$$

by the properties of the f_j.

◊ **Exercise 5.5.** Show that if $\pi : E \longrightarrow M$ is a vector bundle of dimension n, and $P \longrightarrow M$ is its associated principal bundle, then E is equivalent to $P[\mathbb{R}^n]$

◊ **Exercise 5.6.** Show that if E is a vector bundle over M with fibre dimension n and with a Riemannian metric, and a cocycle $\{g_{ji}\}$ so that each map g_{ji} takes values in the orthogonal group $O(n)$ (it is always possible to have such a cocycle, see Theorem 5.6.4 below) then each of the following subspaces of E

(1) $D(E) = \{v \in E \mid \|v\| \leq 1\|$,
(2) $S(E) = \{v \in E \mid \|v\| = 1\|$,

is a fibre bundle over M associated to the vector bundle E (that is, to the principal bundle associated to E) with fibre D^n and S^{n-1} respectively.

These are also called respectively the **disk bundle** and the **sphere bundle** associated to the vector bundle E.

5.2. Construction of vector bundles

We shall describe a method, which is due to Atiyah, of constructing a wide variety of vector bundles.

Let $\mathcal{V}$ be the category whose objects are finite dimensional vector spaces and morphisms are linear maps. Let $\mathcal{F} : \mathcal{V} \longrightarrow \mathcal{V}$ be a covariant functor. This consists of an object function $V \mapsto \mathcal{F}(V)$, where both V and $\mathcal{F}(V)$ are objects of $\mathcal{V}$, and a morphism function $\mathcal{F} : L(V, W) \longrightarrow L(\mathcal{F}(V), \mathcal{F}(W))$, where V and W are vector spaces and $L(V, W)$ is the set of linear maps from V to W, such that if $f : V \longrightarrow W$ and $g : W \longrightarrow X$ are linear maps, then $\mathcal{F}(g \circ f) = \mathcal{F}(g) \circ \mathcal{F}(f)$, and $\mathcal{F}(\mathrm{id}_V) = \mathrm{id}_{\mathcal{F}(V)}$. Note that $L(V, W)$ is a smooth manifold, since it is a vector space isomorphic to $\mathbb{R}^{nm}$ where $n = \dim V$ and $m = \dim W$.

The functor $\mathcal{F}$ is called a **smooth functor** if for every pair of vector spaces V and W, the map $\mathcal{F} : L(V, W) \longrightarrow L(\mathcal{F}(V), \mathcal{F}(W))$ is smooth map between manifolds.

Theorem 5.2.1. *Given a smooth vector bundle $\pi : E \longrightarrow M$ and a smooth covariant functor $\mathcal{F} : \mathcal{V} \longrightarrow \mathcal{V}$, the disjoint union of vector spaces $\mathcal{F}(E) = \cup_{x\in M}\mathcal{F}(E_x)$ is a vector bundle over M.*

We first prove a lemma.

Lemma 5.2.2. *Let $\pi : E \longrightarrow M$ be a vector bundle, E' be a set, and $\pi' : E' \longrightarrow M$ be a surjective set function such that the fibre $E'_x = {\pi'}^{-1}(x)$ is a vector space for each $x \in M$. Let $\lambda : E \longrightarrow E'$ be a bijection such that $\pi = \pi' \circ \lambda$ and $\lambda|E_x : E_x \longrightarrow E'_x$ is a linear map for each $x \in M$. Then there is a unique smooth structure on E' so that $\pi' : E' \longrightarrow M$ becomes a vector bundle and $\lambda : E \longrightarrow E'$ a bundle equivalence.*

Proof. We transport the topological structure of E to E' by means of the bijection λ, by defining the open sets of E' as the images under λ of the open sets of E. Since λ then becomes a homeomorphism, we may also transport the smooth structure of E to E' by means of the homeomorphism λ, as described in Theorem 1.3.1. Then E' gets a unique manifold structure that makes λ into a diffeomorphism. If $\phi : U \times \mathbb{R}^n \longrightarrow \pi^{-1}(U)$ is a VB-chart chart in E, where n is the fibre dimension of E, then the diffeomorphism $\lambda \circ \phi : U \times \mathbb{R}^n \longrightarrow (\pi')^{-1}(U)$ gives VB-chart for E', showing the local triviality of E'. $\square$

Proof of the theorem. The assertion is true if E is the product family $M \times V$, $V \in \mathcal{V}$, with $E_x = \{x\} \times V$. Because, then we have a bijection of $\mathcal{F}(E) = \cup_{x\in M}\mathcal{F}(E_x)$ onto the vector bundle $E' = M \times \mathcal{F}(V)$, and therefore the lemma applies.

The assertion is also true if E is a trivial bundle. In this case, we have a bundle equivalence $\lambda : E \longrightarrow M \times V = E'$ for some $V \in \mathcal{V}$. Then λ defines a bijection $\mathcal{F}(\lambda) : \mathcal{F}(E) \longrightarrow \mathcal{F}(E')$ by $\mathcal{F}(\lambda)(v) = \mathcal{F}(\lambda_x)(v)$, where $\lambda_x = \lambda|E_x : E_x \longrightarrow E'_x$, $x \in M$, and $v \in E_x$. Since $\mathcal{F}(E')$ is a vector bundle, $\mathcal{F}(E)$ is a vector bundle and $\mathcal{F}(\lambda)$ is a bundle equivalence, by the lemma.

The bundle structure on $\mathcal{F}(E)$ does not depend on the choice of the bundle equivalence $\lambda : E \longrightarrow M \times V$. If $\mu : E \longrightarrow M \times W$ is another bundle equivalence, then $\mu \circ \lambda^{-1} : M \times V \longrightarrow M \times W$ is an isomorphism, and can be identified with $\eta : M \longrightarrow L(V, W)$, which is defined by $\eta(x) = \mu \circ \lambda^{-1}|(\{x\} \times V)$. Then η is smooth, since $\mu \circ \lambda^{-1}$ is so. The map

$$\mathcal{F}(\mu \circ \lambda^{-1}) : M \times \mathcal{F}(V) \longrightarrow M \times \mathcal{F}(W)$$

is an isomorphism and can be identified with $\mathcal{F}\circ\eta$. Then $\mathcal{F}(\mu\circ\lambda^{-1})$ is smooth, since $\mathcal{F}$ is a smooth functor and η is smooth. Therefore $\mathcal{F}(\mu\circ\lambda^{-1})$ is a bundle equivalence. It follows that the composition

$$\mathcal{F}(E) \xrightarrow{\mathcal{F}(\lambda)} \mathcal{F}(E') \xrightarrow{\mathcal{F}(\mu\circ\lambda^{-1})} \mathcal{F}(E') \xrightarrow{\mathcal{F}(\mu^{-1})} \mathcal{F}(E)$$

is smooth and it is the identity map on $\mathcal{F}(E)$ connecting two vector bundle structures on it. Thus, for a trivial bundle E, $\mathcal{F}(E)$ can be given a vector bundle structure in a unique way.

Finally, suppose that E is any vector bundle. Then we can cover M by trivialising open sets U such that $E|U$ is trivial, and so $\mathcal{F}(E|U)$ has a unique vector bundle structure. If U and V are two open sets of the covering such that $U\cap V\neq\emptyset$, then $\mathcal{F}(E|(U\cap V))$ has two vector bundle structures $(\mathcal{F}(E|U)) \mid (U\cap V)$ and $(\mathcal{F}(E|V)) \mid (U\cap V)$. These structures must be the same, by the uniqueness of the structure on $\mathcal{F}(E|(U\cap V)$. Thus we have a unique vector bundle structure on $\mathcal{F}(E)$. □

Remark 5.2.3. This theorem holds when $\mathcal{F}$ is a contravariant functor, and also when $\mathcal{F}$ is a functor of several variables covariant in some and contravariant in the others. Here are some examples.

Examples 5.2.4. (1) **Dual bundle.** Let $\mathcal{F}:\mathcal{V}\longrightarrow\mathcal{V}$ be the contravariant functor where $\mathcal{F}(V)$ is the dual vector space V^* and $\mathcal{F}: L(V,W)\longrightarrow L(W^*,V^*)$ sends a linear map A to its adjoint A^*. Then $\mathcal{F}$ is smooth, since it is a continuous linear map. Therefore, if E is a vector bundle over M, so is $E^*=\mathcal{F}(E)$.

In particular, if E is the tangent bundle $\tau(M)$, then $\mathcal{F}(E)$ is the cotangent bundle $\tau(M)^*$ of M.

(2) **Whitney sum of bundles.** Let $\mathcal{F}:\mathcal{V}\times\mathcal{V}\longrightarrow\mathcal{V}$ be the functor covariant in both the variables, where $\mathcal{F}(V,W)=V\oplus W$, and

$$\mathcal{F}: L(V_1,V_2)\times L(W_1,W_2)\longrightarrow L(V_1\oplus W_1, V_2\oplus W_2)$$

is given by $\mathcal{F}(f,g)=f\oplus g$. Then, $\mathcal{F}$ is smooth, since it is continuous and bilinear.

If E and E' are vector bundles over M, then the vector bundle $\mathcal{F}(E,E')$ is the *Whitney sum* $E\oplus E'$ of E and E' (defined earlier in Example 5.1.4). Since $(E\oplus E')_x = E_x\oplus E'_x$ for every $x\in M$, we have $\dim(E\oplus E')=\dim E+\dim E'$, assuming M is connected.

(3) **Bundle of homomorphisms.** Consider the functor $\mathcal{F}:\mathcal{V}\times\mathcal{V}\longrightarrow\mathcal{V}$, which is contravariant in the first variable and covariant in the second, where $\mathcal{F}(V,W)=L(V,W)$ and $\mathcal{F}: L(V_2,V_1)\times L(W_1,W_2)\longrightarrow L(L(V_1,W_1),L(V_2,W_2))$ is given by $\mathcal{F}(f,g)(h)=ghf$.

The resulting vector bundle $\mathcal{F}(E,E')$ corresponding to vector bundles E and E' over M is denoted by $L(E.E')$. Its fibre over $x\in M$ is the vector space $L(E_x, E'_x)$. We shall denote the space of sections of this vector bundle by Hom(E, E′).

(4) **Tensor product of bundles.** The tensor product $E \otimes E'$ of two vector bundles E and E' is a vector bundle. This time one has to consider the functor $\mathcal{F} : \mathcal{V} \times \mathcal{V} \longrightarrow \mathcal{V}$ so that $\mathcal{F}(V, W) = V \otimes W$ and $\mathcal{F}(f, g) = f \otimes g$.

(5) **Bundle of bilinear forms.** We shall describe only a special case. Consider a functor $\mathcal{F}$ which assigns to each vector space V the vector space $S^2(V)$ of symmetric positive definite bilinear forms on V. Define

$$\mathcal{F} : L(V, W) \longrightarrow L(S^2(W), S^2(V))$$

as follows. Let $f \in L(V, W)$, $g \in S^2(W)$ and $v, w \in V$. Then

$$\mathcal{F}(f)(g)(v, w) = g(f(v), f(w)).$$

Then $\mathcal{F}$ is continuous and linear, so it is smooth. Therefore if E is a vector bundle over M, then $\mathcal{F}(E) = S^2(E)$ is a vector bundle over M.

Thus the natural operations on vector spaces, namely, direct sum, tensor product, etc. extend to vector bundles.

♢ **Exercise 5.7.** Show that if E and E' are vector bundles over M and N respectively, then $L(E, E')$ is a vector bundle over $M \times N$ whose fibre over $(x, y) \in M \times N$ is the vector space $L(E_x, E'_y)$.

5.3. Homotopy property of vector bundles

Lemma 5.3.1. *Let $\pi : E \longrightarrow M$ be a vector bundle, and $f : M_1 \longrightarrow M$ be a smooth map, M_1 being a manifold. Then there is a vector bundle*

$$\pi_1 : E_1 \longrightarrow M_1,$$

and a bundle map $(g, f) : (E_1, M_1, \pi_1) \longrightarrow (E, M, \pi)$

$$\begin{tikzcd} E_1 \arrow[r, "g"] \arrow[d, "\pi_1"'] & E \arrow[d, "\pi"] \\ M_1 \arrow[r, "f"] & M \end{tikzcd}$$

such that, for each $x \in M_1$, $g_x : \pi_1^{-1}(x) \longrightarrow \pi^{-1}(f(x))$ is a linear isomorphism. The bundle E_1 is determined uniquely up to equivalence.

The bundle E_1 is called the **induced bundle,** induced from E by f. It is also called the **pull-back** of E by f, and is denoted by $f^*(E)$. The bundle map (g, f) is called the **canonical bundle map** of the induced bundle.

Proof. Define E_1 by $\{(x,\alpha) \in M_1 \times E \mid f(x) = \pi(\alpha)\}$, and π_1 by $\pi_1(x,\alpha) = x$, and g by $g(x,\alpha) = \alpha$. The local triviality of E_1 may be seen as follows: if

$$\phi : U \times \mathbb{R}^n \longrightarrow \pi^{-1}(U)$$

is a VB-chart for E, then $\phi_1 : f^{-1}(U) \times \mathbb{R}^n \longrightarrow \pi_1^{-1}(f^{-1}(U))$ defined by $\phi_1(x,v) = (x, \phi(f(x),v))$ is a VB-chart for E_1. The inverse of ϕ_1 is given by $\phi_1^{-1}(x,\alpha) = (x, p_2 \circ \phi^{-1}(\alpha))$, where $x \in f^{-1}(U), \alpha \in \pi^{-1}(U)$ so that $f(x) = \pi(\alpha)$, and p_2 is the projection onto the second factor. Then g is a diffeomorphism and an isomorphism on each fibre.

To see the uniqueness of E_1, take another vector bundle $\pi' : E' \longrightarrow M_1$ and a bundle morphism $(g', f) : (E', M, \pi') \longrightarrow (E, M, p)$ such that $g' : E' \longrightarrow E$ is a diffeomorphism and an isomorphism on each fibre, and $\pi \circ g' = f \circ \pi'$. Define $h : E' \longrightarrow E_1$ by $h(\alpha') = (\pi'(\alpha'), g'(\alpha'))$, $\alpha' \in E'$. Then $g \circ h = g'$, and so h is a bundle equivalence. □

◊ **Exercises 5.8.** (1) If M_1 is a submanifold of M and $f : M_1 \longrightarrow M$ is the inclusion map, then show that there is an isomorphism $E|M_1 \simeq f^*(E)$.

(2) Show that if E is a vector bundle over M, and $f_1 : M_1 \longrightarrow M$, $f_2 : M_2 \longrightarrow M_1$ are smooth maps, then $(f_1 \circ f_2)^*(E) \cong f_2^*(f_1^*(E))$. Also, if Id_M is the identity map on M, then $(\mathrm{Id}_M)^*(E) \simeq E$

Lemma 5.3.2. *Any vector bundle morphism*

$$(g, f) : (E_1, M_1, \pi_1) \longrightarrow (E_2, M_2, \pi_2)$$

can be factored as $(g, f) = (k, f) \circ (h, \mathrm{Id})$, $g = k \circ h$,

$$\begin{array}{ccccc} E_1 & \xrightarrow{h} & f^*(E_2) & \xrightarrow{k} & E_2 \\ \downarrow{\scriptstyle \pi_1} & & \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi_2} \\ M_1 & \xrightarrow{\mathrm{Id}} & M_1 & \xrightarrow{f} & M_2 \end{array}$$

where (h, Id) *is a homomorphism and* (k, f) *is the canonical bundle map of the induced bundle* $f^*(E_2)$.

Proof. The map $k : f^*(E_2) \longrightarrow E_2$ is given by $k(x,\alpha) = \alpha$, and $h : E_1 \longrightarrow M_1 \times E_2$ by $h(\alpha) = (\pi_1(\alpha), g(\alpha))$. Since $f \circ \pi_1 = \pi_2 \circ g$, Image $h \subset f^*(E_2)$. Then h is linear on each fibre, and $g = k \circ h$. □

We shall now consider homotopy property of vector bundles.

Lemma 5.3.3. *Let M be a manifold, K a submanifold and a closed subset of M, and E a smooth vector bundle over M. Then any smooth section σ of $E|K$ can be extended to a smooth section of E.*

Proof. Locally σ may be looked upon as a map taking values in a vector space. Therefore, by Tietze Extension Theorem, for each point $p \in K$ there is an open neighbourhood U of p and a section η of $E|U$ such that $\eta = \sigma$ on on $U \cap K$. By Theorem 2.2.3, we may suppose that η is smooth. Such open neighbourhoods $\{U_i\}$ together with the open set $M - K$ form a covering of M. Let $\{\lambda_i\}$ be a smooth partition of unity subordinate to this covering, and η_i be smooth section of $E|U_i$ extending $\sigma|U_i \cap K$. Define section σ_i of E by

$$\sigma_i(x) = \begin{cases} \lambda_i(x)\eta_i(x) & \text{if } x \in U_i, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\sum \sigma_i$ is the required extension of σ. □

♢ **Exercise 5.9.** Use Theorem 2.2.3 to prove the following result.

If $s : M \longrightarrow E$ is a continuous section of a smooth vector bundle E over M with a Riemannian metric g such that s is smooth on a closed subset K of M, then, for any positive continuous function $\delta : M \longrightarrow \mathbb{R}$, there exists a smooth section s' of E which agrees with s on K, and

$$g_x(s(x), s'(x)) < \delta(x)$$

for all $x \in M$.

Lemma 5.3.4. *Let E and E' be two vector bundles of the same dimension over a manifold M, and K be a closed submanifold of M. Then any isomorphism $f : E|K \longrightarrow E'|K$ extends to an isomorphism $E|U \longrightarrow E'|U$ over an open neighbourhood U of K.*

Proof. We have a vector bundle $L(E, E')$ over M whose fibre over $p \in M$ is the vector space $L(E_p, E'_p)$ of linear maps from E_p to E'_p. Then a morphism from E to E' is a section of $L(E, E')$. By Lemma 5.3.3, a section of $L(E|K, E'|K)$ extends to a section of $L(E, E')$. Let U be a set of points of M for which this section is an isomorphism. Then U is an open neighbourhood of K, because the set of isomorphisms $E \longrightarrow E'$ is an open subset of $L(E, E')$. □

Theorem 5.3.5. *Let f_0 and f_1 be smooth homotopic maps from a manifold K to a manifold M, and E be a vector bundle over M. Then $f_0^*(E) \simeq f_1^*(E)$.*

Proof. Let $f : K \times \mathbb{R} \longrightarrow M$ be a homotopy, and $\pi : K \times \mathbb{R} \longrightarrow K$ be the projection. There is an obvious isomorphism between the vector bundles $f^*(E)$ and $\pi^* f_t^*(E)$ over the closed submanifold $K \times \{t\}$ (note that $f = f_t \circ \pi$ on $K \times \{t\}$, which is closed in the product topology). Then Lemma 5.3.4 gives an isomorphism between $f^*(E)$ and $\pi^* f_t^*(E)$ over some strip $K \times (t - \delta, t + \delta)$. This means that the isomorphism class of $f_t^*(E)$ is a locally constant function of t. Since $\mathbb{R}$ is connected, it is constant. Therefore $f_0^*(E) \simeq f_1^*(E)$. □

Corollary 5.3.6. *Any vector bundle over a contractible manifold is trivial.*

Proof. Let E be a vector bundle over a manifold M, and $x_0 \in M$. Let $f = \mathrm{id}_M$, and g denote the constant map $M \longrightarrow x_0$. Then f homotopic to g implies $f^*(E) \simeq g^*(E)$. But $g^*(E)$ is a vector bundle over a point, and so it is trivial, Therefore $E = f^*(E)$ is a trivial bundle. $\square$

5.4. Subbundle and quotient bundle

Definition 5.4.1. A vector bundle $\pi' : E' \longrightarrow M$ is a **subbundle** of a vector bundle $\pi : E \longrightarrow M$ if E' is a submanifold of E and $\pi' = \pi|E'$.

Note that if E' is a subbundle of E, then the inclusion map $f : E' \longrightarrow E$ is a monomorphism.

Lemma 5.4.2. *If E and E' are vector bundles over M and $f : E' \longrightarrow E$ is a monomorphism, then $f(E')$ is a subbundle of E, and $f : E' \longrightarrow f(E')$ is a bundle equivalence.*

Proof. It is sufficient to prove that each $x \in M$ has an open neighbourhood U on which $f(E')$ is a subbundle of E. Therefore we may suppose that E and E' are product bundles. Let $E = M \times \mathbb{R}^n$, and, for $x \in M$, let V_x be a subspace complementary to $f(E'_x)$ in $\mathbb{R}^n$. Then $F = M \times V_x$ is a subbundle of E. Define $g : E' \oplus F \longrightarrow E$ by $g(u, v) = f(u) + i(v)$ where $i : F \longrightarrow E$ is the inclusion. Then g_x is an isomorphism. Therefore there is an open neighbourhood U of x in M such that $g|U$ is an isomorphism, and hence a diffeomorphism, by Lemma 5.1.7. Now, E' is a subbundle of $E' \oplus F$. Therefore $g(E') = f(E')$ is a subbundle of $g(E' \oplus F) = E$ on U. The second part follows from Lemma 5.1.7. $\square$

Remarks 5.4.3. The proof shows that

(1) if $f : E' \longrightarrow E$ is a homomorphism, then the set of points $x \in M$ for which f_x is a monomorphism is an open set of M,
(2) locally a subbundle E' of a bundle E is a direct summand of E.

Definition 5.4.4. If E' is a subbundle of E, then the **quotient bundle** E/E' (of E modulo E') is the union of all vector spaces E_x/E'_x with the quotient topology.

Note that since E' is locally a direct summand in E, E/E' is locally trivial, and hence it is a vector bundle.

Proposition 5.4.5. *Let $(f, g) : (E', M', \pi') \longrightarrow (E, M, \pi)$ be a vector bundle morphism of constant rank, that is, f_x has constant rank for all $x \in M'$, Then*

(i) $\mathrm{Ker}\, f = \cup_x \mathrm{Ker}\, f_x$ *is a subbundle of E',*
(ii) $\mathrm{Image}\, f = \cup_x \mathrm{Image}\, f_x$ *is a subbundle of E,*
(iii) $\mathrm{Coker}\, f = \cup_x \mathrm{Coker}\, f_x$ *is a quotient bundle of E.*

Proof. The assertion (ii) implies (iii). We shall first prove (ii). The problem is local, and therefore we assume that $E' = M \times \mathbb{R}^n$. Let $x \in M$, and V_x be a complement of $\operatorname{Ker} f_x$ in $\mathbb{R}^n$. Then $F = M \times V_x$ is a subbundle of E', and the homomorphism $g = f \circ i : F \longrightarrow E$ (i = inclusion) is such that g_x is a monomorphism. Therefore g is a monomorphism in some open neighbourhood U of x. Therefore $g(F)|U$ is a subbundle of $E|U$, by Lemma 5.4.2. Now $g(F) \subset f(E')$, and, since $\dim f(E'_y)$ is constant for all $y \in M$, we have, for all $y \in U$,

$$\dim g(F_y) = \dim g(F_x) = \dim f(E'_x) = \dim f(E'_y).$$

Therefore $g(F)|U = f(E')|U$, and $f(E')$ is a subbundle of E.

We next prove (i). Note that a homomorphism $f : E' \longrightarrow E$ is a monomorphism if and only if its dual $f^* : E^* \longrightarrow {E'}^*$ is an epimorphism. Also f has constant rank implies that f^* has constant rank. Therefore, since ${E'}^* \longrightarrow {E'}^*/f^*(E^*) = \operatorname{Coker} f^*$ is an epimorphism, $(\operatorname{Coker} f^*)^* \longrightarrow ({E'}^*)^*$ is a monomorphism. Now

$$\operatorname{Image} f_x^* = \{\alpha : E'_x \longrightarrow \mathbb{R} \mid \operatorname{Ker} f_x \subset \operatorname{Ker}\ \alpha\}.$$

Therefore $\operatorname{Coker} f_x^*$ can be identified with the subspace of elements $\alpha \in {E'}^*_x$ such that, for some non-zero $v \in E'_x$, $f_x(v) = 0$ but $\alpha(v) \neq 0$. Then there is an isomorphism

$$\eta : \operatorname{Ker} f_x \longrightarrow L(\operatorname{Coker} f_x^*, \mathbb{R}) = (\operatorname{Coker} f_x^*)^*$$

given by $\eta(v)(\alpha) = \alpha(v)$. In fact, we have for each $x \in M$ a natural commutative diagram

$$\begin{tikzcd}
\operatorname{Ker} f_x \arrow[r] \arrow[d, "\eta"'] & E'_x \arrow[d, "\eta"] \\
(\operatorname{Coker} f_x^*)^* \arrow[r] & (E'_x)^*
\end{tikzcd}$$

where the vertical arrows η are isomorphisms. Therefore

$$\operatorname{Ker} f \simeq (\operatorname{Coker} f^*)^*,$$

and, by Lemma 5.4.2, is a subbundle of E'. □

Definition 5.4.6. A sequence of vector bundles $\{E_i\}$ over M connected by homomorphisms $\{f_i\}$ over Id_M

$$\cdots \longrightarrow E_{i-1} \xrightarrow{f_{i-1}} E_i \xrightarrow{f_i} E_{i+1} \longrightarrow \cdots$$

is called an **exact sequence** over M if for each $x \in M$ we have

$$\operatorname{Image}(f_{i-1})_x = \operatorname{Ker}(f_i)_x, \text{ for all } i.$$

In particular, a five-term exact sequence over M:

$$0 \longrightarrow E' \xrightarrow{f} E \xrightarrow{g} E'' \longrightarrow 0,$$

where 0 denotes the vector bundle of dimension 0, is called a **short exact sequence**. Here exactness means that f is a monomorphism, g is an epimorphism, and $\operatorname{Image} f = \operatorname{Ker} g$.

The bundle E'' of the above short exact sequence is called the quotient bundle of the monomorphism f. For the justification of the terminology note that if E' is a subbundle of E with inclusion homomorphism i, then we have the short exact sequence

$$0 \longrightarrow E' \xrightarrow{i} E \xrightarrow{p} E/E' \longrightarrow 0, \tag{3}$$

where p is the quotient homomorphism. Moreover, every monomorphism has a quotient bundle

$$0 \longrightarrow E' \xrightarrow{f} E \xrightarrow{p} E/f(E') \longrightarrow 0,$$

and it is unique up to isomorphism, as the following lemma shows.

Lemma 5.4.7. *If E_1 and E_2 are two quotient bundles of a monomorphism $f : E' \longrightarrow E$, then there is a unique isomorphism $h : E_1 \longrightarrow E_2$ so that the following diagram commutes.*

$$\begin{tikzcd}
0 \arrow[r] & E' \arrow[r, "f"] \arrow[d, "\mathrm{Id}"'] & E \arrow[r, "g_1"] \arrow[d, "\mathrm{Id}"'] & E_1 \arrow[r] \arrow[d, "h"'] & 0 \\
0 \arrow[r] & E' \arrow[r, "f"] & E \arrow[r, "g_2"] & E_3 \arrow[r] & 0
\end{tikzcd}$$

PROOF. The map h is defined as follows. Let $v \in E_1$. Then, by exactness, there is a $u \in E$ such that $g_1(u) = v$. Then set $h(v) = g_2(u)$. If u' is another element of E for which $g_1(u') = v$ also, then $u - u' \in \operatorname{Ker} g_1 = \operatorname{Image} f$, so there is a $w \in E'$ such that $f(w) = u - u'$. This means that $g_2 \circ f(w) = g_2(u - u') = 0$, or $g_2(u) = g_2(u')$, showing that h is well defined. It is easily checked that h is actually an isomorphism. $\square$

$\lozenge$ **Exercise 5.10.** Show that given an epimorphism $g : E \longrightarrow E''$, there is a unique bundle E' that fits into an exact sequence

$$0 \longrightarrow E' \longrightarrow E \xrightarrow{g} E'' \longrightarrow 0.$$

Lemma 5.4.8. *Given a short exact sequence over a manifold M*

$$0 \longrightarrow E' \xrightarrow{f} E \xrightarrow{g} E'' \longrightarrow 0,$$

there is an equivalence $\phi : E \longrightarrow E' \oplus E''$ such that $\phi \circ f$ is the natural inclusion $i : E' \longrightarrow E' \oplus E''$, and $g \circ \phi^{-1}$ is the natural projection $p : E' \oplus E'' \longrightarrow E''$.

$$\begin{tikzcd}
0 \arrow[r] & E' \arrow[r, "f"] \arrow[d, "\mathrm{Id}"'] & E \arrow[r, "g_1"] \arrow[d, "\phi"'] & E'' \arrow[r] \arrow[d, "\mathrm{Id}"'] & 0 \\
0 \arrow[r] & E' \arrow[r, "i"] & E' \oplus E'' \arrow[r, "p"] & E'' \arrow[r] & 0
\end{tikzcd}$$

Proof. Equip E with a fibrewise metric $\langle\ ,\ \rangle_x, x \in M$. Let $f(E')_x^\perp$ be the subspace of E_x orthogonal to the subspace $f(E')_x$

$$f(E')_x^\perp = \{v \in E_x \mid \langle v, w\rangle_x = 0 \text{ for all } w \in f(E')_x\}.$$

Let $f(E')^\perp$ be the union of all $f(E')_x^\perp$, $x \in M$. Then $f(E')^\perp$ is a vector bundle, since it is the kernel of a homomorphism of constant rank, which is the orthogonal projection of E onto $f(E')$. We have then $E = f(E') \oplus f(E')^\perp$, and a short exact sequence

$$0 \longrightarrow f(E') \xrightarrow{i} E \xrightarrow{p} f(E')^\perp \longrightarrow 0.$$

Thus there are two quotient bundles E'' and $f(E')^\perp$ of the monomorphism $f : E' \longrightarrow E$; therefore they are equivalent by Lemma 5.4.7.

If h is the equivalence $f(E')^\perp \longrightarrow E''$, then the required equivalence ϕ is given by

$$E = f(E') \oplus f(E')^\perp \xrightarrow{f \oplus h} E' \oplus E''.$$

‘ □

Remark 5.4.9. The proof of the lemma contains the definition of orthogonal bundle, which may be singled out as follows. If E' is a subbundle of a bundle E with a metric, then the **orthogonal bundle** of E' in E is the bundle $E'^\perp$ whose fibre over $x \in M$ is given by

$$E'^\perp_x = \{v \in E_x \mid \langle v, w\rangle_x = 0, w \in E'_x\}.$$

5.5. Orientation

Let V be a real vector space of finite dimension. Given two ordered bases $\alpha = (\alpha_1, \ldots, \alpha_n)$ and $\beta = (\beta_1, \ldots, \beta_n)$ of V, there is a unique linear map $L : V \longrightarrow V$ such that $L(\alpha_i) = \beta_i$. The columns of the matrix of L are the components of $L(\alpha_1), \ldots, L(\alpha_n)$ with respect to the basis $\alpha_1, \ldots, \alpha_n$. Since $\beta_1, \ldots, \beta_n$ are linearly independent, the matrix of L is non-singular. We say that the bases α and β are equivalently oriented if $\det L > 0$. This defines an equivalence relation and the set of all ordered bases is partitioned into two disjoint classes. We denote the equivalence class of $(\alpha_1, \ldots, \alpha_n)$ by $[\alpha_1, \ldots, \alpha_n]$, and it is called an **orientation of** V. Note that the ordering of the vectors is important here. We pick up a class arbitrarily and assign a positive sign to it, and a negative sign to the other class. The vector space V with an ordered basis α is called positively or negatively oriented depending on which orientation class α belongs. Thus each vector space with positive dimension has precisely two orientations. If the vector space is of dimension zero, we denote its orientation as one of the symbols $+1$ and -1.

Let $J : V \longrightarrow W$ be a linear isomorphism between oriented vector spaces. If α and β are bases of V, and $L : V \longrightarrow V$ is a linear transformation sending α to β, then $J \circ L \circ J^{-1}(J(\alpha)) = J(\beta)$. Since $\det L = \det J \circ L \circ J^{-1}$, the sign of the orientation of $J(\alpha)$ is either the same as the sign of the orientation of α, or opposite. Therefore the isomorphism J is either **orientation preserving** or **orientation reversing**. We may define sign $J = \text{sign } \alpha \cdot \text{sign } J(\alpha)$, where α is a basis of V, so that J is orientation preserving or reversing according as sign J is $+1$ or -1.

The standard orientation of an Euclidean space $\mathbb{R}^n$, $n \geq 1$, is given by any basis whose coordinate matrix has positive determinant. The standard orientation of $\mathbb{R}^0$ is the number $+1$.

Example 5.5.1. If $0 \to V_1 \xrightarrow{f} V \xrightarrow{g} V_2 \to 0$ is an exact sequence of vector spaces, where V_1 and V_2 are given the orientations $\alpha = [\alpha_1, \ldots, \alpha_n]$ and $\beta = [\beta_1, \ldots, \beta_m]$ respectively, then an orientation γ of V is provided by the equivalence ϕ in Lemma 5.4.8 (in the context of vector spaces) as

$$\gamma = [f(\alpha_1), \ldots, f(\alpha_n), \xi_1, \ldots, \xi_m],$$

where $g(\xi_i) = \beta_i$. This is well-defined. Because, if $g(\eta_i) = \beta_i$ also, then

$$[f(\alpha_1), \ldots, f(\alpha_n), \xi_1, \ldots, \xi_m] = [f(\alpha_1), \ldots, f(\alpha_n), \eta_1, \ldots, \eta_m],$$

since the linear isomorphism $L : V \longrightarrow V$ given by $L(f(\alpha_i)) = f(\alpha_i)$ and $L(\xi_i) = \eta_i$ has the matrix

$$\begin{pmatrix} I_n & * \\ 0 & I_n \end{pmatrix}$$

with determinant 1. This orientation γ of V is denoted by $\alpha \oplus \beta$. Note that any two of α, β, and γ determine the third by the rule of signs

$$\text{sign } \gamma = \text{sign } \alpha \cdot \text{sign } \beta.$$

For example, the orientation β of V_2 is determined by sign γ/sign α.

For example, the orientation the direct sum $V_1 \oplus V_2$ is obtained via the exact sequence

$$0 \to V_1 \xrightarrow{i} V_1 \oplus V_2 \xrightarrow{p} V_2 \to 0,$$

where i and p are as in Lemma 5.4.8, as

$$[(\alpha_1, 0), \ldots, (\alpha_n, 0), (0, \beta_1), \ldots, (0, \beta_m)].$$

Similarly, the exact sequence

$$0 \to W \xrightarrow{f} V \xrightarrow{g} V/W \to 0$$

gives the orientation β of the quotient space V/W as

$$\text{sign } \beta = \text{sign } \gamma / \text{sign } \alpha,$$

where γ and α are the orientations of V and W respectively.

Definition 5.5.2. An **orientation of a vector bundle** E over a manifold M is a family

$$\omega = \{\omega_p \mid \omega_p \text{ an orientation of } E_p, p \in M\}$$

such that there exists a VB-atlas $\Phi = \{(U_i, \phi_i)\}$ of M with the condition that if $p \in U_i$, then the isomorphism $\phi_{ip} : (E_p, \omega_p) \longrightarrow (\mathbb{R}^n, \lambda)$, where λ is the standard orientation of $\mathbb{R}^n$, is orientation preserving.

The condition is referred to as a smooth choice of orientations, and it may also be stated in terms of the cocycle $\{g_{ij}\}$ of E as follows: for every pair of indices (i, j) for which $U_i \cap U_j \neq \emptyset$, $g_{ij}(x) \in GL(n, \mathbb{R})$ has strictly positive determinant at every point $x \in U_i \cap U_j$. The atlas Φ is called an **oriented atlas**.

Example 5.5.3. Let $0 \to E_1 \xrightarrow{f} E \xrightarrow{g} E_2 \to 0$ be an exact sequence of vector bundles over M. Then, given orientations $\alpha = \{\alpha_x\}_{x \in M}$ and $\beta = \{\beta_x\}_{x \in M}$ of E_1 and E_2 respectively, an orientation $\gamma = \{\gamma_x\}_{x \in M}$ of E is obtained by setting $\gamma_x = \alpha_x \oplus \beta_x$. As in Example 5.5.1, we have the rule of signs

$$\text{sign } \gamma = \text{sign } \alpha \cdot \text{sign } \beta.$$

A manifold M is called **orientable** if its tangent bundle $\tau(M)$ ia an orientable vector bundle. In this case, the above condition of smooth choice reads as follows. The map

$$[d\phi_i]_p : (\tau(M)_p, \theta_p) \longrightarrow (\mathbb{R}^n, \lambda)$$

is orientation preserving, for some atlas $\Phi = \{(U_i, \phi_i)\}$ of M, or equivalently, the coordinate changes in some atlas Φ have positive Jacobian determinants at all points. A maximal atlas of this type is called an **oriented smooth structure** of M. If ω is an orientation of $\tau(M)$, then we say that (M, ω) is an oriented manifold. In this case M admits another orientation $-\omega$ defined by $(-\omega)_p = -\omega_p, p \in M$.

Note that for any coordinate chart (U, ϕ) in M with $\phi = (x_1, \ldots, x_n)$, the linearly independent vector fields $\partial/\partial x_1, \ldots, \partial/\partial x_n$ on U define an orientation on U. The orientations defined in this way by two coordinate charts (U, ϕ) and (V, ψ) agree on $U \cap V$ if and only if $\det(\psi \circ \phi^{-1}) > 0$, and they define an orientation on $U \cup V$. This consideration leads us to the following alternative definition of orientation of a manifold.

Definition 5.5.4. A manifold M is **orientable** if there is an atlas $\Phi = \{(U_i, \phi_i)\}$ such that whenever $U_i \cap U_j \neq \emptyset$ the Jacobian matrix of $\phi_j \circ \phi_i^{-1}$ has strictly positive determinant at every point of $\phi_i(U_i \cap U_j)$. The atlas Φ is called an oriented atlas. A manifold is oriented if an oriented atlas has been chosen for it.

Examples 5.5.5. (1) Every sphere S^n is orientable. There is an atlas of S^n consisting of two charts $(S^n - \{P\}, \phi)$ and $(S^n - \{Q\}, \psi)$, where P and Q are the north and south poles of S^n. If $\psi \circ \phi^{-1}$ has negative Jacobian determinant, replace ϕ by $\phi' = r \circ \phi$ where $r : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is the reflection $r(u_1, \ldots, u_n) = (-u_1, \ldots, u_n)$. This provides an orientation of S^n.

(2) The real projective space $\mathbb{R}P^n$ is orientable, if n is odd. To see this consider the atlas $\{(U_i, \phi_i)\}$ for $\mathbb{R}P^n$ given in Exercise 1.4(2), p. 8. Change ϕ_i to $\psi_i = (-1)^i \phi_i$. Then $\{(U_i, \psi_i)\}$ is also an atlas for $\mathbb{R}P^n$, and for $j < i$ we have

$$\psi_j \circ \psi_i^{-1}(x_1, \dots, x_n) = \frac{(-1)^{i+j}}{x_{j+1}}(x_1, \cdots, x_j, x_{j+2}, \dots, x_i, 1, x_{i+1}, \dots, x_n).$$

The Jacobian determinant of this map may easily be seen to be

$$(-1)^{(i+j)(n+1)+2}\Big(\frac{1}{x_{j+1}}\Big)^{n+1},$$

which is always positive, if n is odd.

We shall show in Example 5.5.9 below that $\mathbb{R}P^n$ is not orientable if n is even.

(3) Every complex manifold M is orientable. If $\{(U_i, \phi_i)\}$ is an atlas of M, where $\phi_i = (z_1, \dots, z_n)$ and $\phi_j = (w_1, \dots, w_n)$, $z_k = x_k + \sqrt{-1}y_k$, $w_k = u_k + \sqrt{-1}v_k$, $x_k, y_k, u_k, v_k \in \mathbb{R}$ $(1 \le k \le n)$, such that whenever $U_i \cap U_j \neq \emptyset$ the map $\phi_j \circ \phi_i^{-1}$ is holomorphic (or complex analytic), then it follows by induction from the Cauchy-Riemann equations

$$\frac{\partial u_j}{\partial x_i} = \frac{\partial v_j}{\partial y_i}, \quad \frac{\partial u_j}{\partial y_i} = -\frac{\partial v_j}{\partial x_i}$$

(these equations must be satisfied, because $\phi_j \circ \phi_i^{-1}$ is holomorphic) that the Jacobian determinant is

$$\frac{\partial(u_1, v_1, \dots, u_n, v_n)}{\partial(x_1, y_1, \dots, x_n, y_n)} = \left|\frac{\partial(w_1, \dots, w_n)}{\partial(z_1, \dots, z_n)}\right|^2,$$

which is always positive.

Lemma 5.5.6. *If two orientations of a connected orientable manifold M agree (resp. disagree) at a point, then they agree (resp. disagree) at every point of the manifold.*

Proof. We shall show that the set of points of M where the two orientations agree (resp. disagree) is an open set. Suppose that the orientations agree at a point x, and (U, ϕ), (V, ψ) are two coordinate charts around x with $\phi(x) = \psi(x) = 0$ such that $d(\phi^{-1})_0$ preserves the first orientation and $d(\psi^{-1})_0$ preserves the second. Then, if the orientations agree, $d(\psi \circ \phi^{-1})_0 : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is orientation preserving, and so its Jacobian has positive determinant. This means that the determinant is positive in a neighbourhood of x, since it is a continuous function. $\square$

$\Diamond$ **Exercise 5.11.** Let $f : M \longrightarrow N$ be a local diffeomorphism. Then show that an orientation ω in N with domain V induces an orientation θ in M with domain $U = f^{-1}(V)$.

If $f : E_1 \longrightarrow E_2$ is a bundle morphism which is an isomorphism on each fibre, and E_2 has an orientation ω, then there is a unique orientation θ on E_1 such that f is orientation preserving on each fibre. In particular, the pull-back f^*E_2 has a natural orientation which is denoted by $f^*\omega$.

Theorem 5.5.7. *Let a discrete group G act properly discontinuously on a manifold M. Then the quotient manifold M/G is orientable if and only if there is an orientation on M which is preserved by all diffeomorphisms $\eta_g : M \longrightarrow M, g \in G$, given by the action η of G on M.*

Proof. Suppose that θ is an orientation on M/G. Since $\pi : M \longrightarrow M/G$ is a local diffeomorphism, $\omega = \pi^*\theta$ is an orientation on M. Since $\pi \circ \eta_g = \pi$ for every $g \in G$, $\eta_g^*\omega = \eta_g^*\pi^*\theta = \pi^*\theta = \omega$, and the orientation ω is preserved by every η_g.

Conversely, suppose that ω is an orientation on M such that $\eta_g^*\omega = \omega$ for every $g \in G$. Let U be an open set in M such that π is a diffeomorphism on U with inverse $(\pi|U)^{-1} = \lambda$. Then $\lambda^*\omega$ is an orientation on $\pi(U)$. We must show that the local orientations in M/G obtained in this way agree on the intersections of their domains. Now, as shown in the proof of Theorem 1.7.8, if $x \in \pi(U_1) \cap \pi(U_2)$, then on a neighbourhood of x, $\lambda_2 = \eta_g \circ \lambda_1$ for some $g \in G$, where $\lambda_i = (\pi|U_i)^{-1}$, $i = 1, 2$. Since $\eta_g^*\omega = \omega$, we have then $\lambda_2^*\omega = \lambda_1^*\eta_g^*\omega = \lambda_1^*\omega$, and so the two orientations agree. □

Example 5.5.8. As we have seen in Example 1.7.11, the Möbius band is the quotient of $\mathbb{R}^2$ by the action of the proper discontinuous transformation group $\mathbb{Z}$ generated by a single transformation $\phi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ given by $\phi(x, y) = (x + 1, -y)$ A simply computation shows that

$$[d\phi]_p([\partial/\partial x]_p) = [\partial/\partial x]_{\phi(p)}, \quad [d\phi]_p([\partial/\partial y]_p) = -[\partial/\partial y]_{\phi(p)}, p \in \mathbb{R}^2.$$

Therefore ϕ does not preserve the orientation of $\mathbb{R}^2$, and so the Möbius band is not orientable.

Similarly the Klein bottle is not orientable.

Example 5.5.9. The projective space $\mathbb{R}P^n$ is the quotient of S^n modulo the action of the group $\mathbb{Z}_2$ generated by the antipodal map $s : S^n \longrightarrow S^n$ given by $s(p) = -p, p \in S^n$. Let $p = (1, 0, \dots, 0)$. Then in terms of the stereographic charts (U_+, ϕ_+) and (U_-, ϕ_-) at p and $s(p)$ respectively (see Example 1.1.9), the local representation of the map s is given by

$$(x_1, \dots, x_n) \mapsto (-x_1/(x_1^2 + \dots + x_n^2), \dots, -x_n/(x_1^2 + \dots + x_n^2)).$$

A computation of the Jacobian of this at p reveals that for $fori = 2, \dots, n$

$$[ds]_p([\partial/\partial x_1]_p) = [\partial/\partial x_1]_{s(p)}, \text{ and } [ds]_p([\partial/\partial x_i]_p) = -[\partial/\partial x_i]_{s(p)}.$$

Since the coordinate neighbourhood U_+ is connected, it has an orientation ω determined by the ordered basis $\{[\partial/\partial x_1]_p, \dots, [\partial/\partial x_n]_p\}$. The above computations show that $[ds]_p$ transforms ω into $(-1)^{n-1}\omega$ in the other coordinate neighbourhood U_-. Therefore $\mathbb{R}P^n$ cannot be orientable if n is even.

Remark 5.5.10. It can be shown using algebraic topology that a manifold M is orientable if and only if the first Stiefel-Whitney class of its tangent bundle $\omega_1(M)$, which belongs to the cohomology group $H^1(M;\mathbb{Z}_2)$, is zero. For $\mathbb{R}P^n$, the cohomology group $H^1(\mathbb{R}P^n;\mathbb{Z}_2)$ is a cyclic group of order 2. If a is the non-zero element in it, then $\omega_1(\mathbb{R}P^n) = (n+1)a$, where $n+1$ is an integer reduced mod 2. Therefore $\mathbb{R}P^n$ is orientable if and only if n is odd (see Milnor [**32**]).

Theorem 5.5.11. *A manifold M of dimension n is orientable if and only if it has a nowhere vanishing n-form.*

Proof. Suppose that ω is a nowhere vanishing n form on M, and $\Phi = \{(U_i,\phi_i)\}$ is an atlas in M. We shall show that Φ can be modified into an oriented atlas. If $\phi_i : U_i \longrightarrow \mathbb{R}^n$ is a chart in Φ, then $\phi_i^*\lambda$, where $\lambda = du_1\cdots du_n$ is the n-form on $\mathbb{R}^n$ with respect to its standard coordinates $u_1,\ldots,u_n$, is nowhere zero, and so we may write $\phi_i^*\lambda = f_i\cdot\omega$, where $f_i : U_i \longrightarrow \mathbb{R}$ is a non-zero function. Then f_i is either everywhere positive or everywhere negative. By permuting the coordinates in $\mathbb{R}^n$, we may assume that f_i is positive on U_i. For example, an interchange of u_1 and u_2 gives

$$\phi_i^* du_2\cdot du_1\cdot du_3\cdots du_n = -\,\phi_i^* du_1\cdot du_2\cdot du_3\cdots du_n = -f_i\cdot\omega.$$

In this way we get an atlas $\{(U_i,\psi_i)\}$ for which the functions f_i on U_i are positive. Then on $\psi_i(U_i\cap U_j)$, $\psi_i^{*-1}\psi_j^* = (\psi_j\psi_i^{-1})^*$ pullbacks λ onto a positive multiple of itself

$$(\psi_j\psi_i^{-1})^*\lambda = f_j f_i^{-1}\cdot\lambda.$$

By Theorem 3.3.13, $\det(\psi_j\psi_i^{-1})^* = f_j f_i^{-1} > 0$. Thus we get an oriented atlas in M, and so M is orientable.

Conversely, if M has an oriented atlas $\{(U_i,\phi_i)\}$, then

$$(\phi_j\phi_i^{-1})^* du_1\cdots du_n = f\cdot du_1\cdots du_n,$$

for some positive function f. Therefore

$$\phi_j^* du_1\cdots du_n = (\phi_i^* f)(\phi_i^* du_1\cdots du_n).$$

We get from this, on writing $\omega_i = \phi_i^* du_1\cdots du_n$, that $\omega_j = g\cdot\omega_i$, where $g = \phi_i^* f = f\circ\phi_i$ is positive. This happens for every pair of indices i and j. Then, if $\{\lambda_i\}$ is a partition on unity subordinate to the covering $\{U_i\}$, we get the required nowhere vanishing n-form as $\omega = \sum_i \lambda_i\cdot\omega_i$ $\square$

Example 5.5.12 (Product Orientation). If M and N are orientable manifolds, then their product $M\times N$ is an orientable manifold. To see this, suppose that $\{(U_i,\phi_i)\}$ and $\{(V_j,\psi_j)\}$ are atlases of M and N respectively such that the Jacobians $J(\phi_i\circ\phi_k^{-1})$ and $J(\psi_j\circ\psi_\ell^{-1})$ have positive determinants. Then an atlas of $M\times N$ is given by the coordinate charts

$$h_{ij} = \phi_i\times\psi_j : U_i\times V_j \longrightarrow \mathbb{R}^{n+m}, \qquad n = \dim M,\quad m = \dim N.$$

Then $h_{ij} \circ h_{k\ell} = (\phi_i \times \psi_j) \circ (\phi_k \times \psi_\ell) = (\phi_i \circ \phi_k^{-1}) \times (\psi_j \circ \psi_\ell^{-1})$ has Jacobian

$$\begin{pmatrix} J(\phi_i \circ \phi_k^{-1}) & 0 \\ 0 & J(\psi_j \circ \psi_\ell^{-1}) \end{pmatrix}$$

and this has positive determinant. Thus if $\alpha = \{\alpha_1, \ldots, \alpha_n\}$ and $\beta = \{\beta_1, \ldots, \beta_m\}$ are ordered bases of $\tau(M)_x$ and $\tau(N)_y$, then the sign of the ordered basis $\alpha \times \beta = \{\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m\}$ of $\tau(M \times N)_{(x,y)}$ is given by

$$\text{sign } (\alpha \times \beta) = (\text{sign } \alpha) \cdot (\text{sign } \beta)$$

Theorem 5.5.13. *Any vector bundle over a simply connected manifold is orientable.*

PROOF. The proof will follow after the following two observations (1) and (2) for any vector bundle E over a manifold M.

(1) Let $\sigma : I \longrightarrow M$ be a path in M, and ω_0 be a given orientation of $E_{\sigma(0)}$. Then we can move ω_0 along σ to get an orientation of $E_{\sigma(1)}$ in the following sense. Since I is contractible, the pullback $\sigma^* E$ is a trivial bundle. Therefore, since I is connected, $\sigma^* E$ admits a unique orientation ω so that $\omega(\sigma(0)) = \omega_0$. We denote by $\widehat{\sigma}(\omega_0)$ the orientation $\omega(\sigma(1))$ of $E_{\sigma(1)}$.

(2) Let σ and $\tau : I \longrightarrow M$ be two paths such that $\sigma(0) = \tau(0)$, $\sigma(1) = \tau(1)$, and $\sigma \simeq \tau$ rel $\{0, 1\}$. Then $\widehat{\sigma}(\omega_0) = \widehat{\tau}(\omega_0)$. To see this, suppose $h : I \times I \longrightarrow M$ is a homotopy between σ and τ so that $h = \sigma$ on $I \times \{0\}$, and $h = \tau$ on $I \times \{1\}$. Since $I \times I$ is contractible, the pullback $h^* E$ is a trivial bundle. Since $I \times I$ is connected, there is a unique orientation θ on $I \times I$. It follows that the orientation θ at the point $h(1, 0) = \sigma(1) = \tau(1) = h(1, 1) \in M$ is equal to both the orientations $\widehat{\sigma}(\omega_0)$ and $\widehat{\tau}(\omega_0)$, and therefore they are equal.

We now turn to the proof of the theorem. If M is simply connected, then any two paths σ and τ in M from x_0 to x_1 are homotopic. For, the product path $\sigma \cdot \tau^{-1}$ is a closed path, and so $\sigma \cdot \tau^{-1} \simeq *$, where $*$ denotes the constant path on x_0. Therefore $\sigma \simeq \tau$, working in the fundamental groupoid of M. Therefore, starting with a point $p \in M$ with an orientation ω at p, we can move ω to an orientation ω_q at other point $q \in M$ along a path joining p and q so that ω_q does not depend on the path joining p and q. It follows that $\omega = \{\omega_q, q \in M\}$ is an orientation of M. $\square$

$\lozenge$ **Exercise 5.12.** Show that the vector bundle E over a manifold M is orientable if and only if every loop $\sigma : I \longrightarrow M$, $\sigma(0) = \sigma(1)$, preserves the orientation ω_0 of $E_{\sigma(0)}$, that is, $\widehat{}(\omega_0) = \omega_0$.

5.6. Reduction of structure group of a vector bundle

The **Gram-Schmidt orthonormalisation process** says that for any matrix $A \in GL(n, \mathbb{R})$, there is a unique triangular matrix T such that $AT \in O(n)$. We shall prove this result in the following equivalent form:

Theorem 5.6.1. *Let V is a vector space with an inner product $\langle\ ,\ \rangle$, and with a basis $v_1, \dots, v_n$. Then there is a unique set of orthonormal vectors $w_1, \dots, w_n$ such that $L(w_1, \dots, w_i) = L(v_1, \dots, v_i)$ (along with orientation) for $i = 1, \dots, n$, where $L(v_1, \dots, v_i)$ denotes the linear span of $v_1, \dots, v_i$.*

Proof. Define $w_1 = v_1/\|v_1\|$. Then, suppose that $w_1, \dots, w_{k-1}$ have been defined such that they form an orthonormal set, and

$$L(w_1, \dots, w_i) = L(v_1, \dots, v_i)$$

(along with orientation) for $i = 1, \dots, k-1$. If U denotes $L(w_1, \dots, w_{k-1})$, then any $v \in V$ is uniquely expressible as a sum of a vector in U and a vector in $U^\perp$, namely,

$$v = \sum_{i=1}^{k-1} \langle v, w_i \rangle \cdot w_i + \Big(v - \sum_{i=1}^{k-1} \langle v, w_i \rangle \cdot w_i\Big).$$

Therefore we may write $v_k = u + w$, where $u \in U$, and $w \in U^\perp$. Define $w_k = w/\|w\|$. Then $w_1, \dots, w_k$ is an orthonormal set, and $v_k = u + \|w\| \cdot w_k \in L(w_1, \dots, w_k)$. But $w_k = (v_k - u)/\|w\| \in L(w_1, \dots, w_{k-1}, v_k)$. Therefore

$$L(w_1, \dots, w_k) = L(w_1, \dots, w_{k-1}, v_k) = L(v_1, \dots, v_k)$$

(along with orientation). This completes the inductive step, and the proof of the theorem. □

Definition 5.6.2. We say that the structure group $GL(n, \mathbb{R})$ of a vector bundle E with cocycle $\{g_{ij}\}$ may be **reduced** to a subgroup G of $GL(n, \mathbb{R})$, if there is an equivalent cocycle $\{g'_{ij}\}$ (in the sense of Definition 5.1.12) such that each g'_{ij} takes values in G.

An alternative definition of an orientable vector bundle may be given in the following way. A vector bundle is orientable if its structure group $GL(n, \mathbb{R})$ may be reduced to the group $GL_+(n, \mathbb{R})$ consisting of linear isomorphisms of $\mathbb{R}^n$ with positive determinant.

Lemma 5.6.3. *If $E = M \times \mathbb{R}^n$ is a trivial bundle with metric $\langle\ ,\ \rangle$, and $(s_1, \dots, s_n)$ is a frame of E, then there is an orthonormal frame $(s'_1, \dots, s'_n)$ of E in the sense that*

$$\langle s'_i(x), s'_j(x) \rangle = \delta_{ij} \quad (\textit{Kronecker delta}).$$

for each $x \in M$.

Proof. The sections s'_j are obtained by using, for each $x \in M$, the Gram-Schmidt orthonormalisation process, which converts the basis $(s_1(x), \dots, s_n(x))$ of E_x into an orthonormal basis $(s'_1(x), \cdots, s'_n(x))$ of E_x. For each $x \in M$, there is a linear transformation $\lambda(x) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ such that $s'_j(x) = \lambda(x)(s_j(x))$, $j = 1, \dots, n$, where the matrix of $\lambda(x)$ is an upper triangular matrix whose diagonal entries are positive. Thus that each s'_j is a linear combination of $s_1, \dots, s_j$ with real valued smooth functions as scalars, and therefore it is smooth. □

Theorem 5.6.4. *The structure group $GL(n, \mathbb{R})$ of a vector bundle E of dimension n may be reduced to the orthogonal group $O(n)$. It may be reduced to the special orthogonal group $SO(n)$ if and only if E is orientable.*

PROOF. We assign an inner product to E. Let $\{(U_i, \phi_i)\}$ be a VB-atlas of E with cocycle $\{g_{ji}\}$. In view of Definition 5.6.2, it is sufficient to construct another VB-atlas $\{(U_i, \phi'_i)\}$ of E such that the corresponding cocycle $\{g'_{ji}\}$ is equivalent to $\{g_{ji}\}$, and each g'_{ji} takes values in $O(n)$.

There are n sections s_j of E over U_i given by $s_j(x) = \phi_i(x, e_j)$, which form a frame on U_i, where $(e_1, \ldots, e_n)$ is a frame on $\mathbb{R}^n$. Then, by Lemma 5.6.3, there is a map $\lambda_i : U_i \longrightarrow GL(n, \mathbb{R})$ such that $s'_j(x) = \lambda_i(x)(s_j(x))$, and $(s'_1, \ldots, s'_n)$ is an orthonormal frame of E over U_i. Now define $\phi'_i : U_i \times \mathbb{R}^n \longrightarrow \pi^{-1}(U_i)$ by $\phi'_i(x, v) = v_1 s'_1(x) + \cdots + v_n s'_n(x)$, where $v = v_1 e_1 + \cdots + v_n e_n \in \mathbb{R}^n$. This gives a VB-chart (U_i, ϕ'_i), and their collection is a VB-atlas of E. Now, it follows from the definitions that $\phi'_i(x, v) = \phi_i(x, \lambda_i(x)(v))$, or $\phi'_{ix}(v) = \phi_{ix}(\lambda_i(x)(v))$. Therefore

$$g'_{ji}(x) = {\phi'}_{jx}^{-1} \circ \phi'_{ix} = {\lambda_j}^{-1}(x) \circ \phi_{jx}^{-1} \circ \phi_{ix} \circ \lambda_i(x) = \lambda_j^{-1}(x) \circ g_{ji} \circ \lambda_i(x),$$

and so the cocycles $\{g_{ji}\}$ and $\{g'_{ji}\}$ are equivalent by Definition 5.1.12. Finally each $g'_{ji}(x) \in O(n)$, because each $\phi'_{jx} \in O(n)$. Moreover, if $\det g'_{ji}(x) > 0$, then $g'_{ji}(x)$ will be in $SO(n)$. □

5.7. Homology characterisation of orientation

All our homology groups without any mention of the coefficients will have coefficients in the group of integers $\mathbb{Z}$.

Definition 5.7.1. Let U and V be connected open subsets of $\mathbb{R}^n$, and $f : U \longrightarrow V$ be a diffeomorphism. Then the **orientation number** $n(f) = \pm 1$ is defined as follows: $n(f) = +1$ if f is orientation preserving, that is, if $\det J(f)(x) > 0$ for all $x \in U$, and $n(f) = -1$, otherwise.

For example, the identity map, translations $x \mapsto x + a$, and dilations (stretching or shrinking) $x \longrightarrow ax$, $a > 0$, all have orientation number $+1$. On the other hand, any reflection $r : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ through a hyperplane has orientation number -1.

The following lemma follows easily.

Lemma 5.7.2. (a) *If $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a linear isomorphism, then $n(f) =$ sign of* $\det f$.

(b) *If $f : U \longrightarrow V$, and $g : V \longrightarrow W$ are diffeomorphisms between connected open subsets of $\mathbb{R}^n$, then $n(g \circ f) = n(g) \cdot n(f)$.*

(c) *If $U' \subset U$, $V' \subset V$ are connected open subsets of $\mathbb{R}^n$, and $f : U \longrightarrow V$ is a diffeomorphism with $f(U') = V'$, then $n(f) = n(f|U')$.*

Recall now few facts from the homology theory.

The suspension $\Sigma(X)$ of a topological space X is the quotient space of $X \times [0,1]$ in which $X \times \{0\}$ and $X \times \{1\}$ have been identified to single points. Then the projection $X \times [0,1] \longrightarrow [0,1]$ defines a function $h : \Sigma(X) \longrightarrow [0,1]$ such that $h^{-1}(0)$, $h^{-1}(1)$ are single points, and there is a homeomorphism $h^{-1}(t) \approx X$ for each $t \neq 0, 1$. Then $X_1 = h^{-1}(0,1]$ and $X_2 = h^{-1}[0,1)$ are contractible spaces, as may be seen by moving vertically towards $h^{-1}(1)$ and $h^{-1}(0)$ respectively. Then the reduced Mayer-Vietoris sequence of the excisive triad $(\Sigma(X); X_1, X_2)$

$$\to \widetilde{H}_{k+1}(\Sigma(X)) \to \widetilde{H}_k(X_1 \cap X_2) \to \widetilde{H}_k(X_1) \oplus \widetilde{H}_k(X_2) \to \widetilde{H}_k(\Sigma(X)) \to$$

gives the isomorphisms

$$\widetilde{H}_{k+1}(\Sigma(X)) \cong \widetilde{H}_k(X_1 \cap X_2) = \widetilde{H}_k(X \times [0,1]) \cong \widetilde{H}_k(X),$$

where the last isomorphism $\cong$ is given by the homotopy equivalence $X \times [0,1] \simeq X$.

Since for a sphere S^n, $n \geq 0$, we have a homeomorphism $\Sigma(S^n) \approx S^{n+1}$, the above considerations give an isomorphism $\sigma : \widetilde{H}_n(S^n) \longrightarrow \widetilde{H}_{n+1}(S^{n+1})$ We may therefore choose a generator $\alpha_n \in \widetilde{H}_n(S^n)$ such that $\sigma(\alpha_n) = \alpha_{n+1}$.

Let $x \in U \subset \mathbb{R}^n$, U open, and B be an open n-ball centred at x such that $\overline{B} \subset U$. Then there is a natural isomorphism $\widetilde{H}_{n-1}(S^{n-1}) \longrightarrow H_n(U, U - x)$ given by the composition

$$\begin{aligned} \widetilde{H}_{n-1}(S^{n-1}) \leftarrow H_n(D^n, S^{n-1}) &\to H_n(D^n, D^n - 0) \\ &\to H_n(B, B - x) \to H_n(U, U - x), \end{aligned}$$

where all the arrows represent isomorphisms, the first is from the exact sequence of the pair (D^n, S^{n-1}), the second is from the fact that S^{n-1} has the homotopy type of $D^n - 0$, the third is from the translation and dilation homeomorphisms, and the last one is from the excision theorem (see Spanier [**43**], Theorems 3 and 4 in p.188). Then the image of $\alpha_{n-1} \in \widetilde{H}_{n-1}(S^{n-1})$ is a canonical generator α_x.

Theorem 5.7.3. *If $f : U \longrightarrow V$ is a diffeomorphism between connected open subsets of $\mathbb{R}^n$, then, for $x \in U$, the isomorphism*

$$f_* : H_n(U, U - x) \longrightarrow H_n(V, V - f(x))$$

is given by $f_*(\alpha_x) = n(f) \cdot \alpha_{f(x)}$.

Proof. We may reduce the problem to the case when f is a linear map, by finding a linear map homotopic to f in the following way. Suppose, without loss of generality, that $x = 0$ and $f(x) = 0$. Let ℓ be the linear map whose matrix is $Jf(0)$, and let $f - \ell = g$. Then $g(0) = 0$ and $Jg(0) = 0$. Define a homotopy $h_t : (U, U - 0) \longrightarrow (\mathbb{R}^n, \mathbb{R}^n - 0)$ by $h_t(x) = \ell(x) + (1 - t) \cdot g(x)$. Then $h_0 = f$ and $h_1 = \ell$

By a further homotopy, we may suppose that f is an orthogonal transformation, and so $f = r_1 \cdots r_k$, where r_i are reflections[1]. Then $(r_i)_*(\alpha_0) = -\alpha_0 = n(r_i) \cdot \alpha_0$. Therefore

$$f_*(\alpha_0) = n(f) \cdot \alpha_0,$$

by the above lemma. □

We now turn to algebraic description of the orientation of a manifold M. Let $x \in M$ and (U, ϕ) be a coordinate chart about x. Then we have the isomorphisms

$$H_n(M, M - x) \leftarrow H_n(U, U - x) \rightarrow H_n(\phi(U), \phi(U) - \phi(x)),$$

where the first one is the excision isomorphism, and the second one is induced by ϕ. Thus $H_n(M.M - x) \cong \mathbb{Z}$, which has two generators. This group is called the **local homology group** of M at x.

The union of $\widetilde{M}$ of all local homology groups $H_n(M, M - x)$, $x \in M$, is called the **orientation bundle** of M with projection $\gamma : \widetilde{M} \longrightarrow M$ defined by $\gamma^{-1}(x) = H_n(M, M - x)$. The next exercise claims that $\gamma : \widetilde{M} \longrightarrow M$ is a smooth fibre bundle with fibre $\mathbb{Z}$.

♢ **Exercise 5.13.** For each pair (U, z), where U is an open subset of M, and z is a cycle in $Z_n(M, M - U)$, define subsets U_z of $\widetilde{M}$ by

$$U_z = \{[z]_x \in H_n(M.M - x), x \in M\},$$

$[z]_x$ being the image of the homology class [z] in $H_n(M, M - x)$ by the homomorphism $H_n(M, M - U) \longrightarrow H_n(M, M - x)$ induced by inclusion. Then show that

(a) The set of all such U_z is a basis of a topology in $\widetilde{M}$.

(b) The map $\gamma : \widetilde{M} \longrightarrow M$ is continuous, and a local homeomorphism. It is even a covering projection.

Hence conclude that $\gamma : \widetilde{M} \longrightarrow M$ is a smooth fibre bundle with fibre $\mathbb{Z}$. Moreover, the maps $\widetilde{M} \times \widetilde{M} \longrightarrow \widetilde{M}$, given by $u + v = u \pm v$, are smooth whenever they are defined.

Hint. See Dold [**7**], p. 251.

A section ω of the orientation bundle of M may be described as a choice of a generator ω_x of $H_n(M, M - x)$ for each $x \in M$ so that each $x \in M$ has a neighbourhood U and a homology class $\omega_U \in H_n(M, M - U)$ whose image in $H_n(M, M - y)$ is ω_y for each $y \in U$.

Theorem 5.7.4. *A manifold M is orientable if and only if its orientation bundle admits a section.*

---

[1] This is Cartan-Dieudonné theorem: every orthogonal transformation is a product of reflections, see Artin [**1**].

PROOF. Suppose that M is orientable. Then, by Definition 5.5.2, there is an atlas $\mathcal{A}$ of M such that for any two charts (U, ϕ) and (V, ψ) in $\mathcal{A}$ we have $\det(\psi \circ \phi^{-1}) > 0$ on $\phi(U \cap V)$. Then define a section ω of the orientation bundle by $\omega_x = (\phi^{-1})_*(\alpha_{\phi(x)})$, where $(U, \phi) \in \mathcal{A}$, $x \in U$, and $\alpha_{\phi(x)} \in H_n(\phi(U), \phi(U) - \phi(x))$ is the canonical generator used in Theorem 5.7.3. It follows from that theorem and the fact that the change of coordinates has positive determinant that

$$(\psi \circ \phi^{-1})_*(\alpha_{\phi(x)}) = n(\psi \circ \phi^{-1})(\alpha_{\psi(x)}) = \alpha_{\psi(x)}.$$

Therefore the definitions of ω agree nicely on the intersection of two coordinate patches.

Conversely, given a section ω of the orientation bundle of M. we can choose an atlas $\mathcal{A}$ of M consisting of charts (U, ϕ) such that $\omega_x = (\phi^{-1})_*(\alpha_{\phi(x)})$ for $x \in U$ and $\alpha_{\phi(x)} \in H_n(\phi(U), \phi(U) - \phi(x))$. Then by Theorem 5.7.3 the change of coordinates will have positive determinant. □

Theorem 5.7.5. *If K is a compact subset of an n-manifold M without boundary, then*

(1) $H_i(M, M - K) = 0$ *for* $i > n$,
(2) *an element* $\alpha \in H_n(M, M - K)$ *is zero if and only if* $\lambda_*(\alpha) = 0$ *for every inclusion map* $\lambda : (M, M - K) \longrightarrow (M, M - x)$, $x \in K$,
(3) *if* ω *is an orientation of* M, *then there exists a unique* $\omega_K \in H_n(M, M - K)$ *such that* $\lambda_*(\omega_K) = \omega_x$, *for every* $x \in K$.

Note that for $M = K$, the class ω_M is called the **fundamental class of** *M.*

PROOF. The proof may be given in the following two steps:

(1) First show that if the theorem is true for compact subsets K_1, K_2, and $L = K_1 \cap K_2$, then it is true for $K = K_1 \cup K_2$. This may be accomplished by using the following Mayer-Vietoris exact sequence.

$$\cdots \longrightarrow H_{i+1}(M, M - L) \xrightarrow{\partial} H_i(M, M - K)$$

$$\xrightarrow{u} H_i(M, M - K_1) \oplus H_i(M, M - K_2) \xrightarrow{v} H_i(M, M - L) \longrightarrow \cdots,$$

where $u(\alpha) = (\lambda_{1*}(\alpha), \lambda_{2*}(\alpha))$ and $v(\alpha, \beta) = \mu_{1*}(\alpha) - \mu_{2*}(\beta)$, λ_i and μ_i being the inclusion maps

$$(M, M - K) \xrightarrow{\lambda_i} (M, M - K_i) \xrightarrow{\mu_i} (M, M - L),\ i = 1, 2.$$

Then clearly $H_i(M, M-K) = 0$ for $i > n$. The existence of ω_K follows from the fact that $v(\omega_{K_1}, \omega_{K_2}) = 0$. Now if ω'_K is another such class, and $\alpha = \omega_K - \omega'_K$, then each $\lambda_*(\alpha) = 0$ in $H_n(M, M - x)$, and so $u(\alpha) = 0$, which implies that $\alpha = 0$ since $H_{n+1}(M, M - L) = 0$. Therefore ω_K is unique.

(2) Next show that the theorem is true for a compact set within a coordinate neighbourhood U with coordinate system ϕ. This follows at once when $\phi(K)$

is a convex subset of $\mathbb{R}^n$. Because each inclusion $\phi(U)-\phi(x) \longrightarrow \phi(U)-\phi(K)$ is a homotopy equivalence, and so the induced homomorphisms

$$H_i(M, M-K) \longrightarrow H_i(M, M-x)$$

are isomorphisms, $x \in K$. Therefore the theorem holds for compact sets $K \subset U$ for which $\phi(K)$ is a finite union of convex sets.

For an arbitrary compact set $K \subset U$, note that for a class

$$\phi_*(\alpha) \in H_i(\phi(U), \phi(U) - \phi(K)),$$

where $\alpha \in H_*(U, U-K)$, there is a chain $c \in C_i(\phi(U))$ whose image modulo $\phi(U)-\phi(K)$ is a cycle representing $\phi_*(\alpha)$. Then the boundary $d(c)$ is supported by a compact set S which does not intersect $\phi(K)$. We can choose a compact neighbourhood N of $\phi(K)$ in $\phi(U)$ so that N also does not intersect S. This means that there is an $\alpha' \in H_i(\phi(U), \phi(U) - N)$ such that $\lambda_*(\alpha')) = \phi_*(\alpha)$. Now cover $\phi(K)$ by finitely many closed balls $B_1, \ldots, B_m$ so that $B = B_1 \cup \cdots \cup B_m \subset N$. There exists $\beta \in H_i(\phi(U), \phi(U) - N)$ such that $\lambda_*(\beta) = \phi_*(\alpha)$. By the above arguments, we have $\lambda_*(\beta) = 0$ in $H_i(\phi(U), \phi(U) - B)$ for $i > n$, and for $i = n$ we have $\lambda_*(\beta) = 0$ in $H_n(\phi(U), \phi(U) - x)$ for $x \in B$. Therefore $\alpha = 0$, since $H_i(U, U-K) = H_i(M, M-K)$ for $i \geq n$. Finally define $\omega_K = \phi_*^{-1}(\lambda_*(\omega_B))$.

This completes the steps we wanted to consider.

The theorem follows from these steps, because any compact set $K \subset M$ can be written as $K = K_1 \cup \cdots \cup K_r$, where each K_i is a compact subset of some coordinate neighbourhood in M. $\square$

$\Diamond$ **Exercise 5.14.** Prove the following general case of the theorem

For each compact subset K of an oriented n-manifold M with boundary, there exists a unique class $\omega_K \in H_n(M, (M-K) \cup \partial M)$ such that $\lambda_*(\omega_K) = \omega_x$ for every $x \in K \cap (M - \partial M)$, where $\lambda : (M, (M-K) \cup \partial M) \longrightarrow (M, M-x)$ is the inclusion map.

In particular, if M is compact, there is a unique $\omega_M \in H_n(M, \partial M)$ with the above property such that the boundary homomorphism $\partial : H_n(M, \partial M) \longrightarrow H_{n-1}(\partial M)$ maps ω_M onto the fundamental class of ∂M.

Hint. See Spanier [**43**], p. 304.

5.8. Integration of differential forms on manifolds

The support of a real valued smooth function f on $\mathbb{R}^n$ is the closure of a set on which f in not zero

$$\operatorname{supp} f = \overline{\{x \in \mathbb{R}^n \mid f(x) \neq 0\}}.$$

Let $x_1, \ldots, x_n$ be the standard coordinate functions on an open set $U \subseteq \mathbb{R}^n$, and $\omega = f \cdot dx_1 \cdots dx_n$ be an n-form on U, where f has compact support on

U. Then the integral of ω on U is defined to be the Riemann integral of f on U

$$\int_U \omega = \int_U f\, dx_1, \ldots, dx_n.$$

However, there is a difference between the Riemann integral of a function and the integral of a differential form. Where as the Riemann integral does not depend on the ordering of the variables $x_1, \ldots, x_n$, the integral of a form depends on the orientation of $\mathbb{R}^n$ in the following way. If π is a permutation of the sequence $\{1, \ldots, n\}$, then

$$\int_U f\, dx_{\pi(1)}, \ldots, dx_{\pi(n)} = (\text{sign } \pi) \int_U f\, dx_1, \ldots, dx_n.$$

If $\phi : V \longrightarrow U$ is a diffeomorphism changing coordinates from $(y_1, \ldots, y_n)$ to $(x_1, \ldots, x_n)$, then we have from calculus the change of variables formula for Riemann integral as

$$\int_U f\, dx_1 \cdots, dx_n = \int_V (f \circ \phi)\, |J(\phi)|\, dy_1 \cdots dy_n,$$

where $J(\phi) = \det(\partial\phi_i/\partial y_j)$ is the Jacobian determinant of ϕ.

For differential forms, the action of the pull-back $\phi^* : \Omega^1(U) \longrightarrow \Omega^1(V)$ on 1-forms dx_i is given by

$$\phi^* dx_i = \sum_j \frac{\partial \phi_i}{\partial y_j}\, dy_j = d\phi_i,$$

and $\phi^*(dx_1 \wedge \cdots \wedge dx_n) = d\phi_1 \wedge \cdots \wedge d\phi_n = J(\phi)\, dy_1 \wedge \cdots \wedge dy_n$. Therefore, if $\omega = f\, dx_1 \wedge \cdots \wedge dx_n$, we have the change of variables formula for integration of form as

$$\begin{aligned} \int_V \phi^*\omega &= \int_V \phi^*(f\, dx_1 \wedge \cdots \wedge dx_n) \\ &= \int_V (\phi^* f) \cdot \phi^*(dx_1 \wedge \cdots \wedge dx_n) \\ &= \int_V (f \circ \phi) \cdot J(\phi)\, dy_1 \wedge \cdots \wedge dy_n \\ &= \pm \int_V (f \circ \phi) \cdot |J(\phi)|\, dy_1 \wedge \cdots \wedge dy_n = \pm \int_U \omega, \end{aligned}$$

depending on whether the sign of the Jacobian determinant is positive or negative, that is, whether the diffeomorphism ϕ is orientation preserving or reversing.

We shall use this property of transformation to define integration of forms on manifolds. Let M be an orientable manifold of dimension n with or without boundary, and ω an n-form on M with compact support. Take an open covering on M by coordinate neighbourhoods $\{U_\alpha\}$ such that the diffeomorphisms $\phi_\alpha : U_\alpha \longrightarrow V_\alpha \subset \mathbb{R}^n$ or $\mathbb{R}^n_+$, and ϕ_α^{-1} are orientation preserving. Let $\{\rho_\alpha\}$ be a smooth partition of unity subordinate to the covering $\{U_\alpha\}$. Then by the local finiteness property (Property (iii) of Definition 2.1.9), all but finitely many of ρ_α are identically zero on the compact set $\operatorname{supp}\omega$. Thus only finitely many of the forms $\rho_\alpha\omega$ are nonzero, and each has compact support inside a coordinate neighbourhood. Define

$$\int_M \omega = \sum_\alpha \int_{U_\alpha} \rho_\alpha\omega,$$

where $\int_{U_\alpha} \rho_\alpha\omega = \int_{V_\alpha} (\phi_\alpha^{-1})^*(\rho_\alpha\omega)$. The definition may be justified as follows. Since $\sum_\alpha \rho_\alpha(x) = 1$, for every $x \in M$, we have $\sum_\alpha \rho_\alpha\omega = \omega$ and

$$\int_M \omega = \sum_\alpha \int_M \rho_\alpha\omega = \sum_\alpha \int_{U_\alpha} \rho_\alpha\omega.$$

Proposition 5.8.1. *The definition of the integral $\int_M \omega$ is independent of the choice of the oriented atlas $\{(U_\alpha, \phi_\alpha)\}$, and the subordinate smooth partition of unity $\{\rho_\alpha\}$.*

Proof. Let $\{(V_\beta, \psi_\beta)\}$ be another oriented atlas of M, and $\{\rho'_\beta\}$ a smooth partition of unity subordinate to the covering $\{V_\beta\}$. Since $\sum_\beta \rho'_\beta(x) = 1$ for every $x \in M$, $\omega = \sum_\beta \rho'_\beta\omega$. Therefore $\sum_\alpha \int_{U_\alpha} \rho_\alpha\omega = \sum_{\alpha,\beta} \int_{U_\alpha} \rho_\alpha\rho'_\beta\omega$. Now, since $\rho_\alpha\rho'_\beta\omega$ has support in $U_\alpha \cap V_\beta$, $\int_{U_\alpha} \rho_\alpha\rho'_\beta\omega = \int_{V_\beta} \rho_\alpha\rho'_\beta\omega$. Therefore

$$\begin{aligned}\sum_\alpha \int_{U_\alpha} \rho_\alpha\omega &= \sum_{\alpha,\beta} \int_{U_\alpha} \rho_\alpha\rho'_\beta\omega = \sum_{\alpha,\beta} \int_{V_\beta} \rho_\alpha\rho'_\beta\omega \\ &= \sum_\beta \int_{V_\beta} \rho'_\beta\omega.\end{aligned}$$

This completes the proof. □

◊ **Exercise 5.15.** (1) If M and N are oriented manifolds of the same dimension n, $\phi : N \longrightarrow M$ is an orientation preserving diffeomorphism, and ω is a compactly supported smooth n-form on M, then

$$\int_M \omega = \int_N \phi^*\,\omega.$$

(2) Prove the following properties of the integral.

$$\int_M (\omega_1 + \omega_2) = \int_M \omega_1 + \int_M \omega_2, \quad \text{and} \quad \int_M \lambda\,\omega = \lambda \int_M \omega, \text{ where } \lambda \in \mathbb{R}.$$

Theorem 5.8.2 (Stokes' Theorem)**.** *Let M be a compact oriented manifold of dimension n so that its boundary ∂M, which is a manifold of dimension $n-1$, has the boundary orientation. If ω is a smooth $(n-1)$-form om M, then*

$$\int_{\partial M} \omega = \int_M d\omega.$$

Note. The half space $\mathbb{R}^n_+$ has the orientation of the standard basis $e_1, \dots, e_n$ of $\mathbb{R}^n$. Its boundary $\partial\mathbb{R}^n_+ = \{(x_1, \dots, x_n) \in \mathbb{R}^n_+ \mid x_1 = 0\}$ is naturally diffeomorphic to $\{0\} \times \mathbb{R}^{n-1} \equiv \mathbb{R}^{n-1}$, where the diffeomorphism $(x_2, \dots, x_n) \longrightarrow (x_1, \dots, x_{n-1})$ is given by $x_i \mapsto x_{i-1}$ for $2 \le i \le n$. In the boundary orientation of $\partial\mathbb{R}^n_+$ the sign of the ordered basis $(e_2, \dots, e_n)$ is defined to be the sign of the ordered basis $(-e_1, e_2, \dots, e_n)$ in the standard orientation of $\mathbb{R}^n_+$. The definition is designed to make the Stokes' theorem sign free.

Proof. Since the equation is linear in ω, we may suppose that ω has compact support contained in a coordinate neighbourhood U_α and the diffeomorphism $\phi_\alpha : U_\alpha \longrightarrow V_\alpha \subset \mathbb{R}^n$ or $\mathbb{R}^n_+$ (and its inverse $\psi = \phi_\alpha^{-1} : V_\alpha \longrightarrow U_\alpha$) is orientation preserving.

First suppose that V_α is open in $\mathbb{R}^n$, that is, $\psi_\alpha(V_\alpha)$ does not intersect ∂M (see Theorem 1.8.3). Then $\int_{\partial M} \omega = 0$. Therefore we must show that $\int_M d\omega = 0$ also. This will establish the theorem in this case.

We have, by Proposition 3.5.7, $\int_M d\omega = \int_{V_\alpha} \psi_\alpha^* \, d\omega = \int_{V_\alpha} d\psi_\alpha^* \omega = \int_{V_\alpha} d\eta$, where $\eta = \psi_\alpha^* \omega$ is an $(n-1)$-form in $\mathbb{R}^n$. Therefore we may write

$$\eta = \sum_{i=1}^{n} (-1)^{i-1} f_i \, dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_n,$$

where $\widehat{dx_i}$ means that dx_i is omitted from the wedge product. Then

$$\begin{aligned} d\eta &= \sum_{i=1}^{n} (-1)^{i-1} df_i \wedge dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_n \\ &= \Big(\frac{\partial f_1}{\partial x_1} dx_1 + \frac{\partial f_1}{\partial x_2} dx_2 + \cdots + \frac{\partial f_1}{\partial x_n} dx_n\Big) \wedge dx_2 \wedge dx_3 \wedge \cdots \wedge dx_n \\ &\quad - \Big(\frac{\partial f_2}{\partial x_1} dx_1 + \frac{\partial f_2}{\partial x_2} dx_2 + \cdots + \frac{\partial f_2}{\partial x_n} dx_n\Big) \wedge dx_1 \wedge dx_3 \wedge \cdots \wedge dx_n \\ &\quad + \cdots \cdots \cdots \\ &= \Big(\frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} + \cdots + \frac{\partial f_n}{\partial x_n}\Big) dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n. \\ &= \Big(\sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i}\Big) dx_1 \wedge \cdots \wedge dx_n \end{aligned}$$

and

$$\int_{\mathbb{R}^n} d\eta \; = \; \sum_i \int_{\mathbb{R}^n} \frac{\partial f_i}{\partial x_i} \, dx_1 \cdots dx_n. \tag{4}$$

By Fubini's theorem, this integral over $\mathbb{R}^n$ may be computed by an iterated sequence of integrals over $\mathbb{R}^1$ in any order. The integral of the i-th term with respect to x_i is

$$\int_{\mathbb{R}^{n-1}} \left(\int_{-\infty}^{\infty} \frac{\partial f_i}{\partial x_i} dx_i \right) dx_1 \cdots \widehat{dx_i} \cdots dx_n.$$

Now $\int_{-\infty}^{\infty} \dfrac{\partial f_i}{\partial x_i} dx_i = \displaystyle\int_{-\infty}^{\infty} g'(t)dt$, where $g(t) = f_i(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_n)$. Since η has compact support, g vanishes outside any sufficiently large interval $(-a, a)$ in $\mathbb{R}^1$. Therefore, by the fundamental theorem of calculus,

$$\int_{-\infty}^{\infty} g'(t)dt = \int_{-a}^{a} g'(t)dt = g(a) - g(-a) = 0 - 0 = 0.$$

Therefore $\int_M d\omega = 0$.

When $V_\alpha \subset \mathbb{R}^n_+$, the above arguments will work for every term of (4), except the first one, since the boundary of $\mathbb{R}^n_+$ is the set where $x_1 = 0$. The first integral is

$$\int_{\mathbb{R}^{n-1}} \Big(\int_0^{\infty} \frac{\partial f_1}{\partial x_1} dx_1 \Big) dx_2 \cdots dx_n.$$

Now compact support of η implies that f_1 vanishes if x_1 is outside some large interval $[0, a)$. Then $f_1(a, x_2, \ldots, x_n) = 0$, but $f_1(0, x_2, \ldots, x_n) \neq 0$. Therefore, by the fundamental theorem of calculus,

$$\int_M d\omega = \int_{\mathbb{R}^{n-1}} -f_1(0, x_2, \ldots, x_n)\, dx_2 \cdots dx_n.$$

On the other hand, $\int_{\partial M} \omega = \int_{\partial \mathbb{R}^n_+} \eta$. Since $x_1 = 0$ on $\partial\mathbb{R}^n_+$, $dx_1 = 0$ on $\partial\mathbb{R}^n_+$ also. Therefore, if $i > 1$, then the form $(-1)^{i-1} f_i\, dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_n$ restricts to 0 on $\partial\mathbb{R}^n_+$. Therefore the restriction of η to $\partial\mathbb{R}^n_+$ is

$$-f_1(0, x_2, \cdots, x_n)\, dx_2 \cdots dx_n,$$

and its integral over $\partial\, \mathbb{R}^n_+$ is $\int_{\partial M} \omega$, using the definition of the boundary orientation. $\square$

CHAPTER 6

TRANSVERSALITY

This chapter is devoted to elementary transversality theory with some simple applications. The more serious transversality theorem will appear in §8.7. The condition of transversality was first introduced by Thom as a generalisation of the condition of submersion. The transversality condition on a smooth map $f : M \longrightarrow N$ with respect to a submanifold $A \subset N$ ensures that the inverse image $f^{-1}(A)$ is a submanifold of M. In the special case when A is a point $y \in N$, the transversality reduces to the fact that y is a regular value of f, and so, as we already know, $f^{-1}(y)$ is a submanifold of M. It was this result that enabled Thom build up his cobordism theory by associating manifolds to smooth maps. Based on transversality we have the concepts of the degree of a map, the index of a vector field, and the Euler characteristic of a manifold. These objects are homotopy invariants, and thus lead us to the realm of algebraic topology.

6.1. ϵ-neighbourhood of submanifold of Euclidean space

Let M be an n-submanifold of $\mathbb{R}^m$. Then the normal space of M at a point $x \in M$ is the orthogonal complement $\nu(M)_x$ of the tangent space $\tau(M)_x$. Explicitly

$$\nu(M)_x = \{(x, v) \in M \times \mathbb{R}^m \mid v \perp \tau(M)_x\}.$$

The **normal bundle** $\nu(M)$ of M is the disjoint union of all normal spaces $\nu(M)_x$ as x runs over M. The map $\sigma : \nu(M) \longrightarrow M$ given by $\sigma(x, v) = x$ is the projection of the normal bundle.

Theorem 6.1.1. *The normal bundle $\sigma : \nu(M) \longrightarrow M$ is a vector bundle of fibre dimension $m - n$, and so $\nu(M)$ is a manifold of dimension m, and the projection σ is a submersion.*

Proof. The proof uses a simple fact from linear algebra which we want state first. If $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a linear map, then its transpose is a linear map $L^t : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ defined by the inner product relation $\langle Lv, w\rangle = \langle v, L^t w\rangle$, where $v \in \mathbb{R}^n$ and $w \in \mathbb{R}^m$. It follows easily from the definition that the matrix of L^t is the transpose of the matrix of L. If L is surjective, then L^t is injective; moreover, L^t maps $\mathbb{R}^m$ isomorphically onto the orthogonal complement $(\operatorname{Ker} L)^{\perp}$ of $\operatorname{Ker} L$.

A. Mukherjee, *Differential Topology*, DOI 10.1007/978-3-319-19045-7_6

Turning now to the proof of the theorem, note that the submanifold $M \subset \mathbb{R}^m$ admits at each of its point a neighbourhood V in $\mathbb{R}^m$ and a submersion

$$\phi : V \longrightarrow \mathbb{R}^{m-n} \quad (m-n \text{ is the codim } M)$$

such that $U = M \cap V = \phi^{-1}(0)$ (Lemma 1.5.2). Then $\nu(U) = \nu(M) \cap (U \times \mathbb{R}^m)$ is an open subset of $\nu(M)$. If $x \in U$, then $d\phi_x : \mathbb{R}^m \longrightarrow \mathbb{R}^{m-n}$ is an epimorphism and its kernel is $\tau(M)_x$. Its transpose $(d\phi_x)^t$ then maps $\mathbb{R}^{m-n}$ isomorphically onto the orthogonal complement $\nu(M)_x$. Therefore the map $\psi : U \times \mathbb{R}^{m-n} \longrightarrow \nu(U)$, defined by $\psi(x, v) = (x, (d\phi_x)^t(v))$ is bijective. It is, in fact, a diffeomorphism, because its Jacobian matrix

$$\begin{pmatrix} I_n & 0 \\ 0 & S \end{pmatrix},$$

where S is the matrix of $(d\phi)^t$, has rank m. This proves all the assertions of the theorem. $\square$

If M is compact, define for a given $\epsilon > 0$ the open subsets

$$\nu(M, \epsilon) = \{(x, v) \in \nu(M) \mid \|v\| < \epsilon\}, \;\; M(\epsilon) = \{x \in \mathbb{R}^m \mid d(x, M) < \epsilon\}.$$

Here d denotes the standard metric on $\mathbb{R}^m$, and $\| \; \|$ is the standard norm.

It is necessary to modify the definitions when M is non-compact. In this case, define for a given smooth function $\epsilon : M \longrightarrow (0, \infty)$ the above subsets as

$$\nu(M, \epsilon) = \{(x, v) \in \nu(M) \mid \|v\| < \epsilon(x)\},$$

$$M(\epsilon) = \{x \in \mathbb{R}^m \mid d(x, M) < \epsilon(x)\}.$$

We shall refer to $M(\epsilon)$ as an ϵ-neighbourhood of M. For compact M, ϵ may be taken to be a constant function, and we get back the previous definition, as the following lemma shows.

Lemma 6.1.2. *Let M be a submanifold of $\mathbb{R}^m$. Then for any open neighbourhood U of M in $\mathbb{R}^m$ there is a smooth function $\epsilon : M \longrightarrow (0, \infty)$ such that $M(\epsilon) \subset U$. When M is compact, ϵ may be taken constant.*

Proof. For each $x \in M$ choose an open neighbourhood U_x in U and a number $\epsilon_x > 0$ so that the ϵ_x-neighbourhood $U_x(\epsilon_x) \subset U$. This gives an open covering $\{U_x\}$ of M. Choose a subordinate partition of unity $\{\lambda_i\}$. Then the function ϵ is given by $\sum_i \lambda_i \epsilon_i$ (after re-indexing the ϵ_x). $\square$

Lemma 6.1.3. *Let $f : M \longrightarrow N$ be a smooth map, and $A \subset M$, $B \subset N$ be submanifolds such that $df_x : \tau(M)_x \longrightarrow \tau(N)_{f(x)}$ is an isomorphism for every $x \in A$, and $f|A : A \longrightarrow B$ is a diffeomorphism. Then there is an open neighbourhood V of A in M such that $f(V)$ is an open neighbourhood of B in N, and $f|V$ is a diffeomorphism.*

Note that if A is a point, this is just the inverse function theorem.

PROOF. The proof is due to Godement [**10**], p. 150, Since A is second countable, there is a countable covering of A by open sets U_i in M such that $f_i = f|U_i$ is a diffeomorphism. Since B is paracompact, there is a locally finite covering of B by open sets V_i in N which is a refinement of the open covering $\{f(U_i)\}$. By replacing U_i by $f_i^{-1}(V_i)$, where $V_i \subset f(U_i)$, we may suppose that $V_i = f_i(U_i)$. Find another locally finite covering $\{W_i\}$ of B by open sets of N such that $\overline{W}_i \subset V_i$. Let $W = \cup_{i=1}^{\infty} W_i$, and

$$F = \{y \in W \mid y \in \overline{W}_i \cap \overline{W}_j \Rightarrow f_i^{-1}(y) = f_j^{-1}(y)\}.$$

Then $B \subset F$.

We shall show that F contains an open neighbourhood G of B in N. For each $b \in B$, any open neighbourhood G_b of b intersects only finitely many $\overline{W}_i$, so, in particular, b is within a finite number of them, say, $\overline{W}_1, \ldots, \overline{W}_k$. We can take G_b small enough so that G_b is contained in $\cap_{i=1}^{k} \overline{W}_i \subset \cap_{i=1}^{k} V_i$. Now f_i^{-1} is a local inverse of f near b. Then there exists an open neighbourhood H of b such that $f_1^{-1}|H = \cdots = f_k^{-1}|H$, using uniqueness of inverse function and the fact $f_1^{-1}(b) = \cdots = f_k^{-1}(b)$. Then, $\widetilde{G}_b = H \cap G_b$ is an open neighbourhood of b in N, and $\widetilde{G}_b \subset F$. Let $G = \cup_{b \in B} \widetilde{G}_b$. Then $G \subset F$. Now define $g : G \longrightarrow M$ by $g(y) = f_i^{-1}(y)$ if $y \in G \cap W_i$. Then g is well defined, and smooth since locally $g = f_i^{-1}$ for some i. Thus f and g are diffeomorphisms inverse to each other. Then, taking $V = g(G)$, we get what we want. $\square$

Let $\theta : \nu(M) \longrightarrow \mathbb{R}^m$ be the map given by $\theta(x, v) = x + v$.

Theorem 6.1.4 (**ϵ-Neighbourhood Theorem**). *If M is a boundaryless submanifold of $\mathbb{R}^m$, then there is a function $\epsilon : M \longrightarrow (0, \infty)$ such that θ maps $\nu(M, \epsilon)$ diffeomorphically onto the ϵ-neighbourhood $M(\epsilon)$. Moreover, $\pi = \sigma \circ \theta^{-1} : M(\epsilon) \longrightarrow M$ is a submersion and a retraction, where σ is the projection of the normal bundle.*

PROOF. It follows from the proof of Theorem 6.1.1 that the tangent space of the normal bundle $\tau(\nu(M))_{(x,0)}$ splits as

$$\tau(\nu(M))_{(x,o)} = \tau(M \times \{0\})_{(x,0)} \oplus \tau(\{x\} \times \nu(M)_x)_{(x,0)} = \tau(M)_x \oplus \nu(M)_x.$$

Because, if U is an open neighbourhood of x in M, then an open neighbourhood of $(x, 0)$ in $\nu(M)$ is given by the product of two manifolds $U \times \{0\}$ and $\{x\} \times \nu(M)_x$. The restriction of $d\theta$ to $\tau(M)_x$ is just di, where $i : M \longrightarrow \mathbb{R}^m$ is the inclusion map, and therefore rank $(d\theta \,|\, \tau(M)_x) = n$. Again, rank $(d\theta \,|\, \nu(M)_x) = m - n$, since θ is a translation for a fixed x. Thus rank $d\theta = m$, and hence $d\theta$ is an isomorphism, at each point $(x, 0) \in M \times \{0\}$. Since θ also maps $M \times \{0\}$ diffeomorphically onto M, it must map an open neighbourhood of $M \times 0$ in $\nu(M)$ diffeomorphically onto a neighbourhood of M in $\mathbb{R}^m$ (Lemma 6.1.3). By Lemma 6.1.2, any neighbourhood of M in $\mathbb{R}^m$ contains some $M(\epsilon)$. Thus $\theta^{-1} : M(\epsilon) \longrightarrow \nu(M, \epsilon)$ is a diffeomorphism, and hence $\pi = \sigma \circ \theta^{-1}$ is a submersion, because σ is so. Clearly π is a retraction. $\square$

We shall close this section with some applications to the homotopy theory.

Theorem 6.1.5. *Let M and N be manifolds and N have no boundary. Then given any continuous map $f : M \longrightarrow N$ which is smooth on a closed subset K of M, there is a smooth map $g : M \longrightarrow N$ such that $g|K = f|K$, and g is homotopic to f (rel K). (We may have $K = \emptyset$),*

Proof. Embed N is some Euclidean space $\mathbb{R}^k$ as a closed submanifold. Then N inherits a metric from $\mathbb{R}^k$ which induces the same topology on N. Take an ϵ-neighbourhood $N(\epsilon)$ of N in $\mathbb{R}^k$, and let $\pi : N(\epsilon) \longrightarrow N$ be the retraction as given by Theorem 6.1.4. By Theorem 2.2.3, there is a smooth map $h : M \longrightarrow N$ within ϵ of f so that $h(M) \subset N(\epsilon)$, and $h|K = f|K$. Let $g = \pi \circ h$. Then g is smooth, and $g|K = \pi \circ h|K = \pi \circ f|K = f|K$, and g is homotopic to f (rel K) by the homotopy given by

$$H(x,t) = \pi[tf(x) + (1-t)h(x)].$$

The homotopy is well-defined, since, for any $x \in M$ and $0 \leq t \leq 1$, $tf(x) + (1-t)h(x)$ is within ϵ of $f(x)$, and so belongs to $N(\epsilon)$. □

Thus, if M is without boundary, then every element of a homotopy group $\pi_k(M)$ may be represented by a smooth map $S^k \longrightarrow M$. Here is a simple application.

Proposition 6.1.6. *If M is an n-manifold and $n < m$, then any continuous map $f : M \longrightarrow S^m$ is homotopic to a constant map.*

Proof. The map f is homotopic to a smooth map $g : M \longrightarrow S^m$. Since $n < m$, Sard's theorem implies that there is a point $p \in S^m$ which is outside Image g (Corollary 2.3.10). Now $S^m - \{p\}$ is homeomorphic to the contractible space $\mathbb{R}^m$ by the stereographic projection from p onto the tangent space to S^m at $-p$. Therefore the identity map of $S^m - \{P\}$ is homotopic to a constant map. The composition of g and this homotopy gives a homotopy of g to a constant map. □

Theorem 6.1.7. *Let $f : M \longrightarrow N$ be a continuous map of manifolds, and $\epsilon : N \longrightarrow \mathbb{R}$ be a positive continuous function. Then f can be ϵ-approximated to a smooth map $g : M \longrightarrow N$ which is homotopic to f by a homotopy $h : M \times [0,1] \longrightarrow N$ such that*

(1) $h(x,t) = f(x)$ for any $x \in M$ for which $g(x) = f(x)$,

(2) $h_t : M \longrightarrow N$ is an ϵ-approximation to f for any $t \in [0,1]$

Proof. As in Theorem 6.1.5, we embed N in some $\mathbb{R}^k$, and give N the metric induced from $\mathbb{R}^k$. There is an open neighbourhood U of N in $\mathbb{R}^k$, and a smooth retraction $\rho : U \longrightarrow N$. We choose a positive continuous function $\delta : N \longrightarrow \mathbb{R}$ such that the open ball in $\mathbb{R}^k$ with centre at $f(x)$ and radius $\delta(x)$ lies in U, and such that the image of the ball under ρ lies in a ball in N about $f(x)$ of radius less than $\epsilon(x)$. We approximate f by a smooth map

$f_1 : M \longrightarrow N$ within δ, using Theorem 2.2.3, and define $g = \rho \circ f_1$. Then g is smooth, and

$$\|g(x) - f(x)\| = \|\rho(f_1(x)) - f(x)\| < \epsilon(x).$$

The line segment joining $f(x)$ and $g(x)$ lies in U. Then the homotopy $h : M \times [0,1] \longrightarrow N$ defined by $h(x,t) = \rho((1-t)f(x) + tg(x))$ has the properties (1) and (2) of the theorem. □

Complement 6.1.8. If M is compact, N has no boundary, and $\dim N \geq 2\dim M + 1$, and if f is a smooth map whose restriction to a closed subset K of M is an embedding, then we can make f_1 an embedding with $f_1|K = f|K$, using Theorem 2.4.3. Then g will be homotopic to f by a homotopy h_t which is an ϵ-approximation to f for each t. By the stability property (Theorem 2.7.2(4)), h_t will be an embedding for small values of t.

6.2. Transversality

Definition 6.2.1. Let M and N be manifolds, and A be a submanifold of N. Then a smooth map $f : M \longrightarrow N$ is called **transverse to A at a point** $x \in f^{-1}(A)$ (written $f \overline{\pitchfork}_x A$) if

$$df_x(\tau(M)_x) + \tau(A)_{f(x)} = \tau(N)_{f(x)} \text{ (not necessarily a direct sum) .}$$

We say that f is **transverse to** A (written $f \overline{\pitchfork} A$) if either (1) $f^{-1}(A) = \emptyset$, or (2) $f \overline{\pitchfork}_x A$ for every $x \in f^{-1}(A)$.

If S is a subset of $f^{-1}(A)$ and $f \overline{\pitchfork}_x A$ at all points $x \in S$, then we say that f is transverse to A on S. If S is a subset of A and $f \overline{\pitchfork}_x A$ for all $x \in f^{-1}(S)$, then also we say that f is transverse to A on S (this terminology will be used in §8.7).

If $\dim M + \dim A < \dim N$, then the condition (2) is not possible, so in this case $f \overline{\pitchfork} A$ means that $f(M)$ does not intersect A.

The condition (2) is equivalent to saying that df_x induces an epimorphism $\overline{df_x} = \pi \circ df_x : \tau(M)_x \longrightarrow \tau(N)_{f(x)}/\tau(A)_{f(x)}$. Indeed, $\overline{df_x}$ is an epimorphism if and only if for every $v \in \tau(N)_{f(x)}$ there is a $w \in \tau(M)_x$ such that $v - df_x(w) \in \tau(A)_{f(x)}$. Thus f is always transverse to any open subset of N, because then the target of $\overline{df_x}$ is zero. Also if f is a submersion, then it is transverse to any submanifold A of N. If A is a point y in N, then the transversality condition (2) means that y is a regular value of f.

◊ **Exercise 6.1.** Show that the set of points in $f^{-1}(A)$ at which f is transverse to A is an open subset of $f^{-1}(A)$.

Definition 6.2.2. Two submanifolds M_1 and M_2 of a manifold N are called **transverse** (written $M_1 \overline{\pitchfork} M_2$) if the inclusion map $i : M_1 \longrightarrow N$ is transverse to M_2. Then the condition (2) means that, for every $x \in M_1 \cap M_2$,

$$\tau(M_1)_x + \tau(M_2)_x = \tau(N)_x,$$

since di_x is the inclusion of $\tau(M_1)_x$ into $\tau(N)_x$.

In this case we also say that M_1 and M_2 are in a general position in N. It may be seen that a smooth map $f: M \longrightarrow N$ is transverse to a submanifold $A \subset N$ if and only if the graph of f and $M \times A$ are in general position in $M \times N$.

Note that we may interchange the role of M_1 and M_2 in this definition.

Example 6.2.3. Two curves in $\mathbb{R}^2$ which intersect at a point are non-transverse if the curves are tangent to each other at the point.

In all the results that we shall prove in this section, it is assumed that $f^{-1}(A) \neq \emptyset$ whenever $f \pitchfork A$.

Lemma 6.2.4. *Let $f: M \longrightarrow N$ be a smooth map, and A a submanifold of N of* codim k. *Let $x \in f^{-1}(A)$ and U be an open neighbourhood of $f(x)$ in N, and $g: U \longrightarrow \mathbb{R}^k$ a submersion such that $g^{-1}(0) = A \cap U$ (U and g exist by Lemma 1.5.2). Then $f \pitchfork_x A$ if and only if x is a regular point* (or 0 is a regular value) of $g \circ f$.

Proof. We have Ker $dg_{f(x)} = \tau(A)_{f(x)}$, since g is constant on $A \cap U$. Then $f \pitchfork_x A$ if and only if

$$df_x(\tau(M)_x) + \text{Ker } dg_{f(x)} = \tau(N)_{f(x)},$$

or equivalently, $d(g \circ f)_x = dg_{f(x)} \circ df_x$ is an epimorphism. To see this, take any $a \in \mathbb{R}^k$, then, since $dg_{f(x)}$ is an epimorphism, $a = dg_{f(x)}(u)$ for some $u \in \tau(N)_{f(x)}$, and $u = df_x(v) + w$ for some $v \in \tau(M)_x$ and $w \in$ Ker $dg_{f(x)}$, therefore $a = dg_{f(x)} \circ df_x(v)$, and $d(g \circ f)_x$ is an epimorphism.

Conversely, if $d(g \circ f)_x$ is an epimorphism, then, for any $v \in \tau(N)_{f(x)}$, there is a $w \in \tau(M)_x$ such that $dg_{f(x)} \circ df_x(w) = dg_{f(x)}(v)$, so $v - df_x(w) \in$ Ker $dg_{f(x)}$, and the above equality of vector spaces holds. □

Theorem 6.2.5. *If a smooth map $f: M \longrightarrow N$ is transverse to a submanifold A of N, where none of M, N, and A has boundary, then $f^{-1}(A)$ is a submanifold of M, whose codimension in M is equal to the codimension of A in N. In particular, if* $\dim M =$ codim A, *then $f^{-1}(A)$ consists of isolated points only.*

Proof. It is sufficient to show that $f^{-1}(A)$ is locally a submanifold, that is, each point x of $f^{-1}(A)$ has an open neighbourhood V in M such that $V \cap f^{-1}(A)$ is a manifold.

By Lemma 6.2.4, each $x \in f^{-1}(A)$ is a regular point of $g \circ f$, and the open neighbourhood $V = f^{-1}(U)$ of x is such that $V \cap f^{-1}(A) = (g \circ f)^{-1}(0)$. Therefore, by Theorem 1.5.6, $V \cap f^{-1}(A)$ is a submanifold of M of dimension $n - k$, where $n = \dim M$ and $k =$ codim A. □

Corollary 6.2.6. *The intersection of two transverse submanifolds M_1 and M_2 of N, where none of them has boundary, is a submanifold, and*

$$\text{codim }(M_1 \cap M_2) = \text{codim } M_1 + \text{codim } M_2.$$

PROOF. This is a special case of the above theorem. □

Note that the transversality of two submanifolds depends on the dimension of the manifold where they are embedded. For example, the two coordinate axes are transverse in $\mathbb{R}^2$, but they are not in $\mathbb{R}^3$.

Lemma 6.2.7. *If $\pi : B \longrightarrow \mathbb{R}$ is a smooth map with a regular value at* 0, *then $\pi^{-1}[0, \infty)$ is a manifold with boundary $\pi^{-1}(0)$*

PROOF. The set $\pi^{-1}(0, \infty)$ is open in B, so it is a submanifold of B. A point $x \in B$ for which $\pi(x) = 0$ is a regular point, and so it has an open neighbourhood in B where π looks like the canonical submersion. The proof now follows, because the lemma is obviously true for the canonical submersion $\mathbb{R}^n \longrightarrow \mathbb{R}$. □

NOTATION. If M is a manifold with boundary and $f : M \longrightarrow N$ is a smooth map, then ∂f will denote the restriction map $f|\partial M : \partial M \longrightarrow N$.

Theorem 6.2.8. *Let M be a manifold with boundary, N a boundaryless manifold, A a boundaryless submanifold of N, and $f : M \longrightarrow N$ a smooth map. Then, if both f and $\partial f = f|\partial M$ are transverse to A, the inverse image $f^{-1}(A)$ is a neat submanifold of M with boundary*

$$\partial(f^{-1}(A)) = f^{-1}(A) \cap \partial M,$$

and the codimension of $f^{-1}(A)$ in M equals that of A in N.

PROOF. The map $f|\text{Int } M : \text{Int } M \longrightarrow N$ is transverse to A, and therefore $(f|\text{Int } M)^{-1}(A) = f^{-1}(A) \cap \text{Int } M$ is a boundaryless submanifold, by Theorem 6.2.5. Therefore it is necessary only to examine $f^{-1}(A)$ in a neighbourhood of a point $x \in f^{-1}(A) \cap \partial M$ (note that $f^{-1}(A) = (f^{-1}(A) \cap \text{Int } M) \cup (f^{-1}(A \cap \partial M))$. Applying Lemma 1.5.2 to $f(x) \in A$, we get an open neighbourhood P of $f(x)$ in N and a submersion $g : P \longrightarrow \mathbb{R}^k$ ($k = \text{codim } A$) such that $g^{-1}(0) = A \cap P$. Then $(g \circ f)^{-1}(0) = f^{-1}(A) \cap Q$, where $Q = f^{-1}(P)$. Let U be a coordinate neighbourhood of x and $\phi : U \longrightarrow \mathbb{R}^n_+$ be a coordinate system with $\phi(U \cap Q) = V$ an open set in $\mathbb{R}^n_+$ ($n = \dim M$). Let h denote the smooth map $g \circ f \circ \phi^{-1} : V \longrightarrow \mathbb{R}^k$. Then $f^{-1}(A)$ will be a manifold with boundary near x if and only if $\phi(f^{-1}(A) \cap Q \cap U) = h^{-1}(0)$ is a manifold with boundary near $\phi(x) = a$.

We have by Lemma 6.2.4,

$$f \overline{\pitchfork}_x A \Leftrightarrow g \circ f \text{ is regular at } x \Leftrightarrow h \text{ is regular at } a.$$

Extend $h : V \longrightarrow \mathbb{R}^k$ to a smooth map $\widetilde{h} : \widetilde{V} \longrightarrow \mathbb{R}^k$ on an open set $\widetilde{V}$ of $\mathbb{R}^n$. Since $d\widetilde{h}_a = dh_a$, h is regular at a implies $\widetilde{h}$ is regular at a. This means that the intersection of $\widetilde{h}^{-1}(0)$ with some open neighbourhood of the regular point a is a boundaryless submanifold Z of $\mathbb{R}^n$, since $\widetilde{h}$ is a smooth map of the boundaryless manifold $\widetilde{V}$. Without loss of generality, we may suppose that $Z = \widetilde{h}^{-1}(0)$.

As $h^{-1}(0) = Z \cap \mathbb{R}^n_+$, we must show that $Z \cap \mathbb{R}^n_+$ is a manifold with boundary. For this purpose, let $\pi : Z \longrightarrow \mathbb{R}$ be the restriction to Z of the first coordinate function on $\mathbb{R}^n$. Then $Z \cap \mathbb{R}^n_+ = \pi^{-1}[0, \infty)$. By Lemma 6.2.7, this will be a manifold with boundary if 0 is a regular value of π. Suppose that 0 is not a regular value of π. Then $\pi(z) = 0$ and $d\pi_z = 0$ for some point $z \in Z$. Now $\pi(z) = 0$ implies $z \in Z \cap \partial\mathbb{R}^n_+$, and, since $d\pi_z = \pi$ (π being linear), $d\pi_z = 0$ implies that the first coordinate of every vector in the tangent space $\tau(Z)_z$ is zero, or $\tau(Z)_z \subset \tau(\partial R^n_+)_z = \mathbb{R}^{n-1}$. Since $\widetilde{h}^{-1}(0) = Z$, $\tau(Z)_z$ is the kernel of $d\widetilde{h}_z$, and, since $dh_z = d\widetilde{h}_z$, Ker $dh_z = \tau(Z)_z \subset \mathbb{R}^{n-1}$. This means that the maps $dh_z : \mathbb{R}^n \longrightarrow \mathbb{R}^k$ and $d(\partial h)_z : \mathbb{R}^{n-1} \longrightarrow \mathbb{R}^k$ have the same kernel, because $d(\partial h)_z = dh_z|\mathbb{R}^{n-1}$. By transversality conditions both the linear maps are epimorphisms. But the dimension relation for linear maps says that $\dim \operatorname{Ker} dh_z = n - k$, whereas $\dim \operatorname{Ker} d(\partial h)_z = n - 1 - k$. This is a contradiction. Therefore 0 must be a regular value of h. $\square$

$\Diamond$ **Exercise 6.2.** Let M be a manifold with boundary, N a manifold without boundary, A a submanifold of N with boundary, and $f : M \longrightarrow N$ a smooth map. Then, if both f and $\partial f = f|\partial M$ are transverse to A and f is transverse to ∂A, the inverse image $f^{-1}(A)$ is a submanifold of M with boundary

$$\partial(f^{-1}(A)) = [f^{-1}(A) \cap \partial M] \cup f^{-1}(\partial A),$$

and the codimension of $f^{-1}(A)$ in M equals that of A in N.

$\Diamond$ **Exercise 6.3 (Transitivity of transverse maps).** Let $f : M \longrightarrow N$ and $g : N \longrightarrow R$ be smooth maps, and A is a submanifold of R such that $g \pitchfork A$. Then show that $f \pitchfork g^{-1}(A)$ if and only if $(g \circ f) \pitchfork A$.

In particular, if $M = N$ and f is a diffeomorphism, then $(g \circ f) \pitchfork A$.

Theorem 6.2.9 (Transversality Theorem). *Suppose M, N, A, and B are manifolds, where only M has boundary, and A is a submanifold of N. Suppose $F : M \times B \longrightarrow N$ is a smooth map such that both F and ∂F are transverse to A. Suppose, for each $b \in B$, $f_b : M \longrightarrow N$ is the map $f_b(x) = F(x, b)$. Then. for almost all $b \in B$, both f_b and ∂f_b are transverse to A.*

PROOF. By Theorem 6.2.8, $C = F^{-1}(A)$ is a manifold with boundary $\partial C = C \cap \partial(M \times B)$. Let $\pi : M \times B \longrightarrow B$ be the projection onto the second factor. We shall show that (1) if $b \in B$ is a regular value of $\pi|C$, then f_b is transverse to A, and (2) if $b \in B$ is a regular value of $\partial\pi|\partial C$, then ∂f_b is transverse to A. This will complete the proof, because, by Sard's theorem, almost every point b of B is a regular value of both the maps $\pi|C$ and $\partial\pi|\partial C$.

It is sufficient to prove only (1), because (2) follows from the fact that (1) is true for the special case of the boundaryless manifold ∂M.

Let x be any point of $f_b^{-1}(A)$. Then $f_b(x) = F(x,b) = a \in A$, and the transversality condition of F gives that

$$dF_{(x,b)}(\tau(M \times B)_{(x,b)}) + \tau(A)_a = \tau(N)_a.$$

Therefore, there is always a $u \in \tau(M \times B)_{(x,b)}$ for any given $v \in \tau(N)_a$ such that $dF_{(x,b)}(u) - v \in \tau(A)_a$. The problem here is to find a $w \in \tau(M)_x$ such that $df_b(w) - v \in \tau(A)_a$. Since $\tau(M \times B)_{(x,b)} = \tau(M)_x \times \tau(B)_b$, we may write $u = (r,t)$ where $r \in \tau(M)_x$ and $t \in \tau(B)_b$. If b is a regular value of $\pi|C : C \longrightarrow B$, then $d(\pi|C)_{(x,b)} : \tau(C)_{(x,b)} \longrightarrow \tau(B)_b$ is onto, and it is the restriction of $d\pi_{(x,b)} : \tau(M)_x \times \tau(B)_b \longrightarrow \tau(B)_b$ which is just the projection onto the second factor. Therefore for $t \in \tau(B)_b$ we can find a $(s,t) \in \tau(C)_{(x,b)}$, $s \in \tau(M)_x$, which is mapped onto t by $d\pi_{(x,b)}$. Since F maps C onto A, and $F(x,b) = a$, we have $dF_{(x,b)}(s,t) \in \tau(A)_a$. Then $w = r - s \in \tau(M)_x$ is a solution of our problem. To see this, first note that, since $F|(M \times \{b\}) = f_b$, we have $dF_{(x,b)}(w,0) = df_b(w)$. Therefore

$$df_x(w) - v = dF_{(x,b)}(w,0) - v = dF_{(x,b)}[(r,t) - (s,t)] - v$$
$$= [dF_{(x,b)}(r,t) - v] - dF_{(x,b)}(s,t) = [dF_{(x,b)}(u) - v] - dF_{(x,b)}(s,t) \in \tau(A)_a.$$

This completes the proof. □

A consequence of the transversality theorem is that transverse maps are generic, in the sense that almost all C^∞ maps $M \longrightarrow N$ are transverse to any submanifold A of N. We show this first for the case when $N = \mathbb{R}^m$.

Corollary 6.2.10. *Any smooth map $f : M \longrightarrow \mathbb{R}^m$ is arbitrarily close to a smooth map $g : M \longrightarrow \mathbb{R}^m$ which is transverse to any boundaryless submanifold A of $\mathbb{R}^m$. Moreover, f is homotopic to g by a small homotopy.*

Proof. Let B be an open ball about the origin in $\mathbb{R}^m$. Define $F : M \times B \longrightarrow \mathbb{R}^m$ by $F(x,b) = f(x) + b$. For a fixed $x \in M$, F simply translates the ball B, and so it a submersion on $\{x\} \times B$. Then, both F and ∂F are submersions, because the Jacobian matrix of each contains the matrix of this translation as a submatrix. Therefore F and ∂F are transverse to any boundaryless submanifold A of $\mathbb{R}^m$. Then Theorem 6.2.9 implies that for almost every $b \in B$, the map $f_b(x) = f(x) + b$ is transverse to A. Choosing such a b, f may be deformed into the transverse map $g = f_b$ by the homotopy $H : M \times I \longrightarrow \mathbb{R}^m$ given by $H(x,t) = f(x) + tb$. Finally, the map g may be made arbitrarily closed to f by making the radius of B sufficiently small. □

The next theorem deals with the general case.

Theorem 6.2.11 (Transversality Homotopy Theorem). *For any smooth map $f : M \longrightarrow N$, $\partial N = \emptyset$, and any boundaryless submanifold A of N, there is a smooth map $g : M \longrightarrow N$ such that both g and ∂g are transverse to A, and f is homotopic to g.*

Proof. Embed N in some $\mathbb{R}^m$, and let $N(\epsilon) = \{x \in \mathbb{R}^m | d(x, N) < \epsilon(x)\}$ be an ϵ-neighbourhood of N in $\mathbb{R}^m$ with retraction $r : N(\epsilon) \longrightarrow N$, where ϵ is a smooth positive function on N. Let B be the open unit ball in $\mathbb{R}^m$. Define $F : M \times B \longrightarrow N$ by $F(x, b) = r[f(x) + \epsilon(f(x))\dot{b}]$. Since r is a retraction onto M, $F(x, 0) = f(x)$. Also, both F and ∂F are submersions on $\{x\} \times B$ for a fixed x, being the composition of two submersions $b \mapsto f(x) + \epsilon(f(x))b$ and r. Since each point of $M \times B$, and each point of $\partial M \times B$, lies on a submanifold $\{x\} \times B$, both F and ∂F are submersions.

Therefore both F and ∂F are transverse to any boundaryless submanifold A of N, and hence both f_b and ∂f_b are transverse to A for all most all $b \in B$. Finally, f is homotopic to each such f_b by homotopy $H : M \times I \longrightarrow N$ given by $H(x, t) = F(x, tb)$. □

Theorem 6.2.12 (Transversality Extension Theorem). *Let $f : M \longrightarrow N$ be a smooth map, K a closed subset of M, and A a closed subset and a submanifold of N. Let both A and N be without boundary. Let $f \overline{\pitchfork} A$ on K and $\partial f \overline{\pitchfork} A$ on $K \cap \partial M$. Then there is a smooth map $g : M \longrightarrow N$ such that f is smoothly homotopic to g, $g \overline{\pitchfork} A$, $\partial g \overline{\pitchfork} A$, and $f = g$ on a neighbourhood of K.*

Proof. First show that $f \overline{\pitchfork} A$ on a neighbourhood of K by considering the following two cases.

(1) $x \in K$ and $x \notin f^{-1}(A)$, and

(2) $x \in K \cap f^{-1}(A)$.

In case (1), $M - f^{-1}(A)$ is an open neighbourhood of x (as A is closed) on which f is obviously transverse to A. In case (2), there is a neighbourhood U of $f(x)$ in N and a submersion $\phi : U \longrightarrow \mathbb{R}^k$ such that $f \overline{\pitchfork} A$ implies that $\phi \circ f$ is regular at x, and hence in a neighbourhood of x. Thus $f \overline{\pitchfork} A$ on a neighbourhood of every point of K, so $f \overline{\pitchfork} A$ on a neighbourhood U of K.

Next, construct a map $\alpha : M \longrightarrow [0, 1]$ such that $\alpha = 1$ on $M - U$, and $\alpha = 0$ on a neighbourhood of K. The construction of α may be seen easily by applying the Smooth Urysohn's Lemma (2.1.17) to the open neighbourhood $M - K$ of the closed set $M - U$. Define $\beta = \alpha^2$. Then $d\beta_x = 2\alpha(x) \cdot d\alpha_x$. Therefore $d\beta_x = 0$ whenever $\alpha(x) = 0$, that is, $\beta(x) = 0$. Now, the proof of Transversality Homotopy Theorem (6.2.11) gives a smooth map $F : M \times B \longrightarrow N$, B an open ball in some $\mathbb{R}^m$, such that $F(x, 0) = f(x)$, both F and ∂F are submersions, and, for fixed $x \in M$, the map $b \mapsto F(x, b)$ is a submersion $B \longrightarrow N$. Define a map $G : M \times B \longrightarrow N$ by $G(x, b) = F(x, \beta(x) \cdot b)$. Then $G \overline{\pitchfork} A$.

To see this, suppose that $(x, b) \in G^{-1}(A)$ and $\beta(x) \neq 0$. Then the map $B \longrightarrow N$ given by $b \mapsto G(x, b)$ is a submersion, because it is the composition of a diffeomorphism $b \mapsto \beta(x) \cdot b$ and a submersion $b \mapsto F(x, b)$. Therefore G is a submersion at (x, b), and hence $G \overline{\pitchfork} A$ at (x, b) in this case. Next, if $\beta(x) = 0$ the conclusion can be arrived at by computing $dG_{(x,b)}$ at a point

$(v, w) \in \tau(M)_x \times \tau(B)_b = \tau(M)_x \times \mathbb{R}^m$. Note that $G = F \circ H$, where $H : M \times B \longrightarrow M \times B$ is given by $H(x, b) = (x, \beta(x) \cdot b)$. Then

$$dH_{(x,b)}(v, w) = (v, \beta(x) \cdot w + d\beta_x(v) \cdot b) = (v, 0),$$

since $\beta(x) = 0$ and $d\beta_x = 0$, and we have

$$dG_{(x,b)}(v, w) = dF_{(x,0)}(v, 0) = df_x(v),$$

as $F \mid M \times 0 = f$. This implies that Image $dG_{(x,b)}$ = Image df_x. But, if $\beta(x) = 0$, then $\alpha(x) = 0$, so x belongs to the neighbourhood U of K, and therefore $f \overline{\pitchfork} A$ at x. This means that $G \overline{\pitchfork} A$, since Image $dG_{(x,b)}$ = Image df_x, and $G(x, b) = f(x)$ for $x \in U$.

Similarly, it can be shown that $\partial G \overline{\pitchfork} A$. Therefore, by Transversality Theorem, there is a $b \in B$ such that the map $g : M \longrightarrow N$ given by $g(x) = G(x, b)$, and the map ∂g are both transverse to A. Also, f is homotopic to g by homotopy $H : M \times I \longrightarrow N$ given by $H(x, t) = G(x, tb) = F(x, t\beta(x) \cdot b)$. Moreover, if x is in the neighbourhood U of K on which $\alpha = 0$, then $g(x) = G(x, b) = F(x, 0) = f(x)$. $\square$

Corollary 6.2.13. *If for a smooth map $f : M \longrightarrow N$, the restriction to boundary $\partial f : \partial M \longrightarrow N$ is transverse to A, where $\partial N = \emptyset$ and $\partial A = \emptyset$, then there is a smooth map $g : M \longrightarrow N$ homotopic to f such that $\partial f = \partial g$ and $g \overline{\pitchfork} A$.*

PROOF. This is a special case of the previous theorem, since the boundary ∂M is always closed in M. $\square$

6.3. Compact one-manifolds and Brouwer's theorem

We shall show that the only compact connected 1-manifolds are either the closed interval $[0, 1]$ or else the circle S^1, up to diffeomorphism.

Let M be a compact connected 1-manifold. We may suppose that M is a submanifold of $\mathbb{R}^3$, by the Whitney Embedding Theorem.

Recall that a subset C of M is a **parametrised curve** if it is the image of a diffeomorphism $\phi : I \longrightarrow C$, where I is an interval in $\mathbb{R}$ which may be open, or closed, or half-open (finite or infinite). We shall call ϕ a **parametrisation** in M. Its parametric equation is

$$\phi(t) = (\phi_1(t), \phi_2(t), \phi_3(t)),\ t \in I.$$

As t runs over I, $\phi(t)$ traces the curve C. The velocity vector or tangent vector of C at a point $\phi(t)$ is the derivative of $\phi(t)$ at t

$$\phi'(t) = (\phi_1'(t), \phi_2'(t), \phi_3'(t)).$$

As ϕ is a diffeomorphism, $\phi'(t)$ is never zero on I (if $\phi'(t) = 0$, then ϕ cannot be smoothly invertible, by the Inverse Function Theorem).

A **reparametrisation** ψ is obtained by a change of parameter $t = t(\theta)$ (which is a diffeomorphism $t : J \longrightarrow I$, where J is another interval) as $\psi(\theta) = \phi(t(\theta)),\ \theta \in J$. Since $dt/d\theta \neq 0$, either $dt/d\theta > 0$ or $dt/d\theta < 0$ on J, that is, t is strictly increasing or strictly decreasing, by Mean Value Theorem. Therefore, if $dt/d\theta > 0$ (resp. < 0) then t increases (resp. decreases) as θ increases, and $\phi(t)$ and $\psi(\theta)$ trace the same curve C in the same (resp. opposite) direction.

The arc length function $s = s(t) = \int_{t_0}^{t} \|\phi'(t)\| dt$ is a change of parameter, since it has continuous non-zero derivative which is the speed function $\|\phi'(t)\|$. The parametrisation $\psi(s) = \phi(t(s))$ is called a parametrisation by arc length. Its speed is

$$\|d\psi/ds\| = \|d\phi/dt\| \cdot |dt/ds| = 1.$$

Lemma 6.3.1. *If $\phi : I \longrightarrow C$ and $\psi : J \longrightarrow D$ are two parametrisations in M such that $C \cap D$ is connected, then $C \cup D$ is a parametrised curve.*

Proof. We may suppose that ϕ and ψ are parametrisations by arc length. Then $\psi^{-1} \circ \phi$ is a diffeomorphism of a relatively open subset I' of I onto a relatively open subset J' of J with constant derivative $+1$ or -1. This means, since $C \cap D$ is connected, that the graph of $\psi^{-1} \circ \phi$ is a straight line segment of slope ± 1 extending from an edge to another edge of the rectangle $I' \times J'$ (note that since the graph is closed and $\psi^{-1} \circ \phi$ is a local diffeomorphism, the graph cannot begin or end in the interior of the rectangle). If $y = \pm x + c$ is the equation of the line segment, then $\psi^{-1} \circ \phi$ may be extended to a diffeomorphism $\lambda : \mathbb{R} \longrightarrow \mathbb{R}$ given by $\lambda(s) = \pm s + c$. Now define $\sigma : I \cup \lambda^{-1}(J) \longrightarrow C \cup D$ by $\sigma(s) = \phi(s)$ if $s \in I$, and $\sigma(s) = \psi(\lambda(s))$ if $s \in \lambda^{-1}(J)$. It can be checked easily that σ is well-defined, and it is a diffeomorphism (use chain rule). $\square$

Theorem 6.3.2. *If M is a compact connected manifold of dimension one, then M is diffeomorphic either to S^1, or to* $[0, 1]$.

Proof. Let $\mathcal{S}$ be the family of all pairs (I, ϕ), where $\phi : I \longrightarrow \phi(I) \subseteq M$ is a parametrisation in M. Partially order $\mathcal{S}$ by the binary relation

$$(I, \phi) \leq (J, \psi) \text{ if and only if } I \subseteq J \text{ and } \phi = \psi \mid I.$$

Then any linearly ordered subset $(I_1, \phi_1) \leq (I_2, \phi_2) \leq \cdots$ of $\mathcal{S}$ is bounded by an element $(I, \phi) \in \mathcal{S}$, where I is the interval $\cup_i I_i$. and ϕ is the parametrisation given by $\phi \mid I_i = \phi_i$. Therefore by Zorn Lemma, $\mathcal{S}$ contains a maximal parametrised curve C in M and with parametrisation $\psi : I \longrightarrow C$. We may suppose by applying a change of variable, if necessary, that I is one of the intervals $(0, 1), [0, 1), (0, 1]$, or $[0, 1]$. Then, for any sequence $\{t_n\}$ in I converging to 0 (resp. 1), the sequence $\{\psi(t_n)\}$ in C converges to a unique point x_0 (resp. x_1). It follows that the closure of C is $\overline{C} = C \cup \{x_0, x_1\}$, and ψ extends to a smooth map $\widetilde{\psi} : [0, 1] \longrightarrow \overline{C}$ by $\widetilde{\psi}(0) = x_0$ and $\widetilde{\psi}_1(1) = x_1$, of course, $\widetilde{\psi} = \psi$ if x_0 and x_1 are already in C. We consider two cases $x_0 = x_1$, and $x_0 \neq x_1$.

If $x_0 = x_1$, set $\theta = 2\pi s$, where $0 \leq s \leq 1$, define a diffeomorphism $f : S^1 \longrightarrow M$ by

$$f(\cos\theta, \sin\theta) = \widetilde{\psi}(s).$$

Then f is onto, since $f(S^1)$ is compact and open in the connected space M. This proves the first part of the theorem.

If $x_0 \neq x_1$, then the parametrisation ψ must map I onto M, that is, C must be equal to M. If this is not true, then any point $x \in M - C$ would admit a parametrised curve neighbourhood D_x not intersecting C, otherwise C would extend to a larger parametrised curve by the above lemma violating the maximality of C. Then C and $\cup_{x\in M-C} D_x$ would give a separation of M, which is not possible since M is connected. □

Corollary 6.3.3. *The boundary of a compact one- manifold with boundary consists of an even number of points.*

◊ **Exercise 6.4.** Show that a non-compact connected manifold of dimension one is diffeomorphic to an interval.

Theorem 6.3.4. *If M is a compact manifold with boundary, then there is no retraction of M onto ∂M.*

Proof. If $f : M \longrightarrow \partial M$ is a retraction, then, by Sard's theorem, there is a point $x \in \partial M$ which is a regular value of f. Then $f^{-1}(x)$ is a submanifold V of M with boundary

$$\partial V = V \cap \partial M.$$

The codimension of V is the codimension of $\{x\}$, which is $n-1$ if $n = \dim M$. Therefore V is one-dimensional, and it is closed, and so compact. Then ∂V has an even number of points. But, since $f|\partial M = \mathrm{Id}$, ∂V consists of only one point x. This contradiction shows the non-existence of a smooth retraction f. □

Lemma 6.3.5. *There is no continuous retraction of D^n onto S^{n-1}.*

Proof. Let $f : D^n \longrightarrow S^{n-1}$ be a continuous retraction. Consider the continuous map $g : D^n \longrightarrow D^n$ given by

$$\begin{aligned} g(x) &= 2x, && \text{if} \quad 0 \leq \|x\| \leq \tfrac{1}{2}, \\ &= x/\|x\|, && \text{if} \quad \tfrac{1}{2} \leq \|x\| \leq 1. \end{aligned}$$

Then $h = f \circ g$ is also a retraction, and it is smooth on a closed neighbourhood K of S^{n-1}. The map h can be approximated by a smooth map $k : D^n \longrightarrow S^{n-1}$ which agrees with h on K. This means that k is a smooth retraction, which is in contradiction with the above theorem. □

Theorem 6.3.6 (Brouwer fixed-point theorem). *Any continuous map f of D^n to itself has a fixed point.*

Proof. Such a continuous map f without fixed point gives rise to a retraction $g : D^n \longrightarrow S^{n-1}$ which sends $x \in D^n$ to a point where the directed line segment from $f(x)$ to x hits the boundary S^{n-1}. We shall show that this map g is continuous. This will be in contradiction with the above lemma, and our proof will be complete.

Since x lies in between $f(x)$ and $g(x)$ on a line segment, we may write

$$g(x) = rx + (1-r)f(x),$$

where $r \geq 1$. Then g will be continuous, if r is a continuous function of x. Now, since $\|g(x)\| = 1$, the above relation gives a quadratic equation in r

$$r^2\|x - f(x)\|^2 + 2r(x \cdot f(x) - \|f(x)\|^2) + \|f(x)\|^2 - 1 = 0.$$

Solving the quadratic equation, the unique positive root r can be expressed in terms of continuous functions of x. Therefore g is a continuous map, and the proof is complete. □

6.4. Boundary and preimage orientations

Here we describe two standard orientations, namely, boundary orientation and preimage orientation. They will be in force throughout the rest of the chapter.

Boundary orientation. Suppose that $\dim M = n \geq 1$. Then, since

$$\operatorname{codim}\ \partial M = 1,$$

at each point $x \in \partial M$, there are exactly two unit vectors in $\tau(M)_x$ which are orthogonal to $\tau(\partial M)_x$. One of them is inward pointing and the other is outward pointing. Here is their precise definition. At the origin $0 \in \mathbb{R}^n_+$, the unit vector $e_1 = (1, 0, \ldots, 0)$ is the inward pointing normal vector to $\partial\mathbb{R}^n_+$, and $-e_1$ is the outward pointing normal vector. If $\phi : U \longrightarrow \mathbb{R}^n_+$ is a coordinate system with $\phi(x) = 0$, then $(d\phi_x)^{-1}(e_1)$ is the inward pointing normal vector to ∂M at x, and its negative is outward pointing. This distinction between inward and outward directions does not depend on the choice of ϕ. Because if ϕ' is another compatible coordinate system in the oriented atlas of M, then the isomorphism $d(\phi'\phi^{-1})_0$ maps the half space $\mathbb{R}^n_+$ onto itself preserving e_1 (Lemma 1.8.1). The boundary orientation on ∂M is defined as follows. If $\alpha = \{\alpha_1, \ldots, \alpha_{n-1}\}$ is an ordered basis of $\tau(\partial M)_x$, then sign α is the sign of the ordered basis $\{\nu_x, \alpha\} = \{\nu_x, \alpha_1, \ldots, \alpha_{n-1}\}$ of $\tau(M)_x$, where ν_x is the outward normal vector at x. It can be seen easily that this defines an orientation on ∂M. If $\dim M = 1$, then the sign of the orientation of the zero-dimensional vector space $\tau(\partial M)_x$ is the sign of the basis $\{\nu_x\}$ of $\tau(M)_x = \mathbb{R}$, $x \in \partial M$.

Remark 6.4.1. The orientation of $\tau(\partial M)_x$ is obtained from the direct sum decomposition $\tau(M)_x = \langle\nu_x\rangle \oplus \tau(\partial M)_x$. It is not necessary to take the outward unit normal vector ν_x in the definition of the boundary orientation. We might just as well replace ν_x by any outward pointing vector, which is a vector like $r\nu_x + w$, where $r > 0$, and $w \in \tau(\partial M)_x$, and get the same orientation of ∂M.

Example 6.4.2. The standard orientation of the closed unit disk D^2 induces counter-clockwise orientation on the boundary circle S^1.

Example 6.4.3. The standard orientation of the unit interval $0 \le t \le 1$ induces the orientation -1 and $+1$ at the end points 0 and 1 respectively.

Example 6.4.4. Let M be an oriented manifold without boundary, and I be the unit interval $[0, 1]$ with the standard orientation. Then the product $M \times I$ has two boundary components $M_0 = M \times \{0\}$ and $M_1 = M \times \{1\}$, and each of them is diffeomorphic to M. At a point $(x, 0) \in M_0$, the outward normal vector $\nu_{(x,0)}$ is $(0, -1) \in \tau(M)_x \times \tau(I)_0$, and at a point $(x, 1) \in M_1$, the outward normal vector $\nu_{(x,1)}$ is $(0, 1) \in \tau(M)_x \times \tau(I)_1$. Therefore if α is an ordered basis of $\tau(M)_x$, the signs of the induced orientation of M_0 and M_1 are given respectively as

$$\text{sign } (\nu_{(x,0)}, \alpha) = \text{ sign } \alpha \cdot \text{ sign } (-1) = - \text{ sign } \alpha,$$

$$\text{sign } (\nu_{(x,1)}, \alpha) = \text{ sign } \alpha \cdot \text{ sign } (1) = \text{ sign } \alpha.$$

Thus the induced orientations on M_0 and M_1 are opposite, and we may write

$$\partial(M \times I) = M_1 \cup (-M_0).$$

Preimage orientation. Let M, N, and A be oriented manifolds, where A is a submanifold of N, and A and N are without boundary. Let $f : M \longrightarrow N$ be a smooth map with $f \pitchfork A$ and $\partial f \pitchfork A$.transverse to A. Then the manifold $B = f^{-1}(A)$ receives a natural orientation from the orientations on M, N, and A. This orientation on B, which is called the preimage orientation induced by f, is defined as follows.

Let $x \in M$ and $y = f(x) \in A$. Then, $\tau(B)_x = df_x^{-1}(\tau(A)_y)$ is a subspace of the vector space $\tau(M)_x$. Let $\nu(B)_x$ be the orthogonal complement of $\tau(B)_x$ in $\tau(M)_x$ so that

$$\tau(M)_x = \nu(B)_x \oplus \tau(B)_x. \tag{1}$$

Then, $df_x(\tau(M)_x) = df_x(\nu(B)_x) + df_x(\tau(B)_x)$. Substituting this in the transversality condition $\tau(N)_y = df_x(\tau(M)_x) + \tau(A)_y$, we get

$$\tau(N)_y = df_x(\nu(B)_x) \oplus \tau(A)_y. \tag{2}$$

The sum is direct, because the dimensions of both sides are equal (note that, since Ker $df_x \subset \tau(B)_x$, df_x maps $\nu(B)_x$ isomorphically onto its image). The orientations of $\tau(N)_y$ and $\tau(A)_y$ induce an orientation of $df_x(\nu(B)_x)$ via the direct sum decomposition (2), this induces an orientation of $\nu(B)_x$ via the isomorphism df_x, finally, the orientations of $\tau(M)_x$ and $\nu(B)_x$ induce an orientation of $\tau(B)_x$ via the decomposition (1). In this way, we may define orientation on each tangent space $\tau(B)_x$ smoothly, because df_x varies smoothly with x.

With this knowledge, we can summarise the rule for finding the preimage orientation as follows. If α, β, and γ are the orientations of M, A, and N respectively, then the preimage orientation ω of B is given by

$$\text{sign } \omega = \frac{\text{sign } \alpha \cdot \text{sign } \beta}{\text{sign } \gamma}. \tag{3}$$

Remark 6.4.5. Note that in this definition of the preimage orientation, the orthogonality of the complement of $\tau(B)_x$ is unnecessary. In fact, for any complement P of $\tau(B)_x$ in $\tau(M)_x$, the equations $\tau(M)_x = P \oplus \tau(B)_x$ and $df_x(P) \oplus \tau(A)_y = \tau(N)_y$ will define the same preimage orientation on B.

Recall that if M is a manifold with boundary ∂M, and if $f : M \longrightarrow N$ and $\partial f : \partial M \longrightarrow N$ are both transverse to a submanifold A of N, where both N and A are without boundary, then $B = f^{-1}(A)$ is a manifold with boundary

$$\partial B = f^{-1}(A) \cap \partial M.$$

Then ∂B receives two orientations, one as the preimage of A under $\partial f : \partial M \longrightarrow N$, and the other as the boundary of B. The follows lemma shows that these two orientations are the same if codim A is even.

Lemma 6.4.6. $\partial(f^{-1}(A)) = (-1)^{\text{codim } A}(\partial f)^{-1}(A)$.

Proof. At any point $x \in B$, $\tau(\partial B)_x$ is a subspace of $\tau(\partial M)_x$. Let P be a complement of $\tau(\partial B)_x$ in $\tau(\partial M)_x$ so that

$$\tau(\partial M)_x = P \oplus \tau(\partial B)_x \tag{4}$$

Therefore, since $\tau(\partial B)_x = \tau(B)_x \cap \tau(\partial M)_x$, $P \cap \tau(B)_x = 0$. This means that P is also a complement of $\tau(B)_x$ in $\tau(M)_x$, and we have the direct sum decomposition

$$\tau(M)_x = P \oplus \tau(B)_x, \tag{5}$$

since $\dim P + \dim \tau(B)_x = \dim \tau(M)_x$. Thus P is complementary to both $\tau(B)_x$ and $\tau(\partial B)_x$. Now, df_x and $d(\partial f)_x$ agree on P, because $P \subset \tau(\partial M)_x$. Therefore P receives the same orientation by the maps df_x and $d(\partial f)_x$, via the decomposition

$$\tau(N)_{f(x)} = df_x(P) \oplus \tau(A)_{f(x)}.$$

Then decomposition (5) (resp. (4)) defines the preimage orientation of B (resp. ∂B) induced by f (resp. ∂f) (see Remark 6.4.5).

The boundary orientation of ∂B induced from the orientation of B is defined by the decomposition $\tau(B)_x = \langle \nu_x \rangle \oplus \tau(\partial B)_x$, where ν_x is the outward normal to ∂B in B, and $\langle \nu_x \rangle$ denotes the one-dimensional space spanned by it, oriented so that $\{v_x\}$ is a positively oriented basis. This v_x may not be orthogonal to $\tau(\partial M)_x$. But we may suppose that v_x is an outward pointing vector (see Remark 6.4.1) so that the orientations of $\tau(\partial M)_x$ and $\tau(M)_x$ are related by the direct sum relation $\tau(M)_x = \langle v_x \rangle \oplus \tau(\partial M)_x$. Substituting the preimage orientations of B and ∂B into this, we obtain

$$P \oplus \tau(B)_x = \langle \nu_x \rangle \oplus P \oplus \tau(\partial B)_x.$$

Then, further imposition of the boundary orientation of ∂B yields

$$P \oplus \langle \nu_x \rangle \oplus \tau(\partial B)_x = \langle \nu_x \rangle \oplus P \oplus \tau(\partial B)_x = (-1)^k P \oplus \langle \nu_x \rangle \oplus \tau(\partial B)_x,$$

where the last equality is obtained by making $k = \dim P$ number of transpositions to move ν_x from left to right. Therefore, we have $\tau(\partial B)_x = (-1)^k \tau(\partial B)_x$, where on the left hand side $\tau(\partial B)_x$ has the boundary orientation and on the right hand side it has the preimage orientation. Since $\dim P = \text{codim } B = \text{codim } A$, this completes the proof. $\square$

6.5. Intersection numbers, and Degrees of maps

Let M, N, and A be oriented manifolds without boundary, where M is compact, and A is a closed submanifold of N, such that

$$\dim M + \dim A = \dim N. \tag{6}$$

Now, if $f : M \longrightarrow N$ is a smooth map transverse to A, then $f^{-1}(A)$ is a finite set of points, since it is compact and its dimension is zero (see that $\text{codim } f^{-1}(A) = \text{codim } A = \dim M$ by the dimension condition (6)). If $x \in f^{-1}(A)$, then the transversality condition at x and the dimension condition (6) imply that we have a direct sum decomposition

$$df_x(\tau(M)_x \oplus \tau(A)_{f(x)} = \tau(N)_{f(x)}, \tag{7}$$

and df_x is an isomorphism onto its image.

Definition 6.5.1. The preimage orientation at a point $x \in f^{-1}(A)$ is called the **orientation number** of f, and is denoted by $n(f, x)$. This number is $+1$ if in the direct sum decomposition (7), the orientation on $df_x(\tau(M)_x)$ plus the orientation on $\tau(A)_{f(x)}$ (in this order) is the prescribed orientation on $\tau(N)_{f(x)}$, and it is -1 otherwise.

An alternative definition of $n(f, x)$ may be obtained in the following way. The composition $\pi \circ df_x$

$$\tau(M)_x \xrightarrow{df_x} \tau(N)_{f(x)} \xrightarrow{\pi} \tau(N)_{f(x)}/\tau(A)_{f(x)},$$

where π is the canonical projection, is an isomorphism. Then $n(f, x)$ is $+1$ or -1 according to whether this isomorphism $\pi \circ df_x$ preserves or reverses orientation.

The **intersection number** $I(f, A)$ is defined to be the sum of the orientation numbers $n(f, x)$ over all $x \in f^{-1}(A)$.

If M has boundary, and $f : M \longrightarrow N$ is any smooth map, then Theorem 6.2.11 says that there is a smooth map $g : M \longrightarrow N$ homotopic to f such that both g and ∂g are transverse to A. In this case the intersection number

$I(f, A)$ is defined to be the intersection number $I(g, A)$. That this definition in independent of the choice of g will be seen in Lemma 6.5.3 below.

◊ **Exercise 6.5.** Let $f : M \longrightarrow N$ and $g : M \longrightarrow (N - A)$ be smooth maps homotopic in N. Then show that $I(f, A) = 0$.

Theorem 6.5.2 (Extendability Theorem). *Let M, N, and A be oriented manifolds, where M has compact boundary, and both N and A are boundaryless. Let A be a closed subset and a submanifold of N such that*

$$\dim \partial M + \dim A = \dim N. \tag{8}$$

Let $f : \partial M \longrightarrow N$ be a smooth map which extends to a smooth map $g : M \longrightarrow N$ such that both g and ∂g are transverse to A. Then $I(f, A) = 0$.

Note that f may be a map from a component of ∂M into N.

PROOF. In view of the given dimension condition, $g^{-1}(A)$ is a compact oriented one-dimensional manifold with boundary $\partial(g^{-1}(A)) = g^{-1}(A) \cap \partial M = f^{-1}(A)$, which consists of pairs of points with orientation numbers $+1$ and -1. Consequently, $I(f, A) = 0$. □

Lemma 6.5.3. *Let M, N, and A be oriented boundaryless manifolds, where M is compact, and equation (8) holds. Let A be a closed subset and a submanifold of N. Let g_0 and g_1 be smooth maps from M into N which are homotopic and both transverse to A. Then $I(g_0, A) = I(g_1, A)$.*

PROOF. If $G : M \times I \longrightarrow N$ is a homotopy between g_0 and g_1, then $I(\partial G, A) = 0$ by the above lemma. Now, by Example 6.4.4, $\partial(M \times I) = M_1 - M_0$, where M_0 and M_1 are diffeomorphic copies of M, and ∂G agrees with g_0 and g_1 on M_0 and M_1 respectively. Therefore $(\partial G)^{-1}(A) = g_1^{-1}(A) - g_0^{-1}(A)$, and hence

$$I(g_1, A) - I(g_0, A) = I(\partial G, A) = 0.$$

□

In the special case when $\dim M = \dim N$, A is a point $y \in N$, and N is connected, the intersection number $I(f, \{y\})$ is called the **degree of** f, and denoted by $\deg f$. Here y is a regular value of f, and for $x \in f^{-1}(y)$, $n(f, x)$ is $+1$ or -1 according as df_x preserves or reverses orientation. Then

$$\deg\, f = \sum_{x \in f^{-1}(y)} n(f, x).$$

In general, we may allow y to be any point of N (not necessarily a regular value of f), and define $\deg\, f$ by the above formula. This is justified by the following lemma.

Lemma 6.5.4. *The number* $\deg\, f$ *is the same for all* $y \in N$.

PROOF. Given y, we may find a smooth map $g : M \longrightarrow N$ such that y is a regular value of g, and g is homotopic to f. Then $g^{-1}(y)$ is a finite set $\{x_1, \ldots, x_k\}$, say. Because $\dim M = \dim N$, we may invoke the local submersion theorem to find disjoint open neighbourhoods U_i of x_i such that g maps each U_i diffeomorphically onto an open neighbourhood V of y, and $g^{-1}(V) = U_1 \cup \cdots \cup U_k$ (disjoint union). Then, for any $z \in V$, $g^{-1}(z)$ consists of k points, and, for all x' in an U_i, the orientation number $n(g, x')$ is the same. This means that the map $N \longrightarrow \mathbb{Z}$ given by $y' \mapsto I(g, \{y'\})$ is locally constant. Since N is connected, this map must be constant on the whole of N. $\square$

For example, the identity map of M has degree $+1$, and the anti-podal map $S^n \longrightarrow S^n$, sending x to $-x$, has degree $(-1)^{n+1}$. Note that the antipodal map is homotopic to Id if and only if n is odd. For $n = 1$, the homotopy $h_t : S^1 \longrightarrow S^1$ is given by $h_t(x, y) = (x', y')$, where $x^2 + y^2 = 1$, and $x' = x\cos\pi t - y\sin\pi t$, $y' = x\sin\pi t + y\cos\pi t$. The general case can be seen from this.

$\Diamond$ **Exercise 6.6.** Let M, N, and P be connected boundaryless manifolds of the same dimension. Then for smooth maps $f : M \longrightarrow N$ and $g : N \longrightarrow P$ show that $\deg(g \circ f) = \deg f \cdot \deg g$.

Lemma 6.5.5. *Represent the circle S^1 as the set of complex numbers z with $|z| = 1$, and let m be an integer. Then the map f of the circle S^1 onto itself, given by $z \mapsto z^m$, has degree m.*

PROOF. If $m > 0$ (resp. $m < 0$), the image point $f(z)$ moves around the circle in the counterclockwise (resp. clockwise) sense m times as z moves around the circle once in the counterclockwise sense. Therefore the inverse image $f^{-1}(z)$ of a point $z \in S^1$ contains $|m|$ number of points, unless $m = 0$. In terms of the coordinate systems arising from the exponential map $\mathbb{R} \longrightarrow S^1$ given by $\theta \mapsto \exp(i\theta)$. local representation of f is $\theta \mapsto m\theta$. Therefore f is regular everywhere if $m \neq 0$. If $m > 0$, f is orientation preserving, and $\deg f$ is the number of points in $f^{-1}(z)$ which is m. If $m < 0$, f is orientation reversing, and $\deg f$ is $-|m| = m$. If $m = 0$, then f is a constant map, and so $\deg f = 0$. $\square$

Thus, given any integer m, there is a smooth map $S^1 \longrightarrow S^1$ whose degree is m. This may be seen from the following example.

Example 6.5.6. Any smooth map $f : S^1 \longrightarrow S^1$ may be written as

$$f(\exp(i\theta)) = \exp(ig(\theta)),$$

where $g : \mathbb{R} \longrightarrow \mathbb{R}$ is a smooth map. Then, since θ and $\theta + 2\pi$ represent the same point of the circle, we must have $g(\theta + 2\pi) = g(\theta) + 2k\pi$ for all θ. where k is a fixed integer which is positive or negative according to whether f is orientation preserving or reversing. In particular, $g(2\pi) = g(0) + 2k\pi$. Thus as z moves round the circle once, the image $f(z)$ moves round the circle k times. Thus $\deg f = k$.

Theorem 6.5.7. *Two maps $f_0, f_1 : S^1 \longrightarrow S^1$ are homotopic if and only if they have the same degree.*

Proof. We already know that the condition is necessary. To prove the sufficiency, suppose that $\deg f_0 = \deg f_1$. Then, as in Example 6.5.6, we may write, for $j = 0, 1$, f_j as $f_j(\exp(i\theta)) = \exp(ig_j(\theta))$, where $g_j : \mathbb{R} \longrightarrow \mathbb{R}$ are smooth maps such that $g_j(\theta+2\pi) = g_j(\theta)+2k\pi$. Then g_0 and g_1 are homotopic by homotopy

$$g_t = (1-t)g_0 + tg_1, \quad 0 \leq t \leq 1.$$

Then $g_t(\theta + 2\pi) = g_t(\theta) + 2k\pi$ and so $\exp(ig_t(\theta))$ is a homotopy between f_0 and f_1. $\square$

This theorem is a special case of the Hopf degree theorem. One-half of the Hopf degree theorem is contained in the extendability theorem, which in terms of the degree reads as follows.

Theorem 6.5.8. *Let M be a compact oriented manifold which is the boundary of a manifold P, and N another oriented boundaryless manifold of the same dimension as M. If a smooth map $f : M \longrightarrow N$ extends to a smooth map on all of P, then* $\deg f = 0$.

The Hopf degree theorem says that $f : M \longrightarrow S^n$ extends to a smooth map on all of P if and only if $\deg f = 0$. This we will prove in the next section. Meanwhile, let us look at a simple application of Theorem 6.5.8.

Let $p(z) = z^m + a_1 z^{m-1} + \cdots + a_m$ be a monic polynomial with complex coefficients. Consider a family of polynomials of $p_t(z)$, $0 \leq t \leq 1$, given by

$$p_t(z) = tp(z) + (1-t)z^m = z^m + t(a_1 z^{m-1} + \cdots + a_m).$$

Since

$$\lim_{z\to\infty} \frac{p_t(z)}{z^m} = 1,$$

there is a closed ball B of sufficiently large radius such that none of the polynomials $p_t(z)$ vanish on ∂B. Therefore the homotopy $p_t/|p_t| : \partial B \longrightarrow S^1$ is defined for all t, and so $\deg(p/|p|) = \deg(p_0/|p_0|) = m$, by Lemma 6.5.5. This implies that p has at least one zero inside B. Otherwise, the map $p/|p| : \partial B \longrightarrow S^1$ will extend on all of B, and $\deg(p/|p|)$ will be zero, by the extendability property. This proves

Theorem 6.5.9 (Fundamental Theorem of Algebra). *Any non-constant complex polynomial has a zero.*

Let M, N, and A be oriented boundaryless manifolds, where M is a compact submanifold of N. Let A be a closed subset and a submanifold of N of complementary dimension (so that (6) is satisfied). Let $i : M \longrightarrow N$ be the inclusion map, and $M \pitchfork A$. Then the **intersection number** $I(M, A; N)$ is defined to be the number $I(i, A)$. This is the sum of the orientation numbers $+1$ and -1 of the points $x \in M \cap A$, where the number for x is $+1$ if the orientations of M and A (in this order) at x give the orientation of N at x, otherwise the number is -1.

We often write $I(M, A; N)$ simply as $I(M, A)$, when it is not necessary to mention the ambient manifold N.

◊ **Exercise 6.7.** Show that if M and A are compact submanifolds of N of dimensions n and k respectively such that $\dim N = n + k$, then

$$I(M, A) = (-1)^{nk} I(A, M).$$

In particular, if $\dim N = 2 \dim M$, then the self-intersection number $I(M, M)$ is defined. Moreover, if $\dim M$ is odd, then $I(M, M) = 0$.

Definition 6.5.10. Let M be a compact connected manifold without boundary, $\pi : \tau(M) \longrightarrow M$ be the tangent bundle of M, and $i : M \longrightarrow \tau(M)$ be the zero section. Identify M with the zero section. Then the **Euler characteristic** of M, denoted by $\chi(M)$, is defined to be the number $I(M, M) = I(M, M; \tau(M))$.

An alternative definition of the Euler characteristic in terms of the homology of the manifold M appears in Chapter 9 (see Theorem 9.3.12).

Definition 6.5.11. Let $g : M \longrightarrow \tau(M)$ be a smooth vector field transverse to the zero-section $i(M)$. Let $x \in M$ be a zero of g (i.e. $x \in g^{-1}(i(M))$). Then the orientation number $n(g, x)$ is called the **index of the vector field** g **at** x, and is denoted by $\operatorname{Ind}_x g$.

A more general definition of the index appears in Definition 6.6.1.

Lemma 6.5.12. *A compact oriented manifold M without boundary admits a vector field $g : M \longrightarrow \tau(M)$ transverse to the zero-section $i(M)$ such that*

$$\chi(M) = \sum_{x \in g^{-1}(i(M))} \operatorname{Ind}_x g.$$

Proof. Approximate the zero section $i : M \longrightarrow \tau(M)$ by a smooth map

$$f : M \longrightarrow \tau(M)$$

homotopic to i and transverse to the zero section $i(M)$ (Theorem 6.2.11). If the approximation is sufficiently small, then the map $\pi \circ f : M \longrightarrow M$ is a diffeomorphism homotopic to Id (Theorem 2.7.2(5)), and the map $g = f \circ (\pi \circ f)^{-1} : M \longrightarrow \tau(M)$ is a smooth section transverse to the zero section $i(M)$ (Exercise 6.3 in p. 176), and homotopic to i. Therefore

$$\chi(M) = I(i, M) = I(g, M) = \sum_{x \in g^{-1}(i(M))} \operatorname{Ind}_x g.$$

□

Theorem 6.5.13. *If M admits a nowhere vanishing vector field, then its Euler characteristic $\chi(M) = 0$.*

PROOF. If $g : M \longrightarrow \tau(M)$ is a nowhere vanishing vector field, then g is transverse to the zero-section $i(M)$. Also g is homotopic to the zero section i by the homotopy given by $h_t(x) = (1-t)g(x) + t \cdot i(x)$. Therefore

$$\chi(M) = I(i, M) = I(g, M) = 0.$$

$\square$

We may compute $\chi(M)$ in another way. Let $g : M \longrightarrow \tau(M)$ be the vector field transverse to the zero-section, as constructed in Lemma 6.5.12. Let $x_1, \ldots, x_r \in M$ be the zeros of g. Let $\phi_i : \pi^{-1}(U_i) \longrightarrow U_i \times \mathbb{R}^n$ be a VB-chart of $\tau(M)$ over open neighbourhoods U_i of x_i in M so that ϕ_i is orientation preserving. Then the composition $h_i = (\mathrm{proj}) \circ \phi_i \circ g$

$$U_i \xrightarrow{g} \pi^{-1}(U_i) \xrightarrow{\phi_i} U_i \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

has a regular value at $0 \in \mathbb{R}^n$ and $x_i \in h_i^{-1}(0)$ (Lemma 6.2.4). Since ϕ_i is orientation preserving, the orientation number $n(h_i, x_i)$ is the same as the sign of the isomorphism

$$\tau(U_i)_{x_i} \xrightarrow{d\phi_i \circ dg} \tau(U_i \times \mathbb{R}^n)_{y_i} \xrightarrow{\mathrm{proj}} \frac{\tau(U_i \times \mathbb{R}^n)_{y_i}}{\tau(U_i \times \{0\})_{y_i}}, \ y_i = \phi_i(g(x_i)),$$

which is the same as $n(g, x_i)$ (see the alternative definition of the orientation number in Definition 6.5.1). Therefore the Euler characteristic of M is

$$\chi(M) = \sum_i n(h_i, x_i).$$

The method of computation of $\chi(M)$ may be summarised as follows. Take a smooth vector field $f : M \longrightarrow \tau(M)$ transverse to the zero section. At each zero x_i of f, take a coordinate chart $\phi_i : U_i \longrightarrow \mathbb{R}^n$ which preserves orientation. Then the local representation of f is given by the composition

$$\phi_i(U_i) \xrightarrow{\phi_i^{-1}} U_i \xrightarrow{f} \pi^{-1}(U_i) \xrightarrow{d\phi_i} \tau(\phi_i(U_i)).$$

This is a section of the trivial bundle on $\phi_i(U_i)$, and so it defines a smooth map $g_i : \phi_i(U_i) \longrightarrow \mathbb{R}^n$ with a regular value at 0. Then, if $d_i = n(g_i, x_i)$ is the index of the vector field f at x_i, the Euler characteristic of M is

$$\chi(M) = \sum_i d_i.$$

Theorem 6.5.14. $\chi(S^n) = 1 + (-1)^n$ *so that it is* 0 *if* n *is odd, and it is* 2 *if* n *is even.*

PROOF. Let P and $Q = -P$ be the north and the south pole of S^n. Let $U = S^n - P$ and $V = S^n - Q$, and $\phi : U \longrightarrow \mathbb{R}^n$ and $\psi : V \longrightarrow \mathbb{R}^n$ be stereographic projections from P and Q respectively (see Example 1.1.9). Consider the atlas $\{(U, \phi), (V, \psi')\}$, where $\psi' = -\psi$, of S^n. The coordinate changes $\psi' \circ \phi^{-1} = \phi \circ \psi'^{-1} : \mathbb{R}^n - \{0\} \longrightarrow \mathbb{R}^n - \{0\}$ is given by $x \mapsto -x/\|x\|^2$.

Define a section $\lambda : U \longrightarrow \pi^{-1}(U)$ by $\lambda(x) = (x, d\phi_x^{-1}(\phi(x))$, and a section $\mu : V \longrightarrow \pi^{-1}(V)$ by $\mu(x) = (x, d\psi_x'^{-1}(\psi'(x))$. These sections are compatible with respect to the above change of coordinates, and therefore they fit together smoothly to give a global section f of the tangent bundle (note that in the construction of f we have not used orientation of S^n). Moreover, f vanishes only at P and Q. In ϕ coordinates f corresponds to $x \mapsto x$ on $\phi(U)$, and in ψ' coordinates f corresponds to $x \mapsto -x$ on $\psi'(V)$. Since the identity map of $\mathbb{R}^n$ has degree 1, and the anti-podal map has degree $(-1)^n$, $\mathrm{Ind}_P\, f = 1$ and $\mathrm{Ind}_Q\, f = (-1)^n$. This proves the theorem. □

Corollary 6.5.15 (Hairy Ball Theorem). *Every vector field on S^{2n} vanishes somewhere.*

A graphic description of this result says that a hairy ball cannot be combed continuously.

Example 6.5.16. Every sphere S^n of odd dimension n admits a non-zero vector field f, where, for $p = (x_1, \cdots, x_{n+1}) \in S^n$, $f(p)$ is given by

$$f(p) = (-x_2, x_1, -x_3, x_4, \cdots, -x_{n+1}, x_n).$$

This vector is orthogonal to p, and so $f(p) \in \tau(S^n)_p$.

$\Diamond$ **Exercise 6.8.** If M and N are compact oriented manifolds without boundaries, then show that $\chi(M \times N) = \chi(M) \cdot \chi(N)$.

Theorem 6.5.17. *If M is an odd dimensional compact oriented manifold without boundary, then*

$$\chi(M) = 0.$$

The converse is false as may be seen when $M = S^1 \times S^1$.

PROOF. Let us compute $\chi(M)$ using a vector field f, and then using the vector field $-f$. The computations will give

$$\chi(M) = \sum_x \mathrm{Ind}_x\, f = \sum_x \mathrm{Ind}_x\, (-f),$$

where the summations are over all zeros x of f or $-f$. Now, if $\dim M = n$ and x is a zero of f, then

$$\mathrm{Ind}_x\, f = (-1)^n \mathrm{Ind}_x\, (-f).$$

This gives us the theorem, if n is odd. □

6.6. Hopf's degree theorem

The zeros of a vector field $g : M \longrightarrow \tau(M)$ constitute interesting object of study. If $g(x) \neq 0$, then g remains more or less constant in magnitude and direction in a neighbourhood of x. However, if $g(x) = 0$, then the direction may change drastically in a small neighbourhood of the zero x. It may circulate around a circle about x, it may have a sink at x, source at x, or x may be a saddle point; it may spiral into x or out of x; or the change may be even more complicated. Note that a sink (or source) at x is a pattern of vector field in which the directions are pointed towards (or away from) x. The directional variation of g around an isolated zero x is measured by the index $\text{Ind}\,_x g$ of g at x, which is defined as follows.

Definition 6.6.1. If g is a vector field on an n-manifold M, and $x \in \text{Int } M$ is an isolated zero of g (not necessarily a regular point), then $\text{Ind}_x\, g$ is defined as the degree of the map $h : \partial B \longrightarrow S^{n-1}$ given by $h(y) = g(y)/\|g(y)\|$, where ∂B is the boundary of an n-ball B about x.

By Theorem 6.5.8, the definition is independent of the radius of B (provided it is small). We shall see in Proposition 6.6.4 that the definition coincides with the earlier definition of index (Definition 6.5.11 when x is a regular point of g.

Thus in dimension 2, $\text{Ind}_x\, g$ is simply equal to the number of times $g(y)$ rotates completely around the circle S^1 as y moves once around ∂B in the counterclockwise direction; its sign is positive or negative according as the direction of g moves counterclockwise or clockwise. The following figure shows the diagrams of six vector fields with their indices at an isolated zero.

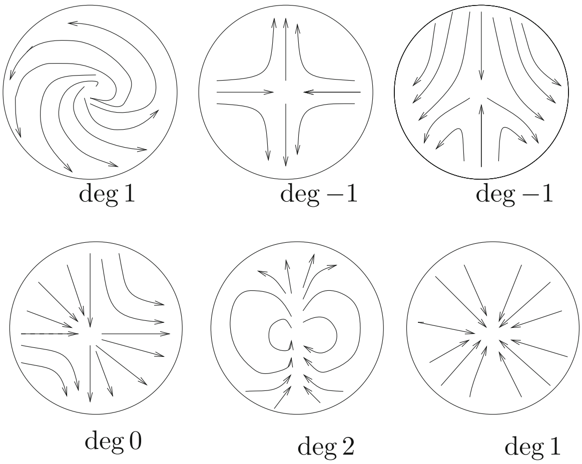


FIGURE 6.1

We shall first prove a homogeneity theorem. For this purpose, we need to know that a diffeotopy is a smooth homotopy $F : M \times I \longrightarrow N$ such that each F_t is a diffeomorphism of M, and also that a diffeotopy F is compactly supported, if each F_t is Id outside a fixed compact subset of M.

Theorem 6.6.2 (Homogeneity Theorem). *If y and z are two points of a connected boundaryless manifold M, then there is a diffeomorphism $h : M \longrightarrow M$ such that $h(y) = z$, and h is diffeotopic to* Id *by a compactly supported diffeotopy.*

PROOF. We call two points y and z of M equivalent if the conditions of the theorem are satisfied for them. Clearly, this defines an equivalence relation on M. Since M is connected, the proof will be complete if we can show that each equivalence class is an open set.

Since each point of M has an open neighbourhood diffeomorphic to $\mathbb{R}^n$, it is enough to show that every point of $\mathbb{R}^n$ sufficiently close to 0 is equivalent to 0. To this end, we shall construct a diffeotopy h_t of $\mathbb{R}^n$ such that $h_0 = \mathrm{Id}$, each h_t is Id outside some specified open n-ball around 0, and $h_1(0)$ is any specified point sufficiently close to 0.

We can construct a smooth function $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}$ such that $\phi(x) > 0$ if $\|x\| < \epsilon$, and $\phi(x) = 0$ if $\|x\| \geq \epsilon$, for any given $\epsilon > 0$ (it may be defined by $\phi(x) = \mathcal{B}(1 - \|x\|^2/\epsilon)$, where $\mathcal{B}$ is a bump function). Then for any $\overline{x} \in \mathbb{R}^n$, the system of differential equations

$$\frac{dx_i}{dt} = c_i\phi(x_1, \ldots, x_n), \quad i = 1, \ldots, n, \ c_i = \text{ any constant },$$

has a unique solution $x = x(t)$ with $x(0) = \overline{x}$ by Picard's theorem (Theorem 3.2.10). The solution is defined for all $t \in \mathbb{R}$, because ϕ has compact support. Then

$$h_t(\overline{x}) = x(t) = \overline{x} + c\int_0^t \phi(x(s))ds,$$

where $c = (c_1, \ldots, c_n)$, gives a one-parameter group of diffeomorphisms of $\mathbb{R}^n$, which is the required diffeotopy. Then $h_0 = \mathrm{Id}$, and

$$h_1(0) = c\int_0^1 \phi(x(t))dt.$$

We may therefore choose c_i so that $h_1(0)$ becomes a prescribed point sufficiently close to 0. □

Corollary 6.6.3. *If M is a connected manifold without boundary of dimension > 1, and $y_1, \ldots, y_r$ and $z_1, \ldots, z_r$ are two sets of distinct points in M, then there is a diffeomorphism $h : M \longrightarrow M$ diffeotopic to* Id *by a compactly supported diffeotopy such that $h(y_i) = z_i$, $i = 1, \ldots, r$.*

PROOF. We use inductive arguments. The case $r = 1$ is the above theorem. Assuming the result for $r - 1$, we construct a compactly supported diffeotopy h'_t of the manifold $M - \{y_r, z_r\}$ (this is connected since $\dim M > 1$) such that $h'_0 = \mathrm{Id}$ and $h'_1(y_i) = z_i$, for $i < r$. The diffeotopy is Id on some neighbourhoods of y_r and z_r, and so it extends to a diffeomorphism of M which fixes y_r and z_r. Then, we apply the theorem to the manifold $M - \{y_1, \ldots, y_{r-1}, z_1, \ldots, z_{r-1}\}$ to obtain a compactly supported diffeotopy h''_t such that $h''_0 = \mathrm{Id}$, $h''_1(y_r) = z_r$, and each h''_t fixes the points y_i and z_i for $i < r$. Then $h''_1 \circ h'_1$ is the required diffeomorphism. □

Proposition 6.6.4. *Let $f : U \longrightarrow \mathbb{R}^n$ be a vector field on an open set U of $\mathbb{R}^n$ with a regular value at $0 \in \mathbb{R}^n$, and an isolated zero at $x \in U$. Let B be a* *closed n-ball in U with centre at $x \in f^{-1}(0)$, and $g : \partial B \longrightarrow S^{n-1}$ be the map $g(y) = f(y)/\|f(y)\|$. Then the index* $\operatorname{Ind}_x f$ *of f at x is equal to* $\deg g$. (*Here ∂B has the boundary orientation of B*).

Proof. We may suppose that $x = 0$ so that $f(0) = 0$. Define a homotopy $F : U \times [0,1] \longrightarrow \mathbb{R}^n$ by

$$\begin{aligned} F(y,t) &= df_0(y) \quad \text{if } t = 0, \\ &= f(ty)/t \quad \text{if } 0 < t \leq 1. \end{aligned}$$

This is well-defined, since $df_0(y) = \lim_{t\to 0} f(ty)/t$. Let $y = (y_1, \ldots, y_n)$. Then we may write f in a neighbourhood of 0 (which we may suppose to be U itself) as

$$f(y) = y_1 g_1(y) + \cdots + y_n g_n(y),$$

where $g_1, \ldots, g_n$ are suitable smooth functions $U \longrightarrow \mathbb{R}^n$ such that $g_i(0) = \partial f/\partial y_i(0)$, $1 \leq i \leq n$, (Lemma 3.1.3). Therefore

$$F(y,t) = y_1 g_1(ty) + \cdots + y_n g_n(ty), \text{ for all } t \in [0,1],$$

and F is a smooth map. Next, define $G : \partial B \times [0,1] \longrightarrow S^{n-1}$ by

$$G(y,t) = \frac{F(y,t)}{\|F(y,t)\|}.$$

This is well-defined, since $F(y,t) = 0$ only at $y = 0$ (which is outside of ∂B), and it is a smooth homotopy from G_0 to $G_1 = g$. Therefore $\deg G_0 = \deg g$.

We shall show that $\operatorname{Ind}_0 F_0 = \operatorname{Ind}_0 f$, and then $\operatorname{Ind}_0 F_0 = \deg G_0$. This will complete our proof. Each of the results follow from the definition of the index. The first one follows, because $F_0 = df_0$ is a linear map, and so $(dF_0)_0 = df_0$. The second one may be seen in the following way. Since $F_0 = df_0$ is a linear isomorphism, it is homotopic through linear isomorphisms to an orthogonal transformation, by Gram-Schmidt orthonormalisation process (§5.6). So we may suppose that F_0 itself is an orthogonal transformation. Then $H : \partial B \times I \longrightarrow R^n$ given by

$$H(y,t) = \left(1 - t + \frac{t}{\|F_0(y)\|}\right) \cdot F_0(y)$$

is a homotopy from $F_0|\partial B$ and G_0. Therefore we may suppose that $F_0|\partial B = G_0$. This implies the second result, because dG_0 at a regular point (and there is only one such point over a regular value of G_0) is the orthogonal transformation F_0, which is dF_0 at any point of U (and, in particular, at 0). □

Proposition 6.6.5. *Let B be a closed n-ball in $\mathbb{R}^n$, $f : B \longrightarrow \mathbb{R}^n$ be a smooth map with a regular value at $0 \in \mathbb{R}^n$ so that $f^{-1}(0) \cap \partial B = \emptyset$, and $g : \partial B \longrightarrow S^{n-1}$ be the map $g(y) = f(y)/\|f(y)\|$. Then*

$$\deg g = \sum_{x \in f^{-1}(0)} \operatorname{Ind}_x f.$$

Proof. Let $f^{-1}(0) = \{x_1, \ldots, x_r\}$, and $B_1, \ldots, B_r$ be disjoint closed n-balls in Int B with centres at $x_1, \ldots, x_r$ respectively. Let $B_0 = B - \cup_i \text{Int } B_i$. Then $\partial B_0 = \partial B \cup_{i=1}^r \partial B_i$, and $f(B_0) \subset \mathbb{R}^n - \{0\}$. Define $h : B_0 \longrightarrow S^{n-1}$ by $h(y) = f(y)/\|f(y)\|$. Then

$$\deg(h|\partial B_0) = \deg(h|\partial B) - \sum_{i=1}^r \deg(h|\partial B_i).$$

The negative sign is due to the fact that the boundary orientation on each ∂B_i is opposite to that of ∂B. Now, $\deg(h|\partial B_0) = 0$ (as $h|\partial B_0$ extends to B_0), $\deg(h|\partial B) = \deg(g|\partial B)$ (as $h = g$ on ∂B), and $\deg(h|\partial B_i) = \text{Ind}_{x_i} f$ by Proposition 6.6.3. Therefore we have the result. □

Theorem 6.6.6. *Any smooth map $f : S^n \longrightarrow S^n$ of degree zero is homotopic to a constant map.*

The proof will be given after the following two lemmas.

Lemma 6.6.7. *If B is a closed n-ball in $\mathbb{R}^n$, M is a manifold without boundary, and $f : (\mathbb{R}^n - \text{Int } B) \longrightarrow M$ is a smooth map such that $f|\partial B$ is homotopic to a constant map, then f extends to a smooth map $\mathbb{R}^n \longrightarrow M$.*

Proof. Let B have its centre at $0 \in \mathbb{R}^n$, and $g_t : \partial B \longrightarrow M$ be a homotopy such that $g_0 =$ constant , and $g_1 = f|\partial B$. Define an extension $h : \mathbb{R}^n \longrightarrow M$ of f by

$$\begin{aligned} h(tx) \quad &= g_t(x), \quad \text{if } \; x \in \partial B, \; 0 \leq t \leq 1, \\ &= f(tx), \quad \text{if } \; x \in \partial B, \; t \geq 1. \end{aligned}$$

The map h may be approximated by a smooth map which agrees with f on the closed set $\mathbb{R}^n - \text{Int } B$, and is homotopic to h, by Proposition 6.1.5. □

Lemma 6.6.8. *Suppose that Theorem* 6.6.6 *is true for $n-1$. Then, for a smooth map $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ with a regular value at $0 \in \mathbb{R}^n$ such that $f^{-1}(0)$ is a finite set, and*

$$\sum_{x \in f^{-1}(0)} \text{Ind}_x f = 0,$$

there is a smooth map $\mathbb{R}^n \longrightarrow \mathbb{R}^n - \{0\}$ which agrees with f outside a compact set.

Proof. Let B be a closed n-ball in $\mathbb{R}^n$ with centre at 0 containing $f^{-1}(0)$ in its interior, and $g : \partial B \longrightarrow S^{n-1}$ be the map given by $g(y) = f(y)/\|f(y)\|$. Then, by Proposition 6.6.4 and the hypotheses of the present lemma,

$$\deg g = \sum_{x \in f^{-1}(0)} \text{Ind}_x f = 0,$$

and g is homotopic to a constant map. The last assertion follows, because S^{n-1} is diffeomorphic to ∂B. Now, the map

$$\partial B \xrightarrow{g} S^{n-1} \hookrightarrow \mathbb{R}^n - \{0\}$$

is homotopic to the map $f|\partial B : \partial B \longrightarrow \mathbb{R}^n - \{0\}$ by the homotopy

$$F : \partial B \times [0,1] \longrightarrow \mathbb{R}^n - \{0\}$$

given by $F(y,t) = tf(y) + (1-t)g(y)$. This is well defined, because $g(y)$ is the unit vector along $f(y)$, and so it cannot change its direction to become equal to $-f(y)$. Therefore $f|\partial B$ is homotopic to a constant map, and so $f|(\mathbb{R}^n - \text{Int } B)$ extends to a smooth map $\mathbb{R}^n \longrightarrow \mathbb{R}^n - \{0\}$ which is f outside a compact set containing B, by Lemma 6.6.7. □

Remark 6.6.9. After we establish Theorem 6.6.6, the lemma will be true for all n.

Proof of Theorem 6.6.6. The proof is by inductive arguments. The case $n = 1$ follows from Theorem 6.5.7 (since the degree of a constant map is zero). Assuming the result for $n-1$, take a smooth map $f : S^n \longrightarrow S^n$ having degree 0. If a is a regular value of f, and $f^{-1}(a) = \{a_1, \ldots, a_r\}$, then

$$\deg f = \sum_{i=1}^{r} n(f, a_i) = 0.$$

Using the homogeneity theorem, we can find a coordinate neighbourhood U in S^n containing $f^{-1}(a)$, and an orientation preserving diffeomorphism $\alpha : \mathbb{R}^n \longrightarrow U$. Let B be the image under α of a closed n-ball with centre 0 in $\mathbb{R}^n$ so that $f^{-1}(a) \subset B \subset U$. Let b be a point in S^n outside $f(U)$. Then $f(U) \subset S^n - \{b\}$. Let $\beta : S^n - \{b\} \longrightarrow \mathbb{R}^n$ be an orientation preserving diffeomorphism such that $\beta(a) = 0$. Consider the smooth map $h = \beta \circ f \circ \alpha$

$$\mathbb{R}^n \xrightarrow{\alpha} U \xrightarrow{f} f(U) \hookrightarrow S^n - \{b\} \xrightarrow{\beta} \mathbb{R}^n.$$

Then 0 is a regular value of h, and $h^{-1}(0) = (\alpha^{-1}(a_1), \ldots, \alpha^{-1}(a_r))$.

$$\deg h = \sum_{i=1}^{r} n(h, \alpha^{-1}(a_i)) = \sum_{i=1}^{r} n(f, a_i) = 0.$$

Therefore Lemma 6.6.8 gives rise to a smooth map $g_1 : U \longrightarrow \mathbb{R}^n - \{0\}$ so that $g_1 = h$ on $U - \text{Int } B$. Then the map $g_2 = \beta^{-1} \circ g_1$ agrees with f on $U - \text{Int } B$. We may therefore define a map $g : S^n \longrightarrow S^n - \{b\}$ by

$$\begin{aligned} g(x) &= f(x) \quad && \text{if } x \in S^n - \text{Int } B, \\ &= g_2(x) && \text{if } x \in U. \end{aligned}$$

The map g is smooth, and homotopic to f by the homotopy $H : S^n \times I \longrightarrow S^n$ defined by

$$H(x,t) = \frac{tf(x) + (1-t)g(x)}{\|tf(x) + (1-t)g(x)\|}.$$

This means that f is homotopic to a constant map, since $S^n - \{b\}$ is contractible. □

Theorem 6.6.10. *If M is a compact connected oriented manifold with $\chi(M) = 0$, then M has a nowhere vanishing vector field.*

Proof. Let $F : M \longrightarrow \tau(M)$ be a smooth vector field transverse to the zero section $i(M)$. By homogeneity theorem, the finite set $F^{-1}(M)$ lies in a coordinate neighbourhood U of a chart (U, ϕ) with $\phi(U) = \mathbb{R}^n$. Let

$$\tau_\phi : \pi^{-1}(U) \longrightarrow \phi(U) \times \mathbb{R}^n$$

be the corresponding chart of $\tau(M)$ (see Theorem 3.2.2). Then the local representation of F with respect to the pair (ϕ, τ_ϕ) is $\tau_\phi \circ F \circ \phi^{-1} : \mathbb{R}^n \longrightarrow \mathbb{R}^n \times \mathbb{R}^n$, and it is a vector field on $\mathbb{R}^n$ transverse to the zero section $\mathbb{R}^n \times \{0\}$. This gives a smooth map $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ (by projection onto the second factor), and we may write

$$F(x) = \tau_\phi^{-1}(\phi(x), f \circ \phi(x)), \ x \in U.$$

Now 0 is a regular value of f, and $\chi(M) = 0$ implies

$$\sum_{x \in f^{-1}(0)} \operatorname{Ind}_x f = 0.$$

Let B be a closed n-ball in $\mathbb{R}^n$ containing $f^{-1}(0)$ in its interior. Then, by Lemma 6.6.7, there is a smooth map $g : \mathbb{R}^n \longrightarrow \mathbb{R}^n - \{0\}$ which agrees with f on $\mathbb{R}^n - \text{Int } B$. We may now define a non-zero vector field $F_1 : M \longrightarrow \tau(M)$ by

$$\begin{aligned} F_1(x) &= \tau_\phi^{-1}(\phi(x), g \circ \phi(x)) && \text{if } x \in \phi^{-1}(B), \\ &= F(x) && \text{if } x \in M - \phi^{-1}(\text{Int } B). \end{aligned}$$

□

Lemma 6.6.11. *Let M be a compact manifold with boundary. Then any smooth map $f : \partial M \longrightarrow \mathbb{R}^m$ extends to a smooth map $F : M \longrightarrow \mathbb{R}^m$.*

Proof. Suppose that M is a submanifold of some Euclidean space $\mathbb{R}^k$. Let U be an ϵ-neighbourhood of ∂M in $\mathbb{R}^k$ with retraction $\pi : U \longrightarrow \partial M$. Then $H = f \circ \pi : U \longrightarrow \mathbb{R}^m$ is an extension of f over U. Let $\phi : \mathbb{R}^k \longrightarrow \mathbb{R}$ be a smooth map such that $\phi = 1$ on ∂M, and supp $\phi \subset U$ (Urysohn's Lemma 2.1.17). Then extend f to a smooth map $G : \mathbb{R}^k \longrightarrow \mathbb{R}^m$ by setting

$$\begin{aligned} G(x) &= \phi(x) \cdot H(x) && \text{if } x \in U, \\ &= 0 && \text{if } x \in \mathbb{R}^k - U. \end{aligned}$$

Then $F = G|M$ is the required extension of f. □

Theorem 6.6.12 (Extension Theorem). *Let M be a compact connected oriented $(n+1)$-manifold with boundary, and $f : \partial M \longrightarrow S^n$ be a smooth map. Then f extends to a smooth map $M \longrightarrow S^n$ if and only if* $\deg f = 0$.

Proof. We already know from Theorem 6.5.8 that the condition is necessary. To see the converse, suppose that $f : \partial M \longrightarrow S^n$ is a smooth map of degree zero. Extend f to a smooth map $F : M \longrightarrow \mathbb{R}^{n+1}$ using Lemma 6.6.10. We may assume that 0 is a regular value of F, by Theorem 6.2.12. By Theorem 6.6.1, we can find an open set U in Int M which contains $F^{-1}(0)$, and which is diffeomorphic to $\mathbb{R}^{n+1}$. Let B be a closed $(n+1)$-ball in U containing

$F^{-1}(0)$ in its interior, and $G : \partial B \longrightarrow S^n$ be the map $G = F/\|F\|$. Then, by Proposition 6.6.4,

$$\deg G = \sum_{x \in F^{-1}(0)} \operatorname{Ind}_x F.$$

Now, since G extends to $M - \operatorname{Int} B$, $\deg G = 0$, by Theorem 6.5.8. Therefore

$$\sum_{x \in F^{-1}(0)} \operatorname{Ind}_x F = 0,$$

and hence there is a smooth map $U \longrightarrow \mathbb{R}^n - \{0\}$ which agrees with $F|U$ outside of a compact set in U, by Lemma 6.6.7. This gives a smooth extension of f mapping M into $\mathbb{R}^n - \{0\}$. Composing this extension with the radial projection $\mathbb{R}^n - \{0\} \longrightarrow S^n$, given by $x \mapsto x/\|x\|$, we get the required extension of f. $\square$

Theorem 6.6.13 (Hopf's Degree Theorem). *Two smooth maps of a connected oriented compact n-manifold M into S^n are homotopic if and only if they have the same degree.*

Proof. Again we need to prove only one part of the theorem. So take two smooth maps f and g from M to S^n which have the same degree. They define a smooth map $h : \partial(M \times I) \longrightarrow S^n$ so that $h|M \times \{0\} = f$ and $h|M \times \{1\} = g$. Then, since boundary orientations on $M \times \{0\}$ and $M \times \{1\}$ are opposite, we have $\deg h = \deg f - \deg g = 0$. Therefore h extends to a smooth map $H : M \times I \longrightarrow S^n$ which is a homotopy between f and g. $\square$

CHAPTER 7

TUBULAR NEIGHBOURHOODS

The concept of a tubular neighbourhood was introduced by Whitney in [**56**] and [**59**]. This sketches a link between vector bundle and topology of manifold depicting how nicely a submanifold is embedded in a manifold. If M is submanifold of N, then a tubular neighbourhood of M is a neighbourhood of M in N which resembles the normal bundle of M in N. Thus a tubular neighbourhood describes the topology of a neighbourhood of the submanifold.

We shall prove that if M is a neat submanifold of N, then there exists a tubular neighbourhood of M in N, and that it is essentially unique (any two of them are related by an isotopy of N). We shall also prove the analogous results for collar neighbourhoods of the boundary ∂M of M. The applications include straightening of corners of manifolds, and constructions of manifolds by gluing process. The main theorems are influenced by the lectures of C.T.C. Wall [**50**].

7.1. Tubular neighbourhood theorems

Definition 7.1.1. Let $f : M \longrightarrow N$ be an immersion. Then we have a bundle monomorphism over M

$$\phi : \tau(M) \longrightarrow f^*\tau(N)$$

obtained by factoring $df : \tau(M) \longrightarrow \tau(N)$ through $f^*\tau(N)$ (see Lemma 5.3.2). The quotient bundle $f^*\tau(N)/\phi(\tau(M))$ of the monomorphism ϕ is called the **normal bundle of the immersion** f.

This also provides the definition of the normal bundle of an embedding. In particular, if M is a submanifold of N, then the normal bundle $\nu(M)$ of M in N is the quotient bundle of $\tau(N)|M$ modulo $\tau(M)$. In view of Lemma 5.4.7, the bundle $\nu(M)$ is equivalent to the orthogonal bundle $\tau(M)^\perp$ of $\tau(M)$ in $\tau(N)$ with respect to a Riemannian metric on N. Thus the fibre $\nu(M)_x$ is the normal space considered earlier in §6.1. Note that the second definition depends on the choice of Riemannian metric on N.

Definition 7.1.2. Let M be a submanifold of N. A **tubular neighbourhood** of M in N consists of a vector bundle $\pi : E \longrightarrow M$, and an embedding $\phi : E \longrightarrow N$ extending the diffeomorphism of the zero section Z onto M induced by π, in other words, $\phi(x, 0) = x$ for $(x, 0) \in Z$.

A. Mukherjee, *Differential Topology*, DOI 10.1007/978-3-319-19045-7_7

Frequently, we call the embedding ϕ or its image in N as tubular neighbourhood of M.

A tubular neighbourhood of an embedding $f : M \longrightarrow N$ is a tubular neighbourhood $\phi : E \longrightarrow N$ of the submanifold $f(M)$ of N. Then ϕ is an extension of f, i.e. $\phi(x, 0) = f(x)$ for $(x, 0)$ in the zero section.

There is a definition for closed tubular neighbourhood, which we shall consider in Definition 7.1.6.

Proposition 7.1.3. *If M is a submanifold of a boundaryless manifold N, then there is an embedding of an open neighbourhood of the zero section of the normal bundle $\nu(M)$ into N extending the projection of the zero section onto M.*

Proof. We have a short exact sequence

$$0 \longrightarrow \tau(M) \longrightarrow \tau(N)|M \longrightarrow \nu(M) \longrightarrow 0,$$

and a splitting

$$\tau(N)|M = \tau(M) \oplus \nu(M).$$

Since N is without boundary, the exponential map exp is defined on an open neighbourhood W of the zero section in $\tau(N)$ into N. Restricting exp to $W \cap \nu(M)$, we get a map $W \cap \nu(M) \longrightarrow N$, which maps the zero section diffeomorphically onto M.

Now refer back to Proposition 4.4.4, where we have shown that the Jacobian matrix of $\exp_p$ is the identity matrix if p belongs to the zero section, and so it is non-singular there. The proof may now be completed by using Lemma 6.1.3. $\square$

Remark 7.1.4. The essence of the proposition is that M has a neighbourhood U in N such that each point q of U is joined to M by a unique geodesic of length $d(q, M)$ which meets M orthogonally, and $(\exp)^{-1}$ is a diffeomorphism of U onto a neighbourhood of the zero-section in $\nu(M)$.

Theorem 7.1.5 (Existence of tubular neighbourhood). *If M is a submanifold of a manifold N without boundary, then there is a tubular neighbourhood of M in N.*

Proof. Denote the normal bundle $\nu(M)$ by E, and its projection onto M by π. Then, by Proposition 7.1.3, there is an open neighbourhood W of the zero section in E which is mapped diffeomorphically by exp. The normal bundle E inherits a fibrewise norm $\| \;\|_x, x \in M$. Let $f : M \longrightarrow \mathbb{R}$ be a positive continuous function such that if v is a vector in the fibre E_x of length less than $f(x)$, then v is contained in W. Then, by Lemma 2.1.15, we can find a positive smooth function $g : M \longrightarrow \mathbb{R}$ such that $0 < g(x) < f(x)$ for all $x \in M$. Define $\psi : E \longrightarrow E$ by

$$\psi(v) = g(\pi(v)) \cdot \frac{v}{1 + \|v\|_{\pi(v)}}.$$

Then ψ is a diffeomorphism mapping the fibre E_x at any point $x \in M$ onto the open disc of radius $g(x)$ in E_x, and thus induces a bundle structure on $\psi(E)$. Since $\psi(E) \subset W$, the map $\exp \circ \psi$ embeds E into N extending the diffeomorphism of the zero section onto M. $\square$

The theorem is also true when N has boundary, in this case we need to assume that M is a neat submanifold of N. We shall prove this result in Theorem 7.2.12 of the next section.

In the definition of tubular neighbourhood, E is just the normal bundle $\nu(M)$ of M in N, which is a vector bundle with fibre $\mathbb{R}^{m-n}$ and structure group $O(m-n)$, where $m-n$ is the codimension of M in N. We shall now restrict this definition bringing in a subbundle of the bundle E.

Let B be the set of vectors $(x,v) \in E$, $v \in E_x$, such that $\|v\|_x \leq 1$. Then B is a submanifold of E with boundary S consisting of all vectors of length exactly equal to 1. The manifolds B and S are respectively the disk bundle and sphere bundle associated to E (see Exercise 5.6 in 142). Let $\pi : B \longrightarrow M$ be the projection $\pi(x,v) = x$. This is not a vector bundle, but a fibre bundle, where the fibre over x is the closed unit $(m-n)$-disc in E_x, and the structure group is $O(m-n)$.

Actually B will be a manifold with corner $\Lambda(B) = S \cap \pi^{-1}(\partial M)$, and so B will not be a smooth manifold, unless $\partial M = \emptyset$. Therefore we should assume that M is without boundary, whenever we consider a normal disk bundle over it.

Definition 7.1.6. A **closed tubular neighbourhood** of M in N is an embedding $\psi : B \longrightarrow N$ of a manifold with boundary such that $\psi|M = \mathrm{Id}_M$, identifying M with the zero section.

Of course, the definition can be extended to the category of manifolds with corner. But this will take us too far away.

Theorem 7.1.7 (**Existence of closed tubular neighbourhood**). *If M is a submanifold of N, both without boundary, then there is a closed tubular neighbourhood of M in N.*

Proof. Let W, f, and g be as in the proof of Theorem 7.1.5. Define

$$\psi : B \longrightarrow N$$

by $\psi(x,v) = \exp(x, g(x)v)$. This is well defined, because $g(x) \neq 0$, and so

$$\|v\|_x \leq 1 \Rightarrow \|g(x)v\|_x \leq g(x) < f(x) \Rightarrow (x, g(x)v) \in W.$$

This completes the proof. $\square$

7.2. Collar neighbourhoods

Much of the results of Chapter 4 extend to a manifold N with boundary. For example, N has a Riemannian metric, the proof is the same as before. The discussion of geodesics also carries through at interior points. But geodesics may not exist at boundary points. At a boundary point, we have two kinds of tangent vectors to N, inward- and outward-pointing tangent vectors. Precisely, in terms of local coordinates, a tangent vector to N has the form $\sum_i \lambda_i \partial/\partial x_i$. A tangent vector is inward- (resp. outward-) pointing if $\lambda_1 > 0$ (resp. $\lambda_1 < 0$). If $\lambda_1 = 0$, then it is a tangent vector to the boundary ∂N; note that such vectors form the image of the inclusion $di : \tau(\partial N) \longrightarrow \tau(N)$, where $i : \partial N \subset N$, and so they are tangent to ∂N. It follows from Proposition 4.4.1 that local geodesics can be constructed only for all inward-pointing tangent directions, and for no outward-pointing ones. For directions tangent to the boundary, local geodesics may or may not exist (see what happens by drawing tangent vectors at boundary points of the closed disk D^2, and the closure of the region $\mathbb{R}^2 - D^2$, each with the standard metric of $\mathbb{R}^2$). Thus the boundary ∂N does not admit a tubular neighbourhood. However, we can have a kind of half tubular neighbourhood of the boundary, which is called a collar neighbourhood. The precise definition is as follows.

Definition 7.2.1. A **collar neighbourhood** of ∂N in N is an embedding

$$\phi : \partial N \times [0, \infty) \longrightarrow N$$

such that $\phi(x, 0) = x$.

Theorem 7.2.2 (Existence of collar neighbourhood). *There exists a collar neighbourhood of ∂N in N.*

Proof. In order to show this, in the spirit of Theorem 7.1.5, all we need is to identify $\partial N \times [0, -\infty)$ with the set of inward-pointing normal vectors to ∂N. The identification is possible, because there is only one such normal direction at each point of ∂N, and so an inward-pointing normal vector is determined by its length. Then, the previous arguments of 7.1.3 and 7.1.5 carry over to this case. □

The proof may seem rather cooked. Let us try an alternative approach to the proof using differential equations in a straightforward way without having recourse to geodesics and the exponential map. Some of these arguments will be used in Theorem 7.3.3.

Lemma 7.2.3. *Let M be a manifold without boundary, and X_0 be the constant unit vector field on $M \times \mathbb{R}$ whose value $(X_0)_{(x,\ t)}$ at any point $(x, t) \in M \times \mathbb{R}$ is tangent to the curve $t \mapsto (x, t)$ at that point. Then, for any smooth vector field X on M, there is a positive smooth function $\epsilon : M \longrightarrow \mathbb{R}$, and a unique smooth map $f : W(\epsilon) \longrightarrow M$, where $W(\epsilon) = \{(x, t) \in M \times \mathbb{R} \mid |t| < \epsilon(x)\}$, such that*

(1) $f(x,0) = x$,
(2) $df_{(x,t)}((X_0)_{(x,\ t)}) = X_{f(x,\ t)}$.

The converse is trivially true: given f with conditions (1) and (2), the vector field X may be defined by (2).

PROOF. Consider an atlas $\{(U_\alpha, \phi_\alpha)\}$ for M. On a coordinate neighbourhood U_α with local coordinates $\phi_\alpha = (x_1, \ldots, x_n)$, the vector field $X_\alpha = X|U_\alpha$ has representation $X_\alpha = a_1 \cdot \partial/\partial x_1 + \cdots + a_n \cdot \partial/\partial x_n$, where $a_i : U_\alpha \longrightarrow \mathbb{R}$ are smooth functions. The vector field X_0 has $n+1$ components $(0, \ldots, 0, 1)$ (with 1 in the $(n+1)$-th slot) in any coordinate neighbourhood of $M \times \mathbb{R}$. Then the condition (2) gives a system of differential equations

$$\frac{\partial f_i}{\partial t}(x,t) = a_i(f_1(x,t), \ldots, f_n(x,t)),$$

where $f_i : U_\alpha \times \mathbb{R} \longrightarrow \mathbb{R}$ are the components of f, with an initial condition corresponding to (1) at $t = 0$.

Solving the system of equations, we get a unique solution

$$f_\alpha : V_\alpha \times (-\epsilon_\alpha, \epsilon_\alpha) \longrightarrow U_\alpha.$$

where $V_\alpha = \phi_\alpha^{-1}(B)$ for some small open n-ball B in $\mathbb{R}^n$, and ϵ_α is some small positive number.

We may suppose that $V_\alpha \subseteq U_\alpha$. Then carrying out the construction for every α, we have $f_\alpha = f_\beta$ on $V_\alpha \cap V_\beta$ by the uniqueness. Also we may get a positive smooth function $\epsilon : M \longrightarrow \mathbb{R}$ as $\epsilon = \sum_\alpha \lambda_\alpha \epsilon_\alpha$, by gluing the constant functions ϵ_a on V_α by a partition of unity $\{\lambda_\alpha\}$ subordinate to the covering $\{V_\alpha\}$. Therefore we may define the required map $f : W(\epsilon) \longrightarrow M$ by $f|V_\alpha = f_\alpha$. □

The proof of the lemma will break down if M has boundary. In this case, if $x \in \partial U_\alpha$, then the solution f_α may not lie in U_α, as its first component may not be ≥ 0. However, if the vector field $X_\alpha = X|U_\alpha$ is inward pointing at any point of ∂U_α, that is, if its first component a_1 is positive, then the solution $f_1(x,t)$ will be positive for small values of $t \geq 0$. To take care of this situation, we need to construct a vector field X on M such that, for each $x \in \partial M$, the vector X_x is the inward pointing.

So, suppose that M is a manifold with boundary, and take an atlas $\{(U_\alpha, \phi_\alpha)\}$ for M. Let Y_α be the vector field on U_α defined by $Y_\alpha = d\phi_\alpha^{-1}(e_1)$, where e_1 is the unit vector along the first coordinate axis. Let $\{\lambda_a\}$ be a partition of unity subordinate to the covering $\{U_\alpha\}$. Then $X = \sum_\alpha \lambda_\alpha Y_\alpha$ is the desired vector field, and working with this X, we may get the following analogue of the above lemma for manifolds with boundary.

Lemma 7.2.4. *Let X be the vector field constructed above on a manifold with boundary M. Then there is a positive smooth function $\epsilon : M \longrightarrow \mathbb{R}$ and a* *smooth map* $f : W_+(\epsilon) \longrightarrow M$*, where* $W_+(\epsilon) = \{(x,t) \in M \times \mathbb{R}_+ \mid t < \epsilon(x)\}$ *satisfying the conditions* (1) *and* (2) *of Lemma* 7.2.3.

ANOTHER PROOF OF THEOREM 7.2.2. Construct the vector field X and the map f of Lemma 7.2.4. Clearly, f maps $\partial M \times \{0\}$ diffeomorphically onto ∂M, and df is an isomorphism at each point of $\partial M \times \{0\}$, because X is inward pointing along ∂M. Therefore by Lemma 6.1.3, f is an embedding of a neighbourhood of $\partial M \times \{0\}$ into M. This neighbourhood contains a $W_+(\epsilon)$ for some positive smooth function ϵ. Then there is a diffeomorphism $W_+(\epsilon) \longrightarrow M \times (0,1)$ given by

$$(x,t) \mapsto \Big(x, \frac{t}{\epsilon(x)}\Big),$$

and a diffeomorphism $M \times [0,1) \longrightarrow M \times [0,\infty)$ given by

$$(x,t) \mapsto \Big(x, \frac{t}{1-t}\Big).$$

This completes the proof of the collar neighbourhood theorem. $\square$

This theorem enables us in most cases to overcome some difficulties arising at the boundary.

Remark 7.2.5. We may take a collar neighbourhood of ∂N as an embedding $\partial N \times [0,1) \longrightarrow N$ (or even $\partial N \times [0,1] \longrightarrow N$) which is Id on ∂N.

Convention. The domain of any collar neighbourhood that we shall consider will always be like $\partial N \times [0,1]$, unless it is explicitly stated otherwise.

A Riemannian structure on N induces a Riemannian structure on ∂N (we suppose that ∂N is connected).

Definition 7.2.6. A Riemannian metric on N is **adapted to the boundary** ∂N if the construction of N-geodesics for directions tangent to ∂N are possible locally, and if such geodesics are curves lying entirely in ∂N, that is, if ∂N is a totally geodesic submanifold of N (see §4.6)..

.

Lemma 7.2.7. *If* M *is a manifold without boundary and with a Riemannian metric, and the half open interval* $[0,\infty)$ *has the standard Euclidean metric, then the product metric on* $N = M \times [0,\infty)$ *is adapted to its boundary.*

PROOF. Let $x_0, \dots, x_n$ be local coordinates in N, x_0 being the coordinate in $[0,\infty)$. Then for the Riemannian metric g_{ij} on N, we have $g_{0j} = \delta_{0j}$(Kronecker delta). Therefore one of the geodesic differential equations in Proposition 4.2.5 simply becomes $d^2x_0/dt^2 = 0$. Thus if $x_0 = 0$ and $dx_0/dt = 0$ initially, then $x_0 = 0$ all along the geodesic. This means that the geodesics are curves lying in ∂N. $\square$

Proposition 7.2.8. *Every manifold N with boundary has a Riemannian metric adapted to the boundary.*

PROOF. The proof consists of constructing a Riemannian metric on N which is a product metric on a small neighbourhood $\partial N \times [0, \epsilon]$ of ∂N. Let $\phi : \partial N \times [0,1] \longrightarrow N$ be a collar neighbourhood of ∂N in N. Let there be given a metric d_1 on N, and a metric d_2 on $\partial N \times [0,1]$ which is the product of some metric on ∂N and the standard metric on $[0,1]$. Then, define a metric d on N as follows.

$$d = \begin{cases} d_1 & \text{outside Image } \phi, \\ d_2 + (d_1 - d_2) \cdot \mathcal{B}(3t-1) & \text{on Image } \phi, \end{cases}$$

where $\mathcal{B}(3t-1)$ is a bump function, which is 0 if $t \leq 1/3$, and 1 if $t \geq 2/3$. Then d is smooth, because the definitions agree in a neighbourhood of $t = 1$. It is a Riemannian structure, because a positive linear combination of positive definite forms is a positive definite form. Moreover, it agrees with d_2 in a neighbourhood of $t = 0$. Therefore, by Lemma 7.2.7, the metric d is adapted to ∂N. □

Definition 7.2.9. We say that a neat submanifold M of N **meets ∂N orthogonally** if the normal vectors to M and ∂N are orthogonal at each point of ∂M.

Lemma 7.2.10. *Let M be a neat submanifold of N, where each of them has boundary. Then N admits a Riemannian metric in which M meets ∂N orthogonally.*

PROOF. The required Riemannian metric on N may be constructed as in Theorem 4.1.3 by fitting together standard metrics on coordinate neighbourhoods in N by means of a partition of unity. Since M is a neat submanifold of N, at a point of ∂M there is a coordinate neighbourhood U in N that contains a coordinate neighbourhood V in M so that $\partial V = V \cap \partial U$. Since the pair (U, V) is like the pair $(\mathbb{R}^m_+, \mathbb{R}^n_+)$, V meets ∂U orthogonally with respect to the standard metric. Now if we fit the partial metrics together, we get a metric on N with respect to which M continues to meet ∂N orthogonally. □

Corollary 7.2.11. *The manifold N has a metric adapted to its boundary in which M meets ∂N orthogonally.*

PROOF. Construct a Riemannian metric on N as in Lemma 7.2.10 so that M meets ∂N orthogonally, and then construct a corresponding collar neighbourhood of ∂N in N. Since at points $x \in \partial M$, a vector at x that is normal to ∂N is tangent to M, the vector is a generator of the collar neighbourhood. Therefore, using this collar neighbourhood in the proof of Lemma 7.2.8, we get a Riemannian metric on N which is adapted to ∂N, and with respect to which M continues to be orthogonal to ∂N. □

Theorem 7.2.12. *If N is a manifold with boundary, and M is a neat submanifold of N, then there exists a tubular neighbourhood of M in N.*

PROOF. The arguments of the corresponding theorem for manifolds without boundary can now easily be carried through in this case. We leave the obvious details to the reader. □

The corresponding result for closed tubular neighbourhood may also be seen to be true.

Theorem 7.2.13. *If N has boundary and M is a submanifold of N without boundary, then there is a closed tubular neighbourhood of M in N.*

We shall need one more result about collar neighbourhood.

Theorem 7.2.14. *If M is a neat submanifold of N, then there is a collar neighbourhood $\psi : \partial N \times [0,1] \longrightarrow N$ of ∂N in N such that $\psi|\partial M \times [0,1]$ is a collar neighbourhood of ∂M in M.*

PROOF. Let $\phi : E \longrightarrow N$ be a tubular neighbourhood of M in N. Choose a Riemannian structure on N (and this induces a metric on M), and give E a product metric structure in the following way. Take a trivialising open covering $\{U_i\}$ of M, and then take product metric structures on the trivial bundles $E|U_i$ (explicitly, the structure is the product of the metric on U_i and the standard metric on the fibre which is a Euclidean space). Since the structure group of the bundle E is the orthogonal group which preserves the standard metric on the fibre, these local structures agree on their intersections, and thus define a global product metric structure on E. Now, as in Proposition 7.2.8, we modify the Riemannian structure on N so that it agrees with the above product metric structure on E in a neighbourhood of M. Then using this metric we construct a collar neighbourhood ψ of ∂N in N as in Theorem 7.2.2. We assert that ψ has the required property. Indeed, since in a neighbourhood of M the metric is the product metric constructed above, geodesics tangent to M are contained in M, as in Lemma 7.2.7. This completes the proof. □

7.3. Isotopy extension theorem

Let M and N be manifolds (possibly with boundary), $\dim M \leq \dim N$. Recall from Definition 2.6.4 that an **isotopy** between two embeddings

$$f, g : M \longrightarrow N$$

is a smooth map $h : M \times \mathbb{R} \longrightarrow N$ such that

(a) for each $t \in \mathbb{R}$, the map $h_t : M \longrightarrow N$, given by $h_t(x) = h(x,t)$, is an embedding,
(b) $h_0 = f$ and $h_1 = g$.

A **diffeotopy** of N is an isotopy $h : N \times \mathbb{R} \longrightarrow N$ such that each h_t is a diffeomorphism of N, and $h_0 = \mathrm{Id}$.

A **level-preserving-isotopy** of M in N is an embedding $f : M \times \mathbb{R} \longrightarrow N \times \mathbb{R}$ such that it preserves the second coordinate $t \in \mathbb{R}$, and such that the map $f_t : M \longrightarrow N$, given by $f_t(p) = \pi_1 \circ f(p,t)$, where $\pi_1 : N \times \mathbb{R} \longrightarrow N$ is the projection onto the first factor, is an embedding. Thus we have

$$f(p,t) = (f_t(p), t).$$

Note that a level-preserving isotopy is an embedding, whereas an isotopy is just a smooth map, and that 'level-preserving isotopy' is not a new concept. Level-preserving smooth maps $f : M \times \mathbb{R} \longrightarrow N \times \mathbb{R}$ correspond bijectively to smooth maps $\overline{f} : M \times \mathbb{R} \longrightarrow N$ via the assignment

$$f(x,t) = (\overline{f}(x,t), t).$$

Thus $\overline{f} = \pi_1 \circ f$, and $\overline{f}_t = f_t$. A simple computation of the Jacobian matrix of f shows that f is an embedding if and only if each $\overline{f}_t$ is so, in other words, f is an level-preserving isotopy if and only if $\overline{f}$ is an isotopy. We may call $\overline{f}$ the **track of** f.

A **level-preserving diffeotopy** f of N, $f : N \times \mathbb{R} \longrightarrow N \times \mathbb{R}$, is defined similarly by requiring that f, and each f_t, be a diffeomorphism, and that $f_0 = \mathrm{Id}$.

The **support** of a diffeomorphism $f : N \longrightarrow N$ is the closure of the set of points $p \in N$ such that $f(p) \neq p$. The **support of a level-preserving isotopy**

$$f : M \times \mathbb{R} \longrightarrow N \times \mathbb{R},$$

or its track $\overline{f} : M \times \mathbb{R} \longrightarrow N$, is the closure of the set of points $p \in M$ for which $f_t(p) \neq f_0(p)$, for some non-zero $t \in \mathbb{R}$. The support of f is denoted by $\mathrm{supp} f$. We have $f_t = f_0$ outside $\mathrm{supp} f$, for all $t \in \mathbb{R}$.

Definition 7.3.1. A level-preserving diffeotopy $g : N \times \mathbb{R} \longrightarrow N \times \mathbb{R}$ is said to **cover** a level-preserving isotopy $f : M \times \mathbb{R} \longrightarrow N \times \mathbb{R}$ if the following diagram commutes.

$$\begin{tikzcd}
M \times \mathbb{R} \arrow[rr, "f_0 \times \mathrm{Id}"] \arrow[dr, "f"'] & & N \times \mathbb{R} \arrow[dl, "g"] \\
& N \times \mathbb{R} &
\end{tikzcd}$$

that is, if

$$g(f_0(x), t) = f(x,t), \text{ or } g_t \circ f_0 = f_t \quad \text{or all } t \in \mathbb{R}.$$

A level-preserving isotopy is called a **strong** if it can be covered by a level preserving diffeotopy.

Let M be a submanifold of N. Then a diffeotopy $f : M \times \mathbb{R} \longrightarrow M$ is covered by a diffeomorphism $g : N \times \mathbb{R} \longrightarrow N$ if $g_t|M = f_t$ for all $t \in \mathbb{R}$.

Proposition 7.3.2. *Any diffeotopy of ∂N can be covered by a diffeotopy of N with support contained in* Int N*.*

Proof. We shall suppose that the diffeotopy f_t of ∂N is normalised so that $f_t = \mathrm{Id}$ for $t \leq \frac{1}{3}$, and $f_t = f_1$ for $t \geq \frac{2}{3}$. (See §2.6; note that for any diffeotopy f_t of ∂N, the diffeotopy $f_{\lambda(t)}$, where $\lambda : \mathbb{R} \to \mathbb{R}$ is the bump function $\lambda(t) = \mathcal{B}(3t-1)$, is of this form.)

Let $\phi : \partial N \times [0,1] \longrightarrow N$ be a collar neighbourhood of ∂N in N. Consider a smooth map $\alpha : \partial N \times \mathbb{R} \times \mathbb{R} \longrightarrow \partial N$, where $\alpha_{(t,s)} : \partial N \longrightarrow \partial N$, $(t,s) \in \mathbb{R} \times \mathbb{R}$ is given by

$$\alpha_{(t,s)} = \begin{cases} \mathrm{Id} & \text{if } s \geq t, \\ f_{t-s} & \text{if } s \leq t. \end{cases}$$

The definitions match together smoothly, since $f_t = \mathrm{Id}$ in a neighbourhood of $t = 0$, making $\alpha_{(t,s)}$ a diffeomorphism. Now a covering diffeotopy g_t of N may be defined by

$$g_t = \mathrm{Id} \text{ outside } \phi(\partial N \times [0,1]), \text{ and}$$

$$g_t(\phi(x,s)) = \phi(\alpha_{(t,s)}(x), s) \text{ otherwise .}$$

The map $g : N \times \mathbb{R} \longrightarrow N$ is smooth. Also g_t covers f_t, because $\alpha_{(t,s)} = f_t$ if $s = 0$. This completes the proof. □

Theorem 7.3.3 (Isotopy extension theorem). *A level-preserving isotopy*

$$f : M \times \mathbb{R} \longrightarrow N \times \mathbb{R}$$

with compact support and $\overline{f}(\mathrm{supp} f \times \{0\}) \subset$ Int N *can be covered by a level-preserving diffeotopy $g : N \times \mathbb{R} \longrightarrow N \times \mathbb{R}$ whose support is also compact and contained in* Int N*. Thus f is a strong isotopy.*

Proof. We may suppose that N is a manifold without boundary. Because, if the theorem is proved in this case, then we obtain a level-preserving diffeotopy g with compact support contained in Int N so that g is Id on a neighbourhood of $\partial N \times \mathbb{R}$, and therefore g can be extended to the rest of $N \times \mathbb{R}$ as identity map.

The proof begins with the observation that level-preserving diffeotopies of N with compact support bear an important relation with vector fields on $N \times \mathbb{R}$ which project to d/dt on $\mathbb{R}$, and whose projections to N vanish outside a compact subset of N. To see this, suppose that X_0 is the vector field on $N \times \mathbb{R}$ which assigns to each point $(p,s) \in N \times \mathbb{R}$ the unit positive tangent vector at the point s to the curve $\sigma : \mathbb{R} \longrightarrow N \times \mathbb{R}$ given by $\sigma(t) = (p,t)$, so that X_0 projects to the zero vector on N, and to the vector d/dt on $\mathbb{R}$ by the projections $\pi_1 : N \times \mathbb{R} \longrightarrow N$ and $\pi_2 : N \times \mathbb{R} \longrightarrow \mathbb{R}$ respectively. This means that $X_0 = d\sigma(d/dt), d\pi_1(X_0) = 0, d\pi_2(X_0) = d/dt$.

Now, given a level-preserving diffeotopy $g : N \times \mathbb{R} \longrightarrow N \times \mathbb{R}$, define a vector field X on $N \times \mathbb{R}$ by $X = dg(X_0)$. Note that X still projects to d/dt on

$\mathbb{R}$, since g is level-preserving. Also $X = X_0$ except at points of some compact set, if g has compact support.

Conversely, any vector field X on $N \times \mathbb{R}$, which is equal to d/dt on $\mathbb{R}$ and equals X_0 outside a compact set K in N, gives rise to a level-preserving diffeotopy $g' : N \times \mathbb{R} \longrightarrow N \times \mathbb{R}$ with compact support. This follows because $X = X_0$ outside the compact set K and X_0 projects to 0 on N, and hence X generates a unique flow $\{\phi_t : N \times \mathbb{R} \longrightarrow N \times \mathbb{R},\ t \in \mathbb{R}\}$, by Theorem 3.2.12, so that

$$\frac{d}{dt}\phi_t(x, s)\Big|_{t=0} = X(\phi_t(x, s)).$$

Since X projects to d/dt on $\mathbb{R}$ and its integral is t, the second coordinate of $\phi_t(p, 0)$ is t, and so $\phi_t(p, 0)$ is of the form

$$\phi_t(p, 0) = (g'_t(p), t),$$

and $g'(p, t) = (g'_t(p), t)$ is the required level-preserving diffeotopy. If we start with g and get the vector field X, and then construct g' using X, then we must have $g = g'$ by uniqueness, since each of them satisfies the equation

$$\frac{\partial}{\partial t}x_i(g(p, t)) = X_i(g(p, t)),$$

where x_i are the local coordinates in $N \times \mathbb{R}$, and X_i are the components of X in this coordinate system. On the other hand, if we start with X and get g', then we have $dg'(X_0) = X$. This establishes a one-to-one correspondence between g and X.

In view of this correspondence, it is sufficient to construct a vector field X on $N \times \mathbb{R}$ such that

(1) $X = X_0$ outside a compact set,
(2) X projects on $\mathbb{R}$ to the constant positive unit vector field d/dt,
(3) $X = df(X_0)$ on $f(M \times \mathbb{R})$.

Note that g covers f if and only if condition (3) holds. Indeed, as described above, the level-preserving isotopy $f : M \times \mathbb{R} \longrightarrow N \times \mathbb{R}$ defines a vector field Y on $f(M \times \mathbb{R})$ by $Y = df(X_0)$. The integral curve of Y, which passes through $(f_0(x), 0)$ at $t = 0$, is $t \mapsto f(x, t)$. Since $X = Y$ on $f(M \times \mathbb{R})$, this is also the integral curve of X passing through the same point at $t = 0$, which is $t \mapsto g(f_0(x), t)$. Therefore $g(f_0(x), t) = f(x, t)$, and g covers f.

Although the construction can be carried out using tubular neighbourhoods, we shall follow a general method in order to take care of the boundary points. For this purpose, we endow N with a Riemannian metric and $N \times \mathbb{R}$ with the product metric. The condition (2) gives the component of X in the direction of $\mathbb{R}$ in a way compatible with the conditions (1) and (3). Therefore we need only to find the component of X in the direction of N. We assert that if we can do this in an open neighbourhood of each point of $f(M \times \mathbb{R})$, then X can be constructed. For, such neighbourhoods together with the complement of the closed set $\overline{f}(\text{supp} f \times \{0\})$ constitute an open covering $\{U_\alpha\}$ of $N \times \mathbb{R}$.

Let $\{\lambda_\alpha\}$ be a partition of unity subordinate to this covering. Then, if X_α is a vector field on U_α satisfying the conditions (1), (2), and (3), the vector field $X = \sum_\alpha \lambda_\alpha X_\alpha$ will be the desired vector field satisfying the same conditions.

On the complement of $\overline{f}(\mathrm{supp} f \times \{0\})$, we take $X_\alpha = X_0$. Since $f(M \times \mathbb{R})$ is a submanifold of $N \times \mathbb{R}$, each point of $f(M \times \mathbb{R})$ admits a coordinate chart $\phi : U \longrightarrow \mathbb{R}^{m+1}$ in $N \times \mathbb{R}$ such that $U \cap f(M \times \mathbb{R}) = \phi^{-1}(\mathbb{R}^{n+1})$ (we suppose that $\dim M = n$ and $\dim N = m$). Then $d\phi(df(X_0)) = \sum a_i \partial/\partial x_i$ in $\mathbb{R}^{n+1}$. We define X by taking the same formula in $\mathbb{R}^{m+1}$ (that is, by taking the a_i independent of the last $m - n$ coordinates). In the case of boundaries, the a_i are only defined on the closed set $\mathbb{R}^{n+1}_+$ in $\mathbb{R}^{n+1}$. By Tietze's Extension Theorem, they can be extended to continuous functions on $\mathbb{R}^{n+1}$. By Theorem 2.2.3, these continuous functions can be approximated by smooth extensions on $\mathbb{R}^{n+1}$, and then extended to $\mathbb{R}^{m+1}$ as above. This completes the proof. □

Corollary 7.3.4. *If M is a compact submanifold of* Int N*, where M may have boundary, then any isotopy of the inclusion $i : M \hookrightarrow$* Int N *can be covered by a diffeotopy of N having compact support.*

Proof. By the theorem, the given isotopy can be covered by a diffeotopy of Int N with compact support. Thus ∂N has a neighbourhood in Int N which is left fixed by the diffeotopy, and so can be extended to N by taking it to be Id on ∂N. □

Proposition 7.3.5. *Let P be a submanifold of a compact manifold M, and N a manifold without boundary. Let $f_0 : M \longrightarrow N$ be an embedding, and $\overline{f}_1 : P \longrightarrow N$ be an embedding isotopic in N to the restriction $\overline{f}_0$ of f_0 to P. Then there is an embedding $f_1 : M \longrightarrow N$ which is isotopic to f_0 such that $f_1|P = \overline{f}_1$.*

Proof. Let $\overline{f}_t : P \longrightarrow N$, $0 \le t \le 1$, be an isotopy between $\overline{f}_0$ and $\overline{f}_1$. Then the map $\overline{f}_t \circ (\overline{f}_0)^{-1} : \overline{f}_0(P) \longrightarrow N$ is an isotopy of the inclusion $\overline{f}_0(P) \subset N$. By Corollary 7.3.4, this extends to a diffeotopy $g_t : N \longrightarrow N$ so that $g_t \circ \overline{f}_0 = \overline{f}_t$. Then $h_t = g_t \circ f_0 : M \longrightarrow N$ is an isotopy such that $h_0 = f_0$ and $h_t|P = \overline{f}_t$. Therefore $f_1 = h_1$ is the required embedding. □

Proposition 7.3.6. *If M is a compact submanifold with boundary of a manifold N without boundary, then there is a submanifold P without boundary of N containing M such that* $\dim P = \dim M$.

Proof. The idea of the proof is to find an isotopy $f : M \times \mathbb{R} \longrightarrow M$ so that $f_0 = \mathrm{Id}$ and $f_k(M) \subset \mathrm{Int}\ M$ for some $k > 0$. Then, since M is compact, the isotopy $i \circ f$, where i is the inclusion $M \hookrightarrow N$, can be covered by a diffeotopy g so that $g_k(M) = f_k(M) \subset \mathrm{Int}\ M$. Then $P = g_k^{-1}(\mathrm{Int}\ M)$ will be the required submanifold of N.

To construct the isotopy f_t, consider a collar neighbourhood

$$\phi : \partial M \times [0,1] \longrightarrow M,$$

and a bump function $\mathcal{B}(t)$. Choose a number k such that

$$0 < k < 1, \text{ and } \partial\mathcal{B}(t)/\partial t < 1 \text{ if } t \leq k.$$

Let $\alpha : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$ be a smooth function such that

(1) $\alpha(0, s) = s$,
(2) $\alpha(t, 0) > 0$ if $t > 0$,
(3) $\alpha(t, s) = s$ if $t \leq k < 1 - \epsilon < s < 1$, for an ϵ such that $0 < \epsilon < 1$, and
(4) $\partial\alpha/\partial s > 0$, if $t - s \leq k$.

The function α may be constructed in the following way. First define it over the interval $[0, k]$ by $\alpha(t, s) = s + \mathcal{B}(t - s)$. Then extend it to a smooth function over the whole $\mathbb{R} \times \mathbb{R}$. Finally define f_t by

$$f_t(x) = x \text{ outside } \phi(\partial M \times [0, 1]),\ f_t(\phi(x, s)) = \phi(x, \alpha(t, s)) \text{ otherwise.}$$

Then the isotopy f pushes the boundary ∂M a little inside M, and the proof is complete. □

◊ **Exercise 7.1.** Construct a proof of the above proposition when N is a manifold with boundary and M is a neat submanifold of N.

Hint. Proceed as above using the part of ∂N not in ∂M.

We shall now improve Theorem 7.3.3 in a particular situation. It is an improvement in that the support of the isotopy need not be contained in the interior of N.

Theorem 7.3.7. *If M is a neat submanifold of N, where both of them may have boundary, then any isotopy of M in N with compact support, which may not be contained in* Int N*, can be covered by a diffeotopy of N with compact support.*

Proof. Since M is a neat submanifold of N, Theorem 7.2.14 gives a collar neighbourhood $\psi : \partial N \times [0, 1] \longrightarrow N$ of ∂N in N which restricts to a collar neighbourhood of ∂M in M.

Let f be an isotopy of M in N with compact support, which may not be in Int N. By Theorem 7.3.3, the isotopy $f|\partial M$ can be covered by a diffeotopy of ∂N. By Proposition 7.3.2, this diffeotopy can be covered by a diffeotopy g of N, using the collar neighbourhood ψ of ∂N in N, so that the support of g is compact and contained in Int N (in fact, in Int $\psi(\partial N \times [0, 1])$). By construction, there is an open neighbourhood U of ∂M in M so that $f(U) \subset \psi(\partial N \times [0, 1])$. This means that the diffeotopy g covers f on U.

To cover f on the remaining portion of M, it is necessary to cover it only on a set of compact support contained in Int N. This can be done by using the methods of Theorem 7.3.3. The diffeotopy of N thus obtained will agree with the previous diffeotopy g on a common part, giving the required covering diffeotopy of f. □

7.4. Uniqueness of tubular neighbourhoods

In this section we shall complete our discussion about tubular neighbourhoods by showing that they are essentially unique.

Definition 7.4.1. Let M be a submanifold of N without boundary. Then two closed tubular neighbourhoods $\phi_1 : B_1 \longrightarrow N$ and $\phi_2 : B_2 \longrightarrow N$ of M in N are *equivalent* (or *isotopic*), if there is a bundle map $\lambda : B_1 \longrightarrow B_2$ over the identity map of M (i.e., an $O(m-n)$-bundle map respecting the group action), and a isotopy from $\phi_2 \circ \lambda$ to ϕ_1 which fixes the zero-section of B_1.

This means that ϕ_1 and ϕ_2 are equivalent if we can fit an isotopy $f_t : B_1 \longrightarrow N$ between $\phi_2 \circ \lambda$ and ϕ_1 so that each f_t is a closed tubular neighbourhood.

We can go one step further and define ϕ_1 and ϕ_2 to be *strongly equivalent* if the isotopy f_t is a strong isotopy, that is, if it can be covered by a diffeotopy of N. Equivalent (and strongly equivalent) collar neighbourhoods, and tubular neighbourhoods, are defined similarly.

We shall show in Theorem 7.4.6 that any two closed tubular neighbourhoods are equivalent. However, for strong equivalence we must assume that M is compact. The result may be false without this assumption, as the following example shows.

Example 7.4.2. Let M be the x-axis and N be $\mathbb{R}^2$. Let B_1 be $\mathbb{R} \times [-1, 1]$, and $B_2 = \{(x, y) \in \mathbb{R}^2 | x^2 + (y-2)^2 \geq 1, -3 \leq y < 3\}$. Then B_1 and B_2 are tubular neighbourhoods of $\mathbb{R}$ in $\mathbb{R}^2$, where the projection of B_1 on $\mathbb{R}$ is the standard projection $(x, t) \mapsto x$, and the projection of B_2 on $\mathbb{R}$ is given by straight lines through the point $(0, 3)$. Clearly B_1 and B_2 are isotopic, but not strongly isotopic.

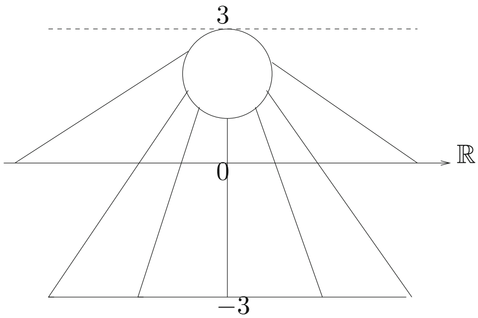


FIGURE 7.1

We shall show in Theorem 7.4.7 that if M is compact, then any two closed tubular neighbourhoods of M in N are strongly equivalent.

Lemma 7.4.3. *Suppose that M is compact. Then any closed tubular neighbourhood $\phi : B \longrightarrow N$ can be extended to a tubular neighbourhood $\phi' : E \longrightarrow N$.*

Proof. Given a closed tubular neighbourhood $\phi : B \longrightarrow N$, define an isotopy $h : B \times \mathbb{R} \longrightarrow N$ by

$$h_t(p,v) = \phi\Big(p, \Big(1 - \frac{\mathcal{B}(2t)}{2}\Big)v\Big),$$

where $\mathcal{B}(t)$ is a bump function, v belongs to the fibre B_p, and $p \in M$, $t \in \mathbb{R}$. Then $h_0 = \phi$, and $h_{1/2}(p,v) = \phi(p, v/2)$. Since B is compact, the isotopy h_t can be covered by a diffeotopy $k_t : N \longrightarrow N$ such that $h_t = k_t \circ \phi$ for all $t \in \mathbb{R}$.

Let $\lambda : \mathbb{R} \longrightarrow \mathbb{R}$ be a smooth map such that $\lambda(t) < 1$ and $d\lambda(t)/dt > 0$ for all $t \in \mathbb{R}$, and $\lambda(t) = t/2$ for $0 \leq t \leq 1$. For example, λ may be taken to be the function given by

$$\lambda(t) = \frac{1}{2}\int_0^t [1 + (e^{-x} - 1)\cdot \mathcal{B}(x-1)]dx.$$

Now use λ to extend the embedding $h_{1/2} : B \longrightarrow N$ to a tubular neighbourhood $\psi : E \longrightarrow N$ in the following way.

$$\psi(p,v) = \phi\Big(p, \lambda(\|v\|)\cdot \frac{v}{\|v\|}\Big).$$

Then, $\phi' = k_{1/2}^{-1} \circ \psi : E \longrightarrow N$ is the required tubular neighbourhood. Note that, if $\|v\| \leq 1$, then $\lambda(\|v\|) = (1/2)\|v\|$, and so $\psi|B = h_{1/2} = k_{1/2} \circ \phi$, and $\phi'|B = \phi$. ☐

Theorem 7.4.4 (Tubular neighbourhood theorem). *Let $\phi_1 : E_1 \longrightarrow N$ and $\phi_2 : E_2 \longrightarrow N$ be two tubular neighbourhood of M in N. Then there is a bundle map $\lambda : E_1 \longrightarrow E_2$ and an isotopy from $\phi_2 \circ \lambda$ to ϕ_1 which is fixed on the zero-section of E_1.*

Proof. It is sufficient to prove the theorem under the assumption that $\phi_1(E_1) \subset \phi_2(E_2)$. Because, if this has been proved, then, since the embedding of M in $\phi_1(E_1) \cap \phi_2(E_2)$ admits a tubular neighbourhood $\phi_3 : E_3 \longrightarrow N$ with $\phi_3(E_3) \subset \phi_1(E_1) \cap \phi_2(E_2)$, there would exist bundle maps $\lambda_1 : E_1 \longrightarrow E_3$ and $\lambda_2 : E_2 \longrightarrow E_3$ such that ϕ_1 is isotopic to $\phi_3 \circ \lambda_1$, and ϕ_2 is isotopic to $\phi_3 \circ \lambda_2$, whence ϕ_1 is isotopic to $\phi_2 \circ \lambda_2^{-1}\lambda_1$.

So suppose that $\phi_1(E_1) \subset \phi_2(E_2)$. Then the map $\psi = \phi_2^{-1} \circ \phi_1 : E_1 \longrightarrow E_2$ is defined, and is an embedding. Consider the family of embeddings $\psi_t : E_1 \longrightarrow E_2$, $0 < t \leq 1$, given by

$$\psi_t(u) = t^{-1}\cdot \psi(t \cdot u),\ u \in E_1,$$

which involves scalar multiplications by t and t^{-1} in the fibres. Then $\psi_1 = \psi$. We shall show that the definition of ψ_t can be extended to $t = 0$, and so ψ_0 can be taken as the bundle map λ of the theorem, and then $\phi_2 \circ \psi_t$ will be the required isotopy between $\phi_2 \circ \lambda$ and ϕ_1.

In terms of local coordinates $x = (x_1, \dots, x_n)$ in M, and coordinates $v = (v_1, \dots, v_k)$, $w = (w_1, \dots, w_k)$ in the fibres of E_1, E_2 respectively, we may write

$$\psi(x,v) = (\alpha(x,v), \beta(x,v)).$$

Then

$$\psi_t(x, v) = (\alpha(x, tv), t^{-1}\beta(x, tv)),$$

and we have $\alpha(x, 0) = x$, and $\beta(x, 0) = 0$, since ψ takes the zero-section of E_1 into that of E_2. Now treating β as a function of v with x as parameter, and using Lemma 3.1.3 to each component of β, we may find smooth maps $\beta_i : \mathbb{R}^k \longrightarrow \mathbb{R}^k$, where $\beta_i(x, 0) = (\partial\beta/\partial v_i)(0)$, $i = 1, \ldots, k$, such that

$$\beta(x, v) = \sum_{i=1}^{k} v_i \beta_i(x, v).$$

Then $t^{-1}\beta(x, tv) = \sum v_i \beta_i(x, tv)$, and therefore

$$\psi_t(x, v) = \Big(\alpha(x, tv), \sum v_i \beta_i(x, tv)\Big),$$

where the left hand side is smooth also at $t = 0$. Therefore the map $\Psi : E_1 \times I \longrightarrow E_2 \times I$ defined by ψ_t is smooth. This will be an isotopy if the Jacobian of ψ_t is non-singular everywhere. This is true if $t \neq 0$, because ψ is an embedding, and scalar multiplications by t and t^{-1} are diffeomorphisms. For $t = 0$, we have

$$\psi_0(x, v) = \Big(x, \sum v_i \beta_i(x, 0)\Big) = \Big(x, \sum v_i \frac{\partial\beta}{\partial v_i}(0)\Big).$$

Therefore ψ_0 induces a linear map on each fibre with matrix $(\partial\beta_j/\partial v_i) = (\partial w_j/\partial v_i)$, which is also the matrix of the partial derivatives of ψ on the zero-section of E_1. This matrix is non-singular, since ψ is an embedding. Therefore ψ_0 maps each fibre isomorphically onto a fibre, and hence it is a bundle map. We may therefore take $\psi_0 = \lambda$. The above arguments also show that Ψ is an isotopy. This completes the proof. □

◊ **Exercise 7.2.** Let M be a submanifold of N, and $f : M \longrightarrow N$ be an embedding. Let $\phi : E \longrightarrow N$ be a tubular neighbourhood of $f(M)$ in N, and E_1 be the pull-back f^*E. Let $\overline{f} : E_1 \longrightarrow E$ be the canonical map of the pull-back, and $\phi_1 = \phi \circ \overline{f} : E_1 \longrightarrow N$. We may call ϕ_1 a tubular neighbourhood of the embedding f. Let $\phi_2 : E_2 \longrightarrow N$ be a tubular neighbourhood of M in N.

Then show that the tubular neighbourhoods ϕ_1 and ϕ_2 are equivalent, that is, there is a bundle map $\lambda : E_1 \longrightarrow E_2$ and an isotopy from $\phi_2 \circ \lambda$ to ϕ_1 which is fixed on the zero-section of E_1.

Theorem 7.4.5. *Let $\phi_1 : E_1 \longrightarrow N$ and $\phi_2 : E_2 \longrightarrow N$ be two tubular neighbourhood of M in N, where the vector bundles E_1 and E_2 have structure group $O(m-n)$. Then the conclusion of Theorem* 7.4.4 *holds with λ an $O(m-n)$-bundle map.*

PROOF. It is sufficient to show that any bundle map $\psi : E_1 \longrightarrow E_2$ is isotopic to an $O(m-n)$-bundle map.

As before, ψ is given in local coordinates by $\psi(x, v) = (x, w)$, where the components of w are $w_i = \sum a_{ij}(x)v_j$. Applying the Gram-Schmidt orthonormalisation process (§5.6)to the row vectors of the matrix $A = (a_{ij})$ of ψ, we may write $A = BC$, where $B = (b_{ij})$ is a triangular matrix with $b_{ij} = 0$ for $i < j$ and $b_{ii} > 0$, and C is an orthogonal matrix. Then none of the matrices $tI + (1-t)B$, for $t \in \mathbb{R}$ and I (the identity matrix), is singular, because each matrix is triangular with non-zero diagonal entries. We have therefore an isotopy $f_t : E_1 \longrightarrow E_2$ given by

$$f_t(x, v) = (x, [tI + (1-t)B]C \cdot v),$$

where v is written as a column vector. Then $f_0 = \psi$, and f_1 takes an orthogonal basis to another, so is an $O(m-n)$-bundle map. □

Theorem 7.4.6. *Let $\phi_1 : B_1 \longrightarrow N$ and $\phi_2 : B_2 \longrightarrow N$ be closed tubular neighbourhoods of M in N. Then there is a bundle map $\lambda : B_1 \longrightarrow B_2$ such that $\phi_2 \circ \lambda$ is isotopic to ϕ_1 by an isotopy which fixes the zero-section.*

Proof. By Lemma 7.4.3, ϕ_1, ϕ_2 extend to tubular neighbourhoods $\overline{\phi}_1 : E_1 \longrightarrow N$, $\overline{\phi}_2 : E_2 \longrightarrow N$. By Theorem 7.4.5, there is an $O(m-n)$-bundle map $\overline{\lambda} : E_1 \longrightarrow E_2$ such that $\overline{\phi}_2 \circ \overline{\lambda}$ is isotopic to $\overline{\phi}_1$. Since $\overline{\lambda}$ maps B_1 into B_2, we can take λ as the restriction of $\overline{\lambda}$. □

Theorem 7.4.7 (Closed tubular neighbourhood theorem). *If M is a compact submanifold of N, then any two closed tubular neighbourhoods of M in N are strongly equivalent.*

Proof. By Theorem 7.3.3, the isotopy constructed in Theorem 7.4.6 is strong. □

Theorem 7.4.8 (Collar neighbourhood theorem). *Any two collar neighbourhoods of ∂M in M are strongly equivalent, provided ∂M is compact.*

Proof. The proof imitates the above arguments closely. However, note that in this case, Theorem 7.4.5 becomes trivial. Because, the normal bundle of ∂M is orientable (being one-dimensional), and so its structure group $O(1)$ reduces to the trivial group $SO(1) = \{1\}$. □

Remark 7.4.9. In the definition of equivalence and strong equivalence of collar neighbourhoods (which we have not stated explicitly, but is analogous to Definition 7.4.1), we may take $\lambda = \text{Id}$. Because $\lambda : M \times [0, 1] \longrightarrow M \times [0, 1]$ is of the form $\lambda(x, t) = (x, \mu(x, t))$, where $\mu : M \times [0, 1] \longrightarrow [0, 1]$ is a smooth map smoothly homotopic to the projection $(x, t) \mapsto t$ (the interval $[0, 1]$ being convex).

Theorem 7.4.10 (Disk theorem). *Let N be a connected manifold of dimension n, perhaps with boundary, and $f_1, f_2 : D^n \longrightarrow N$ be two embeddings of the disk D^n into N as submanifolds with boundary. Then f_1 and f_2 are strongly isotopic, unless N is orientable and f_1, f_2 have opposite orientations* (*that is, one of f_1 and f_2 is orientation preserving and the other is orientation reversing*).

PROOF. Consider the points $p_1 = f_1(0)$ and $p_2 = f_2(0)$ as submanifolds of dimension zero in Int N. Since Int N is connected, p_1 and p_2 can be joined by a smooth path lying in Int N. The path is an isotopy, and it is strong by the isotopy extension theorem. This means that there is a diffeotopy of N mapping p_1 onto p_2. Therefore we may suppose that $p_1 = p_2 = p$, say. Now f_1 and f_2 are closed tubular neighbourhoods of p, and so, by Theorem 7.4.5, there is an orthogonal transformation λ of D^n such that f_1 and $f_2 \circ \lambda$ are strongly isotopic.

Now if $\lambda \in SO(n)$, then f_2 is isotopic, and hence strongly isotopic to $f_2 \circ \lambda$ (λ being homotopic to Id), so we get the theorem in this case, since strong isotopy is an equivalence relation. If $\lambda \notin SO(n)$, and N is orientable, then we have the case excluded by the theorem, because f_1 and f_2 have opposite orientations, and so f_1, f_2 cannot be homotopic. If $\lambda \notin SO(n)$, and N is non-orientable, then through any point of N there is a smooth orientation reversing path in N (see Exercise 5.12, p.157). We can therefore change the sign of the determinant of λ by a strong isotopy round an orientation reversing loop on p. This leads us to the theorem in this case. □

7.5. Manifolds with corner and straightening them

Let M be a manifold with boundary ∂M and corner ΛM. Then straightening the corner ΛM is a construction of a smooth manifold N with boundary and a homeomorphism $h : M \longrightarrow N$ which is a diffeomorphism except on ΛM. Moreover, the construction gives N up to diffeomorphism. This process will be very useful later where in some cases corner may be ignored by the results of this section.

We shall first construct an open neighbourhood of ΛM in M which is analogous to the collar neighbourhood of the boundary.

In a coordinate neighbourhood U in M of a point of ΛM with coordinates $x_1, \ldots, x_n$, the corner $\Lambda U = U \cap \Lambda M$ separates the boundary $\partial U = U \cap \partial M$ into two parts $\{x_2 = 0, x_1 \geq 0\}$ and $\{x_1 = 0, x_2 \geq 0\}$. In general, we define

Definition 7.5.1. The corner ΛM **separates the boundary** ∂M if there exists two submanifolds $\partial_1 M$ and $\partial_2 M$ of ∂M with boundary

$$\partial(\partial_1 M) = \partial(\partial_2 M) = \partial_1 M \cap \partial_2 M = \Lambda M$$

such that $\partial M = \partial_1 M \cup \partial_2 M$.

Suppose ΛM separates ∂M into submanifolds $\partial_1 M$ and $\partial_2 M$. We endow M with a Riemannian metric d. This induces metric d_i on $\partial_i M$, $i = 1, 2$. We use the method of Proposition 7.2.8 to construct a metric on $\partial_i M$ adapted to its boundary ΛM, and thereby obtain a metric on M which is product metric on a collar neighbourhood $U_i = \Lambda M \times [0,1]$ of ΛM in $\partial_i M$, $i = 1, 2$, and which is d outside of U_1 and U_2.

Since ΛM is a submanifold of M of codimension 2, its normal bundle has fibre dimension 2. Consider a fibre bundle over ΛM with fibre consisting of inward-pointing normal vectors of ΛM. Such normal vectors of ΛM are represented in a coordinate neighbourhood in M with coordinates $x_1, \ldots, x_n$ by tangent vectors $\sum_{i=1}^n \lambda_i \cdot \partial/\partial x_i$ of M with $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$. The vectors corresponding to $\lambda_1 = 0$ (resp. $\lambda_2 = 0$) are tangent to $\partial_2 M$ (resp. $\partial_1 M$). Then using the exponential map outside the zero section, and the method of Theorem 7.1.5, we obtain a homeomorphism of the fibre bundle onto a neighbourhood U of ΛM in M, which is a diffeomorphism on $U - \Lambda M$, but not differentiable on ΛM. In particular, if the fibre bundle is trivial, then we have a homeomorphism of U onto $\Lambda M \times \mathbb{R}_+ \times \mathbb{R}_+$. We have proved the following proposition.

Proposition 7.5.2. *If M is a manifold with corner ΛM which separates the boundary ∂M into submanifolds $\partial_1 M$ and $\partial_2 M$, and if the normal bundle of ΛM is trivial, then there is an open neighbourhood U of ΛM in M and a homeomorphism $\phi : U \longrightarrow \Lambda M \times \mathbb{R}_+ \times \mathbb{R}_+$ such that $\phi(x) = (x, 0, 0)$ for $x \in \Lambda M$, and such that ϕ is a diffeomorphism of $U - \Lambda M$ onto $\Lambda M \times \mathbb{R}_+ \times \mathbb{R}_+ - \Lambda M \times \{0\} \times \{0\}$.*

The proposition provides a guide line for the formulation of the problem of straightening corner, which is summarised in the following theorem.

Theorem 7.5.3. *Let M be a smooth $(n-2)$-submanifold without boundary of a topological n-manifold X such that*

(1) *M is a closed subset of X,*
(2) *$X - M$ is a smooth manifold,*
(3) *there is an open neighbourhood U of M in X and a homeomorphism ϕ of U onto $M \times \mathbb{R}_+ \times \mathbb{R}_+$ with $\phi(x) = (x, 0, 0)$ for $x \in M$ such that ϕ is a diffeomorphism of $U - M$ onto $M \times \mathbb{R}_+ \times \mathbb{R}_+ - M \times \{0\} \times \{0\}$.*

Then X can be given a unique smooth structure which agrees with the smooth structures on M and $X - M$.

Proof. There is a homeomorphism $h : \mathbb{R}_+ \times \mathbb{R}_+ \longrightarrow \mathbb{R} \times \mathbb{R}_+$ such that h maps $\mathbb{R}_+ \times \mathbb{R}_+ - \{(0,0)\}$ diffeomorphically onto $\mathbb{R} \times \mathbb{R}_+ - \{(0,0)\}$. For example, h is given in polar coordinates by $h(r, \theta) = (r, 2\theta)$, where $0 \leq \theta \leq \pi/2$.

Let $h' : M \times \mathbb{R}_+ \times \mathbb{R}_+ \longrightarrow M \times \mathbb{R} \times \mathbb{R}_+$ be the homeomorphism given by $h' = \mathrm{Id}_M \times h$. Transporting the product smooth structure on $M \times \mathbb{R} \times \mathbb{R}_+$ to U by means of the homeomorphism $h' \circ \phi$, we get a unique smooth structure on U so that $h' \circ \phi$ is a diffeomorphism. This smooth structure extends the smooth structure on M and agrees with the smooth structure on $X - M$ on the intersection $(X - M) \cap U = U - M$. Thus X has a unique smooth structure which agrees with the smooth structures on M and $X - M$. $\square$

This differentiable structure on X is said to have been obtained by **straightening the angle or corner**.

Example 7.5.4. Consider a region M of $\mathbb{R}^2$ bounded by a closed curve which is smooth everywhere, except at a point p. Then M can be given a smooth structure which agrees with the smooth structure on $M - \{p\}$, by straightening the boundary at p.

Take another point q on the boundary ∂M, and consider it as the union of two curves $\partial_1 M$ and $\partial_2 M$ from q to p. Then by Proposition 7.5.2, there is an open neighbourhood U of p in M and a homeomorphism $\phi : U \longrightarrow \{p\} \times \mathbb{R}_+ \times \mathbb{R}_+$ such that $\phi(p) = (p, 0, 0)$ satisfying the condition (3) of Theorem 7.5.3. Therefore we obtain a smooth structure on M as stated above.

Example 7.5.5. Let us consider the product of two manifolds with boundary M_1 and M_2. Let $X = M_1 \times M_2$ and $M = M_1 \times M_2 - \Lambda(M_1 \times M_2)$. By Theorem 7.2.2, there exist open neighbourhoods U_1 and U_2 of ∂M_1 and ∂M_2 in M_1 and M_2 respectively, and diffeomorphisms $\phi_1 : U_1 \longrightarrow \partial M_1 \times \mathbb{R}_+$ and $\phi_2 : U_2 \longrightarrow \partial M_2 \times \mathbb{R}_+$ such that $\phi_1(x_1) = (x_1, 0)$ and $\phi_2(x_2) = (x_2, 0)$ for $x_1 \in \partial M_1$ and $x_2 \in \partial M_2$. Let $U = U_1 \times U_2$. Then $\phi = \phi_1 \times \phi_2$ is a homeomorphism of U onto $\Lambda(M_1 \times M_2) \times \mathbb{R}_+ \times \mathbb{R}_+$ with the properties (3) of the above theorem. We then obtain a differential structure on $M_1 \times M_2$ by straightening the angle.

$\Diamond$ **Exercise 7.3.** Show that the manifold obtained by straightening the corner of $D^n \times D^m$ is diffeomorphic to the disk D^{n+m}.

The reverse of the process of 'straightening corner' is the process of **introduction of corner**. If N is a manifold with boundary ∂N, and M is a submanifold of ∂N without boundary and of codimension one, then we can construct a tubular neighbourhood of M in ∂N, and then redefine the smooth structure of N so as to have corner along M. The resulting manifold is unique up to diffeomorphism, and if we straighten its corner, then we get back the original manifold N. To describe the process more clearly, consider a Riemannian structure on N, and induced structures on ∂N and M. Then applying the exponential map to sufficiently small normal vectors to M in ∂N, we obtain an embedding $\phi : M \times (-\epsilon, \epsilon) \longrightarrow \partial N$, where ϵ is small and positive. Combining this with a collar neighbourhood $\psi : \partial N \times [0, 1] \longrightarrow N$, it follows that there is a neighbourhood U of M in N and a topological embedding

$$U \longrightarrow M \times \mathbb{R} \times \mathbb{R}_+ \xrightarrow{\mathrm{Id} \times h} M \times \mathbb{R}_+ \times \mathbb{R}_+,$$

where h is the homeomorphism described in the proof of Theorem 7.5.3. This embedding introduces corners along M, keeping the smooth structure on $U - M$ unaltered. We have therefore the following theorem.

Theorem 7.5.6. *If M is a closed submanifold of ∂N of codimension one, then we may introduce corner along M in an essentially unique way. We may retrieve N by straightening the corner.*

7.6. Construction of manifolds by gluing process

In this section we construct manifolds by gluing together manifolds along their boundaries.

Definition 7.6.1. Let M_1 and M_2 be manifolds with boundaries $\partial M_1 = V_1$ and $\partial M_2 = V_2$. Let $h : V_1 \longrightarrow V_2$ be a diffeomorphism. Let N be the quotient space of the disjoint union $M_1 \cup M_2$ by the identification where $x \in V_1$ equals $h(x) \in V_2$ for all $x \in V_1$. This is the adjunction space $N = M_1 \cup_h M_2$, and we say that N is obtained by **gluing M_1 and M_2 along V_1 by h.**

Theorem 7.6.2. *There is a smooth structure on N which induces the given smooth structures on M_1 and M_2. Moreover, if V_1 is compact, then the smooth structure on N is unique up to diffeomorphism.*

Proof. Let $\pi : M_1 \cup M_2 \longrightarrow N$ denote the identification map. We take collar neighbourhoods

$$\phi_1 : V_1 \times [0,1] \longrightarrow M_1, \text{ and } \phi_2 : V_2 \times [0,1] \longrightarrow M_2$$

of V_1 and V_2 in M_1 and M_2. Define $\phi : V_1 \times [-1,1] \longrightarrow N$ by

$$\phi(x,t) = \begin{cases} \pi \circ \phi_1(x,t) & \text{if} \quad t \geq 0, \\ \pi \circ \phi_2(h(x), -t) & \text{if} \quad t \leq 0. \end{cases}$$

The definitions agree at $t = 0$, because V_1 and V_2 are identified by h. It is also clear that ϕ is a topological embedding. Let $j_1 : M_1 \longrightarrow N$ and $j_2 : M_2 \longrightarrow N$ be the inclusion maps (actually these are restrictions of the identification map π). Then a smooth structure can be given to N by transporting the smooth structures on M_1, M_2, and $V_1 \times [-1,1]$ by j_1, j_2, and ϕ respectively. This gives an atlas of N consisting of coordinate charts of M_1, M_2, and $V_1 \times [-1,1]$; they agree on the overlaps. Alternatively, we may get the smooth structure on N by defining a function $f : N \longrightarrow \mathbb{R}$ to be smooth provided $f \circ \pi$ is smooth on $M_1 \cup M_2$ (disjoint union), and $f \circ \phi$ is smooth on $V_1 \times [-1,1]$.

The second part is true, because the only objects of arbitrary choice involved in the definition of the smooth structure on N are the collar neighbourhoods ϕ_1 and ϕ_2, and these are unique up to diffeomorphisms of M_1 and M_2 respectively provided V_1 is compact, by the Collar Neighbourhood Theorem 7.4.8. $\square$

Complement 7.6.3. The proof remains unaffected if M_1 (resp. M_2) has several boundary components, and V_1 (resp. V_2) is a union of some of these components. In this case, N will be a manifold with boundary, which consists of the remaining boundary components of M_1 and M_2.

The following proposition may be used to show the uniqueness of the gluing process, namely, any smooth structure on N, which induce the original smooth structure on M_1 and M_2, is equivalent to the smooth structure constructed above by pasting together collar neighbourhoods of V_1 and V_2.

Proposition 7.6.4. *Suppose, for $i = 1, 2$, that W_i, M_i, N_i, and V_i are manifolds, where $W_i = M_i \cup N_i$, $M_i \cap N_i = \partial M_i = \partial N_i = V_i$, and V_i is compact. Suppose that $f : W_1 \longrightarrow W_2$ is a homeomorphism which maps M_1 and N_1 diffeomorphically onto M_2 and N_2 respectively. Then there is a diffeomorphism $g : W_1 \longrightarrow W_2$ such that $g(M_1) = M_2$, $g(N_1) = N_2$, and $g|V_1 = f|V_1$.*

Proof. For each $i = 1, 2$, let $\phi_i : V_i \times (-1, 1) \longrightarrow U_i \subset W_i$ be a tubular neighbourhood of V_i in W_i with Image $\phi_i = U_i$. This gives collar neighbourhoods α_i of V_i in M_i, and β_i of V_i in N_i by restriction

$$\alpha_i = \phi_i|\phi_i^{-1}(U_i \cap M_i) : V_i \times [0, 1) \longrightarrow M_i,$$

$$\beta_i = \phi_i|\phi_i^{-1}(U_i \cap N_i) : V_i \times (-1, 0] \longrightarrow N_i.$$

Denote by $f[\alpha_1]$ the collar neighbourhood $f \circ \alpha_1 \circ (f^{-1} \times \mathrm{Id})$ of V_2 in M_2:

$$V_2 \times [0, 1) \xrightarrow{f^{-1}\times\mathrm{Id}} V_1 \times [0, 1) \xrightarrow{\alpha_1} M_1 \xrightarrow{f} M_2.$$

By Theorem 7.4.8, the collar neighbourhoods $f[\alpha_1]$ and α_2 are strongly equivalent, since V_2 is compact. So $f[\alpha_1]$ and α_2 are isotopic, and the isotopy is covered by a diffeotopy Γ_t of M_2. We have then $\Gamma_1 \circ f[\alpha_1] = \alpha_2$. This means, after writing $g_1 = \Gamma_1 \circ f$, that we have a commutative diagram

$$\begin{tikzcd}
V_1 \times [0,1) \arrow[r, "\alpha_1"] \arrow[d, "f\times\mathrm{Id}"'] & M_1 \arrow[d, "g_1"] \\
V_2 \times [0,1) \arrow[r, "\alpha_2"] & M_2
\end{tikzcd}$$

Note that $g_1|V_1 = f|V_1$. Similarly, working with the collar neighbourhoods β_i, we may obtain a diffeomorphism $g_2 : N_1 \longrightarrow N_2$, which makes the following diagram commutative

$$\begin{tikzcd}
V_1 \times (-1,0] \arrow[r, "\beta_1"] \arrow[d, "f\times\mathrm{Id}"'] & N_1 \arrow[d, "g_2"] \\
V_2 \times (-1,0] \arrow[r, "\beta_2"] & N_2
\end{tikzcd}$$

where $g_2|V_1 = f|V_1$. A conjunction of these diagrams gives rise to a commutative diagram

$$\begin{tikzcd}
V_1 \times (-1,1) \arrow[r, "\phi_1"] \arrow[d, "f\times\mathrm{Id}"'] & M_1 \cup N_1 \arrow[d, "g_1\cup g_2"] \\
V_2 \times [0,1) \arrow[r, "\phi_2"] & M_2 \cup N_2
\end{tikzcd}$$

Clearly the map $g = g_1 \cup g_2$ is smooth, and it is the required diffeomorphism. $\square$

Complement 7.6.5. The proof remains valid even if we allow $V_i = M_i \cap N_i$ to be a compact submanifold of both ∂M_i and ∂N_i for each $i = 1, 2$.

Proposition 7.6.6. *Let (P, A), (P', A'), and (Q, B) be manifold pairs, where A, A', and B are compact submanifolds of ∂P, $\partial P'$, and ∂Q respectively. let $h : B \longrightarrow A$ and $h' : B \longrightarrow A'$ be diffeomorphisms such that the diffeomorphism $h' \circ h^{-1} : A \longrightarrow A'$ extends to a diffeomorphism $\phi : P \longrightarrow P'$. Then there is a diffeomorphism $\psi : Q \cup_h P \longrightarrow Q \cup_{h'} P'$.*

PROOF. This is a special case of Proposition 7.6.4, where $M_1 = M_2 = Q$, $N_1 = P$, $N_2 = P'$, $V_1 = B = h(B)$, $V_2 = B = h'(B)$, $W_1 = Q \cup_h P$, and $W_2 = Q \cup_{h'} P'$. Define $f : W_1 \longrightarrow W_2$ by $f = \phi$ on P and $f = \mathrm{Id}$ on Q. Then all the conditions of Proposition 7.6.4 are satisfied, and so there exists a diffeomorphism $\psi : W_1 \longrightarrow W_2$ such that $\psi(P) = P'$, $\psi(Q) = Q$ and $\psi|A = f|A$. □

The proposition may also be stated in a slightly different way:

Corollary 7.6.7. *If $h : B \longrightarrow P$ and $h' : B \longrightarrow P'$ are embeddings, and if there exists a diffeomorphism $\phi : P \longrightarrow P'$ such that $\phi \circ h = h'$, then $Q \cup_h P$ is diffeomorphic to $Q \cup_{h'} P'$.*

PROOF. $h' \circ h^{-1}$ extends to ϕ. □

Corollary 7.6.8. *If $h,\ h' : B \longrightarrow P$ are isotopic embeddings, then $Q \cup_h P$ is diffeomorphic to $Q \cup_{h'} P$.*

PROOF. An isotopy $\psi_t : B \longrightarrow P$ from h to h' can be covered by a diffeotopy $\phi_t : P \longrightarrow P$ so that $\phi_t \circ h = \psi_t$. Therefore the proof follows from the last corollary. □

Definition 7.6.9. The **double of a manifold** M is the manifold obtained by gluing M to itself along ∂M by the identity map.

The inverse of gluing process is the cutting process. We consider the simplest case.

Definition 7.6.10. Suppose that N is a manifold (which may have boundary), and V is a submanifold of codimension 1 which separates N, in the sense that $N - V$ has just two components with closures M_1 and M_2, and $\partial M_1 = \partial M_2 = V$ (so that if N is cut along V, it will fall into two pieces) (cf. Definition 7.5.1). In this case, we say that M_1 and M_2 are obtained by **cutting N along V**.

Note that M_1 and M_2 are uniquely determined by N and V (compactness of V is not needed here), and they have the induced structure of a smooth manifold. The following result shows that gluing and cutting are inverse operations.

Proposition 7.6.11. *If N is obtained by gluing M_1 and M_2 along $V_1 = \partial M_1$, and $\pi : M_1 \cup M_2 \longrightarrow N$ is the identification map, then we get back M_1 and M_2 by cutting N along $\pi(V_1)$. Conversely, if N and V are connected, V is compact, and V separates N into submanifolds M_1 and M_2, then we may recover N by gluing M_1 and M_2 along V.*

PROOF. The first part is obvious from the definition of gluing. For the second part, take a tubular neighbourhood $\phi : V \times [-1, 1] \longrightarrow N$ of V in N. By restriction, this gives collar neighbourhoods of V in M_1 and M_2. Using these collar neighbourhoods in the gluing process as described in Theorem 7.6.2, we certainly recover N. □

As an application of the gluing process, we now describe the process of attaching a handle.

Definition 7.6.12. Let M be a manifold of dimension $(n - 1)$, and $f : S^{r-1} \times D^{n-r} \longrightarrow M - \partial M$ be a smooth embedding. Then $M[f]$ denotes the manifold obtained by deleting from M the interior of the image $f(S^{r-1} \times D^{n-r})$ and then gluing in $D^r \times S^{n-r-1}$ by $g = f|S^{r-1} \times S^{n-r-1}$, that is,

$$M[f] = [M - \text{Int } f(S^{r-1} \times S^{n-r-1})] \cup_g [D^r \times S^{n-r-1}].$$

This operation of replacing Int (Image f) by $D^r \times S^{n-r-1}$ is called a **surgery** of type $(n, n - r)$, and $M[f]$ is called a **spherical modification** of M of type $(n, n - r)$ We say that $M[f]$ is obtained from M by **attaching a handle**.

As describe before, $M[f]$ has a unique smooth structure which induces the original smooth structures on $M - \text{Int } f(S^{r-1} \times S^{n-r-1})$ and $D^r \times S^{n-r-1}$.

Proposition 7.6.13. *If $f_0, f_1 : S^{r-1} \times D^{n-r} \longrightarrow M - \partial M$ are isotopic embeddings, then $M[f_0]$ is diffeomorphic to $M[f_1]$.*

PROOF. By the isotopic extension theorem, there is a diffeomorphism $\phi : M \longrightarrow M$ such that $\phi \circ f_0 = f_1$. Let us write $V_i = M - \text{Int } f_i(S^{r-1} \times D^{n-r})$, for $i = 0, 1$. Then $\partial V_i = f_i(S^{r-1} \times S^{n-r-1})$, and $g_i = f_i^{-1}|\partial V_i : \partial V_i \longrightarrow S^{r-1} \times S^{n-r-1}$ is an embedding. We have

$$M[f_0] = V_0 \cup_{g_0} (D^r \times S^{n-r-1}) \text{ and } M[f_1] = V_1 \cup_{g_1} (D^r \times S^{n-r-1}).$$

Note that the diffeomorphism $g_1^{-1} \circ g_0 : \partial V_0 \longrightarrow \partial V_1$ extends to the diffeomorphism $\phi : V_0 \longrightarrow V_1$. Therefore the proof follows from Proposition 7.6.6. □

As second application of the gluing process, we describe the connected sum of two manifolds.

Definition 7.6.14. Let M_1 and M_2 be connected n-manifolds without boundaries, and $f_1 : D^n \longrightarrow M_1$, $f_2 : D^n \longrightarrow M_2$ smooth embeddings. Then the **connected sum** $M_1 \# M_2$ is the manifold obtained by gluing $M_1 -$ Int (Image f_1) and $M_2 -$Int (Image f_2) along the boundary $f_1(S^{n-1})$ by $f_2 \circ f_1^{-1}$

Note that the connected sum $M^n \# S^n$ is diffeomorphic to M^n, because to construct $M^n \# S^n$, we simply delete an n-disk from M^n and replace it by another good one.

Theorem 7.6.15. *The connected sum $M_1 \# M_2$ is determined uniquely up to diffeomorphism by M_1 and M_2, unless M_1, M_2 are both oriented, in which case there are two determinations of $M_1 \# M_2$.*

PROOF. By the Disk Theorem (7.4.10), each of the embeddings f_1 and f_2 is determined uniquely up to strong equivalence, and a possible change of orientation. Then, given f_1 and f_2, the result of gluing is unique up to diffeomorphism, by Theorem7.6.2. This proves the theorem, provided we do not give orientation a consideration. Note that if f_1 and f_2 are replaced by $f_1 \circ g$ and $f_2 \circ g$, where g is an orientation reversing diffeomorphism of D^n, then the connected sum remains unaltered.

Now, if both M_1 and M_2 are non-orientable, the theorem follows. If only one of them, say M_2, is orientable, then also we have the uniqueness of the connected sum, because it does not matter whether f_2 is orientation preserving or reversing by the above possibility of simultaneous reversal. If both M_1 and M_2 are orientable, then the result of gluing has two possible cases. □

◊ **Exercise 7.4.** Show that the connected sum $M_1 \# M_2$ is orientable if and only if both M_1 and M_2 are orientable.

The connected sum can also be defined for manifolds with boundary. However, this operation is somewhat different. Let M_1, M_2 be connected n-manifolds with connected boundaries, and let $f_i : D^{n-1} \longrightarrow \partial M$, be embeddings, $i = 1, 2$. Introduce a corner along $f_i(S^{n-2})$, and then glue the $f_i(D^{n-1})$ together by $f_2 \circ f_i^{-1}$. The resulting manifold is denoted by $M_1 + M_2$. The following results are immediate.

Proposition 7.6.16. *The manifold $M_1 + M_2$ is determined uniquely by M_1 and M_2 up to diffeomorphism, unless ∂M_1 and ∂M_2 are both oriented, in which case we have two determination of $M_1 + M_2$.*

PROOF. Again the proof uses the Disk Theorem exactly as for Theorem 7.6.15. □

◊ **Exercises 7.5.** Show that

(a) $M + D^n \cong M,\ n = \dim M$,
(b) $\partial(M_1 + M_2) = \partial M_1 \# \partial M_2$.

CHAPTER 8

SPACES OF SMOOTH MAPS

Many problems of differential topology can be formulated as problems about spaces of smooth maps between manifolds, and their associated jet spaces. For example, the problems of transversality which we considered earlier can be rephrased in terms of jets. This chapter is devoted to topics and results related to spaces of maps and jet spaces. The main aim is to prove Thom's transversality theorem, and find some of its applications. One of the results gives a better perspective on Whitney's embedding theorem which we have proved already in Chapter 2.

8.1. Spaces of Jets

We use multi-indices to denote higher order partial derivatives of a function $f : U \longrightarrow \mathbb{R}$, where U is an open subset of $\mathbb{R}^n_+$ with coordinates $x_1, \ldots, x_n$. A multi-index is an n-tuple of non-negative integers $\alpha = (\alpha_1, \ldots, \alpha_n)$, and its order is the integer $|\alpha| = \alpha_1 + \cdots + \alpha_n$. Then we write $\partial_\alpha f$ for the partial derivative of f of order $|\alpha|$

$$\frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} f.$$

If $\alpha = (0, \ldots, 0)$, then $\partial_\alpha f = f$. We denote the first order partial derivatives $\partial f / \partial x_i$ by $\partial_i f$. Recall that f is a C^r function, if $\partial_\alpha f : U \longrightarrow \mathbb{R}$ exists and is continuous for each multi-index α with $|\alpha| \leq r$. If the range of f is $\mathbb{R}^m$, and its components are $f_1, \ldots, f_m$, then its partial derivatives are maps from U to $\mathbb{R}^m$ given by

$$\partial_\alpha f = (\partial_\alpha f_1, \ldots, \partial_\alpha f_m).$$

Let $L(\mathbb{R}^n, \mathbb{R}^m)$ denote the vector space of linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$, and for non-negative integral values of k, $L^k(\mathbb{R}^n, \mathbb{R}^m)$ denote the following vector spaces: $L^0(\mathbb{R}^n, \mathbb{R}^m) = \mathbb{R}^m$, $L^1(\mathbb{R}^n, \mathbb{R}^m) = L(\mathbb{R}^n, \mathbb{R}^m)$, and

$$L^k(\mathbb{R}^n, \mathbb{R}^m) = L(\mathbb{R}^n, L^{k-1}(\mathbb{R}^n, \mathbb{R}^m)), \quad \text{for} \quad k \geq 1.$$

The first total derivative of a differentiable map $f : U \longrightarrow \mathbb{R}^m$, U open in $\mathbb{R}^n_+$, is a map $Df : U \longrightarrow L(\mathbb{R}^n, \mathbb{R}^m)$, where, for $x \in U$, the matrix of the linear map $Df(x)$ relative to standard bases is the Jacobian matrix of f at x, $(\partial_j f_i(x))$, $i = 1, \ldots, m$, $j = 1, \ldots, n$. The second total derivative

$$D^2 f : U \longrightarrow L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^m))$$

A. Mukherjee, *Differential Topology*, DOI 10.1007/978-3-319-19045-7_8

is the map $D^2f(x) = D(Df)(x)$. Writing $Df(x) = g(x) = (g_{ij}(x))$, where the component $g_{ij}(x)$ is $\partial_j f_i(x)$, $Dg(x)$ has the matrix $(\partial_k g_{ij}(x))$. Since $\partial_k g_{ij}(x) = \partial_k(\partial_j f_i)(x) = \partial_{(k,j)} f_i(x)$, this matrix contains all the second order partial derivatives of the components f_i of f. The k-th order total derivative of f

$$D^k f : U \longrightarrow L^k(\mathbb{R}^n, \mathbb{R}^m)$$

is defined inductively by $D^k f = D(D^{k-1}f)$. The range of $D^k f$ may be identified with the vector space $L(\mathbb{R}^n \times \cdots \times \mathbb{R}^n, \mathbb{R}^m)$ of k-multilinear maps from $\mathbb{R}^n \times \cdots, \times \mathbb{R}^n$ (k times) to $\mathbb{R}^m$ by means of a linear isomorphism λ given by $\lambda(\phi)(u_1, \ldots, u_k) = \phi(u_1)(u_2)\cdots(u_k)$, $\phi \in L^k(\mathbb{R}^n, \mathbb{R}^m)$, $u_i \in \mathbb{R}^n$. Then

$$D^k f : U \longrightarrow L(\mathbb{R}^n \times \cdots, \times \mathbb{R}^n, \mathbb{R}^m)$$

is given as follows. Let $u_i = (u_{ij}) \in \mathbb{R}^n$, $1 \leq i \leq k$, $1 \leq j \leq n$. Then the s-th component of $D^k f(x)(u_1, \ldots, u_k)$, for $s = 1, \ldots, m$, is

$$[D^k f(x)(u_1, \ldots, u_k)]_s = \sum \frac{\partial^k f_s(x)}{\partial x_{\beta_1} \cdots \partial x_{\beta_k}} \cdot u_{1,\beta_1} \cdots u_{k,\beta_k},$$

where the sum is over all permutations $(\beta_1, \ldots, \beta_k)$ of $(1, \ldots, n)$ with repetitions taken k at a time. The map f is C^r if $D^k f : U \longrightarrow L^k(\mathbb{R}^n, \mathbb{R}^m)$ is continuous for all $k \leq r$.

Let $f : U \longrightarrow \mathbb{R}$ be a smooth function, where U is open in $\mathbb{R}^n_+$ with coordinates $x = (x_1, \ldots, x_n)$. If $\alpha = (\alpha_1, \ldots, \alpha_n)$, let x^α denote $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, and $\alpha!$ denote the integer $\alpha_1! \cdots \alpha_n!$. Then the formal Taylor's series of f about a point $x_0 \in U$ is the infinite series $T_{x_0}(f) = \sum_\alpha a_\alpha x^\alpha$, where $a_\alpha = (1/\alpha!)\partial_\alpha f(x_0)$. The series may not be convergent. Of course, if f is analytic at x_0, the series converges to $f(x + x_0)$ in some neighbourhood of x_0, and it may be differentiated term by term any number of times in that neighbourhood and the coefficients a_α are then as given above. Let $T^r_{x_0}(f)$ denote the polynomial obtained by truncating the series $T_{x_0}(f)$ at the degree r term

$$T^r_{x_0}(f) = \sum_{|\alpha| \leq r} a_\alpha x^\alpha.$$

This is called the r-jet of f at x_0.

For a smooth map $f : U \longrightarrow \mathbb{R}^m_+$, U open in $\mathbb{R}^n_+$, with components $f_1, \ldots, f_m$, the r-jet of f at $x_0 \in U$ is given by

$$T^r_{x_0}(f) = (T^r_{x_0}(f_1), \ldots, T^r_{x_0}(f_m)).$$

It is convenient to look at the things in terms of the total derivatives. The formal Taylor's series of f at x_0 may also be written as

$$T_{x_0}(f) = \sum_{k=0}^{\infty} \frac{D^k f(x_0)(x^{(k)})}{k!},$$

where $x = (x_1, \ldots, x_n)$ and $x^{(k)} = (x, \ldots, x)$ (k times). Then the r-jet of f is obtained by truncating the series at the degree r term.

If $f : M \longrightarrow N$ is a smooth map between manifolds (possibly with boundary), then its partial and total derivatives are defined in terms of the derivatives of its local representative, for example, if $p \in M$, then

$$D^k f(p) = D^k(\psi \circ f \circ \phi^{-1})(\phi(p)),$$

where (U, ϕ) is a coordinate chart around p and (V, ψ) is a coordinate chart around $f(p)$.

With these preliminaries at hand, we now proceed to give a better definition of r-jet in the context of manifold.

Definition 8.1.1. Let $C^\infty(M, N)$ denote the set of all smooth maps $M \longrightarrow N$, f and $g \in C^\infty(M, N)$, and p be a point in M. Then f is said to have **contact of order** r with g at p, written $f \sim_r g$, if $D^k f(p) = D^k g(p)$ for every $k = 0, 1, \ldots, r$.

$\Diamond$ **Exercise 8.1.** (a) Show that $f \sim_r g$ at p if and only if $f(p) = g(p)$, and, with respect to some local coordinates at p and $f(p)$, all partial derivatives of order $\leq r$ of f and g agree at p.

(b) Conclude that the definition is independent of the choice of the local coordinate systems, by showing that if $f : \mathbb{R}^n_+ \longrightarrow \mathbb{R}^m_+$ is a smooth map such that $f(0) = 0$, and all partial derivatives of f of order $\leq r$ vanish at $0 \in \mathbb{R}^n$, and $\phi : R^n_+ \longrightarrow \mathbb{R}^n_+$, $\psi : \mathbb{R}^m_+ \longrightarrow \mathbb{R}^m_+$ are diffeomorphisms keeping 0 fixed, then $\psi \circ f \circ \phi$ has all partial derivatives of order $\leq r$ zero at 0.

The relation $f \sim_r g$ is an equivalence relation in the set $C^\infty(M, N)$. The equivalence class of f is denoted by $j^r_p(f)$, and it is called the **jet of order** r (or r**-jet**) of f with source p and target $f(p) = q$. Let $J^r_p(M, N)_q$ denote the set of all r-jets of maps $M \longrightarrow N$ with source p and target q. The set

$$J^r(M.N) = \cup_{(p,q) \in M \times N} J^r_p(M, N)_q \text{ (disjoint union)}$$

is called the set of r-jets of C^∞ maps $M \longrightarrow N$.

We have then the following maps.

- The **source map** $\sigma : J^r(M, N) \longrightarrow M$ given by $j^r_p(f) \mapsto p$.
- The **target map** $\tau : J^r(M, N) \longrightarrow N$ given by $j^r_p(f) \mapsto f(p)$.
- The r**-jet map** $j^r(f) : M \mapsto J^r(M, N)$ given by $p \mapsto j^r_p(f)$.

Note that $J^0(M, N) = M \times N$, and $j^0(f)(p) = (p, f(p))$. Therefore Image $j^0(f)$ $= \Gamma(f)$, the graph f. Also $J^r_p(M, N)_q = \sigma^{-1}(p) \cap \tau^{-1}(q)$.

Lemma 8.1.2. *Let $U \subset \mathbb{R}^n_+$ and $V \subset \mathbb{R}^m_+$ be open sets. Let $f_1, f_2 : U \longrightarrow V$ and $g_1, g_2 : V \longrightarrow \mathbb{R}^l_+$ be C^∞ maps such that $f_1 \sim_r f_2$ at $p \in U$ and $g_1 \sim_r g_2$ at $q = f_1(p) = f_2(p) \in V$. Then $g_1 \circ f_1 \sim_r g_2 \circ f_2$ at p.*

Proof. The proof is by induction on r. The case $r = 1$ follows, since we have by chain rule

$$\begin{aligned} D(g_1 \circ f_1)(p) &= (Dg_1 \circ f_1)(p) \cdot Df_1(p) \\ &= (Dg_2 \circ f_2)(p) \cdot Df_2(p) = D(g_2 \circ f_2)(p), \end{aligned}$$

where dot denotes composition of linear maps. Assuming the result for $r-1$, and then using Leibnitz's formula

$$\begin{aligned} D^r(g\circ f)(p) &= D^{r-1}((Dg\circ f)\cdot Df)(p) \\ &= \textstyle\sum_{0\le k\le r-1}\binom{r-1}{k} D^k(Dg\circ f)(p)\cdot D^{r-k}f(p) \end{aligned}$$

we get the result for r. $\square$

The correspondence $N \longrightarrow J^r(M,N)$ for fixed M is functorial. A C^r map $g: N_1 \longrightarrow N_2$ induces a map $g_*: J^r(M,N_1) \longrightarrow J^r(M,N_2)$ defined by $g_*(j^r_p(f)) = j^r_p(g\circ f)$. It follows from the above lemma that g_* is well-defined, and that $(\mathrm{id}_N)_* = \mathrm{id}_{J^r(M,N)}$, and $(g_1\circ g_2)_* = (g_1)_*\circ(g_2)_*$. Thus if g is a diffeomorphism, then g_* is a bijection.

On the other hand, a diffeomorphism $h: M_2 \longrightarrow M_1$ induces a map $h^*: J^r(M_1,N) \longrightarrow J^r(M_2,N)$ given by $h^*(j^r_p(f)) = j^r_q(f\circ h)$, where $q = h^{-1}(p)$. Again we have $(\mathrm{id}_M)^* = \mathrm{id}_{J^r(M,N)}$, and $(h_1\circ h_2)^* = (h_2)^*\circ(h_1)^*$, and so h^* is a bijection.

◊ **Exercise 8.2.** Show that if $g: N_1 \longrightarrow N_2$ is a smooth map and $h: M_2 \longrightarrow M_1$ is a diffeomorphism, then the following diagram is commutative.

$$\begin{tikzcd} J^r(M_1,N_1) \arrow[r, "h^*"] \arrow[d, "g_*"'] & J^r(M_2,N_1) \arrow[d, "g_*"] \\ J^r(M_1,N_2) \arrow[r, "h^*"] & J^r(M_2,N_2) \end{tikzcd}$$

As justified earlier, we may denote by $L^k(\mathbb{R}^n,\mathbb{R}^m)$ the space of k-multilinear maps $L(\mathbb{R}^n\times\cdots\times\mathbb{R}^n,\mathbb{R}^m)$. Let $L^k_s(\mathbb{R}^n,\mathbb{R}^m)$ denote the subspace of $L^k(\mathbb{R}^n,\mathbb{R}^m)$ consisting of symmetric k-multilinear maps. Then the space

$$P^r(n,m) = L(\mathbb{R}^n,\mathbb{R}^m)\times L^2_s(\mathbb{R}^n,\mathbb{R}^m)\times\cdots\times L^r_s(\mathbb{R}^n,\mathbb{R}^m)$$

is the space of all polynomials of degree $\le r$ with constant term equal to zero defined on $\mathbb{R}^n$ with values in $\mathbb{R}^m$. We my write

$$P^r(n,m) = P^r_n\oplus\cdots\oplus P^r_n \quad (m \text{ summands}),$$

where P^r_n is the set of all polynomial functions in n variables $x=(x_1,\dots,x_n)$ of degree $\le r$ with no constant term taking values in $\mathbb{R}$. A typical element of P^r_n looks like

$$\sum_{0<|\alpha|\le r} a_\alpha x^\alpha, \qquad a_\alpha\in\mathbb{R}.$$

Thus the functions $x\mapsto x^\alpha$, for $0<|\alpha|\le r$, form a canonical basis of the vector space.

The vector space $P^r(n,m)$ is a manifold. We have canonical identifications

$$J^r_0(\mathbb{R}^n,\mathbb{R}^m)_0 = P^r(n,m), \quad J^r(\mathbb{R}^n,\mathbb{R}^m) = \mathbb{R}^n\times\mathbb{R}^m\times P^r(n,m).$$

It will follow from the next lemma that if M and N are n- and m-manifolds respectively, where N has no boundary, then every set of jets $J^r_p(M,N)_q$ is in one-one correspondence with $J_0^r(\mathbb{R}^n, \mathbb{R}^m)_0$ via the coordinate charts around the points p and q. Therefore the vector space structure on $J_p^r(M, N)_q$ depends on the choice of coordinate charts, unless $r = 1$ in which case we get a canonical vector space structure on it (see Exercise 8.12 in p. 253).

Lemma 8.1.3. *Let $U \subset \mathbb{R}^n_+$ and $V \subset \mathbb{R}^m$ are open subsets. Then there is a canonical bijection*

$$h_{U,V} : J^r(U, V) \longrightarrow U \times V \times P^r(n, m)$$

given by $h_{U,V}(\xi) = (p, f(p), Df(p), \dots, D^r f(p))$, where ξ is an r-jet represented by a smooth map $f : U \longrightarrow V$ with source $\sigma(\xi) = p \in U$.

Proof. This map is obviously well-defined and injective. To see that this is also surjective, corresponding to an element

$$\eta = (p, q, \lambda_1, \dots, \lambda_r) \in U \times V \times P^r(n, m), \lambda_k \in L^k_s(\mathbb{R}^n, \mathbb{R}^m),$$

define a map $f : U \longrightarrow \mathbb{R}^m$ by

$$f(x) = q + \lambda_1(x - p) + \frac{1}{2!}\lambda_2((x - p)^{(2)}) + \dots + \frac{1}{r!}\lambda_r((x - p)^{(r)}),$$

where $(x - p)^{(k)}$ is $(x - p, \cdots, x - p)$ (k times).

Then f is smooth, and $f(p) = q$. By continuity, there is an open neighbourhood W of p in U such that $f(W) \subset V$. If ξ denotes the equivalence class $j_p^r(f|W)$, then $\xi \in J^r(U, V)$, and we have $h_{U,V}(\xi) = \eta$. $\square$

Remark 8.1.4. The proof uses the fact that the r-jet of f at x has a canonical representation by the Taylor's series of f at x of order r. This polynomial maps from $\mathbb{R}^n$ to $\mathbb{R}^m$ is uniquely determined by the derivatives of f at x of order $\leq r$. Conversely, any element of $P^r(m, n)$ arises from a unique r-jet $j_x^r f$.

The proof will break down if V is assumed to be an open subset of the half space $\mathbb{R}^m_+$. In this case $h_{U,V}$ may not be surjective. For example, if $U = \mathbb{R}^2_+$ $V = \mathbb{R}^2_+$, and $r = 1$, then the element $(0, 0, \mathrm{Id}_{\mathbb{R}^2}) \in U \times V \times L(\mathbb{R}^2, \mathbb{R}^2)$ does not belong to Image $h_{U,V}$.

Proposition 8.1.5. *Let M and N be manifolds of dimension n and m respectively (possibly with boundary), and r be an integer ≥ 0. Then there is a unique topology $\mathcal{T}_r$ on $J^r(M, N)$ such that for any pair of coordinate charts (U, ϕ) in M and (V, ψ) in N with $\phi(U) = U'$ and $\psi(V) = V'$, the set $J^r(U, V)$ is an open set in $\mathcal{T}_r$, and the injective map*

$$k_{U,V} : J^r(U, V) \longrightarrow U' \times V' \times P^r(n, m)$$

given by $k_{U,V} = h_{U',V'} \circ (\phi^{-1})^* \circ \psi_*$ is a topological embedding.

Explicitly, $k_{U,V}$ is the composition

$$J^r(U, V) \xrightarrow{\psi_*} J^r(U, V') \xrightarrow{(\phi^{-1})^*} J^r(U', V') \xrightarrow{h_{U',V'}} U' \times V' \times P^r(n, m),$$

and it maps $j^r_p f$ onto

$$\big(\phi(p), \psi\circ f(p), D(\psi\circ f\circ\phi^{-1})(\phi(p)), \cdots, D^r(\psi\circ f\circ\phi^{-1})(\phi(p))\big).$$

This map is not a homeomorphism, unless $\partial N = \emptyset$.

PROOF. First note that the pair of coordinate charts (U,ϕ), (V,ψ) give a unique topology $\mathcal{T}(\phi,\psi)$ on $J^r(U,V)$ such that $k_{U,V}$ is a homeomorphism onto its image.

Next, note that if $(U,\phi),(U_1,\phi_1)$ are coordinate charts in M and (V,ψ), (V_1,ψ_1) are coordinate charts in N with $U\cap U_1\neq\emptyset$ and $V\cap V_1\neq\emptyset$, then

$$J^r(U,V)\cap J^r(U_1,V_1)$$

is open in both the spaces

$$(J^r(U,V),\mathcal{T}(\phi,\psi)) \text{ and } (J^r(U_1,V_1),\mathcal{T}(\phi_1,\psi_1)).$$

This follows from the fact that $J^r(U,V)\cap J^r(U_1,V_1) = J^r(U\cap U_1, V\cap V_1)$, and that if A and B denote the spaces

$$k_{U,V}\big(J^r(U\cap U_1, V\cap V_1)\big) \text{ and } \mathrm{k_{U_1,V_1}\big(J^r(U\cap U_1, V\cap V_1)\big)}$$

respectively, then

$$A = [\phi(U\cap U_1)\times\psi(V\cap V_1)\times P^r(n,m)]\cap \mathrm{Image}\,(k_{U,V}),$$

$$B = [\phi_1(U\cap U_1)\times\psi_1(V\cap V_1)\times P^r(n,m)]\cap \mathrm{Image}\,(k_{U_1,V_1}).$$

Finally, note that

$$\mathcal{T}(\phi,\psi)|J^r(U\cap U_1, V\cap V_1) = \mathcal{T}(\phi_1,\psi_1)|J^r(U\cap U_1, V\cap V_1).$$

This follows from the following commutative diagram, where all the arrows are homeomorphisms:

$$\begin{tikzcd}
J^r(U\cap U_1, V\cap V_1) \arrow[r, "\mathrm{Id}"] \arrow[d, "k_{U,V}"'] & J^r(U\cap U_1, V\cap V_1) \arrow[d, "k_{U_1,V_1}"] \\
A \arrow[r, "k_{U_1,V_1}\circ k_{U,V}^{-1}"] & B
\end{tikzcd}$$

$\square$

Proposition 8.1.6. *Consider the space $J^r(M,N)$ with the $\mathcal{T}_r$ topology. Then*

(1) *The map $\sigma\times\tau : J^r(M,N)\longrightarrow M\times N$ is a continuous surjective map. Consequently, each of σ and τ is a continuous surjective map.*
(2) *If $f: M\longrightarrow N$ is a smooth map, then the map*

$$j^r(f): M\longrightarrow J^r(M,N),$$

given by $x\mapsto j^r_x(f)$, is a continuous injective map.

(3) *For any two integers r and s with $r \geq s \geq 0$, the map*

$$\pi^r_s : J^r(M,N) \longrightarrow J^s(M,N),$$

obtained by omitting derivatives of order $> s$, is a continuous surjective map.

PROOF. The proof of (1) follows from the following commutative diagram

$$\begin{array}{ccc}
J^r(U,V) & \xrightarrow{k_{U,V}} & \mathrm{Image}\,(k_{U,V}) \\
\big\downarrow{\scriptstyle \sigma\times\tau} & & \big\downarrow{\scriptstyle \mathrm{proj}} \\
U \times V & \xrightarrow{\phi\times\psi} & \phi(U) \times \psi(V)
\end{array}$$

because the horizontal maps are homeomorphisms, and the projection map on the right hand side is continuous.

Similarly, the proof of (2) follows from the following commutative diagram

$$\begin{array}{ccc}
U & \xrightarrow{j^r(f)} & J^r(U,V) \\
\big\downarrow{\scriptstyle \phi} & & \big\downarrow{\scriptstyle k_{U,V}} \\
\phi(U) & \xrightarrow{\lambda} & \mathrm{Image}\,(k_{U,V})
\end{array}$$

where λ is given by

$$x \mapsto \big(x, \psi f \phi^{-1}(x), D(\psi \circ f \circ \phi^{-1})(x), \dots, D^r(\psi \circ f \circ \phi^{-1})(x)\big).$$

The proof of (3) is analogous to that of (1). $\square$

$\Diamond$ **Exercise 8.3.** Show that (a) if $g : N_1 \longrightarrow N_2$ is a smooth map, then the map

$$g_* : J^r(M, N_1) \longrightarrow J^r(M, N_2),$$

given by $j^r_p(f) \mapsto j^r_p(g \circ f)$, is continuous.

(b) if $h : M_2 \longrightarrow M_1$ is a diffeomorphism, then the map

$$h^* : J^r(M_1, N) \longrightarrow J^r(M_2, N),$$

given by $j^r_p(f) \mapsto j^r_q(f \circ h),\ q = h^{-1}(p)$, is continuous.

Let M^k denote the k-fold Cartesian product $M \times \cdots \times M$. Let $\xi_1, \dots, \xi_k$ be r-jets in $J^r(M,N)$ given by $\xi_i = j^r_{x_i}(f_i)$, where $f_i \in C^\infty(M,N)$ and $x_i \in M$, for $i = 1, \dots, k$. Then we write $\xi_1 \times \cdots \times \xi_k$ for the r-jet of $f_1 \times \cdots \times f_k$ with source $(x_1, \dots, x_k) \in M^k$. This r-jet belongs to $J^r(M^k, N^k)$.

Proposition 8.1.7. *The map $\gamma : J^r(M,N)^k \longrightarrow J^r(M^k, N^k)$, given by*

$$\gamma(\xi_1, \dots, \xi_k) = \xi_1 \times \cdots \times \xi_k,$$

is continuous.

PROOF. We shall sketch the proof for $k = 2$. The proof for the general case is essentially the same.

A local representation of γ is $k_{U_1 \times U_2, V_1 \times V_2} \circ \gamma \circ (k_{U_1,V_1}^{-1} \times k_{U_2,V_2}^{-1})$, in terms of coordinate charts in M, N, and $M \times N$ (see Proposition 8.1.5). This map is the restriction of a map h, whose domain is

$$\phi_1(U_1) \times \psi_1(V_1) \times P^r(n,m) \times \phi_2(U_2) \times \psi_2(V_2) \times P^r(n,m),$$

and target is Image $k_{U_1,V_1} \times$ Image k_{U_2,V_2}. If the variables in

$$\phi_1(U_1) \times \psi_1(V_1) \times \phi_2(U_2) \times \psi_2(V_2)$$

are held fixed, then h becomes a linear map which varies continuously with these variables. Consequently h is continuous, and the proposition is proved in this case. $\square$

$\Diamond$ **Exercise 8.4.** If $r = 0$, then show that the map

$$\sigma \times \tau : J^0(M,N) \longrightarrow M \times N$$

is a homeomorphism.

Proposition 8.1.8. *The space $J^r(M,N)$ with the topology $\mathcal{T}_r$ is a Hausdorff space.*

PROOF. Take any two distinct elements $j_p^r f$ and $j_q^r g$ in $J^r(M,N)$. Then the following cases come up for consideration:

Case 1. If $p \neq q$, then there are coordinate neighbourhoods U and U' in M around p and q such that $U \cap U' = \emptyset$. Then, for any two coordinate neighbourhoods V and V' in N around $f(p)$ and $g(q)$, $J^r(U,V)$ and $J^r(U',V')$ are disjoint open neighbourhoods of $j_p^r f$ and $j_q^r g$.

Case 2. If $p = q$, but $f(p) \neq g(q)$, then $j_p^r f$ and $j_q^r g$ can be separated using an argument similar to Case 1.

Case 3. If $p = q$ and $f(p) = g(q)$, then take coordinate charts (U,ϕ) in M around p and (V,ψ) in N around $f(p)$. Then, since $j_p^r f \neq j_q^r g$, $k_{U,V}(j_p^r f)$ and $k_{U,V}(j_q^r g)$ are distinct, and so they have disjoint open neighbourhoods in the Hausdorff space Image $(k_{U,V})$. Then the inverse images by $k_{U,V}$ of these open neighbourhoods separate $j_p^r f$ and $j_q^r g$. $\square$

Proposition 8.1.9. *The space $J^r(M,N)$ with the topology $\mathcal{T}_r$ is a paracompact space, and hence it is metrisable.*

PROOF. Any open set $J^r(U,V)$ of $J^r(M,N)$ is metrisable, hence paracompact, since it is homeomorphic to a subset of the metric space

$$\mathbb{R}^n \times \mathbb{R}^m \times P^r(n,m).$$

Since $M \times N$ is a manifold, it has a countable locally finite open covering $\{W_j\}$ by coordinate neighbourhoods, so that each $W_j = U_j \times V_j$, where U_j and V_j are coordinate neighbourhoods in M and N (Lemma 2.1.3). There is another open covering $\{A_j\}$ of $M \times N$ such that $\overline{A}_j \subset W_j$ for each j (Lemma 2.1.4). Then $B_j = (\sigma \times \tau)^{-1}(\overline{A}_j) \subset J^r(U_j, V_j)$, and so B_j is paracompact, being a closed subset of a paracompact space. Therefore if $\mathcal{U}$ is any open covering of $J^r(M,N)$, then the open covering $B_j \cap \mathcal{U}$ of B_j admits a locally finite refinement, say $\mathcal{U}_j$. Then the open covering $\cup_j \mathcal{U}_j$ is a locally finite refinement of $\mathcal{U}$, and so $J^r(M,N)$ is paracompact. $\square$

Remark 8.1.10. [$\mathcal{T}_\infty$ **topology**] The continuous maps

$$\pi^r_s : J^r(M,N) \longrightarrow J^s(M,N),\ r \geq s,$$

define a system of spaces $(J^r(M,N), \mathcal{T}_r)$. The projective limit of the system (in the category of Hausdorff topological spaces) is denoted by $(J^\infty(M,N), \mathcal{T}_\infty)$. If $f \in C^\infty(M,N)$ and $x \in M$, then

$$j^\infty_x(f) = (j^0_x(f), j^1_x(f), \ldots, j^r_x(f), \ldots)$$

is an element of $J^\infty(M,N)$; it is called a **jet of infinite order**. By the properties of projective limit, for each $r \geq 0$ there is a map $\pi^\infty_r : J^\infty(M,N) \longrightarrow J^r(M,N)$ such that $\pi^r_s \circ \pi^\infty_r = \pi^\infty_s$ for $r \geq s$. The topology $\mathcal{T}_\infty$ is the smallest topology on $J^\infty(M,N)$ such that each π^∞_r is continuous.

Then we have continuous source and target maps

$$\sigma : J^\infty(M,N) \longrightarrow M, \quad \text{and} \quad \tau : J^\infty(M,N) \longrightarrow N$$

defined by $\sigma = p_1 \circ \pi^\infty_0$ and $\tau = p_2 \circ \pi^\infty_0$, where $p_1 : M \times N \longrightarrow M$ and $p_2 : M \times N \longrightarrow N$ are the projections. Also the map $j^\infty(f) : M \longrightarrow J^\infty(M,N)$ sending x to $j^\infty_x(f)$ is continuous, because $\pi^\infty_r \circ j^\infty(f) = j^r(f)$. Finally, the map

$$j^\infty : C^\infty(M,N) \longrightarrow C^0(M, J^\infty(M,N))$$

sending f to $j^\infty(f)$ is injective, because $j^r(f) = \pi^\infty_r \circ j^\infty(f)$, and j^r is injective (note that $j^r(f) = j^r(g) \Rightarrow j^r_x(f) = j^r_x(g)) \Rightarrow f(x) = g(x)$ for all x).

We shall show later in Proposition 8.6.4 that if $\partial N = \emptyset$, then the domain and range of j^∞ can be topologised so that j^∞ becomes a continuous map.

8.2. Weak and strong topologies

The **Whitney C^r topology** on $C^\infty(M,N)$ is defined by means of a basis consisting of all sets of the form

$$B_r(G) = \{f \in C^\infty(M,N) | j^r f(M) \subset G\},$$

where G is an open subset of the space of r-jets $J^r(M,N)$ with the $\mathcal{T}_r$ topology.

This topology is also called the **strong** or **fine C^r topology**. It is the topology induced on $C^\infty(M,N)$ by the injective map

$$j^r : C^\infty(M,N) \longrightarrow C^0(M, J^r(M,N)),$$

where the set of continuous maps $C^0(M, J^r(M,N))$ is given a topology whose basic open sets are

$$B_r'(G) = \{g \in C^0(M, J^r(M,N)) \mid g(M) \subset G\},$$

where G is open in $J^r(M,N)$. Then $B_r(G) = (j^r)^{-1}(B_r'(G))$.

The **weak** or **coarse** C^r **topology** on $C^\infty(M,N)$ is obtained in the same way when we endow the set $C^0(M, J^r(M,N))$ with the compact-open topology instead. This topology on $C^\infty(M,N)$ is generated by a subbasis consisting of sets of the form

$$\{f \in C^\infty(M,N) \mid j^r f(K) \subset G\},$$

where K is a compact subset of M and G is an open subset of $J^r(M,N)$. As the names suggest, the strong C^r topology contains the weak C^r topology. This we shall prove below.

Let $C(X,Y)$ denote the set of continuous maps between topological spaces X and Y. A basis $\mathcal{A}$ of the strong topology in $C(X,Y)$ consists of sets of the form $A(V) = \{f \in C(X,Y) \mid f(X) \subset V\}$, where V is an open set in Y.

Lemma 8.2.1. *The strong topology of the space $C(X,Y)$ has an alternative basis $\mathcal{B}$ consists of sets of the form*

$$B(W) = \{f \in C(X,Y) \mid \Gamma(f) \subset W\},$$

where W is an open set in $X \times Y$, and $\Gamma(f)$ denotes the graph of f.

PROOF. On one hand, we have $A(V) \subset B(X \times V)$, and on the other hand, we have $B(W) \subset A(\pi(W))$, where π is the projection $X \times Y \longrightarrow Y$, which is an open map. $\square$

A subbasis $\mathcal{C}$ of the weak topology in $C(X,Y)$ consists of sets of the form

$$M(K,V) = \{f \in C(X,Y) \mid f(K) \subset V\},$$

where $K \subset X$ is compact, and $V \subset Y$ is open. This topology is also called the uniform convergence topology on compact sets.

Lemma 8.2.2. *$\mathcal{C} \subset \mathcal{B}$.*

PROOF. Let $f \in M(K,V)$, and $W = ((X-K) \times Y) \cup (f^{-1}(V) \times V)$. Then $f \in B(W) \subseteq M(K,V)$. $\square$

Let $\mathcal{K} = \{K_i\}$ be a locally finite family of closed sets of X, $\mathcal{V} = \{V_i\}$ be a family of open sets of Y, and $\mathcal{G} = \{G_i\}$ be a family of open sets of $X \times Y$. Let

$$B'(\mathcal{K}, \mathcal{G}) = \{f \in C(X,Y) \mid \Gamma(f|K_i) \subset G_i \text{ for all } i\}, \text{ and}$$

$$B''(\mathcal{K}, \mathcal{V}) = \{f \in C(X,Y) \mid f(K_i) \subset V_i \text{ for all } i\}.$$

Let $\mathcal{B}'$ be the family of sets of the form $B'(\mathcal{K}, \mathcal{G})$, and $\mathcal{B}''$ be the family of sets of the form $B''(\mathcal{K}, \mathcal{V})$. Then

Lemma 8.2.3. (i) *$\mathcal{B}'$ is a basis of the strong topology in $C(X,Y)$,*

(ii) *$\mathcal{B}''$ is a basis of the strong topology in $C(X,Y)$, if X is a paracompact regular space.*

PROOF. (i) We have $\mathcal{B} \subset \mathcal{B}'$, because $B(W) \subset B'(\{X\},\{W\})$, where W is open in $X \times Y$. Again given a $B'(\mathcal{K},\mathcal{G}) \in \mathcal{B}'$, define open sets W_i of $X \times Y$ by $W_i = G_i \cup ((X - K_i) \times Y)$. Then $W = \cap_i W_i$ is an open set in $X \times Y$, because the closed sets $X \times Y - W_i \subset K_i \times Y$ form a locally finite family, and so their union is a closed set. Then $\mathcal{B}'(\mathcal{K},\mathcal{G}) \subset B(W)$.

(ii) Given $B''(\mathcal{K},\mathcal{V}) \in \mathcal{B}''$, define a family of open sets $\mathcal{G} = \{G_i\}$ of $X \times Y$ by $G_i = (X \times V_i) \cup (X - K_i) \times Y$. Then $B''(\mathcal{K},\mathcal{V}) = B'(\mathcal{K},\mathcal{G})$. Next, suppose that $f \in B(W)$, W open in $X \times Y$. Then $\Gamma(f) \subset W$. Since f is continuous, for each $x \in X$ there exist open neighbourhood U_x of x in X and an open neighbourhood $V_{f(x)}$ of $f(x)$ in Y such that $U_x \times V_{f(x)} \subset W$ and $f(U_x) \subset V_{f(x)}$. Since X is paracompact and regular, there is a locally finite closed refinement $\mathcal{K} = \{K_i\}$ of the covering $\{U_x\}_{x \in X}$ of X. For each i, choose $x_i \in X$ such that $K_i \subset U_{x_i}$. Then $f \in B''(\{K_i\},\{V_{f(x_i)}\})$. □

Lemma 8.2.4. *If X is a compact Hausdorff space, then the weak and strong topologies in $C(X,Y)$ coincide.*

PROOF. We have already seen that $\mathcal{C} \subset \mathcal{B}$. Next note that, since X is paracompact and regular, $\mathcal{B}''$ is a basis of the strong topology. Take an element $B''(\mathcal{K},\mathcal{V})$ in $\mathcal{B}''$. Since X is compact, there exists a finite set $\{i_1, i_2, \ldots, i_n\}$ of indices such that

$$B''(\mathcal{K},\mathcal{V}) = \cap_{j=1}^{n} M(K_i, V_{i_j}).$$

□

Yet another basis of the strong topology in $C(X,Y)$ is given by the following proposition.

Proposition 8.2.5. *Let X be a paracompact and (Y,d) be a be a pseudo-metric space. Let $\epsilon : X \longrightarrow \mathbb{R}$ be a positive continuous function. Then the sets of the form*

$$B(f,\epsilon) = \{g \in C(X,Y) \mid d(f(x),g(x)) < \epsilon(x) \;\; \textit{for all} \;\; x \in X\}$$

constitute a neighbourhood basis of $f \in C(X,Y)$ in the strong topology.

PROOF. Take a set $B(f,\epsilon)$, and define an open set W of $X \times Y$ by

$$W = \{(x,y) \in X \times Y \mid d(f(x),y) < \epsilon(x)\}.$$

This set is open in $X \times Y$, because the map $h : X \times Y \longrightarrow \mathbb{R}$ given by $h(x,y) = \epsilon(x) - d(f(x),y)$ is continuous and $W = h^{-1}(0,\infty)$.

Then, for each $x \in X$, $(x,f(x)) \in W$, i.e. $\Gamma(f) \subset W$. Moreover, we have $B(W) = B(f,\epsilon)$, because

$$g \in B(W) \Leftrightarrow (x,g(x)) \in W \;\forall\; x \Leftrightarrow d(f(x),g(x)) < \epsilon(x) \;\forall\; x \Leftrightarrow g \in B(f,\epsilon).$$

Next we show that any $B(W)$ containing f contains a neighbourhood $B(f,\epsilon)$. Let $f \in B(W)$. Then for any $x \in X$, there exist an open neighbourhood U_x of x in X, and a number $\epsilon_x > 0$ such that

$$f(U_x) \subset S(f(x), \epsilon_x), \quad \text{and} \quad U_x \times S(f(x), \epsilon_x) \subset W,$$

where $S(y', \epsilon) = \{y \in Y \mid d(y, y') < \epsilon\}$. Again, since $f : U_x \longrightarrow S(f(x), \epsilon_x)$ is continuous, there exists an open neighbourhood U'_x of x such that

$$U'_x \subset U_x \;\text{ and }\; f(U'_x) \subset S(f(x), \frac{1}{2}\epsilon_x).$$

Using paracompactness of X, we find a locally finite open refinement $\mathcal{V} = \{V_i\}_{i\in I}$ of the covering $\mathcal{U}' = \{U'_x\}_{x\in X}$ so that for each $i \in I$ there is an $x_i \in X$ with $V_i \subset U'_{x_i}$. Then we employ a partition of unity to construct a positive continuous function $\epsilon : X \longrightarrow \mathbb{R}_+$ such that $\epsilon(x) < \frac{1}{2}\epsilon_{x_i}$ for every $x \in V_i$ and every $i \in I$.

Now let $g \in B(f, \epsilon)$. Then, for $x \in X$, there is an $i(x) = j \in I$ such that $x \in V_j \subset U'_{x_j}$ and

$$d(f(x_j), g(x)) \leq d(f(x_j), f(x)) + d(f(x), g(x)) < \frac{1}{2}\epsilon_{x_j} + \frac{1}{2}\epsilon_{x_j} = \epsilon_{x_j}.$$

Thus $(x, g(x)) \in U'_{x_j} \times S(f(x_j), \epsilon_{x_j}) \subset W$ for every $x \in X$. Therefore $\Gamma(g) \subset W$, and $g \in B(W)$. □

Since $J^r(M,N)$ is metrisable, we can choose a metric d_r on $J^r(M,N)$ which is compatible with its topology. Then the strong C^r topology in $C^\infty(M,N)$ can be described using this metric d_r, as the following lemma shows.

Lemma 8.2.6. *A neighbourhood basis of $f \in C^\infty(M,N)$ in the strong C^r topology is given by the sets of the form*

$$S(f, \epsilon, r) = \{g \in C^\infty(M,N) \mid d_r(j^r f(x), j^r g(x)) < \epsilon(x) \text{ for all } x \in M\},$$

where $\epsilon : M \longrightarrow \mathbb{R}$ is a positive continuous function.

Proof. The proof follows immediately from Proposition 8.2.5 □

Remark 8.2.7. The metric d_r involved in $S(f, \epsilon, r)$ is the natural metric on the basic open set $J^r(U,V)$ of $J^r(M,N)$. This is the metric induced from the standard metric on the subset Image $k_{U,V}$ in certain Euclidean space, and may be described as follows.

Let (U, ϕ) and (V, ψ) be coordinate charts in M and N, and ξ and η be r-jets in $J^r(U,V)$ with source x and y in U, represented respectively by smooth maps f and g from U to V. Then $d_r(\xi, \eta) = d_r(j^r f(x), j^r g(y))$ is the maximum of the numbers $\|\phi(x) - \phi(y)\|$, $\|\psi \circ f(x) - \psi \circ f(y)\|$, and,

$$\|D^k(\psi \circ f \circ \phi^{-1})(\phi(x)) - D^k(\psi \circ g \circ \phi^{-1})(\phi(y))\|, \; k \leq r.$$

Consider a family of subsets of $C^\infty(M,N)$ consisting of sets of the form

$$\mathcal{M}_r(f,(U,\phi),(V,\psi),K,\epsilon),$$

where $f \in C^\infty(M,N)$, (U,ϕ) and (V,ψ) are coordinate charts in M and N respectively, $K \subset U$ is a compact set with $f(K) \subset V$, and ϵ is a positive number. This set consists of maps $g \in C^\infty(M,N)$ such that $g(K) \subset V$ and

$$\|\partial_\alpha(\psi\circ f\circ\phi^{-1})(a) - \partial_\alpha(\psi\circ g\circ\phi^{-1}(a)\| < \epsilon$$

for all $a \in \phi(K)$, and all α with $|\alpha| \le r$. To simplify notation, we shall denote this set by $\mathcal{M}_r(f,U,V,K,\epsilon)$

Proposition 8.2.8. *The family of sets $\mathcal{M}_r(f,U,V,K,\epsilon)$ form a neighbourhood subbasis of the weak C^r topology in $C^\infty(M,N)$.*

Proof. Let G be an open set of $J^r(M,N)$. Define the set $M(K,G)$ by

$$M(K,G) = \{g \in C^\infty(M,N) \mid j^r g(K) \subset G\}.$$

Let $f \in M(K,G)$. By compactness of K, there exist following objects:

- charts (U_i,ϕ_i), $i=1,\dots,k$, of M,
- charts (V_i,ψ_i) $i=1,\dots,k$, of N such that $f(U_i) \subset V_i$,
- compact sets K_i $i=1,\dots,k$, of M with $K_i \subset U_i$, and $K = \cup_{i=1}^k K_i$.

Also for each i we have the map

$$k_{U_i,V_i} : J^r(U_i,V_i) \longrightarrow \phi_i(U_i)\times\psi_i(V_i)\times P^r(n,m)$$

given as in Proposition 8.1.5. Since $k_{U_i,V_i}((j^r f)(K_i)) = H_i$ is compact, and contained in the open set $k_{U_i,V_i}(G\cap J^r(U_i,V_i))$ of Image k_{U_i,V_i}, the number

$$\epsilon_i = d(H_i, \mathbb{R}^n\times\mathbb{R}^m\times P^r(n,m) - A_i)$$

is positive, where A_i is an open set of $\mathbb{R}^n\times\mathbb{R}^m\times P^r(n,m)$ such that

$$A_i \cap \text{Image } k_{U_i,V_i} = k_{U_i,V_i}(G\cap J^r(U_i,V_i)).$$

Then

$$f \in \cap_{i=1}^k \mathcal{M}_r(f,U_i,V_i,K_i,\epsilon_i) \subset M(K,G).$$

Next, we shall show that for each set $\mathcal{M}_r(f,U,V,K,\epsilon)$, there is an open set G in $J^r(U,V)$ such that

$$\mathcal{M}_r(f,U,V,K,\epsilon) = M(K,G).$$

Let $g \in \mathcal{M}_r(f,U,V,K,\epsilon)$, and let η be the positive number

$$\eta = \epsilon - \sup\Big(\|\partial_\alpha(\psi\circ f\circ\phi^{-1})(a) - \partial_\alpha(\psi\circ g\circ\phi^{-1}(a)\|, a\in\phi(K), |\alpha|\le r\Big).$$

By compactness of K, and continuity of $j^r g : M \longrightarrow J^r(M,N)$, there exist points $x_j \in K$ and open neighbourhoods W_j of x_j in U, $j=1,\dots,s$, such that $K \subset \cup_{j=1}^s W_j$, and

$$k_{U,V}\Big(j^r g(W_j)\Big) \subset B\Big(k_{U,V}(j^r g(x_j)), \eta/2\Big) = B_j,$$

where B_j is the open $\eta/2$-ball about $k_{U,V}(j^r g(x_j))$, for $j = 1, \dots, s$. Let

$$G_g = \bigcup_{j=1}^{s} k_{U,V}^{-1}\Big(\phi(W_j) \times \mathbb{R}^m \times P^r(n,m) \cap B_j\Big),$$

and $G = \cup\{G_g \mid g \in \mathcal{M}_r(f, U, V, K, \epsilon)\}$. Then G is open in $J^r(U, V)$, and therefore in $J^r(M, N)$. Moreover, $\mathcal{M}_r(f, U, V, K, \epsilon) = M(K, G)$. This completes the proof. $\square$

Next, consider a family of subsets $\mathcal{N}_r(f, \Phi, \Psi, \mathcal{K}, \epsilon)$ of $C^\infty(M, N)$, where $f \in C^\infty(M, N)$, and

(i) $\Phi = \{(U_i, \phi_i)\}_{i \in I}$ is a locally finite family of coordinate charts of M,
(ii) $\Psi = \{(V_i, \psi_i)\}_{i \in I}$ is a family of coordinate charts of N,
(iii) $\mathcal{K} = \{K_i\}_{i \in I}$ is a family of compact sets such that $K_i \subset U_i$ and $f(K_i) \subset V_i$ for every $i \in I$,
(iv) $\epsilon = \{\epsilon_i\}_{i \in I}$ is a family of positive real numbers.

The set $\mathcal{N}_r(f, \Phi, \Psi, \mathcal{K}, \epsilon)$ is called a strong C^r neighbourhood of f; it consists of smooth maps $g : M \longrightarrow N$ such that $g(K_i) \subset V_i$ for every $i \in I$, and

$$\| \, \partial_\alpha(\psi_i \circ f \circ \phi_i^{-1})(a) - \partial_\alpha(\psi_i \circ g \circ \phi_i^{-1})(a) \, \| < \epsilon_i,$$

for all $a \in \phi_i(K_i)$ and all α with $|\alpha| \le r$.

Note that in these inequalities we may replace ϵ_i by $\epsilon(\phi_i^{-1}(a))$, where $\epsilon : M \longrightarrow \mathbb{R}$ is a positive continuous function so that ϵ_i is the positive lower bound of ϵ on the compact set K_i.

Proposition 8.2.9. *The family of strong C^r neighbourhoods*

$$\mathcal{N}_r(f, \Phi, \Psi, \mathcal{K}, \epsilon)$$

form a neighbourhood basis of the strong C^r topology in $C^\infty(M, N)$.

Proof. We have

$$\mathcal{N}_r(f, \Phi, \Psi, \mathcal{K}, \epsilon) = \bigcap_{i \in I} \mathcal{M}_r(f, U_i, V_i, K_i, \epsilon_i).$$

As described in the previous proposition, for every $i \in I$, there exists an open set G_i of $J^r(M, N)$ such that

$$\mathcal{M}_r(f, U_i, V_i, K_i, \epsilon_i) = M(K_i, G_i).$$

Let $G = \cap_{i \in I}(G_i \cup \sigma^{-1}(M - K_i))$. Then G is an open set of $J^r(M, N)$, since its complement is the union of a locally finite family of closed sets. It is easy to see that

$$\mathcal{N}_r(f, \Phi, \Psi, \mathcal{K}, \epsilon) = B_r(G),$$

where $B_r(G) = \{f \in C^\infty(M, N) \mid j^r f(M) \subset G\}$.

Next, suppose that a $B_r(G)$ is given, and $g \in B_r(G)$. Take any atlas $\Psi' = \{(V_j, \psi_j)\}_{j \in J}$ of N. Then there is a locally finite atlas $\Phi = \{U_i, \phi_i\}_{i \in I}$ of M, and a family of compact sets $\mathcal{K} = \{K_i\}_{i \in I}$ such that each $K_i \subset U_i$ and

$\cup_i \operatorname{Int} K_i = M$. For each $i \in I$, pick $\lambda(i) \in J$ so that $f(U_i) \subset V_{\lambda(i)}$, and write $V_{\lambda(i)} = V_i$ and $\Psi = \{V_i, \psi_i\}_{i \in I}$. Then take the family of numbers $\epsilon = \{\epsilon_i\}$ with

$$\epsilon_i = d\big(k_{U_i,V_i}(j^r g(K_i)), \mathbb{R}^n \times \mathbb{R}^m \times P^r(n,m) - A_i\big) > 0,$$

where A_i ia an open set of $\mathbb{R}^n \times \mathbb{R}^m \times P^r(n,m)$ such that

$$A_i \cap \text{Image } K_{U_i,V_i} = k_{U_i,V_i}(G \cap J^r(U_i, V_i)).$$

It is easy to check that $g \in \mathcal{N}_r(g, \Phi, \Psi, \mathcal{K}, \epsilon) \subset B_r(G)$. $\square$

Difference between the weak and the strong topologies on $C^\infty(M,N)$ is shown in the following example.

Example 8.2.10. Suppose that M and N are manifolds, where M is compact. Take a sequence of points $\{y_n\}$ in N such that $y_n \to y$ and $y_n \neq y$ for all n. Then the sequence of constant maps $g_n \in C^\infty(M,N)$ defined by $g_n(x) = y_n$, $x \in M$, converges to the constant map $g \in C^\infty(M,N)$ defined by $g(x) = y$, $x \in M$, in the weak C^r topology. To see this, take any neighbourhood $\mathcal{M}_r = \mathcal{M}_r(g, (U,\phi), (V,\psi), K, \epsilon)$ of g. Let $x \in K$. Then $y \in V$, as $g(K) \subset V$. Therefore, there is an integer $n_0 > 0$ such that for $n \geq n_0$, $y_n \in V$, and

$$\|\psi \circ g_n \circ \phi^{-1}(a) - \psi \circ g \circ \phi^{-1}(a)\| = \|\psi(y_n) - \psi(y)\| < \epsilon, \ a = \phi(x).$$

So $g_n \in \mathcal{M}_r$ if $n \geq n_0$. Note that the sequence g_n is convergent in the strong C^r-topology also, since M is compact.

If M is not compact, then the sequence $\{g_n\}$ does not converge to g in the strong C^r topology. This may be seen in the following way. Since M is paracompact and non-compact, it is not sequentially compact, and therefore there is a sequence of points $\{x_n\}$ in M without a limit point. Choose Φ, Ψ as in (i), (ii) above, and a family of compact sets $\mathcal{K} = \{K_i\}$ so that $K_i \subset U_i$, and $g(K_i) \subset V_i$ for every i. By Tietze extension theorem, there is a positive continuous function $\epsilon : M \longrightarrow \mathbb{R}$ such that

$$\epsilon(x_n) = \frac{1}{2} d(j^r g_n(x_n), j^r g(x_n)) > 0,$$

for every n. Then $g_n \notin \mathcal{N}_r(g, \Phi, \Psi, \mathcal{K}, \epsilon)$ for every n.

A characteristic property of the compact-open topology on the set of continuous maps $C(X,Y)$ is that if X is compact and Y is metric, then $C(X,Y)$ inherits a metric from that of Y given by $d(f,g) = \sup_{x \in X} d(f(x), g(x))$ (this is finite since X is compact), and the compact-open topology on $C(X,Y)$ coincides with the metric topology; moreover, if Y is a complete metric space, then $C(X,Y)$ becomes a complete metric space. Now, if M is compact, then the strong C^r topology on $C^\infty(M,N)$ comes from the compact-open topology on $C^0(M, J^r(M,N))$; moreover, $J^r(M,N)$ is a metric space. Therefore, if M is compact, then the strong C^r topology on $C^\infty(M,N)$ is metrisable.

If M is not compact, then the strong C^r topology on $C^\infty(M,N)$ is not metrisable. In fact, the space is not even first countable. This may be seen by a contradiction. Let $\{W_n\}$ be a countable neighbourhood basis of f in

$C^\infty(M,N)$. then, for each n there is a positive continuous function $\epsilon_n : M \longrightarrow \mathbb{R}$ such that $S(f,\epsilon_n,r) \subset W_n$. Now take a sequence of points $\{x_n\}$ in M that has no limit point, and define a positive continuous function $\epsilon : M \longrightarrow \mathbb{R}$ such that $\epsilon(x_n) < \epsilon_n(x_n)$ for every n. Then, since $\{W_n\}$ is a neighbourhood basis of f, there is an n such that $W_n \subset S(f,\epsilon,r)$. This implies that $S(f,\epsilon_n,r) \subset S(f,\epsilon,r)$, a contradiction.

$\Diamond$ **Exercise 8.5.** Let X be a paracompact space, and Y a metrisable space. Show that a sequence $\{f_n\}$ in $C(X,Y)$ converges to $f \in C(X,Y)$ in the strong topology if and only if there exists a compact set $K \subset X$ such that $f_n = f$ for all but finitely many n, and the sequence $\{f_n|K\}$ converges uniformly to $f|K$.

In the sequel we shall not consider the weak C^r topology in an essential way, and by C^r topology we shall always mean the strong C^r topology, unless it is explicitly stated otherwise.

8.3. Continuity of maps between spaces of smooth maps

Proposition 8.3.1. *Let X, Y, and Z be Hausdorff spaces, where Z is locally compact and paracompact. Let $\lambda : X \longrightarrow Z$ and $\mu : Y \longrightarrow Z$ be continuous maps, and $X \times_Z Y$ be the subspace of $X \times Y$ given by*

$$X \times_Z Y = \{(x,y) \in X \times Y \mid \lambda(x) = \mu(y)\}.$$

Let $A \subset X$ and $B \subset Y$ be subsets such that $\lambda|A$ and $\mu|B$ are proper maps. Let U be an open neighbourhood of $A \times_Z B$ in $X \times_Z Y$. Then there exist an open neighbourhood V of A in X and an open neighbourhood W of B in Y such that $V \times_Z W \subset U$.

Proof. Let S denote the subspace $X \times_Z Y$. Then S is closed in $X \times Y$, since $S = (\lambda \times \mu)^{-1}(\Delta Z)$, where ΔZ is the diagonal in $Z \times Z$. Therefore $U \cup (X \times Y - S)$ is open in $X \times Y$. For any $z \in Z$, the sets $A_z = A \cap \lambda^{-1}(z)$ and $B_z = B \cap \mu^{-1}(z)$ are compact subsets of X and Y respectively, since $\lambda|A$ and $\mu|B$ are proper maps. Also $A_z \times B_z \subset A \times_Z B \subset U$. Therefore $A_z \times B_z$ is contained in the open set $U \cup (X \times Y - S)$ of $X \times Y$.

Then, by a theorem of A.D. Wallace (see Kelley [**19**], p.142)[1], there exist an open neighbourhood V_z of A_z in X and an open neighbourhood W_z of B_z in Y such that $V_z \times W_z \subset U \cup (X \times Y - S)$.

The set $G_z = Z - (\lambda(A - V_z) \cup \mu(B - W_z))$ is open in Z, because the set $\lambda(A - V_z) \cup \mu(B - W_z)$ is closed in Z, $\lambda|A$ and $\mu|B$ being proper (Corollary 2.1.25). Then the collection $\{G_z\}$, $z \in Z$, is an open covering of Z. Since Z is

[1] This theorem states that if A and B are compact subsets of topological spaces X and Y respectively, and U is a neighbourhood of $A \times B$ in $X \times Y$, then there are neighbourhoods V of A and W of B such that $V \times W \subset U$.

paracompact, there is a locally finite refinement $\{H_\alpha\}$ so that for each index α there is a point $\alpha(z) \in Z$ with $H_\alpha \subset G_{\alpha(z)}$. Let

$$V_\alpha = V_{\alpha(z)} \cup \lambda^{-1}(Z - H_\alpha) \text{ and } W_\alpha = W_{\alpha(z)} \cup \mu^{-1}(Z - H_\alpha),$$
$$V = \cap_\alpha V_\alpha \text{ and } W = \cap_\alpha W_\alpha.$$

We shall verify that V and W satisfy the requirements of the proposition.

Firstly, $A \subset V$, because $A \subset V_\alpha$ for each α. This follows, because, if $a \in A$ and α is any index, then

$$\begin{aligned}&a \in V_{\alpha(z)} \Rightarrow a \in V_\alpha, \text{ and}\\ &a \notin V_{\alpha(z)} \Rightarrow \lambda(a) \notin G_{\alpha(z)} \Rightarrow \lambda(a) \notin H_\alpha \Rightarrow a \in \lambda^{-1}(Z - H_\alpha) \subset V_\alpha,\end{aligned}$$

and therefore $a \in V_\alpha$ anyway for any α. Similarly, $B \subset W$.

Secondly, V is open in X. This may be justified as follows. Let $v \in V$. Since $\{H_\alpha\}$ is locally finite, there is an open neighbourhood U' of $\lambda(v)$ and a finite number of H_α, say, $H_{\alpha_1}, \cdots, H_{\alpha_r}$ such that $U' \cap H_{\alpha_i} \neq \emptyset$. Let

$$V' = \lambda^{-1}(U') \cap V_{\alpha_1} \cap \cdots \cap V_{\alpha_r}.$$

This is an open neighbourhood of v. Then

$$U' \cap H_\alpha = \emptyset \Rightarrow \lambda(V') \cap H_\alpha = \emptyset \Rightarrow V' \subset \lambda^{-1}(Z - H_\alpha) \subset V_\alpha.$$

Again, if $U' \cap H_\alpha \neq \emptyset$, then $\alpha = \alpha_i$ for some i, and so

$$\begin{aligned}U' \cap G_{\alpha(z)} \neq \emptyset \quad &\Rightarrow \quad U' \not\subset \lambda(A - V_{\alpha(z)})\\ &\Rightarrow \quad V' \subset \lambda^{-1}(U') \not\subset A - V_{\alpha(z)} \Rightarrow V' \subset V_{\alpha(z)} \subset V_\alpha.\end{aligned}$$

Thus $V' \subset \cap_\alpha V_\alpha = V$, and V is open. Similarly, W is open in Y.

Finally, $V \times_Z W \subset U$. To see this, suppose $(v, w) \in V \times_Z W$, and $z = \lambda(v) = \mu(w)$. Choose α such that $z \in H_\alpha$. Then $v \in V_\alpha - \lambda^{-1}(Z - H_\alpha) \subset V_{\alpha(z)}$. Similarly, $w \in W_{\alpha(z)}$. Therefore $(v, w) \in V_{\alpha(z)} \times W_{\alpha(z)} \subset U \cup (X \times Y - S)$. But $(v, w) \notin X \times Y - S$, as $\lambda(v) = \mu(w)$. So $(v, w) \in U$. □

Proposition 8.3.2. *If M, N, and R are manifolds, then the map*

$$\gamma : J^r(M, N) \times_N J^r(N, R) \longrightarrow J^r(M, R),$$

whose domain consists of pairs of jets (ξ, η) such that $\tau(\xi) = \sigma(\eta)$, and which is defined by $(j^r_p(f), j^r_{f(p)}(g)) \mapsto j^r_p(g \circ f)$, is continuous with respect to the $\mathcal{T}_r$ topology on the jet spaces.

Proof. The proof may be seen from a commutative diagram which we describe below. Let (U, ϕ), (V, ψ), and (W, θ) be coordinate charts in M, N, and R respectively. Then the commutative diagram is

$$\begin{tikzcd} J^r(U,V) \times_V J^r(V,W) \arrow[r, "\gamma"] \arrow[d, "k_{U,V} \times k_{V,W}"'] & J^r(U,W) \arrow[d, "k_{U,W}"] \\ \mathrm{Image}\,(k_{U,V}) \times \mathrm{Image}\,(k_{V,W}) \arrow[r, "\gamma'"] & \mathrm{Image}\,(k_{U,W}) \end{tikzcd}$$

The map γ' is the restriction of a map

$$(\mathbb{R}^n \times \mathbb{R}^m \times P^r(n,m)) \times (\mathbb{R}^m \times \mathbb{R}^p \times P^r(m,p)) \longrightarrow \mathbb{R}^n \times \mathbb{R}^p \times P^r(n,p)$$

given by $((a, b, u_1, \cdots, u_r), (b, c, v_1, \cdots v_r)) \mapsto (a, c, w_1, \cdots, w_r)$, where

$$w_k = \sum_{s=1}^{k} \sum_{|\alpha|=k} n_k(\alpha) \cdot v_s \cdot u_{\alpha_1} \cdots u_{\alpha_s}, \quad 1 \leq k \leq r,$$

where $\alpha = (\alpha_1, \ldots, \alpha_s)$ is an s-tuple of positive integers with $|\alpha| = k$, and $n_k(\alpha)$ is certain integer which depends on α and k. The definition is designed from the chain rule formula:

$$D^k(g \circ f)(x) = \sum_{s=1}^{k} \sum_{|\alpha|=k} n_k(\alpha) D^s g(f(x)) \cdot D^{\alpha_1} f(x) \cdots D^{\alpha_s} f(x),$$

where $1 \leq k \leq r$, and α, $n_k(\alpha)$ are as above. This result may be proved by induction, using Leibnitz's formula (see the expression of it is given in the proof of Lemma 8.1.2). □

Definition 8.3.3. The C^∞ **topology** on $C^\infty(M, N)$ is the topology whose basis is $\mathcal{B} = \cup_{r=0}^{\infty} \mathcal{B}_r$, where $\mathcal{B}_r$ is the set all open sets of $C^\infty(M, N)$ in the C^r topology.

This is indeed a well-defined basis, because if $r \leq s$, and $\pi_r^s : J^s(M, N) \longrightarrow J^r(M, N)$ is the canonical projection sending $j^s f(x)$ onto $j^r f(x)$, then $B_r(G) = B_s((\pi_r^s)^{-1} G)$ for every open set G in $J^r(M, N)$, and therefore $\mathcal{B}_r \subset \mathcal{B}_s$. In other words, the C^∞ topology is the projective limit of the C^r topologies.

◊ **Exercise 8.6.** Show that the following sets

$$\{f \in C^\infty(M, N) \mid j^\infty(f)(M) \subset G\},$$

where G is an open set in $J^\infty(M, N)$ with the $\mathcal{T}_\infty$ topology,. constitute a basis of the C^∞ topology on $C^\infty(M, N)$

Proposition 8.3.4. *If M, N, and R are manifolds, and $Prop(M, N)$ denotes the space of proper smooth maps from M to N, then the composition map*

$$Prop(M, N) \times C^\infty(N, R) \longrightarrow C^\infty(M, R),$$

given by $(f, g) \mapsto g \circ f$ is continuous in the C^∞ topology.

In particular, if M is compact, then the composition map

$$C^\infty(M, N) \times C^\infty(N, R) \longrightarrow C^\infty(M, R)$$

is continuous.

Proof. Suppose that $f \in C^\infty(M, N)$, $g \in C^\infty(N, R)$, and G is an open set in $J^r(M, R)$ so that

$$g \circ f \in B_r(G) = \{h \in C^\infty(M, R) \mid j^r h(M) \subset G\}.$$

Then, $\gamma^{-1}(G)$ is open in $J^r(M,N) \times_N J^r(N,R)$, by Proposition 8.3.2. We assert that there is an open neighbourhood V of $j^r f(M)$ in $J^r(M,N)$ and an open neighbourhood W of $j^r g(N)$ in $J^r(N,R)$ such that $V \times_N W \subset \gamma^{-1}(G)$, or $\gamma(V \times_N W) \subset G$. This will imply that if $f' \in B_r(V)$ and $g' \in B_r(W)$, then $g' \circ f' \in B_r(G)$, and the composition map will be continuous in the C^r topology for all r, and hence in the C^∞ topology.

In view of Proposition 8.3.1, the assertion will follow if we take $X = J^r(M,N)$, $Y = J^r(N,R)$, $\lambda =$ the target map τ on X, $\mu =$ the source map σ on Y, $A = j^r f(M)$, and $B = j^r g(N)$, provided $\lambda|A$ and $\mu|B$ are proper maps. The first map is proper, because $j^r f$ is a homeomorphism onto its image, and so for any compact set $K \subset N$, $\lambda^{-1}(K) = j^r f(f^{-1}(K))$, which is compact if f is proper. The second map is proper, since $\mu \circ j^r g$ is the identity map on N. This completes the proof. $\square$

Remark 8.3.5. If M is not compact, then the composition map

$$C^\infty(M,N) \times C^\infty(N,R) \longrightarrow C^\infty(M,R)$$

may not be continuous. Suppose that M is non-compact and N is compact. Let $\{z_n\}$ be a convergent sequence in R with $\lim z_n = z$, where $z_n \neq z$ for every n. Consider the maps g_n, and $g : N \longrightarrow R$ given by $g_n(y) = z_n$ and $g(y) = z$, where $y \in N$. It has been shown in Example 8.2.10 that the sequence g_n converges to g in the weak topology, and hence in the strong topology, since N is compact. But for any $f \in C^\infty(M,N)$ the sequence $g_n \circ f$ may not be convergent in the strong topology. Thus in this case the composition map is not continuous.

$\Diamond$ **Exercise 8.7.** If $\phi : N \longrightarrow R$ is a smooth map, then show that the map

$$\phi_\# : C^\infty(M,N) \longrightarrow C^\infty(M,R),$$

given by $f \mapsto \phi \circ f$, is continuous in the C^∞ topology.

Hints. The map $\phi_* : J^r(M,N) \longrightarrow J^r(N,R)$, given by $j^r_p(f) \mapsto j^r_p(\phi \circ f)$, is continuous (see Exercise 8.3(a) in p.231), and $\phi_\#^{-1}(B_r(G)) = B_r(\phi_*^{-1}(G))$ for any open set $G \subset J^r(N,R)$.

Proposition 8.3.6. *Let M, N, and R be manifolds. Then the map*

$$\gamma : C^\infty(M,N) \times C^\infty(M,R) \longrightarrow C^\infty(M,N \times R)$$

defined by $(f,g) \mapsto f \times g$ is a homeomorphism in the C^∞ topology.

Proof. Let $\pi_1 : N \times R \longrightarrow N$ and $\pi_2 : N \times R \longrightarrow R$ be projections. These induce continuous maps $(\pi_1)_\# : C^\infty(M,N \times R) \longrightarrow C^\infty(M,N)$ and $(\pi_2)_\# : C^\infty(M,N \times R) \longrightarrow C^\infty(M,R)$ given by $(\pi_i)_\#(h) = \pi_i \circ h$, $i = 1,2$, (see Exercise 8.7 above).

The map is γ is a bijection, and its inverse is the continuous map

$$(\pi_1)_\# \times (\pi_2)_\# : C^\infty(M,N \times R) \longrightarrow C^\infty(M,N) \times C^\infty(M,R).$$

The proof will be complete, if we prove that $(\pi_1)_\# \times (\pi_2)_\#$ is an open map.

Let $h \in C^\infty(M, N \times R)$. Find an open set $G \subset J^r(M, N \times R)$ such that $h \in B_r(G)$. Since we have an identification

$$J^r(M, N \times R) \equiv J^r(M, N) \times_M J^r(M, R),$$

we may consider G as an open set in $J^r(M, N) \times_M J^r(M, R)$. Since the source maps $\sigma : J^r(M, N) \longrightarrow M$ and $\sigma : J^r(M, R) \longrightarrow M$ (denoted by the same symbol!) are proper (see the proof of Proposition 8.3.4), there is an open neighbourhood U of $j^r(\pi_1 \circ h)(M)$ in $J^r(M, N)$, and an open neighbourhood V of $j^r(\pi_2 \circ h)(M)$ in $J^r(M, R)$ such that $U \times_M V \subset G$ (Proposition 8.3.1). This means that $B_r(U) \times B_r(V) \subset (\pi_1)_\# \times (\pi_2)_\#(B_r(G))$. Therefore $(\pi_1)_\# \times (\pi_2)_\#$ is an open map. $\square$

$\lozenge$ **Exercise 8.8.** Show that the operations of addition and multiplication of smooth functions

$$C^\infty(M, \mathbb{R}) \times C^\infty(M, \mathbb{R}) \longrightarrow C^\infty(M, \mathbb{R}),$$

given respectively by $(f, g) \mapsto f + g$ and $(f, g) \mapsto f \cdot g$, are continuous in the C^∞ topology.

Recall that M^k is the k-fold Cartesian product of M.

Proposition 8.3.7. *Let $\lambda : C^\infty(M, N)^k \longrightarrow C^\infty(M^k, N^k)$ be the map given by $(f_1, \ldots, f_k) \mapsto f_1 \times \cdots \times f_k$, where*

$$(f_1 \times \cdots \times f_k)(x_1, \ldots, x_k) = (f_1(x_1), \ldots, f_k(x_k)).$$

Then λ is continuous.

Proof. We give a proof for $k = 2$. The general case may be seen easily from this special case.

Let $f,\ g \in C^\infty(M, N)$, and consider the continuous map

$$\gamma : J^r(M, N)^2 \longrightarrow J^r(M^2, N^2)$$

of Proposition 8.1.7. Then for an open neighbourhood G of $J^r(M^2, N^2)$ such that $f \times g \in B_r(G) \subset C^\infty(M^2, N^2)$, the open set $\lambda^{-1}(G)$ contains $j^r(f)(M) \times j^r(g)(M)$. Therefore there exist open sets U and V in $J^r(M, N)$ such that

$$j^r(f)(M) \times j^r(g)(M) \subset U \times V \subset \gamma^{-1}(G).$$

This means that $f \times g \in B_r(U) \times B_r(V) \subset \lambda^{-1}(B_r(G))$, and λ is continuous. $\square$

8.4. Spaces of immersions and embeddings

Theorem 8.4.1. *The subspace* $\mathrm{Imm}(M, N)$ *of immersions of M in N is open in the space $C^\infty(M, N)$ with the C^∞ topology.*

PROOF. It is sufficient to prove the theorem for C^1 topology of $C^\infty(M, N)$. So for an immersion $f \in C^\infty(M, N)$ take a C^1 neighbourhood

$$\mathcal{N} = \mathcal{N}_1(f, \Phi, \Psi, \mathcal{K}, \epsilon)$$

of f. Here $\Phi = \{(U_i, \phi_i)\}$ is a locally finite atlas of M, $\Psi = \{(V_i, \psi_i)\}$ is an atlas of N, $\mathcal{K} = \{K_i\}$ is a family of compact sets so that $M = \cup_i \text{Int } K_i$, and $K_i \subset U_i$, $f(K_i) \subset V_i$ for each i, and $\epsilon = \{\epsilon_i\}$ is a family of positive numbers. A map $g : M \longrightarrow N$ belongs to $\mathcal{N}$ if $g(K_i) \subset V_i$, $\|\psi_i \circ f \circ \phi_i^{-1}(a) - \psi_i \circ g \circ \phi_i^{-1}(a)\| < \epsilon_i$, and

$$\|D(\psi_i \circ f \circ \phi_i^{-1})(a) - D(\psi_i \circ g \circ \phi_i^{-1})(a)\| < \epsilon_i$$

for all a in $\phi_i(K_i)$. We shall choose the family ϵ so that every element of $\mathcal{N}$ becomes an immersion.

Let $\dim M = n$ and $\dim N = m$, $n \leq m$. Then the set $M_n(m, n)$ of $m \times n$ matrices of rank n is an open set of the space of linear maps $L(\mathbb{R}^m, \mathbb{R}^n) = \mathbb{R}^{mn}$, because $M_n(m, n)$ is the inverse image of the set $\mathbb{R} - 0$ by the continuous map which takes a matrix into the sum of the squares of its minors of order n. Since f is an immersion, $D(\psi_i \circ f \circ \phi_i^{-1})(a) \in M_n(m, n)$ for all $a \in \phi_i(K_i)$. Also, the set

$$L_i = \{D(\psi_i \circ f \circ \phi_i^{-1})(a) | a \in \phi_i(K_i)\}$$

is a compact set in $M_n(m, n)$, because it is the image of the compact set $\phi_i(K_i)$ under the continuous map $a \mapsto D(\psi_i \circ f \circ \phi_i^{-1})(a)$. Therefore there exists an $\epsilon_i > 0$ such that if $T \in L_i$ and $S \in L(\mathbb{R}^m, \mathbb{R}^n)$ with $\|T - S\| < \epsilon_i$ then $S \in M_n(m, n)$. This choice of the family $\{\epsilon_i\}$ gives rise to a C^1 neighbourhood $\mathcal{N}$ of f which consists only of immersions. Consequently $\text{Imm}(M, N)$ is open. $\square$

Theorem 8.4.2. *The subspace $Sub(M, N)$ of submersions of M in N is open in the space $C^\infty(M, N)$ with the C^∞ topology.*

PROOF. The proof is exactly similar. $\square$

$\lozenge$ **Exercise 8.9.** Modify the proof of Theorem 8.4.1 to get the following result.

Let K be a compact subset of M. Then the space $\text{Imm}_K(M, N)$ of smooth maps $f : M \longrightarrow N$ such that $f|K$ is an immersion is open in $C^\infty(M, N)$ with the C^∞ topology.

Theorem 8.4.3. *The subspace of injective immersions of M in N is open in the space $C^\infty(M, N)$ with the C^∞ topology.*

PROOF. By Theorem 8.4.1 any injective immersion $f : M \longrightarrow N$ admits a C^1 neighbourhood $\mathcal{N} = \mathcal{N}_1(f, \Phi, \Psi, \mathcal{K}, \epsilon)$, consisting only of immersions, where $\Phi = \{(U_i, \phi_i)\}$, $\Psi = \{(V_i, \psi_i)\}$, $\mathcal{K} = \{K_i\}$, and $\epsilon = \{\epsilon_i\}$ are as in that theorem.

Our first problem is to find a C^1 neighbourhood within $\mathcal{N}$ consisting only of immersions g such that $g|K_i$ is injective for all i. If this be not possible, then we would find an index i and a sequence $g_n \in \mathcal{N}$ such that $g_n|K_i$ is not injective, and if $\overline{f} = \psi_i \circ f \circ \phi_i^{-1}$ and $\overline{g}_n = \psi_i \circ g_n \circ \phi_i^{-1}$, then

$$\|\overline{f}(a) - \overline{g}_n(a)\| < \epsilon_i/2^n \text{ and } \|D(\overline{f})(a) - D(\overline{g}_n)(a)\| < \epsilon_i/2^n$$

for all $a \in \phi_i(K_i)$. Then $g_n \longrightarrow f$ and $D(g_n) \longrightarrow D(f)$ uniformly on K_i, but for each n there exist distinct points x_n and y_n in K_i such that $g_n(x_n) = g_n(y_n)$. Since K_i is compact, we may assume by passing to subsequences and renumbering that $x_n \longrightarrow x$ and $y_n \longrightarrow y$, where $x, y \in K_i$. Then $g_n(x) = g_n(y)$. Now, since $f(x_n) \longrightarrow f(x)$, then $g_n(x_n) \longrightarrow f(x)$. Similarly $g_n(y_n) \longrightarrow f(y)$. Therefore $f(x) = f(y)$, and so $x = y$, since f is injective. This shows that $g_n|K_i$ is injective, a contradiction.

We have proved that an injective immersion f has a C^1 neighbourhood $\mathcal{N}_1 \subset \mathcal{N}$ consisting of immersions which are injective on each K_i. Next, we shall show that there is a C^0 neighbourhood $\mathcal{N}_0$ of f such that $\mathcal{N}_0 \cap \mathcal{N}_1$ consists of immersions which are injective on the whole of M. For this purpose, choose a covering of M by compact sets C_i such that $C_i \subset \text{ Int } K_i$ for each i, and a metric d in N. Let

$$\delta_i = d(f(C_i), f(M - \text{ Int } K_i)).$$

Then $\delta_i > 0$, since f is a homeomorphism into (and so the sets are disjoint), we can adjust the things so that $\delta_i \neq \delta_j$ for $i \neq j$. We may therefore find a continuous function $\delta : M \longrightarrow \mathbb{R}_+$ such that $\delta(x) < \delta_i/2$ if $x \in C_i$. Now define

$$\mathcal{N}_0 = \{g \in C^\infty(M, N) \mid |f(x) - g(x)| < \delta(x) \text{ for all } x \in M\}.$$

Take any $g \in \mathcal{N}_0 \cap \mathcal{N}_1$, and two distinct points $x \in C_i$, $y \in C_j$ such that $g(x) = g(y)$. Then supposing $\delta_i < \delta_j$, we have by the triangle inequality that

$$d(f(x), f(y)) \leq d(f(x), g(x)) + d(g(y), f(y)) < \delta_i/2 + \delta_j/2 < \delta_j.$$

On the other hand, since g is injective on K_j, $x \notin K_j$. Therefore $d(f(x), f(y)) \geq \delta_j$, by the definition of δ_j. Thus we get a contradiction, showing that g must be injective everywhere. □

Corollary 8.4.4. *If M is a compact manifold, then the space of embeddings Emb (M, N) is open in the space $C^\infty(M, N)$ with the C^∞ topology.*

Proof. We have seen in Exercise 1.17, p.20 that injective immersions on a compact manifold are embeddings. □

Proposition 8.4.5. *The set of proper maps Prop (M, N) is open in the space $C^\infty(M, N)$ with the C^∞ topology.*

Proof. We shall show that each $f \in \text{Prop } (M, N)$ admits an open neighbourhood consisting only of proper maps. For any such f, $f(M)$ is a closed subset of the paracompact space N, and so $f(M)$ is paracompact. Therefore $f(M)$ admits a countable locally finite open covering $\{U_i\}$. By shrinking lemma, there is a covering of $f(M)$ by compact sets $\{K_i\}$ such that $K_i \subset U_i$. Then $S_i = f^{-1}(K_i)$ is a compact covering of M so that $f(S_i) \subset U_i$. Then there is an open neighbourhood $\mathcal{N}$ of f in $C^\infty(M, N)$ such that if $g \in \mathcal{N}$ then $g(S_i) \subset U_i$ for all i. Such a g must be proper. Because, if K is any compact subset of N, then K intersects only a finite number of U_i, and so the closed subset $g^{-1}(K)$ of M is covered by a finite number of compact sets S_i, and hence $g^{-1}(K)$ is compact. □

Corollary 8.4.6. *If M is compact, then the set of closed embeddings of M into N is open in $C^\infty(M, N)$ with the C^∞ topology.*

PROOF. An embedding f such that $f(M)$ is closed is proper. □

◊ **Exercise 8.10.** Show that if K is a compact subset of M, then the space $\mathrm{Emb}_K(M, N)$ of smooth maps $f : M \longrightarrow N$ such that $f|U$ is an embedding for some open neighbourhood U of K in M (depending on f) is open in the space $C^\infty(M, N)$ with the weak C^∞ topology, and hence with the strong C^∞ topology.

Definition 8.4.7. The **limit set** $L(f)$ of a continuous map $f : M \longrightarrow N$ is the set of points $y \in N$ such that y is the limit point of a sequence $\{f(x_n)\}$ for some sequence $\{x_n\}$ in M having no limit point in M.

Lemma 8.4.8. *$L(f) \subset f(M)$ if and only if $f(M)$ is closed in N.*

PROOF. The 'if' part follows right from the definition, For the 'only if' part, suppose that $y \in \overline{f(M)}$. Then there is a sequence $\{y_n\}$ in $f(M)$ such that $\lim y_n = y$, and $y_n = f(x_n)$ for some $x_n \in M$. If the sequence $\{x_n\}$ has no limit point in M, then $y \in L(f) \subset f(M)$. If the sequence $\{x_n\}$ has a limit point x in M, then $y = f(x)$ by the continuity of f. Thus $y \in f(M)$ anyway, and therefore $f(M)$ is closed. □

Lemma 8.4.9. *A continuous map $f : M \longrightarrow N$ is a homeomorphism into if and only if f is injective and $L(f) \cap f(M) = \emptyset$.*

PROOF. If $M \longrightarrow f(M)$ is a homeomorphism, then for every $y \in f(M)$ and every sequence $\{x_n\}$ in M such that $y = \lim f(x_n)$, we have $f^{-1}(y) = \lim x_n$. Then $y \neq L(f)$, and $L(f) \cap f(M) = \emptyset$.

To prove the converse, we need to show that the map $f^{-1} : f(M) \longrightarrow M$ is continuous, or $\overline{f(C)} \cap f(M) = f(C)$ for any closed set C in M. So suppose that $y \in \overline{f(C)} \cap f(M)$. Then there is a sequence $\{y_n\}$ in $f(C)$ with $\lim y_n = y$. Let $x_n = f^{-1}(y_n) \in C$. The sequence $\{x_n\}$ must have a limit point x in M , otherwise y would belong to $L(f)$ which is not possible, since $L(f) \cap f(M) = \emptyset$. Then $x \in C$ since C is closed, and $f(x) = \lim y_n$ since f is continuous. This means that $y = f(x)$, since y is also a limit point of the sequence $\{y_n\}$ and N is Hausdorff. This proves $\overline{f(C)} \cap f(M) \subseteq f(C)$. Since the reverse inclusion is obvious, we have $\overline{f(C)} \cap f(M) = f(C)$. □

Lemma 8.4.10. *There is a smooth map $f : M \longrightarrow \mathbb{R}$ with $L(f) = \emptyset$.*

Proof. Take a countable locally finite open covering $\{U_k\}$ of M so that $\overline{U}_k$ is compact. Then take two more open coverings $\{V_k\}$ and $\{W_k\}$ of M with compact closures $\overline{V}_k$, $\overline{W}_k$ so that

$$W_k \subset \overline{W}_k \subset V_k \subset \overline{V}_k \subset U_k.$$

For each k choose a smooth function $\phi_k : M \longrightarrow \mathbb{R}$ such that $\phi_k(\overline{W}_k) = 1$, $\phi_k(M - V_k) = 0$, and $0 \leq \phi_k \leq 1$. Define a smooth function $f : M \longrightarrow \mathbb{R}$ by $f(x) = \sum_j j\phi_j(x)$. Then f has empty limit set. To see this, note that if a sequence $\{x_k\}$ in M has no limit point in M, then only finitely many x_k can lie in any compact subset of M. Therefore, given an integer s there is an integer r_s such that $x_k \notin \overline{W}_1 \cup \overline{W}_2 \cup \cdots \cup \overline{W}_s$ for $k \geq r_s$. Thus if $k \geq r_s$, then there is an integer $j > s$ such that $x_k \in \overline{W}_j$, so $f(x_k) \geq j > s$. Hence the sequence $\{f(x_k)\}$ has no limit point. □

Theorem 8.4.11 (Whitney's embedding theorem). *Any manifold M of dimension n can be embedded in $\mathbb{R}^{2n+1}$ as a closed subset.*

Proof. Let $f : M \longrightarrow \mathbb{R} \subset \mathbb{R}^{2n+1}$ be a smooth map such that $L(f) = \emptyset$, and $\epsilon : M \longrightarrow \mathbb{R}$ be the constant function $\epsilon(x) = 1$. We shall show that any injective immersion $g : M \longrightarrow \mathbb{R}^{2n+1}$ which is a C^0 ϵ-approximation of f is an embedding with $f(M)$ as a closed set. Suppose that the injective immersion g is such that $|f(x) - g(x)| < 1$ for all $x \in M$. Let $\{x_k\}$ be a sequence in M having no limit point in M. Then, given any integer $s+1$, there is an integer r_{s+1} such that $f(x_k) \geq s+1$ if $k > r_{s+1}$. Then $|g(x_k)| \geq |f(x_k)| - 1 > s$. Thus the sequence $\{g(x_k\}$ cannot have a limit point, so $L(g) = \emptyset$. This shows that g is a topological embedding with $f(M)$ as a closed set. Therefore g is an embedding. □

8.5. Baire property of the space of smooth maps

A subset of a topological space X is called **residual** if it is the intersection of a countable family of dense open sets of X. A topological space X is called a **Baire space** if every residual set of it is dense in X. It is a standard fact that a complete metric space is a Baire space. Therefore, since $C^\infty(M, N)$ is a complete metric space when M is compact, it is a Baire space. This result is also true when M is not compact, as the following theorem shows.

Theorem 8.5.1. *If M and N are manifolds, then $C^\infty(M, N)$ is a Baire space in the C^∞ topology.*

Proof. Let $U_1, U_2, \cdots$ be a sequence of dense open sets, and V any non-empty open set of the C^∞ topology of $C^\infty(M, N)$. We have to show that the intersection $\cap_{k=1}^\infty U_k$ contains a point of V. The proof is given in the following three steps.

Step 1. We construct a decreasing sequence of open sets

$$W_0 \supseteq W_1 \supseteq \cdots \supseteq W_k \supseteq \cdots,$$

of $C^\infty(M,N)$ in the C^∞ topology such that

$$\overline{W}_0 \subset V, \text{ and } \overline{W}_k \subset U_k \cap W_{k-1} \text{ for } k \geq 1.$$

Since V is open in the C^∞ topology, there is an open set X_0 in $J^{r_0}(M,N)$, for some $r_0 \geq 0$, such that the open set $B_{r_0}(X_0)$ is non-empty and $\overline{B_{r_0}(X_0)} \subset V$. We take $W_0 = B_{r_0}(X_0)$.

Since U_1 is dense, the open set $U_1 \cap W_0 \neq \emptyset$, and we can therefore find an open set X_1 in some $J^{r_1}(M,N)$ such that

$$\overline{B_{r_1}(X_1)} \subset U_1 \cap W_0.$$

So we take $W_1 = B_{r_1}(X_1)$.

If $k \geq 2$ and W_{k-1} is chosen, the denseness of U_k shows that the open set

$$U_k \cap W_{k-1} \neq \emptyset.$$

Therefore we can find an open set X_k in some $J^{r_k}(M,N)$ such that

$$\overline{B_{r_k}(X_k)} \subset U_k \cap W_{k-1}.$$

We therefore take $W_k = B_{r_k}(X_k)$.

Step 2. We construct by induction a sequence of positive continuous functions $\epsilon_k : M \longrightarrow \mathbb{R}_+$ for each integer $k \geq 1$ such that $\epsilon_1 < 1$ and $\epsilon_k = \min(\epsilon_{k-1}, 2^{-k})$.

For a sequence of maps $f_1, f_2, \ldots, f_k, \ldots$ in $C^\infty(M,N)$ (which we shall construct in a moment little later), define open sets $E_{s,k}$ and F_k for integers $s \geq 0$, and $k \geq 1$ by

$$E_{s,1} = W_0, \text{ and } E_{s,k} = S(f_{k-1}, \epsilon_k, s) \text{ if } k > 1,$$

$$F_1 = W_0, \text{ and } F_k = \cap_{s=0}^{k} E_{s,k} \text{ if } k > 1.$$

Here $S(f_{k-1}, \epsilon_k, s)$ is the neighbourhood of f_{k-1} as given in Lemma 8.2.6. Let $G_k = W_{k-1} \cap F_k$. Then G_k is also an open set.

Now we define the sequence $\{f_n\}$ in the following way. Take f_1 as any element in the non-empty open set $U_1 \cap G_1 = U_1 \cap W_0$. If $k \geq 2$, and f_{k-1} has been chosen, use f_{k-1} to construct G_k, and then choose f_k as any element of the open set $U_k \cap G_k$, which is non-empty since U_k is dense.

Since f_k belongs to G_k, and hence to F_k, we have

$$d_s(j^s f_{k-1}(x), j^s f_k(x)) < \epsilon_k(x) < 2^{-k}$$

for all $x \in M$, $k > 1$, and $0 \leq s \leq k$. This shows that for each $x \in M$ the sequence $j^s f_1(x), j^s f_2(x), \ldots,$ is a Cauchy sequence in the complete metric space $J^s(M,N)$, and hence converges to an element $g^s(x) \in J^s(M,N)$. Since $j^0 f_k(x) = (x, f_k(x))$, we may define a continuous map $g : M \longrightarrow N$ so that $g^0(x) = (x, g(x))$. Then $\lim_{k\to\infty} f_k = g$. If we can show that g is smooth, then the proof of the theorem will be finished. Indeed, each f_k is in W_0, and therefore their limit $g \in \overline{W}_0$. Now, if $k > s$, then $f_k \in W_s$, and therefore the limit g is in $\overline{W}_s \subset U_s$. Since s is arbitrary, g belongs to both $\overline{W}_0$ and the intersection $\cap_{k=1}^{\infty} U_k$. Thus $V \cap_{k=1}^{\infty} U_k \neq \emptyset$, and we are done.

Step 3. We now prove that g is smooth. The proof uses the following fact about uniformly convergent sequences : If $f_k : U \longrightarrow \mathbb{R}^m$, U open in $\mathbb{R}^n$, is a sequence of smooth maps such that (1) for some point $x_0 \in U$ the sequence $\{f_k(x_0)\}$ converges in $\mathbb{R}^m$, (2) for every $x \in U$ there is a compact neighbourhood K of x contained in U such that the sequence of first total derivatives $\{Df_k\}$ converges uniformly in K to a map f, then the sequence $\{f_k\}$ converges uniformly in K to g with first total derivative $Dg(x) = f(x)$ for every $x \in U$ (the proof is in Dieudonné [**6**], (8.6.3), p.163).

Using this result in the present context, we can show by induction on s (and the above result of Dieudonné is for $s = 1$) that if for each $s \geq 0$ the sequence $D^s f_k$ converges uniformly to g^s, then $g^s = D^s g^0$. Of course, this has to be proved in a local situation by choosing charts around x and $g(x)$ and compact neighbourhoods K and L of these points contained in the respective coordinate neighbourhoods so that $g(K) \subset L$. □

8.6. Smooth structures on jet spaces

If N is without boundary, then the maps $k_{U,V}$ of Proposition 8.1.5 will be homeomorphisms between open sets of $J^r(M, N)$ and

$$\mathbb{R}^n_+ \times \mathbb{R}^m \times P^r(n, m).$$

We shall show that these homeomorphisms constitute a smooth atlas of $J^r(M, N)$, giving it a structure of a smooth manifold. By Remark 8.1.4, the assertion may be false if N has boundary.

Lemma 8.6.1. *Let $g : U \longrightarrow U'$ be a diffeomorphism and $h : V \longrightarrow V'$ be a smooth map, where U, U' are open subsets of $\mathbb{R}^n_+$ and V, V' open subsets of $\mathbb{R}^m$. Then the map*

$$h_{U',V'} \circ (g^{-1})^* \circ h_* \circ h^{-1}_{U,V} : U \times V \times P^r(n, m) \longrightarrow U' \times V' \times P^r(n, m)$$

is smooth

$$\begin{tikzcd}
U \times V \times P^r(n,m) \arrow[r, "h^{-1}_{U,V}"] \arrow[d] & J^r(U,V) \arrow[r, "h_*"] & J^r(U,V') \arrow[d, "\mathrm{Id}"] \\
U' \times V' \times P^r(n,m) & J^r(U',V') \arrow[l, "h_{U',V'}"'] & J^r(U,V') \arrow[l, "(g^{-1})^*"']
\end{tikzcd}$$

Proof. As in Proposition 8.1.3, we may write down the map explicitly as follows. Let $\eta = (x_0, y_0, \lambda_1, \ldots, \lambda_r) \in U \times V \times P^r(n, m)$, $\lambda_k \in L^k_s(\mathbb{R}^n, \mathbb{R}^m)$, and

$\xi = j^r_{x_0}(f) \in J^r_{x_0}(U, V)_{y_0}$, where $f : U \longrightarrow \mathbb{R}^m$ is the map

$$f(x) = y_0 + \sum_{k=1}^{r} \frac{1}{k!}\lambda_k(x - x_0, \cdots, x - x_0).$$

Then $h_{U,V}(\xi) = \eta$.

Denote the map $h \circ f \circ g^{-1} : U' \longrightarrow V'$ by ϕ. Then

$$h_{U',V'} \circ (g^{-1})^* \circ h_* \circ h^{-1}_{U,V}(\eta)$$

becomes equal to

$$(g(x_0), h(y_0), D\phi(g(x_0)), \ldots, D^r\phi(g(x_0))).$$

This map is smoothly dependent on η, because each $D^k\phi(g(x_0))$ is so. The last statement follows using chain rule and induction on $|\alpha|$ from the fact that $\eta \mapsto \partial_\alpha\phi(g(x_0))$ is smooth for each $|\alpha| \le r$. Note that each $\partial_\alpha\phi(g(x_0))$ varies smoothly with η, because it is the sums and products of partial derivatives of the components h_i, f_i, g_i^{-1}, which are of the form

$$\partial_\beta\partial_j h_i(y_0),\ \partial_\beta\partial_j f_i(x_0),\ \partial_\gamma\partial_j g_i^{-1}(g(x_0)),$$

where β and γ are multi-index so that $|\gamma| \le |\beta| - 1$ and $|\beta| \le |\alpha|$. $\square$

Theorem 8.6.2. *Let M be an n-manifold, and N be an m-manifold without boundary. Then we have the following results.*

(1) *The space $J^r(M, N)$ is a smooth manifold of dimension*

$$n + m + \dim P^r(n, m),$$

and with boundary $\partial J^r(M, N) = \sigma^{-1}(\partial M)$.

Thus if M is also without boundary, then $J^r(M, N)$ is a manifold without boundary.

(2) *Each of the maps $\sigma : J^r(M, N) \longrightarrow M$, $\tau : J^r(M, N) \longrightarrow N$ is a submersion.*

(3) *If $g : N_1 \longrightarrow N_2$ is a smooth map, so is the map*

$$g_* : J^r(M, N_1) \longrightarrow J^r(M, N_2).$$

If $h : M_2 \longrightarrow M_1$ is a diffeomorphism, so is

$$h^* : J^r(M_1, N) \longrightarrow J^r(M_2, N).$$

(4) *If $f : M \longrightarrow N$ is a smooth map, then $j^r(f) : M \longrightarrow J^r(M, N)$ is an embedding.*

Proof. (1) Let (U, ϕ) and (V, ψ) be coordinate charts in M and N respectively with $\phi(U) = U'$ and $\psi(V) = V'$. Then a coordinate chart of $J^r(M, N)$ is given by the homeomorphism

$$k_{U,V} = h_{U',V'} \circ (\phi^{-1})^* \circ \psi_* : J^r(U, V) \longrightarrow U' \times V' \times P^r(n, m).$$

The change of coordinates on the overlap of two coordinate neighbourhoods in $J^r(M,N)$ may be obtained in terms of pairs of compatible coordinate charts (U,ϕ), (U_1,ϕ_1), and $(V.\psi)$, (V_1,ψ_1) in M and N respectively as follows.

$$\begin{aligned} &k_{U_1,V_1}\circ k_{U,V}^{-1}\\ &\qquad = h_{U_1',V_1'}\circ(\phi_1^{-1})^*\circ(\psi_1)_*\circ(\psi_*)^{-1}\circ\phi^*\circ h_{U',V'}^{-1}\\ &\qquad = h_{U_1',V_1'}\circ(\phi\circ\phi_1^{-1})^*\circ(\psi_1\circ\psi^{-1})_*\circ h_{U',V'}^{-1},\end{aligned}$$

by the commutative property described in Exercise 8.2, in p.228. This map is smooth by Lemma 8.6.1.

The second part is left as in exercise.

(2) Local representation of σ is

$$\begin{tikzcd} J^r(U,V) \arrow[r, "\sigma"] \arrow[d, "k_{U,V}"'] & U \arrow[d, "\phi"] \\ U'\times V'\times P^r(n,m) \arrow[r] & U' \end{tikzcd}$$

Let $\eta=(x_0,y_0,\lambda_1,\dots,\lambda_r)\in U'\times V'\times P^r(n,m)$. Then

$$\begin{aligned}\phi\circ\sigma\circ k_{U,V}^{-1}(\eta) &= \phi\circ\sigma\circ(\psi_*)^{-1}\circ\phi^*\circ h_{U',V'}^{-1}(a)\\ &= \phi\circ\sigma\circ j^r(\psi^{-1}\circ f\circ\phi)\\ &= x_0,\end{aligned}$$

where f is the representative of $h_{U',V'}^{-1}(\eta)$ as given in the proof of Lemma 8.1.3. and the last line follows from the facts that $j^r(\psi^{-1}\circ f\circ\phi)$ has source $\phi^{-1}(x_0)$. Therefore σ is a smooth map, and a submersion, being the projection onto the first factor.

Similarly, the local representation of τ is $\psi\circ\tau\circ k_{U,V}^{-1}(\eta)=y_0$. Thus τ is also a smooth map and a submersion.

(3) The local representation of g_* is

$$\begin{tikzcd} J^r(U,V) \arrow[r, "g_*"] \arrow[d, "k_{U,V}"'] & J^r(U,W) \arrow[d, "k_{U,W}"] \\ U'\times V'\times P^r(n,m) \arrow[r] & U'\times W'\times P^r(n,m) \end{tikzcd}$$

where (U,ϕ), (V,ψ), and (W,θ) are charts in M, N_1, and N_2 respectively with $g(V)\subset W$, $\phi(U)=U'$, $\psi(V)=V'$, and $\theta(W)=W'$. Then

$$\begin{aligned} &k_{U,W}\circ g_*\circ k_{U,V}^{-1}\\ &\qquad = h_{U',W'}\circ(\phi^{-1})^*\circ\theta_*\circ g_*\circ(\psi^{-1})_*\circ\phi^*\circ h_{U',V'}^{-1}\\ &\qquad = h_{U',W'}\circ(\theta\circ g\circ\psi^{-1})_*\circ h_{U',V'}^{-1}\end{aligned}$$

which is smooth by Lemma 8.6.1 (with $h=\mathrm{id}$).

The second part follows similarly.

(4) The local representation of $j^r(f)$ is

$$\begin{tikzcd} U \arrow[r, "j^r(f)"] \arrow[d, "\phi"'] & J^r(U,V) \arrow[d, "k_{U,V}"] \\ U' \arrow[r] & U' \times V' \times P^r(n,m) \end{tikzcd}$$

Then, if $\phi(x) = x'$, we have

$$\begin{aligned} &k_{U,V} \circ j^r(f) \circ \phi^{-1}(x') \\ &\quad = k_{U,V} \circ j^r_x(f) \\ &\quad = h_{U',V'} \circ (\phi^{-1})^* \circ \psi_* \circ j^r_x(f)) \\ &\quad = h_{U',V'} \circ j^r_{\phi(x)}(\psi \circ f \circ \phi^{-1}) \\ &\quad = (\phi(x), \psi f(x), D\lambda(x'), \ldots, D^r\lambda_(x')), \end{aligned}$$

where $\lambda = \psi \circ f \circ \phi^{-1}$. This is smooth, as each $D^k\lambda$ is a smooth map, being sum of the partial derivatives of the components λ_i.

Thus $j^r(f)$ is a smooth map. Also it is an embedding, since it is a homeomorphism onto its image, the continuous inverse being $\sigma|j^r(f)(M)$, where σ is the source map. $\square$

Remark 8.6.3. (**Jet bundle**). We have proved more than what is stated in part (1) of the theorem, namely, $J^r(M,N)$ is a fibre bundle over $M \times N$ with projection $\sigma \times \tau$ and fibre $P^r(n,m)$. In general, $J^r(M,N)$ is not a vector bundle, since there is no natural addition of jets in the fibre $J^r_p(M,N)_q$.

$\Diamond$ **Exercise 8.11.** Show that $\sigma : J^r(M,\mathbb{R}^m)_0 \longrightarrow M$ is a vector bundle, where the addition of jets in the fibre $J^r_p(M,\mathbb{R}^m)_0$ is given by the addition of maps representing the jets.

$\Diamond$ **Exercise 8.12.** Consider the vector bundle $L(\tau(M),\tau(N)) \longrightarrow M \times N$ whose fibre over $(x,y) \in M \times N$ is $L(\tau_x(M),\tau_y(N))$ (see Exercise 5.7 in 145), and the fibre bundle $\sigma \times \tau : J^1(M,N) \longrightarrow M \times N$. Show that the map

$$\phi : J^1(M,N) \longrightarrow L(\tau(M),\tau(N)),$$

which sends an 1-jet $j^1_x(f)$ into the linear map $df_x : \tau_x(M) \longrightarrow \tau_{f(x)}(N)$, is a fibre preserving diffeomorphism. Hence conclude that $J^1(M,N)$ may be considered as a vector bundle over $M \times N$.

$\Diamond$ **Exercise 8.13.** Show that, if $J^1(M,\mathbb{R})_0$ is the space of 1-jets with target 0, then $\sigma : J^1(M,\mathbb{R})_0 \longrightarrow M$ is a vector bundle, which is canonically equivalent (as vector bundles) to the cotangent bundle $\tau(M)^*$.

Similarly, the space $J^1_0(\mathbb{R},M)$ of 1-jets with source 0 can be identified with the tangent bundle $\tau(M)$.

◊ **Exercise 8.14.** Show that the canonical projection

$$\pi_k^r : J^r(M,N) \longrightarrow J^k(M,N),\ 0 \le k < r,$$

which sends $j_x^r(f)$ into $J_x^k(f)$, is a fibre bundle with fibre

$$L_s^{k+1}(\mathbb{R}^n, \mathbb{R}^m) \times \cdots \times L_s^r(\mathbb{R}^n, \mathbb{R}^m).$$

.

Proposition 8.6.4. *If M and N are manifolds, where N is without boundary, then the map $j^r : C^\infty(M,N) \longrightarrow C^\infty(M, J^r(M,N))$ defined by $f \mapsto j^r f$ is continuous in the C^∞ topology.*

PROOF. Since for each smooth map $f : M \longrightarrow N$, the map $j^r f : M \longrightarrow J^r(M,N)$ is smooth (Lemma 8.6.2(4)), we may define a map

$$\lambda_{r,k} : J^{r+k}(M,N) \longrightarrow J^k(M, J^r(M,N))$$

in the following way. If ξ is an $(r+k)$-jet in $J^{r+k}(M,N)$, represented by $f : M \longrightarrow N$, then $\lambda_{r,k}(\xi) = j^k(j^r f)(x)$. The definition does not depend on the choice of f, and gives a continuous map. It can be shown easily that we have a commutative diagram

$$\begin{tikzcd}
M \arrow[d, "J^{r+k} f"'] \arrow[dr, "j^k(j^r f)"] & \\
J^{r+k}(M,N) \arrow[r, "\lambda_{r,k}"] & J^k(M, J^r(M,N))
\end{tikzcd}$$

Now, if G is an open set in $J^k(M, J^r(M,N))$, then $H = \lambda_{r.k}^{-1}(G)$ is an open set in $J^{r+k}(M,N)$. Then $B_k(G)$ is open in $C^\infty(M, J^r(M,N)$ and $B_{r+k}(H)$ is open in $C^\infty(M,N)$. It can be shown trivially using the above commutativity that $(j^r)^{-1} B_k(G) = B_{r+k}(H)$. This completes the proof. □

8.7. Thom's Transversality Theorem

Convention. Unless it is stated explicitly otherwise, a manifold will mean a manifold without boundary.

Proposition 8.7.1. *Let M and N be manifolds, A a submanifold of N, and C a closed subset of A. Then the set*

$$\mathcal{T}(A,C) = \{f \in C^\infty(M,N) | f \pitchfork A \text{ on } C\}$$

is an open subset of $C^\infty(M,N)$ in the C^1 topology, and hence in the C^∞ topology.

Proof. Define a subset U of $J^1(M,N)$ in the following way. Let ξ be a 1-jet with source x and target y represented by a smooth map $f : M \longrightarrow N$. Then $\xi \in U$ if and only if either (1) $y \notin C$, or (2) $y \in C$ and $(df)_x(\tau(M)_x) + \tau(A)_y = \tau(N)_y$. These conditions imply that the set

$$B_1(U) = \{f \in C^\infty(M,N) | j^1 f(M) \subset U\}$$

is equal to the set $\mathcal{T}(A,C)$. Therefore it is necessary only to show that U is open.

We shall show that the set $V = J^1(M,N) - U$ is closed in $J^1(M,N)$. Jets in V satisfy neither (1) nor (2). Targets of these jets belong to C, and their representatives do not intersect A transversely at any point of M. Take a convergent sequence of 1-jets $\xi_1, \ldots, \xi_j, \ldots$ in V converging to an 1-jet ξ. Then it is required to show that $\xi \in V$.

Suppose that ξ has source p and target q, and that $q_j \in C$ are the targets of the 1-jets ξ_j. Then, since the target map $\tau : J^1(M,N) \longrightarrow N$ is continuous and C is closed, $q \in C$.

We may assume by choosing suitable charts around p and q that $M = \mathbb{R}^n$, $N = \mathbb{R}^m$, $A = \mathbb{R}^k$, and p is the origin 0 in $\mathbb{R}^n$. Let $\pi : \mathbb{R}^m \longrightarrow \mathbb{R}^m/\mathbb{R}^k = \mathbb{R}^{m-k}$ be the projection. Then, by Lemma 6.2.4, a map $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ does not intersect A transversely at 0 if and only if $\pi \circ f$ is not a submersion at 0. This is equivalent to saying that the linear map $\pi \circ (df)_0$ does not belong to the set

$$S = \{\phi \in L(\mathbb{R}^n, \mathbb{R}^{m-k}) \mid \operatorname{rank} \phi < m-k\},$$

which is a closed subset of $L(\mathbb{R}^n, \mathbb{R}^{m-k})$. Now the set $\mathbb{R}^n \times C \times L(\mathbb{R}^n, \mathbb{R}^m)$ is a closed subset of

$$\mathbb{R}^n \times \mathbb{R}^m \times L(\mathbb{R}^n, \mathbb{R}^m) = J^1(\mathbb{R}^n, \mathbb{R}^m),$$

if C is a closed subset of $\mathbb{R}^k$. The map

$$\lambda : \mathbb{R}^n \times C \times L(\mathbb{R}^n, \mathbb{R}^m) \longrightarrow L(\mathbb{R}^n, \mathbb{R}^{m-k}),$$

given by $\lambda(x,y,\phi) = \pi \circ \phi$, is continuous. Therefore $\lambda^{-1}(S)$ is closed in $J^1(\mathbb{R}^n, \mathbb{R}^m)$. Then, V becomes equal to $\lambda^{-1}(S)$, because

$$\eta = (x, y, (dg)_x) \in V$$

for some $g : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ if and only if $y \in C$ and g does not intersect A transversely at 0, that is, if and only if $\lambda(\eta) \in S$. Since V is closed in this local situation, we have $\xi \in V$. This proves that $\mathcal{T}(A,C)$ is open in $C^\infty(M,N)$. $\square$

Remark 8.7.2. In particular, if the submanifold A is itself a closed subset of N, then the set $\{f \in C^\infty(M,N) \mid f \pitchfork A\}$ is open.

This result may be false if A is not assumed to be closed. For example, the map $f : S^1 \longrightarrow \mathbb{R}^3$ given by $(x,y) \mapsto (x,y,0)$, where S^1 is the circle $x^2+y^2=1$, is transverse to the submanifold $A = \{(x,0,0) \in \mathbb{R}^3 \mid 0 < x < 1\}$, which is not closed. But an arbitrarily small perturbations of f given by $(x,y) \mapsto (x-\epsilon, y, 0)$ intersects A, and is thus not transverse to A.

$\Diamond$ **Exercise 8.15.** Show that Proposition 8.7.1 remains true if M and N are assumed to be manifolds with boundary and A a neat submanifold of N.

Proposition 8.7.3. *Let M and N be manifolds, Z a submanifold of $J^r(M, N)$, and W a closed subset of Z. Then the set*

$$\mathcal{T}(Z, W) = \{f \in C^\infty(M, N) \mid j^r(f) \pitchfork Z \text{ on } W\}$$

is an open subset of $C^\infty(M, N)$ in the C^∞ topology.

PROOF. By Proposition 8.7.1, the set

$$S = \{g \in C^\infty(M, J^r(M, N)) \mid g \pitchfork Z \text{ on } W\}$$

is open in $C^\infty(M, J^r(M, N))$. Since the map

$$j^r : C^\infty(M, N) \longrightarrow C^\infty(M, J^r(M, N))$$

is continuous in the C^∞ topology (Proposition 8.6.4),

$$(j^r)^{-1}(S) = \mathcal{T}(Z, W)$$

is open in $C^\infty(M, N)$. $\square$

Let Z be a submanifold of $J^r(M, N)$. A coordinate neighbourhood W in Z is called **good** if

(1) the closure $\overline{W}$ of W in $J^r(M, N)$ is compact, and $\overline{W} \subset Z$,

(2) $\sigma(\overline{W})$ (resp. $\tau(\overline{W})$) is contained in a coordinate neighbourhood U (resp. V) in M (resp. N), where σ (resp. τ) is the source (resp. target) map.

To describe this in one line, W is the restriction to Z of a suitable coordinate neighbourhood of $J^r(M, N)$ of the form $J^r(U, V)$.

Proposition 8.7.4. *Let Z be a submanifold of $J^r(M, N)$, and W be a good coordinate neighbourhood in Z. Then the set*

$$\mathcal{T}(M, Z, \overline{W}) = \{f \in C^\infty(M, N) \mid j^r(f) \pitchfork Z \text{ on } \overline{W}\}$$

is a dense open subset of $C^\infty(M, N)$ in the C^∞ topology.

PROOF. The proof consists of showing that any arbitrary C^∞ map

$$f : M \longrightarrow N$$

can be deformed slightly so that $j^r(f)$ becomes transverse to Z on $\overline{W}$.

Let (U, ϕ) and (V, ψ) be charts in M and N respectively so that $\sigma(\overline{W}) \subset U$ and $\tau(\overline{W}) \subset V$. Since $\overline{W}$ is compact, it is possible to construct C^∞ functions

$$\rho : \mathbb{R}^n \longrightarrow [0, 1], \quad \text{and} \quad \eta : \mathbb{R}^m \longrightarrow [0, 1] \ (n = \dim M, \ m = \dim N)$$

such that

$$\rho = 1 \text{ on a neighbourhood of } \phi \circ \sigma(\overline{W}), \text{ and } \operatorname{supp} \rho \subset \phi(U),$$

$$\eta = 1 \text{ on a neighbourhood of } \psi \circ \tau(\overline{W}), \text{ and } \operatorname{supp} \eta \subset \psi(V).$$

Let P' be the space of all polynomial maps $\mathbb{R}^n \longrightarrow \mathbb{R}^m$ of degree $\leq r$. Then, $P' = \mathbb{R}^m \times P^r(n,m)$, where $P^r(n,m)$ is the vector space of polynomials of degree $\leq r$ with no constant term, the constant terms are taken care of by the factor $\mathbb{R}^m$.

For each $p \in P'$, define a map $g_p : M \longrightarrow N$ by

$$\begin{aligned} g_p(x) &= f(x) \qquad \text{if } x \notin U \cup f^{-1}(V) \\ &= \psi^{-1}\Big(\rho(\phi(x)) \cdot \eta(\psi(f(x))) \cdot p(\phi(x)) + \psi(f(x))\Big) \qquad \text{otherwise.} \end{aligned}$$

This map is smooth, since ρ and η are so. It is obtained by deforming f in the neighbourhood U and keeping it unaltered outside U. Thus $g_p = f$ outside U for all $p \in P'$. Define a C^∞ map $G' : M \times P' \longrightarrow J^r(M,N)$ by $G'(x,p) = j^r(g_p)(x)$.

The strategy of the proof consists of finding an open neighbourhood P of 0 in P' so that $G = G'|M \times P : M \times P \longrightarrow J^r(M,N)$ is transverse to Z on $\overline{W}$. In this case, an earlier transversality theorem (Theorem 6.2.9) guarantees that the set of points $\{p \in P \mid G_p \overline{\pitchfork} Z \text{ on } \overline{W}\}$ is dense in P, where G_p is the map $M \longrightarrow J^r(M,N)$ defined by $G_p(x) = j^r(g_p)(x)$. Then, we can find a sequence of points $p_1, p_2, \cdots$ in P converging to 0 so that, for any given f, $j^r(g_{p_k}) \overline{\pitchfork} Z$ on $\overline{W}$. Since $g_0 = f$ and $g_p = f$ outside U, we will have $\lim_{k\to\infty} g_{p_k} = f$ in $C^\infty(M,N)$, so that every neighbourhood of f contains a g_p for $p \in P$. Therefore the set $\mathcal{T}(M, Z, \overline{W})$ will be dense in $C^\infty(M,N)$.

We now construct P so that G satisfies the above transversality condition. Define a positive number ϵ by

$$\epsilon = \frac{1}{2} \min\{d(\text{supp } \eta, \mathbb{R}^m - \psi(V)),\ d(\psi \circ \tau(\overline{W}), \eta^{-1}[0,1))\},$$

where d is the standard metric in $\mathbb{R}^m$. Then define

$$P = \{p \in P' \mid \|p(a)\| < \epsilon \text{ for all } a \in \text{supp } \rho\},$$

where $\|\ \|$ denotes the norm in $\mathbb{R}^m$. Then P is an open neighbourhood of 0 in P'. We shall show that if $G(x,p) \in \overline{W}$, then G is a local diffeomorphism in a small neighbourhood of $(x,p) \in M \times P$. Then G will be a submersion in a neighbourhood of (x,p) provided $G(x,p) \in \overline{W}$, and so G will be transverse to any submanifold of $J^r(M,N)$ on $\overline{W}$. This will complete the proof of the proposition.

So suppose that $G(x,p) = j^r(g_p)(x) \in \overline{W}$. Then $x \in \sigma(\overline{W})$ and $g_p(x) \in \tau(\overline{W})$. Then

$$\begin{aligned} \|\psi(f(x) - \psi(g_p(x)\| &= \|\rho(\phi(x)) \cdot \eta(\psi(f(x)) \cdot p(\phi(x))\| \\ &\leq \|p(\phi(x))\| < \epsilon. \end{aligned}$$

This first inequality follows because $\|\rho\| \leq 1$ and $\|\eta\| \leq 1$. The second inequality follows from the definition of P, because $\phi(x) \in \text{supp } \rho$. We have then by the definition of ϵ that $\psi(f(x)) \notin \eta^{-1}([0,1))$, because $\psi(g_p(x)) \in \psi \circ \tau(\overline{W})$, and therefore $\psi(f(x)) \in \eta^{-1}([0,1))$ would imply

$$d(\psi(g_p(x)), \psi(f(x))) \geq d(\psi \circ \tau(\overline{W}), \eta^{-1}([0,1)) > \epsilon.$$

Therefore

$$g_p(x) = \psi^{-1}[p(\phi(x)) + \psi(f(x))].$$

This relation holds for all x' in a neighbourhood of x, since $\rho = 1$ on a neighbourhood of $\phi(\sigma(\overline{W}))$, and also for all p' in a neighbourhood of p in P.

The local representation of G is given by the commutative diagram

$$\begin{tikzcd}
U \times P' \arrow[r, "G"] \arrow[d, "\phi\times\mathrm{Id}"'] & J^r(U,V) \arrow[d, "k_{U,V}"] \\
\phi(U) \times P' \arrow[r, "H"] & \phi(U) \times \psi(V) \times P^r(n,m)
\end{tikzcd}$$

where H is the map

$$(a,p) \mapsto (a, \psi \circ g_p \circ \phi^{-1}(a), D(\psi \circ g_p \circ \phi^{-1})(a), \cdots, D^r(\psi \circ g_p \circ \phi^{-1})(a)).$$

If $a \in \phi \circ \sigma(\overline{W})$ and $p \in P$, then $\psi \circ g_p \circ \phi^{-1}(a) = p(a) + \psi \circ f \circ \phi^{-1}(a)$. Therefore the local representation of g_p in a small neighbourhood B of a in $\phi \circ \sigma(\overline{W})$ is given by $\psi \circ g_p \circ \phi^{-1} = p + \psi \circ f \circ \phi^{-1}$, and any two such local representations of g_p and $g_{p'}$, where p' is in a small neighbourhood of p, represent the same r-jet if and only if $p = p'$ on $\phi(B)$. This means that $H(a,p) = H(a',p')$ if $a = a'$ and $p = p'$, or H is injective. A point η near $H(a,p)$ determines an r-jet ξ near $G(\phi^{-1}(a), p)$ in $J^r(U,V)$, with source x', say. Then there is a unique polynomial p' of degree $\leq r$ such that $\xi = j^r(g_{p'})(x')$, and so $H(\phi(x'), p') = \eta$. It can be checked easily that the map $\eta \mapsto (\phi(x').p')$ is smooth, and is the local inverse of H. Consequently, G is a local diffeomorphism.

The openness of $\mathcal{T}(M, Z, \overline{W})$ follows from Proposition 8.7.3. $\square$

Theorem 8.7.5 (Thom's Transversality Theorem). *If M and N are manifolds, and Z is a submanifold of $J^r(M,N)$, then the set*

$$\mathcal{T}(Z) = \{f \in C^\infty(M,N) \mid j^r f \pitchfork Z\}$$

is a residual subset of $C^\infty(M,N)$ in the C^∞ topology, and hence a dense subset of $C^\infty(M,N)$.

Note that if Z is closed, then $\mathcal{T}(Z)$ is also open, by Proposition 8.7.1.

Proof. Since Z is a submanifold of $J^r(M,N)$, we can find a countable covering of Z by good coordinate neighbourhoods $W_1, \ldots, W_k, \ldots$. Then $\mathcal{T}(Z) = \cap_{k=1}^\infty \mathcal{T}(M, Z, \overline{W}_k)$. Each set $\mathcal{T}(M, Z, \overline{W}_k)$ is open and dense. Therefore $\mathcal{T}(Z)$ is residual in $C^\infty(M,N)$. $\square$

Corollary 8.7.6. *Suppose that I is a countable set such that for each $i \in I$, Z_i is a closed submanifolds of $J^{r_i}(M,N)$. Then the set*

$$\{f \in C^\infty(M,N) \mid j^{r_i}(f) \mathrel{\overline{\pitchfork}} Z_i,\ i \in I\}$$

is dense in $C^\infty(M,N)$. Moreover, if the index set I is finite, then the set is also open in $C^\infty(M,N)$.

Proof. The proof follows from the transversality theorem, and the fact that $C^\infty(M,N)$ is a Baire space. □

◊ **Exercise 8.16.** Let M and N be manifolds, and Z a submanifold of $J^r(M,N)$ such that $\overline{\sigma(Z)}$ is contained in an open set U of M. Let $f : M \longrightarrow N$ be a smooth map, and V an open neighbourhood of f in $C^\infty(M,N)$. Then show that there is a smooth map $g : M \longrightarrow N$ in V such that $j^r(g) \mathrel{\overline{\pitchfork}} Z$, and $g = f$ outside U.

Theorem 8.7.7 (Elementary Transversality Theorem). *If M and N are manifolds, and Z is a submanifold of N, then the set*

$$\mathcal{T} = \{f \in C^\infty(M,N) \mid f \mathrel{\overline{\pitchfork}} Z\}$$

is dense in $C^\infty(M,N)$. Moreover, if Z is closed, then this set is also open.

Proof. The projection $\pi : M \times N \longrightarrow N$ is a submersion, so $\pi^{-1}(Z)$ is a submanifold of $M \times N = J^0(M,N)$. Then the set

$$\mathcal{T}' = \{f \in C^\infty(M,N) \mid j^0(f) \mathrel{\overline{\pitchfork}} \pi^{-1}(Z)\}$$

is dense in $C^\infty(M,N)$ by Theorem 8.7.5 for $r = 0$. Therefore any subset of $C^\infty(M,N)$ containing $\mathcal{T}'$ will also be dense in $C^\infty(M,N)$. Thus the proof will be complete if we can show that $\mathcal{T}' \subset \mathcal{T}$, that is, $j^0(f) \mathrel{\overline{\pitchfork}} \pi^{-1}(Z)$ at $x \in M$ implies $f \mathrel{\overline{\pitchfork}} Z$ at x. To this end, suppose that $j^0(f) : M \longrightarrow M \times N$ is transverse to $\pi^{-1}(Z)$ at $x \in M$. Then either $j^0(f)(x) \notin \pi^{-1}(Z)$, in which case $f(x) \notin Z$, or else $j^0(f)(x) \in \pi^{-1}(Z)$ and

$$d(j^0(f))_x \tau(M)_x + \tau(\pi^{-1}(Z))_{(x,f(x))} = \tau(M \times N)_{(x,f(x))},$$

which implies after applying $(d\pi)_{(x,f(x))}$ to both sides that

$$(df)_x T_x(M) + T_{f(x)}Z = T_{f(x)}N.$$

Therefore $f \mathrel{\overline{\pitchfork}} Z$ at x.

The second part of the theorem is obvious. □

◊ **Exercise 8.17.** Let M, N, and Z be as in the above theorem. Let U and V be open sets of M with $\overline{U} \subset V$. Let $f : M \longrightarrow N$ be a smooth map, and W an open neighbourhood of f in $C^\infty(M,N)$. Then show that there is a smooth map $g \in W$ such that $g = f$ on U, and $g \mathrel{\overline{\pitchfork}} Z$ outside V.

8.8. Multi-jet transversality

We now describe a generalisation of Thom's transversality theorem to multi-jets.

Let M and N be manifolds. Recall that M^k stands for the k-fold Cartesian product of M. Let $M^{(k)}$ denote the following subset of M^k:

$$\{(x_1, \ldots, x_k) \in M^k \mid x_i \neq x_j, 1 \leq i < j \leq k\}.$$

If $\pi_{ij} : M^k \longrightarrow M^2$ is the map $(x_1, \ldots, x_k) \mapsto (x_i, x_j)$, and Δ is the diagonal in M^2, then

$$M^k - M^{(k)} = \cup\ (\pi_{ij})^{-1}(\Delta) \text{ for } 1 \leq i < j \leq k.$$

Therefore $M^{(k)}$ is an open subset of M^k.

Let $\sigma^k = \sigma \times \cdots \times \sigma : J^r(M,N)^k \longrightarrow M^k$ be the product of the source map $\sigma : J^r(M,N) \longrightarrow M$. Then the set $J^r(M,N)^{(k)} = (\sigma^k)^{-1}(M^{(k)})$ is called the **space of multi-jets**. It is an open subset of $J^r(M,N)^k$.

If $f : M \longrightarrow N$ is a smooth map, then the map

$$(j^r(f))^k : M^k \longrightarrow J^r(M,N)^k,$$

which sends $(x_1, \ldots, x_k) \mapsto (j^r(f)(x_1), \cdots, j^r(f)(x_k))$, is smooth. Its restriction to $M^{(k)}$ is the map $j^r(f)^{(k)} : M^{(k)} \longrightarrow J^r(M,N)^{(k)}$ This is well-defined, since $j^r(f)$ is injective.

Proposition 8.8.1. *If N is a manifold without boundary, Z is a submanifold of $J^r(M,N)^{(k)}$, and W is a compact subset of Z, then the set*

$$\mathcal{T}(W) = \{f \in C^\infty(M,N) \mid (j^r(f))^{(k)} \overline{\pitchfork}\ Z \text{ on } W\}$$

is open in $C^\infty(M,N)$.

Proof. We shall consider only the case $k = 2$. The general case may be seen easily from this.

For a point $x = (p_1, p_2) \in M^{(2)}$, and for $i = 1, 2$, choose disjoint open neighbourhoods U_i of p_i in M, and open neighbourhoods V_i of p_i such that $\overline{V}_i \subset U_i$. Let

$$U(x) = U_1 \times U_2,\ V(x) = \overline{V}_1 \times \overline{V}_2, \text{ and}$$

$$\mathcal{T}(x) = \{f \in C^\infty(M,N) \mid (j^r(f))^{(2)} \overline{\pitchfork}\ Z \text{ on } W \cap (\sigma^2)^{-1}(V(x))\}.$$

We shall show that $\mathcal{T}(x)$ is open in $C^\infty(M,N)$. This will end the proof. Because, the open sets Int $(V(x))$, $x \in \sigma^2(W)$, form a covering of the compact set $\sigma^2(W)$, so there is a finite subcovering indexed by some finite set of points $x_1, \ldots, x_s \in M^{(2)}$ so that

$$\mathcal{T}(W) = \cap_{i=1}^s \mathcal{T}(x_i).$$

To show $\mathcal{T}(x)$ is open, consider the set

$$\mathcal{T} = \{g \in C^\infty(M^2, J^r(M,N)^2) \mid g\ \overline{\pitchfork}\ Z \text{ on } W \cap (\sigma^2)^{-1}(V(x))\},$$

and the map

$$\alpha : C^\infty(M,N) \longrightarrow C^\infty(M^2, J^r(M,N)^2)$$

given by $\alpha(f) = (j^r(f))^2$. The map α can be written as the composition

$$C^\infty(M,N) \xrightarrow{\delta} C^\infty(M,N)^2 \xrightarrow{j^r \times j^r} C^\infty(M, J^r(M,N))^2$$
$$\xrightarrow{\eta} C^\infty(M^2, J^r(M,N)^2),$$

where δ is the diagonal map $f \mapsto (f,f)$, j^r is the continuous map of Proposition 8.6.4, and η is continuous by Proposition 8.3.7. Therefore α is a continuous map.

Next note that the set $V(x)$ does not intersect the diagonal Δ in M^2. Consequently, $(j^r(f))^{(2)} \overline{\pitchfork} Z$ on $W \cap (\sigma^2)^{-1}(V(x))$ if and only if $(j^r(f))^2 \overline{\pitchfork} Z$ on $W \cap (\sigma^2)^{-1}(V(x))$. This means that $\alpha^{-1}(\mathcal{T}) = \mathcal{T}(x)$, and consequently, $\mathcal{T}(x)$ is open, since α is continuous, and T is open by the arguments of Proposition 8.7.1. $\square$

Theorem 8.8.2 (Multi-jet Transversality theorem). *Let N be a manifold without boundary, and Z be a submanifold of $J^r(M,N)^{(k)}$. Then the set*

$$\mathcal{T}(Z) = \{f \in C^\infty(M,N) \mid (j^r(f))^{(k)} \overline{\pitchfork} Z\}$$

is residual in $C^\infty(M,N)$. Moreover, if Z is compact, then the set is open.

Proof. Again we shall consider only the case $k = 2$. The proof follows by modifying the proof of Thom's transversality theorem. We will explain only the main steps.

Choose open sets W_j in Z such that the closure $\overline{W}_j$ of W_j in $J^r(M,N)^{(2)}$ is compact, and $\overline{W}_j \subset Z$. Next choose disjoint coordinate neighbourhoods U_{j1}, U_{j2} in M, and coordinate neighbourhoods V_{j1}, V_{j2} in N such that $\overline{U}_{j1}$, $\overline{U}_{j2}$ are compact and disjoint, and

$$\pi(\overline{W}_j) \subset U_{j1} \times U_{j2} \times V_{j1} \times V_{j2},$$

where

$$\pi : J^r(M,N)^{(2)} \longrightarrow M^{(2)} \times N^2 \quad (\text{ not } N^{(2)})$$

is the projection $(\xi, \eta) \mapsto (\sigma(\xi), \sigma(\eta), \tau(\xi), \tau(\eta))$. Let

$$\mathcal{T}(W_j) = \{f \in C^\infty(M,N) \mid (j^r(f))^{(2)} \overline{\pitchfork} Z \text{ on } \overline{W}_j\}.$$

It is required to show that each $\mathcal{T}(W_j)$ is open and dense. This will complete the proof, because, covering Z with a countable collection of such W_j, we have

$$\mathcal{T}(Z) = \cap_{j=1}^\infty \mathcal{T}(W_j).$$

The openness of $\mathcal{T}(W_j)$ follows from Proposition 8.8.1, since $\overline{W}_j$ is compact. Therefore we need only to worry about the density of $\mathcal{T}(W_j)$.

Following the proof of Proposition 8.7.4, we can find, for each $i = 1, 2$, a deformation which deforms f in U_{ji} keeping it unaltered outside U_{ji}. These give maps

$$g_{p_1} : M \longrightarrow N, \text{ and } g_{p_2} : M \longrightarrow N,$$

where p_1, p_2 belong to the space P' of polynomial maps of degree $\leq r$ from $\mathbb{R}^n$ to $\mathbb{R}^m$. Note that g_{p_1} and g_{p_2} are defined by means of coordinate charts $\phi_i : U_{ji} \longrightarrow \mathbb{R}^n$ and $\psi_i : V_{ji} \longrightarrow \mathbb{R}^m$, and functions $\rho_i : \mathbb{R}^n \longrightarrow [0,1]$ and $\eta_i : \mathbb{R}^m \longrightarrow [0,1]$, for $i = 1, 2$, such that

$$\rho_i = 1 \text{ on a neighbourhood of } \phi_i \circ \sigma(\overline{W}_j), \text{ and supp } \rho_i \subset \phi_i(U_{ji}),$$

$$\eta_i = 1 \text{ on a neighbourhood of } \psi_i \circ \tau(\overline{W}_j), \text{ and supp } \eta_i \subset \psi_i(V_{ji}).$$

To get a simultaneous deformation, define, for each $p = (p_1, p_2) \in P' \times P'$, a smooth map $g_p : M \longrightarrow N$ by

$$g_p(x) = \begin{cases} g_{p_1}(x) & \text{if } x \in U_{j1} \text{ and } f(x) \in V_{j1}, \\ g_{p_2}(x) & \text{if } x \in U_{j2} \text{ and } f(x) \in V_{j2}, \\ f(x) & \text{otherwise.} \end{cases}$$

Note that

$$g_{p_i}(x) = \psi_i^{-1}\Big(\rho_i(\phi_i(x)) \cdot \eta_i(\psi_i(f(x))) \cdot p_i(\phi_i(x)) + \psi_i(f(x))\Big).$$

Define as before two numbers ϵ_1 and ϵ_2 for the pairs (U_{j1}, V_{j1}) and (U_{j2}, V_{j2}) respectively

$$\epsilon_i = \frac{1}{2} \min\Big(d(\text{supp } \eta_i, \mathbb{R}^m - \psi_i(V_{ji})), d(\psi_i \circ \tau_i(\overline{W}_j), \eta_i^{-1}[0,1))\Big).$$

Write

$$P_i = \{p \in P' \mid \|p_i(a)\| < \epsilon_i, \text{ for all } a \in \text{supp } \rho_i\},$$

and let $P = P_1 \times P_2$. Define $G : M^{(2)} \times P \longrightarrow J^r(M,N)^{(2)}$ by $G(x,p) = (j^r(g_p))^{(2)}(x)$. Since $\overline{U}_{j1} \cap \overline{U}_{j2} = \emptyset$, it will follow that G is a local diffeomorphism. The remainder of the proof can be completed without any essential change. $\square$

$\Diamond$ **Exercise 8.18.** Show that if Z is a submanifold of $J^r(M_1, N_1) \times J^r(M_2, N_2)$, then the set of maps $(f_1, f_2) \in C^\infty(M_1, N_1) \times C^\infty(M_2, N_2)$ such that $j^r f_1 \times j^r f_2$ is transverse to Z is dense in $C^\infty(M_1, N_1) \times C^\infty(M_2, N_2)$. Moreover, if Z is compact, then the set is open.

8.9. More results on Whitney's immersion and embedding

Let ξ be an 1-jet in $J^1(M,N)$ with source x and target y represented by a smooth map $f : M \longrightarrow N$. Then ξ defines a unique linear map $(df)_x : T_xM \longrightarrow T_yN$. Then the rank of ξ is defined by rank $\xi = \text{rank } (df)_x$.

Let $\dim M = n$ and $\dim N = m$, and $k = \min(n, m)$. For $0 \leq r \leq k$, define

$$S_r = \{\xi \in J^1(M,N) \mid \text{ rank } \xi = k - r\},$$

The set

$$S = \cup_{r=1}^k S_r$$

is called the **singular set**. A jet ξ is called a **singular jet** if $\xi \in S$. This is the set of 1-jets of rank $< k$.

Lemma 8.9.1. *A smooth map f is an immersion if and only if*

$$j^1(f)(M) \cap S = \emptyset, \text{ that is, if } j^1(f) \text{ is not a singular jet.}$$

PROOF. The proof is clear. $\square$

Lemma 8.9.2. *If $\partial N = \emptyset$, Then the set S_r is a submanifold of $J^1(M,N)$ of codimension*

$$(n-k+r)(m-k+r),$$

where n, m, and k are as above.

PROOF. A coordinate chart of $J^1(M,N)$ is given by

$$k_{U,V} : J^1_U(M,N)_V \longrightarrow U \times V \times L(\mathbb{R}^n, \mathbb{R}^m),$$

where U and V are coordinate neighbourhoods in M and N respectively, and $L(\mathbb{R}^n,\mathbb{R}^m)$ is the space of linear maps of $\mathbb{R}^n$ into $\mathbb{R}^m$. Clearly, under the homeomorphism $k_{U,V}$, S_r corresponds to $U \times V \times M_{k-r}(m,n)$, where $M_{k-r}(m,n)$ is the space of $m \times n$ matrices of rank $= k-r$. As we have seen in Example 1.2.7, $\dim M_{k-r}(m,n) = (k-r)(m+n-k+r)$. Therefore

$$\dim S_r = n + m + (k-r)(m+n-k+r),$$

and codim $S_r = mn - (k-r)(m+n-k+r) = (n-k+r)(m-k+r)$. $\square$

Here is an alternative proof of Theorem 8.4.1, when $\partial N = \emptyset$.

Lemma 8.9.3. *The space of immersions Imm (M,N) is an open subset of $C^\infty(M,N)$ in the C^∞ topology.*

PROOF. Clearly, S_0 is open in $J^1(M,N)$, and, by Lemma 8.9.1

$$B_1(S_0) = \{f \in C^\infty(M,N) \mid j^1(f)(M) \subset S_0\} = Imm(M,N),$$

where $B_1(S_0)$ is a basic open set of the C^1 topology of $C^\infty(M,N)$ as described in §8.2. Therefore Imm (M,N) is open subset of $C^\infty(M,N)$ in the C^1 topology, and hence in the C^∞ topology. $\square$

Theorem 8.9.4. *Let n and m be the dimensions of M and N respectively, and N has no boundary. Then, if $m \geq 2n$, Imm (M,N) is an open dense subset of $C^\infty(M,N)$.*

PROOF. If $m \geq 2n$, then $k = \min(n,m) = n$. Therefore

$$\mathit{codim}\ S_r = (n-k+r)(m-k+r) = r(m-n+r) \geq r(n+r) \geq n+1 > n,$$

if $r \geq 1$, and so $\dim M < \mathit{codim}\ S_r$ for $r \geq 1$. This means that $j^1(f) \overline{\pitchfork} S_r$ if and only if $j^1(f)(M) \cap S_r = \emptyset$, $r \geq 1$. Thus, in view of Lemma 8.9.1, f is an immersion if and only if $j^1(f) \overline{\pitchfork} S$. The theorem now follows from Thom's Transversality Theorem 8.7.5. $\square$

Theorem 8.9.5. *If* $\dim N \geq 2\dim M + 1$*, and* $\partial N = \emptyset$*, then*

(i) *the set* $\mathrm{Emb}(M, N)$ *of embeddings of* M *in* N *is dense and open in the space* $C^\infty(M, N)$ *with the* C^∞ *topology, provided* M *is compact.*

(ii) *if* M *is not compact, then the set* $\mathrm{Emb}(M, N)$ *is dense in the subspace* $\mathrm{Prop}(M, N)$ *of proper maps of* M *in* N *with the induced* C^∞ *topology.*

Proof. (i) Since M is compact, embeddings are the same as injective immersions. Therefore we need only show that under the given conditions the set $\mathrm{Inj}(M, N)$ of injective maps of M in N is residual in $C^\infty(M, N)$, because the set of immersions $\mathrm{Imm}(M, N)$ is open and dense in $C^\infty(M, N)$, and hence the intersection $\mathrm{Inj}\,(M, N) \cap\ \mathrm{Imm}(M, N)$ is residual, and hence dense, in $C^\infty(M, N)$.

Now a map $f : M \longrightarrow N$ is injective if and only if the image of

$$(j^0(f))^{(2)} : M^{(2)} \longrightarrow J^0(M, N)^{(2)}$$

does not intersect the set $Z = (\tau^{(2)})^{-1}(\Delta(N))$, where $\tau^{(2)}$ is the restriction to $J^0(M, N)^{(2)}$ of $\tau^2 : J^0(M, N)^2 \longrightarrow N^2$, and $\Delta(N)$ is the diagonal in N^2. The set Z is a submanifold of $J^0(M, N)^{(2)}$, because τ^2 is a submersion, and $J^0(M, N)^{(2)}$ is open in $J^0(M, N)^2$. Since

$$\operatorname{codim} Z = \operatorname{codim} \Delta(N) = \dim N > 2\dim M = \dim M^2 = \dim M^{(2)},$$

$(j^0(f))^{(2)} \pitchfork Z$ if and only if $(j^0(f))^{(2)}(M^{(2)})$ does not intersect Z. Thus f is injective if and only if $(j^0(f))^{(2)} \pitchfork Z$. The proof of density now follows from the multi-jet transversality theorem (Theorem 8.8.2) (for $r = 0$). The openness follows from Corollary 8.4.4.

(ii) A map is an embedding if and only if it is an injective immersion and a proper map onto its image. This follows, because any embedding is clearly injective immersion and proper onto its image, and if f is injective immersion and proper onto its image, then it is a homeomorphism onto its image, and hence an embedding.

We shall show in the next chapter using Morse theory that M may be written as a countable union of an increasing family of compact compact manifolds with boundary $\{M_i\}$

$$M_1 \subset \cdots \subset \cdots \subset M_i \subset \cdots \subset \cdots M.$$

Then by (i), the set

$$S_i = \{f \in C^\infty(M, N) \mid f|M_i \in \mathrm{Emb}(M_i, N)\}$$

is a dense open set. Therefore the intersection $S = \cap_{i=0}^{\infty} S_i$ is still dense. Then any open neighbourhood of a proper map g in $\mathrm{Prop}(M, N)$ contains an element of S which is clearly an injective immersion, and hence an embedding. $\square$

Remark 8.9.6. It can be shown by remodelling our earlier results that the sets S_i are open and dense, when the M_i are just compact sets. Thus the proof of the theorem may be completed without using the Morse Theory. We invite the reader to try this method of proof as an interesting exercise.

CHAPTER 9

MORSE THEORY

This chapter introduces Morse theory, which is of fundamental importance to differential topology. A Morse function $f : M \longrightarrow \mathbb{R}$ is a smooth function having only simplest possible critical points, and in Morse theory one studies the relationship between the number of critical points of f and certain homological invariants of M such as Euler-Poincaré characteristic and Betti numbers. This relationship gives information about the topology of M from the critical points of f which are easy to compute. Conversely, when the topology of M is well understood, one can infer the existence of critical points of a complicated f which is difficult to ascertain. The relationship is obtained by a decomposition of M into level sets $f^{-1}(x)$ of f, which provides a CW complex homotopically equivalent to M.

The Morse theory has a variety of applications. Some of them are Morse's proof of the fact that any two points on a sphere S^n, with any Riemannian structure, can be joined by infinitely many geodesics, Bott's proof of his famous periodicity theorem on the homotopy groups of certain Lie groups, Milnor's celebrated constructions of spheres with different differentiable structures, and Smale's solution of Poincaré conjecture in dimensions ≥ 5. We shall not consider all these topics. The main interests of this book are the applications of Morse theory, namely, the topics related to h-cobordism theorem and Poincaré conjecture. We shall take up these topics in Chapter 10.

9.1. Morse functions

Recall that a point $p \in M$ is a critical point of a smooth function $f : M \longrightarrow \mathbb{R}$, if the derivative map of f at p, $df_p : \tau(M)_p \longrightarrow \tau(\mathbb{R})_{f(p)} = \mathbb{R}$, is not surjective, or equivalently, df_p is the zero map. In terms of local coordinates $x_1, \ldots, x_n$ about p, the condition becomes $df_p = \sum(\partial f/\partial x_i)(p)dx_i = 0$. Thus p is a critical point of f if all the first order partial derivatives of f vanish at p.

A critical point p of f is called **non-degenerate** if for some local coordinate system $x = (x_1, \ldots, x_n)$ about p, the matrix of the second order partial derivatives of f at p

$$H(f)(p) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(p)\right)$$

A. Mukherjee, *Differential Topology*, DOI 10.1007/978-3-319-19045-7_9

is nonsingular. This matrix is called the **Hessian** of f at p relative to the coordinate system x. It determines a symmetric bilinear form $H(f)(p) : \tau(M)_p \times \tau(M)_p \longrightarrow \mathbb{R}$ by

$$H(f)(p)(u,v) = \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(p) u_i v_j, \quad u, v \in \tau(M)_p,$$

and a quadratic form

$$Q(f)(p) : \tau(M)_p \longrightarrow \mathbb{R}$$

by $Q(f)(p)(v) = H(f)(p)(v,v)$.

The definition of nondegeneracy is independent of the choice of coordinate system. This follows from the chain rule. Suppose that $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ are two coordinate systems at a critical point p, where p corresponds to the origin $0 \in \mathbb{R}^n$ in both the coordinate systems. Suppose that H_x (resp. H_y) is the Hessian of f at p relative to the coordinate system x (resp. y), and J is the Jacobian matrix of x relative to y evaluated at 0. Then by the chain rule we have $H_y = J^t \cdot H_x \cdot J$, using the fact that all $(\partial f/\partial x_j)(p) = 0$. Therefore H_y is nonsingular if and only if H_x is so.

The non-degeneracy may be seen in terms of transversality in the following way. The smooth map $f : M \longrightarrow \mathbb{R}$ gives rise to a 1-form df which is a section of the cotangent bundle, $df : M \longrightarrow \tau(M)^*$, given by $x \mapsto df_x \in \tau(M)^*_x$. A point p is a critical point of f if df_p is the zero vector 0_p in $\tau(M)^*_p$. If Z is the zero section of $\tau(M)^*$, then the inverse image $df^{-1}(Z)$ is the set of critical points of f. Note that Z is a submanifold of $\tau(M)^*$ diffeomorphic to M.

At a critical point p of f, the derivative map of $df : M \longrightarrow \tau(M)^*$ at p is a linear map $d(df)_p : \tau(M)_p \longrightarrow \tau(\tau(M)^*)_{0_p}$. Now the tangent space $\tau(\tau(M)^*)_{0_p}$ splits canonically into the direct sum of two subspaces: the 'horizontal' space, tangent to Z, which we identify with $\tau(M)_p$, and the 'vertical' space, tangent to the fibre $\tau(M)^*_p$, which we identify with $\tau(M)^*_p$. The composition L of $d(df)_p$ followed by the projection onto the vertical space gives rise to a linear map $\tau(M)_p \longrightarrow \tau(M)^*_p$. The matrix of L relative to some coordinate system $x = (x_1, \ldots, x_n)$ about p is $H(f)(p)$, because the matrices of the projection and the map $d(df)_p$ are respectively

$$(0\ I_n) \text{ and } \begin{pmatrix} I_n \\ H(f)(p) \end{pmatrix}.$$

Therefore p is non-degenerate critical point of f if and only if L maps $\tau(M)_p$ isomorphically onto $\tau(M)^*_p$, or equivalently if and only if

$$\text{Image } d(df)_p$$

is disjoint from the horizontal space at p, that is, df is transverse to the zero section Z at p.

This result may be described in the language of 'jet' as follows. Recall from §8.9 that $J^1(M,\mathbb{R})$ is the union of two disjoint submanifolds

$$S_0 = \{\xi \in J^1(M,\mathbb{R}) | \text{rank } \xi = 1\}, \text{ and } S_1 = \{\xi \in J^1(M,\mathbb{R}) | \text{rank } \xi = 0\},$$

of which S_0 is open, and so S_1 is closed. Also $\dim S_1 = n+1$ if $\dim M = n$.

Proposition 9.1.1. *Let $f : M \longrightarrow \mathbb{R}$ be a smooth function with associated 1-jet map*

$$j^1(f) : M \longrightarrow J^1(M,\mathbb{R}).$$

Then a point $p \in M$ is a critical point of f if and only if $j^1(f)(p) \in S_1$. Moreover, p is a non-degenerate critical point if and only if $j^1(f)$ is transverse to S_1 at p.

Proof. The first part follows right from the definition

$$j^1(f)(p) \in S_1 \Leftrightarrow \mathrm{rank\, j^1(f)(p) = rank\ df_p = 0 \Leftrightarrow df_p = 0.}$$

The second part also follows easily. Recall from §8.6 that if $\phi : U \longrightarrow \mathbb{R}^n$ is a coordinate chart about p with $\phi(U) = V$, then we may identify $J^1(U,\mathbb{R})$ with $V \times \mathbb{R} \times L(\mathbb{R}^n,\mathbb{R})$ by means of the diffeomorphism $j^1_p(f) \mapsto (\phi(p), f(p), df_p)$, and so the projection $\pi : J^1(U,\mathbb{R}) \longrightarrow L(\mathbb{R}^n,\mathbb{R})$ is a submersion and $\pi^{-1}(0) = S_1$. Therefore, in view of Lemma 6.2.4, $j^1(f) \pitchfork_p S_1$ if and only if $\pi \circ j^1(f)$ is a submersion at p, that is, if and only if the map $U \longrightarrow \mathbb{R}^n$ given by

$$x \mapsto \Big(\frac{\partial f}{\partial x_1}(x), \cdots, \frac{\partial f}{\partial x_n}(x)\Big)$$

is a submersion at p, that is, if and only if its matrix $(\partial^2 f/\partial x_i \partial x_j)(p)$ is nonsingular. $\square$

Corollary 9.1.2. *Non-degenerate critical points of $f : M \longrightarrow \mathbb{R}$ are isolated (that is, a non-degenerate critical point has a neighbourhood in M containing no other critical point of f).*

Proof. Since codim $S_1 = \dim M$, the proof follows from Theorem 6.2.5. $\square$

A smooth function f is called a **Morse function**, if all its critical points are non-degenerate. The following theorem shows that there are plenty of Morse functions on a compact manifold without boundary.

Theorem 9.1.3. *For any compact manifold M without boundary, Morse functions $M \longrightarrow \mathbb{R}$ form a dense open subset of $C^\infty(M,\mathbb{R})$.*

Proof. The proof follows from Thom's transversality theorem (8.7.5). $\square$

Theorem 9.1.4. *The set of Morse functions on a manifold with distinct critical values is residual in $C^\infty(M,\mathbb{R})$.*

PROOF. Let S be the subset

$$(S_1 \times S_1) \cap J^1(M,\mathbb{R})^{(2)} \cap (\tau^{(2)})^{-1}(\Delta(\mathbb{R}))$$

of $J^1(M,\mathbb{R})^{(2)}$, where S_1 is the submanifold of singular jets in $J^1(M,\mathbb{R})$. The elements of S are pairs of singular jets with the same target.

Locally $J^1(M.\mathbb{R})^{(2)}$ is like $(\mathbb{R}^n \times \mathbb{R}^n - \Delta(\mathbb{R}^n)) \times (\mathbb{R} \times \mathbb{R}) \times L(\mathbb{R}^n,\mathbb{R})^{(2)}$, and S is like $(\mathbb{R}^n \times \mathbb{R}^n - \Delta(\mathbb{R}^n)) \times \Delta(\mathbb{R}) \times \{(0,0)\}$. Thus S is a closed submanifold of codimension $2n+1$, where $n = \dim M$. Then multi-jet transversality theorem (8.8.2) implies that the set

$$\{f : M \longrightarrow \mathbb{R} \mid (j^1(f))^{(2)} \,\overline{\pitchfork}\, S\}$$

is residual in $C^\infty(M,\mathbb{R})$. The transversality condition means

$$(j^1(f))^{(2)}(M \times M - \Delta(M)) \cap S = \emptyset,$$

because codim $S >$ dim M^2. Also p and q are distinct critical points of f implies $(j^1 f)^{(2)}(p,q) \in S_1 \times S_1 \cap J^1(M,\mathbb{R})^{(2)}$, and $(j^1(f))^{(2)}(p,q) \notin S$ means $f(p) \neq f(q)$, that is, the critical values are distinct. This completes the proof. $\square$

Corollary 9.1.5. *A smooth map $f : M \longrightarrow \mathbb{R}$, which has no critical point in a closed subset K of M, can be approximated by a Morse function $g : M \longrightarrow \mathbb{R}$ with distinct critical values such that g agrees with f, and g has no critical point in a neighbourhood of K.*

PROOF. Since $j^1(f)\overline{\pitchfork}S_1$ on K, we can approximate f by a Morse function $f_1 : M \longrightarrow \mathbb{R}$ with no critical point on a neighbourhood of K, by the extension theorem (6.2.12). Then $(j^1(f_1))^{(2)} \,\overline{\pitchfork}\, S$ on K, and we can approximate f_1 by a smooth function $g : M \longrightarrow \mathbb{R}$ such that $(j^1(g))^{(2)} \,\overline{\pitchfork}\, S$ and $g = f$ on a neighbourhood of K, by the extension theorem again. $\square$

If p is a non-degenerate critical point of $f : M \longrightarrow \mathbb{R}$, then the **index** of p is defined to be the number of negative eigenvalues of the Hessian of f at p. It is the largest dimension of a subspace of $\tau(M)_p$ on which the quadratic form $Qf(p)$ is negative definite (i.e. takes on negative values).

We now describe our next result. If (U,ϕ) is a coordinate chart around a point $p \in M$ with $\phi(p) = 0$, then Taylor's formula of order 2 of the local representation $F = f \circ \phi^{-1}$ is

$$\begin{aligned} F(u) &= F(0) + DF(0)(u) + \tfrac{1}{2}D^2F(0)(u,u) + E(u) \\ &= F(0) + \nabla F(0) \cdot u + \tfrac{1}{2}Q(F)(0)(u) + E(u) \end{aligned}$$

with an error term $E(u)$ satisfying $E(u)/\|u\|^2 \to 0$ as $u \to 0$. Pulling this formula back on M, we find that in a small neighbourhood of p, $f(x)$ is approximately equal to $f(p) + \nabla f(p)(x) + \frac{1}{2}Q(f)(p)(x)$. The following lemma, known as Morse Lemma, asserts that if p is a non-degenerate critical point of f, then one can choose the chart (U,ϕ) suitably so that $f(x)$ is actually equal to $f(p) + \frac{1}{2}Q(f)(p)(x), x \in U$, where the higher order terms do not appear. In this case the chart (U, ϕ) is called a **Morse chart** about p.

Lemma 9.1.6 (Morse Lemma). *If p is a non-degenerate critical point of f of index r, then there is a coordinate system $x_1, \ldots, x_n$ about p in which*

$$f(x) = f(p) - x_1^2 - \cdots - x_r^2 + x_{r+1}^2 + \cdots + x_n^2.$$

PROOF. Because of the local nature of the lemma, we may assume that the point p is the origin in $\mathbb{R}^n$. Then, by Lemma 3.1.3, there are smooth functions $g_1, \ldots, g_n$ such that $g_i(0) = (\partial f/\partial x_i)(0) = 0$ (since 0 is a critical point), and $f(x) = f(0) + \sum_i x_i g_i(x)$ in a small neighbourhood of 0. Applying the lemma again to each g_i, we get $g_i(x) = \sum_j x_j \overline{h}_{ij}(x)$ for some smooth functions $\overline{h}_{ij}$ with $\overline{h}_{ij}(0) = (\partial g_i/\partial x_j)(0) = (\partial^2 f/\partial x_i \partial x_j)(0)$. Therefore

$$f(x) = f(0) + \sum_{i,j} x_i x_j \overline{h}_{ij}(x) = \sum_{i,j} x_i x_j h_{ij}(x)$$

where $h_{ij} = (1/2)(\overline{h}_{ij} + \overline{h}_{ji})$, and $h_{ij}(0) = (1/2)\overline{h}_{ij}(0)$ is nonsingular, since 0 is a non-degenerate critical point. Since the matrix $H = (h_{ij}(0))$ is symmetric, it may be diagonalised. This means that there is an orthogonal matrix A such that $A^t H A$ is the diagonal matrix $D =$ diag $(\lambda_1, \ldots, \lambda_n)$ of the eigenvalues of H. Then the change of coordinates $x = Ay$ transforms $\sum h_{ij} x_i x_j$ to $\sum \lambda^2 y_i^2$:

$$\sum_{i,j} h_{ij} x_i x_j = x^t H x = y^t A^t H A y = y^t D y = \sum_i \lambda_i^2 y_i^2.$$

Since H is symmetric and nonsingular, its eigenvalues are non-zero real numbers.

Therefore we may suppose that $h_{11}(0) \neq 0$, and hence $h_{11}(x) \neq 0$ in a small neighbourhood of 0, by continuity. Apply a change of coordinates from $(x_1, \ldots, x_n)$ to $(y_1, x_2, \ldots, x_n)$ by

$$y_1 = \sum_{i=1}^{n} x_i h_{1i}/\sqrt{|h_{11}|}.$$

The Jacobian matrix of the transformation is non-zero in a neighbourhood of 0, since

$$\frac{\partial y_1}{\partial x_1}(0) = \pm 1, \quad \frac{\partial y_1}{\partial x_i}(0) = 0 \text{ if } i > 1.$$

Then

$$\begin{aligned} y_1^2 &= |h_{11}|\Big(x_1 + h_{11}^{-1} \textstyle\sum_{i=2}^n x_i h_{1i}\Big)^2 \\ &= \pm\Big(h_{11} x_1^2 + 2 \textstyle\sum_{i=2}^n x_1 x_i h_{1i} + h_{11}^{-1} (\sum_{i=2}^n x_i h_{1i})^2\Big). \end{aligned}$$

The sign is $+$ or $-$ according as $h_{11} > 0$ or < 0. Also we have

$$\begin{aligned} f(x) &= \textstyle\sum_{i,j} x_i x_j h_{ij} \\ &= h_{11} x_1^2 + 2 \textstyle\sum_{i=2}^n x_1 x_i h_{1i} + \sum_{i,j=2}^n x_i x_j h_{ij}. \end{aligned}$$

From these two expressions

$$f = \pm y_1^2 + \sum_{i,j=2}^{n} x_i x_i h_{ij} - h_{11}^{-1}\Big(\sum_{i=2}^{n} x_i h_{1i}\Big)^2.$$

This becomes after simplification $f = \pm y_1^2 + \sum_{i,j=2}^{n} x_i x_j h'_{ij}$, where the second term is a quadratic form in variables $x_2, \dots, x_n$. We now repeat the reduction, and eventually obtain the required result. □

9.2. Critical levels and attaching handles

The prototypes of theorems of this section are due to Milnor [**29**] and Palais [**37**].

For $a \in \mathbb{R}$, the set $f^{-1}(a)$ is called a **level set** of f. If a is a regular value of f, then it follows from the inverse function theorem that $f^{-1}(a)$ is a smooth submanifold of M of codimension one. Moreover, the set

$$M_a = \{x \in M \mid f(x) \le a\}$$

is a smooth submanifold on M whose boundary is the level set $f^{-1}(a)$ (see Lemma 6.2.7)

The following popular model of Morse theory may be used to illustrate various concepts and results. Think of the world as $N = M \times \mathbb{R}$ with $M \times 0$ as sea level, and the projection $\pi : N \longrightarrow \mathbb{R}$, $\pi(x,t) = t$, as height above sea level. Then the graph $\Gamma(f)$ of $f : M \longrightarrow \mathbb{R}$ becomes a mountain over M, and f represents altitude, since $\pi(x, f(x)) = f(x)$. The level set $f^{-1}(a)$ of f is the intersection of $\Gamma(f)$ and the level surface $\pi^{-1}(a)$. The critical points of f are mountain peaks, pits, and passes; they are the points $x \in M$ where the tangent plane $\tau(M)_x$ is horizontal (see Example 9.2.19 below).

We shall suppose that M has a complete Riemannian structure. Although the case of significant interest to us will be when M is compact, we shall start with a general situation where the smooth function $f : M \longrightarrow \mathbb{R}$ satisfies the following condition (called Condition C of Palais and Smale [**38**], [**37**]).

Condition C. If $\{x_k\}$ is a sequence of points in M such that $|f(x_k)|$ is bounded, and $\|df_{x_k}\|$ converges to 0, then $\{x_k\}$ has a convergent subsequence converging to a point $p \in M$.

This means that $\|df_p\| = 0$ by continuity, and so p is a critical point of f.

We shall also suppose that f is bounded below.

Of course if M is compact, then Condition C is automatically satisfied for any choice of the Riemannian metric on M. Condition C was introduced in order to study Morse theory not only for non-compact manifolds, but also for infinite dimensional manifolds. A manifold of latter type is far from being locally compact[1], and therefore it is not possible to have a real valued proper

---

[1]A manifold modelled on a Banach space is finite dimensional if and only if it is locally compact, by a theorem of F. Riez.

function on it. However, there are many infinite dimensional manifolds where Condition C is still satisfied. Here is a simple consequence of Condition C.

Lemma 9.2.1. *If $f : M \longrightarrow \mathbb{R}$ is a smooth function satisfying Condition C, and $\mathcal{C}$ is the set of critical points of f, then $f|\mathcal{C}$ is a proper map.*

Proof. By Condition C, any sequence of critical points of f in the set $f^{-1}([c,d])$ has a convergent subsequence whose limit must be a critical point of f, and so belongs to $f^{-1}([c,d]) \cap \mathcal{C}$. Therefore the set $f^{-1}([c,d]) \cap \mathcal{C}$ is compact for every closed interval $[c,d]$, and hence $f|\mathcal{C}$ is proper. $\square$

Corollary 9.2.2. *The set $f(\mathcal{C})$ of critical values of f is a closed subset of $\mathbb{R}$.*

Proof. Any proper map is closed. $\square$

Choose a Riemannian metric $\tau(M) \times \tau(M) \longrightarrow \mathbb{R}$ on M, and denote the inner product and the corresponding norm in any tangent space by $\langle X, Y \rangle$ and $\|X\| = \langle X, Y \rangle^{1/2}$ respectively. The Riemannian structure induces an isomorphism $\mathfrak{X}(M) \longrightarrow \Omega^1(M)$ between the vector space $\mathfrak{X}(M)$ of vector fields on M and the vector space $\Omega^1(M)$ of 1-forms on M by the correspondence $X \mapsto i_X$, where $i_X(Y) = \langle X, Y \rangle$ for every vector field $Y \in \mathfrak{X}(M)$. In particular, for the smooth function f on M, there is a unique vector field ∇f on M corresponding to the differential 1-form df such that $i_{\nabla f} = df$. The vector field ∇f is called the gradient of f relative to the given Riemannian structure. It is characterised by the fact that, for any vector field X on M, $\langle \nabla f, X \rangle = df(X) = Xf$, which is the directional derivative of f along X. If $\alpha : \mathbb{R} \longrightarrow M$ is a smooth curve in M, then

$$\left\langle \nabla f, \frac{d\alpha}{dt} \right\rangle = df\left(\frac{d\alpha}{dt}\right) = \frac{d(f \circ \alpha)}{dt}.$$

Note further that, if X happens to be tangent to a level surface $f^{-1}(c)$, then $Xf = 0$, so that at each regular point $x \in M$, ∇f is orthogonal to the level surface through x.

Let ϕ_t denote maximal flow generated by the vector field $-\nabla f$. Then, for each $x \in M$, $\phi_t(x)$ is defined on an interval $a(x) < t < b(x)$ containing 0, and $t \mapsto \phi_t(x)$ is the maximal solution of the initial value problem

$$\frac{d}{dt}\phi_t(x) = -(\nabla f)_{\phi_t(x)}, \qquad \phi_0(x) = x.$$

Then

$$\frac{d}{dt} f(\phi_t(x)) = -(\nabla f)f = -\|\nabla f\|^2.$$

Therefore $f(\phi_t(x))$ decreases monotonically in t. Since f is bounded below by our assumption, $f(\phi_t(x))$ tends to a limit as $t \to b(x)$.

If we give our mountain model N the product metric, then the negative gradient vector field $-\nabla f$ points in the downhill direction orthogonal to level surfaces, and represents the direction of quickest descent. Roughly speaking, the flow $\{\phi_t\}$ generated by $-\nabla f$ gives the way in which a thick liquid will flow under the action of gravity.

Referring back to our original discussion, We shall show that for each $x \in M$, $b(x) = \infty$ so that $\phi_t(x)$ is defined for all $t > 0$. This property may be expressed by saying that the family $\{\phi_t\}$ is a **positive semigroup**.

Lemma 9.2.3. *If $\sigma : (a, b) \longrightarrow M$ is a C^1 curve of finite length, then* Image(σ) *is relatively compact.*

PROOF. Since $\int_a^b \|\sigma'(t)\|dt < \infty$, given an $\epsilon > 0$ there is a partition

$$a = t_0 < t_1 < \cdots < t_k = b$$

of the interval $(a.b)$ such that each

$$\int_{t_i}^{t_{i+1}} \|\sigma'(t)\|dt < \epsilon.$$

Then it follows by the definition of the metric on M that the closure of Image(σ) is the union of finitely many sets of diameter $< \epsilon$. Therefore the closure is totally bounded, and it is complete since M is complete. Consequently, Image(σ) is relatively compact. Note that a metric space is compact if and only if it is complete and totally bounded. $\square$

Proposition 9.2.4. *Let X be a smooth vector field on M with maximal integral curve $\sigma : (a, b) \longrightarrow M$, where $0 \in (a, b)$. Then $b < \infty$ implies*

$$\int_0^b \|X_{\sigma(t)}\|dt = \infty,$$

and $a > -\infty$ implies

$$\int_a^0 \|X_{\sigma(t)}\|dt = \infty.$$

PROOF. Since σ is maximal, $b < \infty$ implies that $\sigma(t)$ has no limit point in M as $t \to b$. Then Lemma 9.2.3 implies that σ must have infinite length. Since $\sigma'(t) = X_{\sigma(t)}$, we have $\int_0^b \|X_{\sigma(t)}\|dt = \infty$. The proof of the other part is similar. $\square$

Corollary 9.2.5. *A smooth vector field of bounded length generates a one-parameter group of diffeomorphisms of M.*

PROOF. This is clear because we must have $b = \infty$ and $a = -\infty$. For example, if $b < \infty$, then $\|X\| \le c < \infty$ would imply

$$\int_0^b \|X_{\sigma(t)}\|dt \le bc < \infty,$$

contradicting Proposition 9.2.4. By a similar argument $a > -\infty$ is also not possible $\square$

Theorem 9.2.6. *The flow $\{\phi_t\}$ generated by $-\nabla f$ is a positive semigroup, that is, ϕ_t is defined on the whole of M for all $t > 0$. Moreover, for any $x \in M$, the limit*

$$\lim_{t\to\infty} \phi_t(x)$$

is a critical point of f.

Proof. Write $g(t) = f(\phi_t(x))$. Then if B denotes the greatest lower bound of f, then for all $T < b(x)$ we have

$$B \leq g(T) = g(0) + \int_0^T g'(t)dt = g(0) - \int_0^T \|\nabla f_{\phi_t(x)}\|^2 dt. \tag{1}$$

Then by Schwarz inequality

$$\begin{aligned} \int_0^{b(x)} \|\nabla f_{\phi_t(x)}\| dt &\leq \sqrt{b(x)} \left(\int_0^{b(x)} \|\nabla f_{\phi_t(x)}\|^2 dt \right)^{\frac{1}{2}} \\ &= \sqrt{b(x)} (g(0) - g(b(x)))^{\frac{1}{2}}, \end{aligned}$$

which is less than or equal to $\sqrt{b(x)}(g(0) - B)^{1/2}$. This is finite if $b(x)$ is so. But $b(x)$ is finite implies that $\int_0^{b(x)} \|\nabla f_{\phi_t(x)}\| dt = \infty$, by Proposition 9.2.4. Therefore we must have $b(x) = \infty$, and then taking $T \to \infty$ we get from (1)

$$\int_0^{\infty} \|\nabla f_{\phi_t(x)}\|^2 dt \leq g(0) - B.$$

This means that $\|\nabla f_{\phi_t(x)}\|$ must not be bounded away from zero as $t \to \infty$, i.e. we cannot have $\|\nabla f_{\phi_t(x)}\| > K > 0$ for some constant K as $t \to \infty$, otherwise $\int_0^\infty \|\nabla f_{\phi_t(x)}\|^2 dt$ would be infinite. This completes the proof the first part.

For the second part note that, since $f(\phi_t(x))$ is bounded, Condition C implies that $\phi_t(x)$ has a critical point of f as limit point when $t \to \infty$. □

Complement 9.2.7. The proof shows that as $t \to b(x)$ either

$$f(\phi_t(x)) \to \infty$$

or else $b(x)$ must be ∞ and $\phi_t(x)$ has a critical point of f as limit point as $t \to \infty$.

A similar arguments will show that as $t \to a(x)$ either $f(\phi_t(x)) \to -\infty$ or else $a(x)$ must be $-\infty$ and $\phi_t(x)$ has a critical point of f as limit point as $t \to -\infty$.

Corollary 9.2.8. *If $x \in M$ is not a critical point of f, then there is a critical point p of f such that $f(p) < f(x)$.*

Proof. There passes through x an integral curve $\phi_t(x)$ which tends to a critical point of p as $t \to \infty$. Since $f(\phi_t(x))$ is a strictly decreasing function of t, it follows that $f(p) < f(x)$. □

Theorem 9.2.9. *The function f attains its infimum B, that is, there is a critical point p of f such that $f(p) = B$.*

PROOF. Choose a sequence $\{x_k\}$ with $f(x_k) \to B$. Then by Corollary 9.2.8, for each k there is a critical point p_k such that $f(p_k) < f(x_k)$. This means that the sequence $\{f(p_k)\}$ also tends to the limit B. Therefore the sequence $\{f(p_k)\}$ is bounded, and by Condition C a subsequence of $\{p_k\}$ converges to a critical point p of f. It follows that $f(p) = B$. $\square$

Recall that M_a denotes the manifold $f^{-1}((-\infty, a])$.

Theorem 9.2.10. *If $f : M \longrightarrow \mathbb{R}$ is a smooth function on a manifold M satisfying Condition C, and if $[a, b]$ is an interval which contains no critical value of f, then*

(a) *M_b is diffeomorphic to M_a, and ∂M_b is diffeomorphic to ∂M_a,*

(b) *there is a deformation $F : M \times \mathbb{R} \longrightarrow M$ such that, for each t, F_t is a diffeomorphism and maps any level surface $f^{-1}(c)$ diffeomorphically on to a level surface $f^{-1}(c')$, where $c, c' \in [a, b]$ with $c \geq c'$; also $F_0 = \mathrm{Id}$ and $F_1(M_b) = M_a$. Moreover, there is a compact neighbourhood K of $[a, b]$ in $\mathbb{R}$ such that F_t is* Id *outside $M - f^{-1}(K)$ for all $t \in \mathbb{R}$.*

The part (b) of the theorem says that F is a diffeotopy of M which deforms M_b onto M_a pushing down the levels of f, while keeping the complement of $f^{-1}(K)$ fixed.

PROOF. (a) By assumption $[a, b] \cap f(\mathcal{C}) = \emptyset$, where $\mathcal{C}$ is the set of all critical points of f. Since $f(\mathcal{C})$ is closed set (Corollary 9.2.2), the interval $[a - \epsilon, b + \epsilon]$ also contains no critical values of f for some $\epsilon > 0$. The vector field $X = -\nabla f/||\nabla f||^2$ is defined on the set $M - \mathcal{C}$, since $||\nabla f|| \neq 0$ there. We have $Xf = -1$, since $(\nabla f)f = ||\nabla f||^2$. If $g : \mathbb{R} \longrightarrow \mathbb{R}$ is a smooth function vanishing on a neighbourhood of $f(\mathcal{C})$, then $Y = (g \circ f)X$ is a smooth field on M which vanishes on a neighbourhood of $\mathcal{C}$, and $Yf = -(g \circ f)$. We shall suppose that $g : \mathbb{R} \longrightarrow \mathbb{R}$ is a non-negative smooth function which is identically equal to one on a neighbourhood of $[a, b]$ and zero outside $[a - \epsilon/2, b + \epsilon/2]$. Then the vector field Y vanishes outside the set $f^{-1}([a - \epsilon/2, b + \epsilon/2])$.

We shall show that the vector field Y on M has bounded length. Then, by Corollary 9.2.5, Y will generate a one-parameter group of diffeomorphisms of M. We have $\|Y\| = \frac{1}{\|\nabla f\|}|g \circ f|$. Since g is bounded, and $|g \circ f|$ vanishes outside $f^{-1}([a - \epsilon/2, b + \epsilon/2])$, it is sufficient to show that $\frac{1}{\|\nabla f\|}$ is bounded on $f^{-1}([a - \epsilon/2, b + \epsilon/2])$. If this is not true, or equivalently if $\|\nabla f\|$ is not bounded away from zero on $f^{-1}([a - \epsilon/2, b + \epsilon/2])$, then there would exist a sequence $\{x_k\}$ in $f^{-1}([a - \epsilon/2, b + \epsilon/2])$ with $\|\nabla f_{x_k}\| \to 0$, and we would find a subsequence of $\{x_k\}$ converging to a critical point p of f. Then $f(p) \in [a - \epsilon/2, b + \epsilon/2]$, contradicting the assumption that the interval $[a - \epsilon, b + \epsilon]$ contains no critical value of f. Thus Y generates a one-parameter group of diffeomorphisms ϕ_t on M such that $\phi_t = \mathrm{Id}$ outside $f^{-1}([a - \epsilon/2, b + \epsilon/2])$.

Let $\sigma(t,c)$ be the unique solution of the ordinary differential equation

$$\frac{d\sigma}{dt} = -g(\sigma(t)), \text{ with initial value } \sigma(0,c) = c.$$

Then it follows from the choice of g that $\sigma(t,c)$ must satisfy the conditions

$$\sigma(t,c) = \begin{cases} c-t & \text{if} \quad c \in [a,b] \text{ and } c-t \geq a, \\ c & \text{if} \quad c < a-\epsilon/2 \text{ or } c > b+\epsilon/2. \end{cases}$$

Since

$$\frac{d}{dt}\Big(f \circ \phi_t(x)\Big) = \Big\langle \frac{d\phi_t(x)}{dt}, \nabla f \Big\rangle = Y_{\phi_t(x)} f = -g(f \circ \phi_t(x),$$

it follows that $f(\phi_t(x)) = \sigma(t, f(x))$, by the uniqueness of solution. Then $\phi_t(x) \in f^{-1}(\sigma(t,f(x))$, and so $\phi_t(f^{-1}(c)) = f^{-1}(\sigma(t,c))$, showing that ϕ_t maps a level surface onto a level surface. In particular, if $f(x) = c \leq b$, then $\phi_{b-a}(x) = f^{-1}(\sigma(b-a,c)) = f^{-1}(c-b+a) \subseteq f^{-1}(a)$, and $\phi_{b-a}(f^{-1}(b)) = f^{-1}(\sigma(b-a,b)) = f^{-1}(a)$. Therefore ϕ_{b-a} maps M_b diffeomorphically onto M_a, and $f^{-1}(b)$ onto $f^{-1}(a)$ This proves the first part of the theorem.

(b) Define the deformation F by $F(x,t) = \phi_{(b-a)t}(x)$, and take the compact neighbourhood K of $[a,b]$ to be $[a-\epsilon/2, b+\epsilon/2]$. Then all the requirements may easily be seen to hold using the arguments in the part (a). □

The above theorem shows that if we do not pass through a critical value while moving from a to b, then the level sets $f^{-1}(a)$ and $f^{-1}(b)$ are diffeomorphic, so are M_a and M_b, where the diffeomorphisms can be realised as translations along the integral curves of the negative gradient vector field $-\nabla f$ on M.

We shall now examine how the diffeomorphism type of M_a changes when we pass through a critical value between a and b. For this purpose, we introduce the notion of a handle and the operation of attaching a handle to a manifold.

Definition 9.2.11. A **handle of dimension** n **and index** r denoted by $\mathcal{H}^n_r$, is the Cartesian product $D^r \times D^{n-r}$, where D^r is the closed disk in $\mathbb{R}^r$ with centre at the origin. This is a manifold with corner $S^{r-1} \times S^{n-r-1}$, and with boundary

$$\partial\mathcal{H}^n_r = (S^{r-1} \times D^{n-r}) \cup (D^r \times S^{n-r-1}).$$

Sometimes we call a handle of index r an r-handle.

Definition 9.2.12. Let M and N be manifolds of the same dimension n. Then N is said to be obtained from M by attaching a handle $\mathcal{H}^n_r$ if there is a homeomorphism ϕ of $\mathcal{H}^n_r$ onto a closed subset K of N such that

(a) $N = M \cup K$,
(b) $\phi \mid S^{r-1} \times D^{n-r}$ is a diffeomorphism onto $\partial M \cap K$,
(c) $\phi \mid (\text{Int } D^r) \times D^{n-r}$ is a diffeomorphism onto $N - M$.

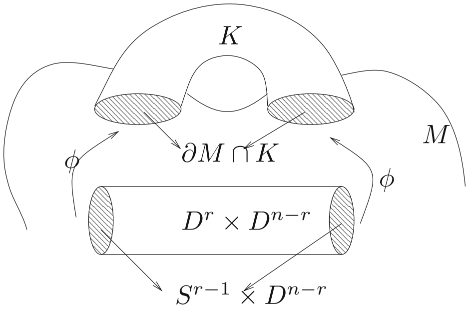


FIGURE 9.1

The map ϕ is called the **attaching map**.

The smooth structure that we give to N is obtained by the process of 'straightening the angle' along the corners so that the handle $\mathcal{H}_r^n$ is smoothly embedded in N. As we know this smooth structure is well-defined up to diffeomorphism.

Theorem 9.2.13. *Let $f : M \longrightarrow \mathbb{R}$ be a Morse function, which is bounded below and satisfies the Condition C. Let c be a critical value of f so that there is only one critical point p of index r in $f^{-1}([c-\epsilon, c+\epsilon])$. Then $M_{c+\epsilon}$ is diffeomorphic to the manifold obtained by attaching a handle $\mathcal{H}_r^n$ of index r to $M_{c-\epsilon}$.*

PROOF. The case $r = n$ is trivial. Here p is a maximum point of f,

$$M_{c+\epsilon} = M \quad \text{and} \quad \mathrm{M_{c-\epsilon} = M - Int \;\; D^n},$$

and so M is diffeomorphic to

$$(M - \text{Int} \;\; D^n) \cup D^n,$$

which is $M_{c-\epsilon}$ with a handle of index n attached. The case $r = 0$ is equally trivial. This is the case $r = n$ turned upside down.

We therefore suppose $0 < r < n$, and construct a proof of the theorem in the following four steps.

Plan of the proof. We identify a neighbourhood of the critical point p with an open neighbourhood U of the origin $\underline{0}$ in $\mathbb{R}^n = \mathbb{R}^r \times \mathbb{R}^{n-r}$ by means of a Morse chart so that p gets identified with $\underline{0}$. We may also assume that the critical value c is equal to $0 \in \mathbb{R}$. We choose $\epsilon > 0$ small enough so that 0 is the only critical value of f in $[-2\epsilon, 2\epsilon]$, or $\underline{0}$ is the only critical point of f with $|f(\underline{0})| \leq 2\epsilon$. We next make ϵ still smaller, if necessary, so that the closed disk of radius $2\sqrt{\epsilon}$, $D^n(2\sqrt{\epsilon})$, is contained in U. We shall denote the boundary of $D^n(2\sqrt{\epsilon})$ by $S^{n-1}(2\sqrt{\epsilon})$

Then in U, f is given by

$$f(x, y) = -\| \; x \; \|^2 + \| \; y \; \|^2, \quad x \in \mathbb{R}^r, \; y \in \mathbb{R}^{n-r}.$$

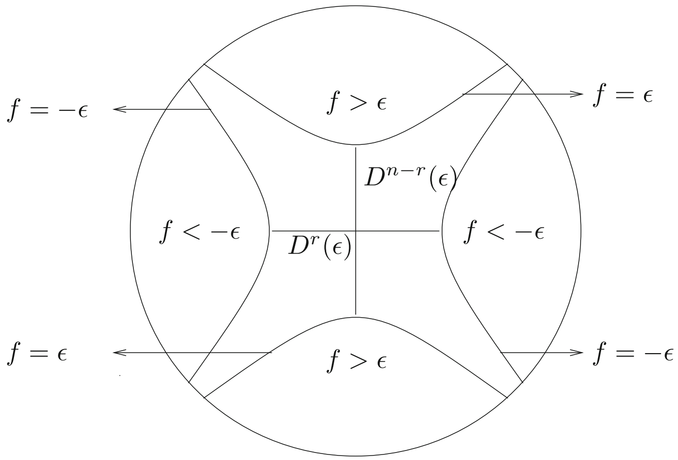


FIGURE 9.2

We shall construct a smooth function $g : M \longrightarrow \mathbb{R}$ such that

(1) $g \leq f$ everywhere on M,
(2) $g \leq -\epsilon \Rightarrow f < \epsilon$,
(3) $f \geq \epsilon \Rightarrow g = f$,
(4) the point $p = \underline{0}$ is the only critical point of g in $g^{-1}([-2\epsilon, \epsilon])$ with critical value $g(\underline{0}) = -\frac{3\epsilon}{2}$, and so g has no critical value in the interval $[-\epsilon, \epsilon]$.

Then, if N is the manifold $M_{-\epsilon}(g)$, we have by (1) and (2)

$$M_{-\epsilon}(f) \leq N \leq M_{\epsilon}(f).$$

Also (1) implies $M_\epsilon(f) \leq M_\epsilon(g)$, and (3) implies its converse $M_\epsilon(g) \leq M_\epsilon(f)$ (because if $g \leq \epsilon$ then either $f \geq \epsilon$ or $f \leq \epsilon$, and in the first case $f = g$). Moreover, the interval $[-\epsilon, \epsilon]$ contains no critical value of g, so by Theorem 9.2.10, there is an isotopy of M which deforms $M_\epsilon(f) = M_\epsilon(g)$ onto $M_{-\epsilon}(g) = N$.

It will turn out that N is the union $N = M_{-\epsilon}(f) \cup K$, where K is the closure of the set $\{(x, y) \in U \mid f(x, y) > -\epsilon, \text{ and } g(x, y) < -\epsilon\}$. We shall show that there is a diffeomorphism ϕ of the handle H^n_r onto K which maps $S^{r-1} \times D^{n-r}$ onto $\partial(M_{-\epsilon}(f)) \cap K$. This will complete the proof.

Step 1. We now go into the construction of the function g. Take a smooth non-increasing function $\lambda : \mathbb{R} \longrightarrow \mathbb{R}$ such that $0 \leq \lambda \leq 1$, $\lambda = 1$ when $t \leq \frac{1}{2}$, $\lambda > 0$ when $t < 1$, and $\lambda = 0$ when $t \geq 1$, and such that the derivative $\lambda' \leq 0$. We may take $\lambda(t) = \mathcal{B}(2 - 2t)$, where $\mathcal{B}$ is a bump function. Then define a function $g : U \longrightarrow \mathbb{R}$ by

$$g(x, y) = f(x, y) - \frac{3\epsilon}{2}\lambda\left(\frac{\|y\|^2}{\epsilon}\right).$$

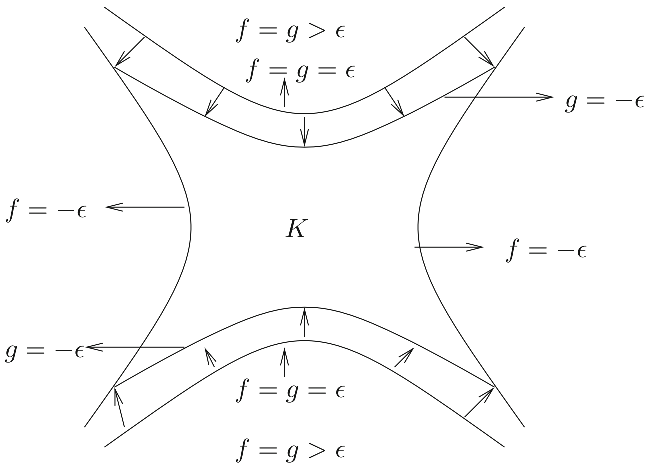


FIGURE 9.3

Then g satisfies (1) and (2), because $0 \le \lambda \le 1$. Also (3) holds, because

$$f(x,y) = -\|x\|^2 + \|y\|^2 \ge \epsilon \Rightarrow \frac{\|y\|^2}{\epsilon} \ge \frac{\|x\|^2}{\epsilon} + 1 \ge 1 \Rightarrow \lambda(\|y\|^2/\epsilon) = 0.$$

Two more properties of g are as follows.

(5) If $(x,y) \in U$, $f(x,y) \ge -2\epsilon$, and $f(x,y) \ne g(x,y)$, then $(x,y) \in D^n(2\sqrt{\epsilon}) \subset U$. This holds, because

$$\|x\|^2 + \|y\|^2 = 2\|y\|^2 + \|x\|^2 - \|y\|^2 = 2\|y\|^2 - f(x,y) < 2\epsilon + 2\epsilon = 4\epsilon.$$

(6) If $(x,y) \in S^{n-1}(2\sqrt{\epsilon})$, and $f(x,y) \ge -2\epsilon$, then $f = g$. This holds, because if $\|x\|^2 + \|y\|^2 = 4\epsilon$ and $-\|x\|^2 + \|y\|^2 \ge -2\epsilon$, then $\|y\|^2 \ge \epsilon$, and so $f = g$.

The properties (5) and (6) says that $f = g$ on

$$\big(U - \operatorname{Int} D^n(2\sqrt{\epsilon})\big) \cap f^{-1}([-2\epsilon, \infty)).$$

Therefore we may extend g to the remaining portion of $f^{-1}([-2\epsilon,\infty))$ by setting it equal to f outside U. This function on $f^{-1}([-2\epsilon,\infty))$ may further be extended to all of M in the following way. Let $\{U_i\}$ be an open covering of $f^{-1}([-2\epsilon,\infty))$ by open sets in M. This together with the open sets U and $M - f^{-1}([-2\epsilon,\infty))$ form an open covering $\mathcal{U}$ of M. Then the family of functions

$$f|U_i, \quad g|U, \quad f|(M - f^{-1}([-2\epsilon,\infty)))$$

may be glued together using a smooth partition of unity subordinate to the covering $\mathcal{U}$. The resulting function g will be smooth and satisfy the inequality $g \le f$.

Having constructed the extended function $g : M \longrightarrow \mathbb{R}$, let us now see how g satisfies the property (4) also. First note that if S is the set $g^{-1}([-2\epsilon,\epsilon])$, then

g has no critical point in $S-U$. This follows, because $g = f$ on $f^{-1}([-2\epsilon,\infty))-U$,

$$S - U \subset g^{-1}([-2\epsilon,\infty)) - U = f^{-1}([-2\epsilon,\infty)) - U,$$

and f has no critical point in $f^{-1}([-2\epsilon, 2\epsilon])$. Next note that in U

$$dg = -2xdx + \Big[2 - 3\lambda'\Big(\frac{\|y\|^2}{\epsilon}\Big)\Big]ydy.$$

This vanishes only at the origin, since the derivative $\lambda' \le 0$. Thus g has no critical point in S, and we get (4) for g.

Step 2. The property (3) implies the following equality of sets

$$\{(x, y) \in U \mid f(x, y) \ge \epsilon\} = \{(x, y) \in U \mid g(x, y) \ge \epsilon\}.$$

Therefore, taking complements of the sets, we get $M_\epsilon(f) = M_\epsilon(g)$. In the case of interest to us,

$$N = M_{-\epsilon}(g) = \{(x, y) \in U \mid g(x, y) \le -\epsilon\},$$

and K is the closure of the set

$$\{(x, y) \in U \mid f(x, y) > -\epsilon \text{ and } g(x, y) < -\epsilon\}.$$

Then $K \subset N$. Also, since $g \le f$, $M_{-\epsilon}(f) \subset M_{-\epsilon}(g)$. Therefore $N \supseteq M_{-\epsilon}(f) \cup K$. Since $f = g$ outside U where $f \ge -\epsilon$, it follows that $N = M_{-\epsilon}(f) \cup K$. The proof of the theorem will be complete, if we construct a homeomorphism of $D^r \times D^{n-r}$ onto K satisfying the conditions (b) and (c) of Definition 9.2.12.

Define a map $\phi : \mathbb{R}^r \times \mathbb{R}^{n-r} \longrightarrow \mathbb{R}^r \times \mathbb{R}^{n-r}$ by setting $\phi(u, v) = (x, y)$ where

$$x = [\epsilon\sigma(\|u\|^2)\cdot\|v\|^2 + \epsilon]^{\frac{1}{2}}u, \quad y = [\epsilon\sigma(\|u\|^2)]^{\frac{1}{2}}v.$$

Here σ is some real valued positive smooth function of a real variable which we shall construct so that ϕ maps $D^r \times D^{n-r}$ homeomorphically onto K satisfying the conditions (b) and (c) of Definition 9.2.12. But meanwhile note that, whatever may be the smooth function σ, ϕ is a homeomorphism with inverse $\psi(x, y) = (u, v)$ given by

$$u = [\epsilon + \|y\|^2]^{-\frac{1}{2}}x, \quad y = \Big[\epsilon\sigma\Big(\frac{\|x\|^2}{\epsilon + \|y\|^2}\Big)\Big]^{-\frac{1}{2}}v.$$

We now go over to the construction of σ. Recall that $\lambda(t)$ is the function $\mathcal{B}(2-2t)$, whose graph is given in the following figure.

Then the function $\lambda(t)/(1+t)$ is a smooth function with a strictly negative derivative on $[0, 1)$. Let $\mu : [0, 1) \longrightarrow [0, 1]$ denote the smooth inverse function of this, and $\eta : [0, 1] \longrightarrow [0, 2/3]$ be the function $\eta(t) = \frac{2}{3}(1-t)$. Then we define σ to be the smooth function $\mu\circ\eta$. Clearly $\sigma(0) = \frac{1}{2}$, $\sigma(1) = 1$, and σ has a positive derivative on $[0, 1)$ (so it is strictly increasing). Moreover, it satisfies the equation

$$\frac{\lambda(\sigma(t))}{1 + \sigma(t)} = \frac{2}{3}(1 - t).$$

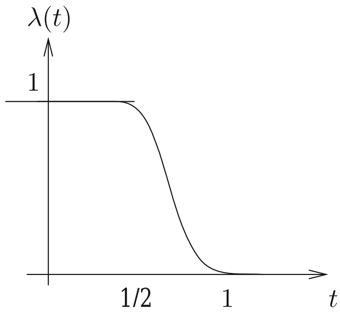


FIGURE 9.4

Thus $\sigma(t)$ is the unique solution of the above equation. The construction of σ is shown in the following figure.

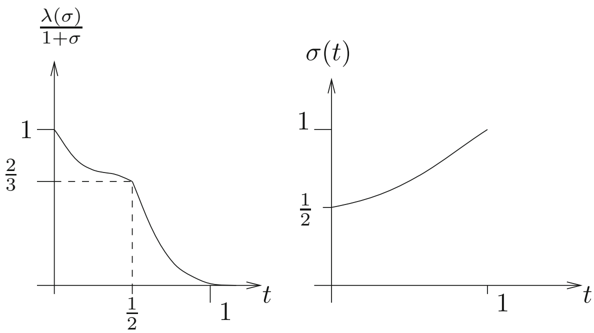


FIGURE 9.5

The following lemma shows that ϕ maps $D^r \times D^{n-r}$ into K.

Lemma 9.2.14. *Let $F, G : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be the functions defined by*

$$F(x,y) = -[\epsilon\sigma(x^2)y^2 + \epsilon]x^2 + [\epsilon\sigma(x^2)]y^2,$$

$$G(x,y) = F(x,y) - \frac{3\epsilon}{2}\lambda(\sigma(x^2)y^2)$$

(note that, if $(u,v) \in D^r \times D^{n-r}$ and $\phi(u,v) = (x,y)$, then

$$f(x,y) = F(\|x\|, \|y\|) \text{ and } g(x,y) = G(\|x\|, \|y\|)\ ).$$

Then in the square region $[0,1] \times [0,1]$, we have

$$F(x,y) \geq -\epsilon, \text{ and } G(x,y) \leq -\epsilon.$$

PROOF. For any fixed $x \in [0,1)$, the region $[0,1) \times [0,1]$ contains only one critical point of F at $y = 0$, where F is minimum and the minimum value is $-\epsilon x^2$ which is $> -\epsilon$. Also at $x = 1$, $F = -\epsilon$. Therefore $F(x,y) \geq -\epsilon$ everywhere on $[0,1] \times [0,1]$.

Similarly, for a fixed $x \in [0,1)$, the region $[0,1) \times [0,1]$ contains only one critical point of G at $y = 0$, where G is minimum. Therefore G must attain its maximum when $x = 1$. Now on the line $x = 1$

$$G = -\epsilon - \frac{3\epsilon}{2}\lambda(y^2) < -\epsilon.$$

Therefore $G(x,y) < -\epsilon$ everywhere on $[0,1] \times [0,1]$. This completes the proof of the lemma. $\square$

Thus $\phi(D^r \times D^{n-r}) \subset K$.

Step 3. Next we shall show that ψ (which is the inverse of ϕ) maps K into $D^r \times D^{n-r}$. For this purpose, we need the following lemma.

Lemma 9.2.15. *Let $\overline{F}, \overline{G} : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be the functions defined by*

$$\overline{F}(x,y) = -x^2 + y^2 \text{ and } \overline{G}(x,y) = \overline{F}(x,y) - \frac{3\epsilon}{2}\lambda\Big(\frac{y^2}{\epsilon}\Big)$$

(so that, in U, $f(x,y) = \overline{F}(\|x\|, \|y\|)$ and $g(x,y) = \overline{G}(\|x| \, \|y\|))$. Then in the region which is the closure of the set

$$R = \{(x,y) \in \mathbb{R}^2 \mid \overline{F}(x,y) \geq -\epsilon,\ \overline{G}(x,y) \leq -\epsilon\}$$

we have

$$y^2 \leq \epsilon\sigma\Big(\frac{x^2}{\epsilon + y^2}\Big).$$

Proof. The proof consists of showing that the function $h : \mathbb{R}^2 \longrightarrow \mathbb{R}$ defined by

$$h(x,y) = y^2 - \epsilon\sigma\Big(\frac{x^2}{\epsilon + y^2}\Big)$$

is ≤ 0 in R. This follows by simple computations. For fixed x, h has only one critical point at $y = 0$ where it has a minimum. Therefore h must attain its maximum on the boundary ∂R, which consists of two curves

$$C_1 = \{(x,y) \in R \mid \overline{F}(x,y) = -\epsilon,\ \overline{G}(x,y) < -\epsilon\},$$

$$C_2 = \{(x,y) \in R \mid \overline{F}(x,y) > -\epsilon,\ \overline{G}(x,y) = -\epsilon\}.$$

We shall show that $h < 0$ on C_1, and $h = 0$ on C_2.

On C_1, we have $\overline{G} < \overline{F}$, so $(-3\epsilon/2)\lambda(y^2/\epsilon) < 0$, or $\lambda(y^2/\epsilon) > 0$, which implies $y^2 < \epsilon$ by the definition of λ. On the other hand, since $\overline{F}(x,y) = -x^2 + y^2 = -\epsilon$, we have $x^2/(\epsilon + y^2) = 1$, and so $\sigma(x^2/(\epsilon + y^2)) = 1$. Therefore $h(x,y) = y^2 - \epsilon < 0$.

On C_2, we have $\overline{G} < \overline{F}$ again, and so $y^2 < \epsilon$ as before. Also $\overline{G}(x,y) = -\epsilon$ implies

$$\frac{x^2}{\epsilon + y^2} = 1 - \Big(\frac{3}{2}\Big)\frac{\lambda(\frac{y^2}{\epsilon})}{1 + \frac{y^2}{\epsilon}}\ ,$$

after simplification. From this expression we may conclude that y^2/ϵ cannot be $< 1/2$. For

$$\frac{y^2}{\epsilon} < \frac{1}{2} \Rightarrow \lambda(\frac{y^2}{\epsilon}) = 1 \text{ and } 1 + \frac{y^2}{\epsilon} < \frac{3}{2} \Rightarrow \frac{x^2}{\epsilon + y^2} < 0,$$

which is not possible. Therefore y^2/ϵ must lie in $[1/2, 1)$, which is the range of σ. Let α be the point in $[0, 1]$ such that $\sigma(\alpha) = y^2/\epsilon$. Therefore

$$\frac{x^2}{\epsilon + y^2} = 1 - \left(\frac{3}{2}\right)\frac{\lambda(\sigma(\alpha))}{1 + \sigma(\alpha)} = 1 - \left(\frac{3}{2}\right)\left(\frac{2}{3}\right)(1 - \alpha) = \alpha,$$

by the definition of σ, and hence

$$h(x, y) = y^2 - \epsilon\sigma\left(\frac{x^2}{\epsilon + y^2}\right) = \epsilon\sigma(\alpha) - \epsilon\sigma(\alpha) = 0,$$

and so h vanishes on C_2. This completes the proof of the lemma. $\square$

We shall now use the lemma to show that ψ maps K into $D^r \times D^{n-r}$. Let $(x, y) \in K$, and $\psi(x, y) = (u, v)$. Then $F(\|x\|, \|y\|) = -\|x\|^2 + \|y\|^2 \geq -\epsilon$ and $G(\|x\|, \|y\|) \leq -\epsilon$. The first inequality gives

$$\frac{\|x\|^2}{\epsilon + \|y\|^2} \leq 1,$$

which means that $\|u\|^2 \leq 1$, and so $u \in D^r$. Also, by the above lemma,

$$\frac{\|y\|^2}{\epsilon\sigma\left(\frac{\|x\|^2}{\epsilon + \|y\|^2}\right)} \leq 1,$$

which means that $\|v\|^2 \leq 1$, and so $v \in D^{n-r}$.

Step 4. Thus ϕ maps $D^r \times D^{n-r}$ homeomorphically onto K. Since σ is smooth and has positive derivative on $[0, 1)$, it follows that ϕ is a diffeomorphism on $(\text{Int } D^r) \times D^{n-r}$. On $S^{r-1} \times D^{n-r}$, ϕ reduces to

$$\phi(u, v) = ([\epsilon\|v\|^2 + \epsilon]^{\frac{1}{2}} u, \epsilon^{\frac{1}{2}} v),$$

which is clearly a diffeomorphism onto

$$\{(x, y) \in U \mid f(x, y) = -\epsilon, g(x, y) \leq -\epsilon\} = K \cap \partial(M_{-\epsilon}(f)).$$

This completes the proof of the fact that $M_{c+\epsilon}$ is diffeomorphic to $M_{c-\epsilon}$ with a handle of index r attached. $\square$

Remark 9.2.16. The above process of attaching handle to $M_{-\epsilon}$ takes place in a small coordinate neighbourhood (the domain of a Morse chart) of a critical point p. If there are several critical points at the same critical level, then it is possible to carry through the above constructions independently in disjoint neighbourhoods of the critical points. This consideration leads us to the following definition and theorem.

Definition 9.2.17. Suppose there is a sequence of manifolds

$$M = N_0, N_1, \dots, N_k = N$$

such that N_j is obtained from N_{j-1} by attaching a handle of index r_j with attaching map ϕ_j, $j = 1, \dots, k$. Then, if the images of the ϕ_j are disjoint, we say that N is obtained from M by attaching handles of indices $r_1, \dots, r_k$ with attaching maps $\phi_1, \dots, \phi_k$ disjointly.

Theorem 9.2.18. *Let $f : M \longrightarrow \mathbb{R}$ be a Morse function, which is bounded below and satisfies the Condition C. Let $c \in (a, b)$ be the only critical value of f in $[a, b]$ such that $p_1, \dots, p_k$ are all the critical points of f of indices $r_1, \dots, r_k$ respectively at the critical level $f^{-1}(c)$. Then M_b is obtained from M_a by attaching handles of indices $r_1, \dots, r_k$ disjointly.*

Example 9.2.19 (**The height function on the torus**)**.** The torus T is obtained as the surface of revolution in $\mathbb{R}^3$ by rotating the circle

$$x^2 + (y - 2)^2 = 1$$

about the x-axis. Then the height function $f : T \longrightarrow \mathbb{R}$ given by $f(x, y, z) = z$ is a Morse function with four non-degenerate critical points: a minimum at $(0, 0, -3)$, a maximum at $(0, 0, 3)$, and two saddle points at the points $(0, 0, -1)$, and $(0, 0, 1)$ on the inner circle. The indices of the saddle points are both 1, and those of the minimum and the maximum are 0 and 2 respectively. Figure 9.6 below shows how the torus can be build up step by step starting with a disk (i.e. 0-handle), then attaching two 1-handles, and finally completing the torus with a 2-handle.

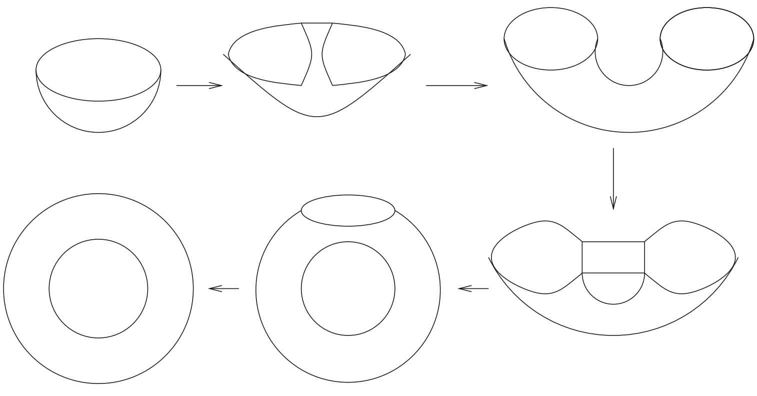


FIGURE 9.6

$\Diamond$ **Exercise 9.1** (**Riemann surface of genus** g)**.** Let g be an integer ≥ 1. Take g copies of the above torus $T_1, \dots, T_g$. Let a_k and b_k be respectively the minimum and the maximum points of T_k. Remove a small disk D_k around a_k for $k = 2, \cdots, g$, and remove a small disk E_k around b_k for $k = 1, \cdots, g - 1$.

Stick the tori together by gluing ∂E_k and ∂D_{k+1}, for $k = 1, \cdots, g-1$. The resulting surface Σ_g is the connected sum

$$T_1 \# \cdots \# T_k \# \cdots \# T_g$$

(see Definition 7.6.14). We shall call this surface a Riemann surface of genus g (g is the number of holes in Σ_g).

Define the same height function on Σ_g as for T of the above Example 9.2.19. Show that the height function has $2g+2$ non-degenerate critical points comprised of a minimum, a maximum, and $2g$ saddle points. What are the indices of the critical points?

9.3. Morse inequalities

Lemma 9.3.1. *The set $D^r \times \{0\} \cup S^{r-1} \times D^{n-r}$ is a strong deformation retract of $D^r \times D^{n-r}$.*

Proof. If X is a convex subset of $\mathbb{R}^n$ and A is a closed subset of X, then any retraction $\rho : X \longrightarrow A$ is a strong deformation retract of X onto A, where a strong deformation retraction h is defined by $h(x,t) = (1-t)x + t\rho(x)$.

Therefore it is enough to construct a retraction

$$\rho : D^r \times D^{n-r} \longrightarrow D^r \times \{0\} \cup S^{r-1} \times D^{n-r}.$$

Such a ρ may be defined by

$$\begin{aligned} \rho(x,y) &= \left(\tfrac{2x}{2-\|y\|},\ 0\right) && \text{if } \|x\| \leq 1 - \tfrac{\|y\|}{2}, \\ &= \left(\tfrac{x}{\|x\|},\ [\ \|x\| + 2\|y\| - 2\ ]\tfrac{y}{\|y\|}\right) && \text{otherwise.} \end{aligned}$$

□

The diagram of the retraction ρ is given in the following Figure 9.7.

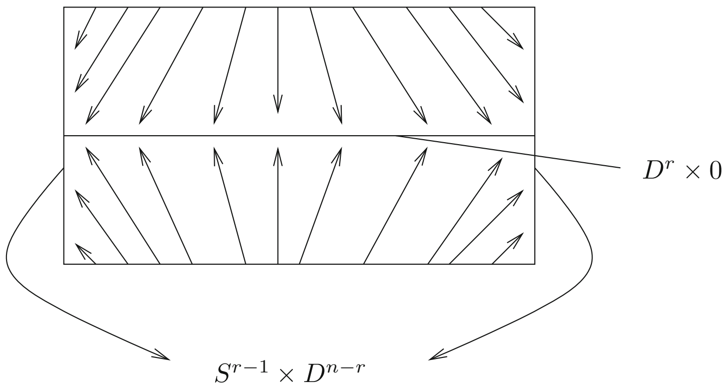


Figure 9.7

We recall the concept of attaching a cell to a topological space.

Definition 9.3.2. Let X be a topological space with a closed subspace Y. Let $G : D^r \longrightarrow X$ be a continuous map onto a closed subset of X, and $g = G|S^{r-1}$. We write $G(D^r) = e^r$ and $\partial e^r = e^r \cap Y$. Then X is said to be obtained from Y by attaching an r-cell by means of g, written $X = Y \cup_g e^r$, if

(1) $X = Y \cup e^r$,
(2) G maps $D^r - S^{r-1}$ homeomorphically onto $e^r - Y$,
(3) $g(S^{r-1}) = \partial e^r$.

The map G is called the **characteristic map**, and g is called the **attaching map**.

Note that X may be reconstructed as the quotient space of the disjoint union $Y \cup D^r$ modulo an equivalence relation R defined in the following way: if $x, y \in Y \cup D^r$, then $x\,R\,y$ if one of the following is true: (a) $x = y$, (b) $x, y \in S^{r-1}$ and $g(x) = g(y)$, (c) $x \in S^{r-1}$ and $y = g(x) \in Y$.

All our homology groups that we shall consider will have coefficients in $\mathbb{R}$. Although all the results will hold equally well for any coefficient field $\mathbb{F}$.

Lemma 9.3.3. *If X is obtained from Y by attaching an r-cell, then*

$$H_q(X,Y) = \begin{cases} \mathbb{R} & \text{if } q = r, \\ 0 & \text{if } q \neq r. \end{cases}$$

Proof. First recall that $\widetilde{H}_{q-1}(S^{r-1}) = \mathbb{R}$ if $q = r$, and it is 0 if $q \neq r$ (see for example, Spanier [**43**], Theorem 4.6.6, p.190). These are also the homology groups of (D^r, S^{r-1}), because D^r is contractible, and so the boundary homomorphisms in the exact homology sequence of the pair (D^r, S^{r-1}) are isomorphisms

$$\partial_* : H_q(D^r, S^{r-1}) \cong \widetilde{H}_{q-1}(S^{r-1}).$$

Next, note that in view of the condition (2) of Definition 9.3.2, we have a relative homeomorphism between the pairs $(e^r, \partial e^r)$ and (D^r, S^{r-1}), and therefore the homology groups $H_q(e^r, \partial e^r)$ and $H_q(D^r, S^{r-1})$ are isomorphic. Again, we have an excision isomorphism between $H_q(X, Y)$ and $H_q(e^r, \partial e^r)$. Combining these isomorphisms we have

$$H_q(X,Y) \cong H_q(e^r, \partial e^r) \cong H_q(D^r, S^{r-1}).$$

This completes the proof. □

Theorem 9.3.4. *If M and N are manifolds with boundary so that N arises from M by attaching a handle of index r, then N has as a strong deformation retract a closed subspace $X = M \cup_g e^r$, which is obtained from M by attaching an r-cell e^r. In particular, N has the homotopy type of X, and so*

$$H_q(N,M) = \begin{cases} \mathbb{R}, & \text{if } q = r, \\ 0, & \text{otherwise.} \end{cases}$$

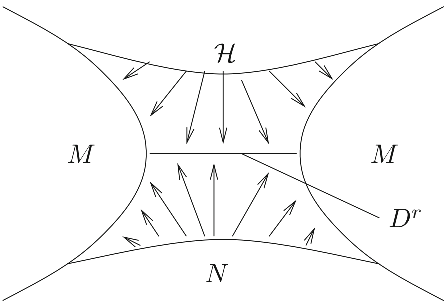


FIGURE 9.8

PROOF. Let $\phi : D^r \times D^{n-r} \longrightarrow K$ be the map with which the handle K is attached to M for obtaining N. Define $G : D^r \longrightarrow e^r$ by $G = \phi|(D^r \times \{0\})$, and $g = G|S^{r-1}$. Then the deformation retract ρ' of $N = M \cup_\phi K$ onto $M \cup_g e^r$ is given by

$$\rho' = \begin{cases} \mathrm{Id} & \text{on } M, \\ \phi \circ \rho \circ \phi^{-1} & \text{on } K, \end{cases}$$

where ρ is the deformation retraction $\rho : D^r \times D^{n-r} \longrightarrow D^r \times \{0\} \cup S^{r-1} \times D^{n-r}$ of Lemma 9.3.1. $\square$

Corollary 9.3.5. *If N is obtained from M by attaching handles of indices $r_1, \dots, r_k$ disjointly, then N has as a strong deformation retract a closed subspace*

$$X = M \cup_{\phi_1} e^{r_1} \cup_{\phi_2} \cdots \cup_{\phi_k} e^{r_k},$$

obtained from M by attaching cells $e^{r_1}, \cdots, e^{r_k}$ disjointly.

We next recall the definition of CW pair of spaces.

Definition 9.3.6. A **relative CW-complex** or **CW-pair** is a pair of topological space (X, A) together with a sequence of closed subspaces

$$A = X_0 \subseteq X_1 \subseteq \cdots \subseteq X_m = X$$

and continuous maps $g_j : S^{r_j - 1} \longrightarrow X_j$, $j = 0, 1, \dots, m-1$, where $r_i \leq r_j$ if $i \leq j$, such that X_{j+1} is homeomorphic to $X_j \cup_{g_j} e^{r_j}$. The maximum dimension of the cells in $X - A$ is called the **dimension** of (X, A).

The sequence of maps g_j is called a **cell decomposition** of (X, A). In the case where X_{j+1} has the homotopy type of $X_j \cup_{g_j} e^{r_j}$, the pair (X, A) is said to have the homotopy type of a CW-pair, and the sequence of maps g_j is called a **homotopy cell decomposition** of (X, A).

In each case, either for a cell decomposition or for homotopy cell decomposition, we denote by ν_j the total number of cells of dimension j, that is, the number of cells $e^{r_0}, \dots, e^{r_{m-1}}$ with $r_j = j$. On the other hand, for a Morse function $f : M \longrightarrow \mathbb{R}$, we define the numbers $\mu_r(f)$ to be the total number of critical points of f of index r, $0 \leq r \leq \dim M$. In general, $\mu_r(f;a,b)$ will denote the total number of critical points of f of index r in $f^{-1}(a,b)$. Then from Theorem 9.2.18 and Corollary 9.3.5, we have the following theorem.

Theorem 9.3.7. *Let M be a manifold, and $f: M \longrightarrow \mathbb{R}$ a Morse function, which is bounded below and satisfies the Condition C, with regular values $a < b$. Then the pair (M_b, M_a) has a homotopy cell decomposition having exactly $\mu_r(f;a,b)$ cells of dimension r, that is, $\nu_r = \mu_r(f;a,b)$.*

Corollary 9.3.8. *Any compact manifold M has the homotopy type of a CW-pair.*

PROOF. For any Morse function $f: M \longrightarrow \mathbb{R}$, there is a homotopy cell decomposition of M with $\nu_r = \mu_r(f)$. □

Let (X, A) be a pair of topological spaces having the homotopy type of a CW-pair. The **qth Betti number** $b_q = b_q(X, A)$ is the dimension of the qth homology group $H_q(X, A)$ with coefficients in $\mathbb{R}$. By the homology theory of CW-pairs, b_q is the number of q-cells in $X - A$. The **homological Euler characteristic** $\chi = \chi(X, A)$ is the alternating sum of the Betti numbers

$$\chi(X, A) = \sum_q (-1)^q b_q(X, A).$$

Define for each q the numbers

$$S_q(X, A) = \sum_{j=0}^{q} (-1)^{q-j}\ b_j(X, A).$$

Then $S_q = b_q - S_{q-1}$, and if $\dim(X, A) = n$, then $S_n = (-1)^n \chi$.

Proposition 9.3.9. *For a sequence of spaces*

$$X_0 \subseteq X_1 \subseteq \cdots \subseteq X_n,$$

where each pair (X_j, X_{j-1}) has the homotopy type of a CW-pair, we have

$$S_q(X_n, X_0) \leq \sum_{j=1}^{n} S_q(X_j, X_{j-1}),$$

$$\chi(X_n . X_0) = \sum_{j=1}^{n} \chi(X_j, X_{j-1}).$$

These relations are expressed by saying that S_q is **subadditive***, and χ is* **additive***.*

PROOF. It is sufficient to show by induction that for a triple of spaces (X, Y, Z), $Z \subseteq Y \subseteq X$, where each pair of spaces arising out of it has the homotopy type of a CW-pair, we have

$$S_q(X, Z) \leq S_q(X, Y) + S_q(Y, Z), \text{ and } \chi(X, Z) = \chi(X, Y) + \chi(Y, Z).$$

The exact homology sequence of the triple of spaces (X, Y, Z)

$$\cdots \xrightarrow{\partial_{k+1}} H_k(Y,Z) \xrightarrow{i_k} H_k(X,Z) \xrightarrow{j_k} H_k(X,Y) \xrightarrow{\partial_k} H_{k-1}(Y,Z) \xrightarrow{i_{k-1}} \cdots$$

breaks into short exact sequences

$$0 \longrightarrow \operatorname{Image}(\partial_{k+1}) \longrightarrow H_k(Y,Z) \longrightarrow \operatorname{Image}(i_k) \longrightarrow 0,$$

$$0 \longrightarrow \operatorname{Image}(i_k) \longrightarrow H_k(X,Z) \longrightarrow \operatorname{Image}(j_k) \longrightarrow 0,$$

$$0 \longrightarrow \operatorname{Image}(j_k) \longrightarrow H_k(X,Y) \longrightarrow \operatorname{Image}(\partial_k) \longrightarrow 0.$$

These respectively imply by the dimension property of vector spaces that

$$b_k(Y,Z) = \dim \operatorname{Image}(\partial_{k+1}) + \dim \operatorname{Image}(i_k),$$

$$b_k(X,Z) = \dim \operatorname{Image}(i_k) + \dim \operatorname{Image}(j_k),$$

$$b_k(X,Y) = \dim \operatorname{Image}(j_k) + \dim \operatorname{Image}(\partial_k).$$

Subtraction of the sum of first and third equation from the second gives

$$b_k(X,Z) - b_k(Y,Z) - b_k(X,Y) = -\dim \operatorname{Image}(\partial_{k+1}) - \dim \operatorname{Image}(\partial_k).$$

Then, multiplication of both sides by $(-1)^{q-k}$, and summation from $k = 0$ to $k = q$ gives (since $\partial_0 = 0$)

$$S_q(X,Z) - S_q(Y,Z) - S_q(X,Y) = -\dim \operatorname{Image}(\partial_{q+1}) \leq 0.$$

This gives the subadditivity of S_q.

Similarly multiplying both sides by $(-1)^k$, summing from $k = 0$ to $k = n$, where $n = \dim(X, A)$,, and using the fact that $\partial_{n+1} = 0$, we get the additivity of χ. $\square$

Theorem 9.3.10. *Let (X, A) have the homotopy type of a CW-pair admitting a homotopy cell decomposition with ν_r cells of dimension r. Then, if $b_q = b_q(X, A)$ is the qth Betti number of (X, A), we have*

$$b_0 \leq \nu_0,$$

$$b_1 - b_0 \leq \nu_1 - \nu_0,$$

$$\cdots\cdots\cdots$$

$$b_q - b_{q-1} + \cdots + (-1)^q b_0 \leq \nu_q - \nu_{q-1} + \cdots + (-1)^q \nu_o.$$

Moreover,

$$\chi(X,A) = \sum_j (-1)^j\ b_j = \sum_j (-1)^j\ \nu_j.$$

Proof. Let $A = X_0 \subseteq X_1 \subseteq \cdots \subseteq X_n = X$ with $X_{j+1} = X_j \cup_{g_j} e^{r_j}$ be the cell decomposition of (X, A). Since $b_k(X_{j+1}, X_j) = \delta_{k,r_j}$ (Kronecker delta), by Lemma 9.3.3, it follows that $\sum_{j=0}^{n-1} b_k(X_{j+1}, X_j) = \nu_k$. Therefore

$$\sum_{j=0}^{n-1} S_q(X_{j+1}, X_j) = \sum_{j=0}^{n-1} \sum_{k=0}^{q} (-1)^{q-k} b_k(X_{j+1}, X_j) = \sum_{k=0}^{q} (-1)^{q-k}\ \nu_k,$$

and

$$\sum_{j=0}^{n-1} \chi(X_{j+1}, X_j) = \sum_{j=0}^{n-1} \sum_{k=0}^{q} (-1)^k b_k(X_{j+1}, X_j) = \sum_{k=0}^{q} (-1)^k \nu_k.$$

The proof is now immediate from the additivity of χ and the subadditivity of S_q. □

Corollary 9.3.11. *Let $a < b$ be regular values of a Morse function $f : M \longrightarrow \mathbb{R}$ satisfying the Condition C. Let $\mu_r = \mu_r(f, a, b)$ be the number of critical points of index r of f in $f^{-1}(a, b)$, and $b_r = b_r(M_b, M_a)$ be the rth Betti number of (M_b, M_a). Then we have the following* **Morse inequalities***:*

$$b_0 \le \mu_0,$$

$$b_1 - b_0 \le \mu_1 - \mu_0,$$

$$\cdots\cdots\cdots$$

$$b_q - b_{q-1} + \cdots + (-1)^q b_0 \le \mu_q - \mu_{q-1} + \cdots + (-1)^q \mu_o.$$

Moreover,

$$\chi(M) = \sum_j (-1)^j b_j = \sum_j (-1)^j \mu_j.$$

Finally, we have the **weak Morse inequalities***:*

$$b_q \le \mu_q.$$

The number $\chi(M)$ is called the **homological** *(or* **Euler-Poincaré characteristic***) of M.*

Proof. The results are immediate from Theorem 9.3.10 and Theorem 9.3.7. The weak Morse inequalities follow by adding two consecutive Morse inequalities. □

The homological Euler characteristic of a manifold M is the alternating sum of the Betti numbers of M. Also in Chapter 6 we have defined the Euler characteristic of a compact oriented manifold as the sum of indices of zeros of a vector field (Definition 6.5.11, Lemma 6.5.12). The following classical result of Hopf says that the two notions of Euler characteristics are essentially the same.

Theorem 9.3.12 (Hopf). *The homological Euler characteristic of a compact orientable manifold without boundary is equal to its Euler characteristic.*

Proof. The Euler characteristic is given by any vector field X with finite number of zeros. We take $X = \frac{1}{2}\nabla f$, where $f : M \longrightarrow \mathbb{R}$ is a Morse function. If $(x_1, \ldots, x_n)$ is the coordinate system in a Morse chart about a critical point $p = (0, \ldots, 0)$ of index k, then

$$f(x) = -x_1^2 - \cdots - x_k^2 + x_{k+1}^2 + \cdots + x_n^2,$$

and so

$$X = (-x_1, \ldots, -x_k, x_{k+1}, \ldots, x_n).$$

Let Y (resp. Z) be the vector field on $\mathbb{R}^k$ (resp. $\mathbb{R}^{n-k}$) defined by $Y(y) = -y$ (resp. $Z(z) = z$). Then $X = Y \times Z$, and we have by simple computations of Jacobian determinants that

$$\mathrm{Ind_0X} = \mathrm{Ind_0Y} \cdot \mathrm{Ind_0Z},$$

and

$$\mathrm{Ind_0Y} = (-1)^{\mathrm{k}}, \quad \mathrm{Ind_0Z} = 1$$

so that $\mathrm{Ind_0X} = (-1)^{\mathrm{k}}$. Therefore the sum of the indices of zeros of X is $\sum_k (-1)^k \mu_k$, where μ_k is the number of critical points of f of index k. By Corollary 9.3.11, this is the homological Euler characteristic of M. □

9.4. Perfect Morse functions

Let $f : M \longrightarrow \mathbb{R}$ be a Morse function on a compact manifold. Then the **Morse polynomial** $\mathcal{M}_t(f)$ of f is defined by

$$\mathcal{M}_t(f) = \sum_p t^{\lambda_p},$$

where the sum is over all non-degenerate critical points p of f, and λ_p is the index of the critical point p. We may also write

$$\mathcal{M}_t(f) = \sum_k \mu_k t^k,$$

where μ_k is the number of critical points of f of index k. On the other hand, the **Poincaré polynomial** $\mathcal{P}_t(M)$ of the manifold M is

$$\mathcal{P}_t(M) = \sum_{k=0}^{n} \dim H_k(M; \mathbb{R})\ t^k = \sum_{k=0}^{n} b_k\ t^k,$$

where b_k is the k-th Betti number, and $n = \dim\ M$. Note that

$$\mathcal{P}_{-1}(M) = b_0 - b_1 + \cdots + (-1)^n b_n$$

is the Euler-Poincaré characteristic of M.

In view of the Morse inequalities $b_k \leq \mu_k$, we have $\mathcal{P}_t(M) \leq \mathcal{M}_t(f)$. Because $\mathcal{M}_{-1}(f) = \mathcal{P}_{-1}(M)$ (Corollary 9.3.11), $1 + t$ is a factor of $\mathcal{M}_t(f) - \mathcal{P}_t(M)$. Therefore we may write

$$\mathcal{M}_t(f) - \mathcal{P}_t(M) = (1 + t) \cdot R(t),$$

where $R(t)$ is the polynomial $\sum_{k=0}^{n-1} a_k\, t^k$, with $a_0 = \mu_0 - b_0$, and $a_k = \sum_{i=0}^{k} (-1)^{k-i}(\mu_i - b_i)$ for $1 \leq k \leq n - 1$.

Definition 9.4.1. A Morse function $f : M \longrightarrow \mathbb{R}$ on a compact manifold is called a **perfect Morse function** if all the Morse inequalities are equalities, or equivalently if $\mathcal{M}_t(f) = \mathcal{P}_t(M)$, or $R(t) = 0$.

Example 9.4.2. Consider the basic example of the height function on the Riemann surface Σ_g (this includes the torus when $g = 1$), as given in Exercise 9.1 in p.285. We have seen before $\mu_0 = 1,\ \mu_1 = 2g,\ $ and $\mu_2 = 1$. Again, since Σ_g is connected, we have $b_0 = 1$, and, since Σ_g is oriented, we have $b_2 = 1$ (these are fundamental facts of algebraic topology – see Spanier [**43**]). Then $\mathcal{M}_{-1}(f) = \mathcal{P}_{-1}(\Sigma_g)$ implies $\mu_1 = b_1$, and so the height function is perfect.

The arguments show that if f is any Morse function on Σ_g with a unique minimum and a unique maximum, then f is perfect.

Theorem 9.4.3 (Morse lacunary principle). *If the polynomial $\mathcal{M}_t(f)$ has no consecutive powers of t, then the Morse function f is perfect.*

Proof. Since $\mu_k = 0$ implies $b_k = 0$, it follows that if k is the first non-zero power in $\mathcal{M}_t(f)$, then it is also the first non-zero power in $R(t)$. Thus a_k occurs in the product

$$(1+t)\cdot R(t) = \sum_{j=0}^{n-1} a_j(t^j + t^{j+1}),$$

and so, if $R(t) \neq 0$, then t^{k+1} also occurs in the product $(1+t)\cdot R(t)$. However, this is not possible, because by hypothesis t^{k+1} does not occur in $\mathcal{M}_t(f)$, and hence in $\mathcal{P}_t(M)$ either, otherwise this would violate the inequality $\mu_{k+1} \geq b_{k+1}$. Hence $R(t)$ must be zero, and so f is perfect. □

Corollary 9.4.4. *If $f : M \longrightarrow \mathbb{R}$ is a Morse function on a compact manifold such that all odd numbers μ_{2k+1} are zero, then all odd Betti numbers are also zero, and for even Betti numbers we have $b_{2k} = \mu_{2k}$.*

Example 9.4.5 (Homology of complex projective space). Recall from Exercise 1.5 in p.9 that the complex projective space $\mathbb{C}P^n$ is the quotient space of $\mathbb{C}^{n+1} - \{0\}$ by the action of the group $\mathbb{C}^* = \mathbb{C} - \{0\}$. In other terms, it is the quotient of $\mathbb{C}^{n+1} - \{0\}$ under the equivalence relation $z \sim \lambda z$, $z \in \mathbb{C}^{n+1} - \{0\}$, $\lambda \in \mathbb{C}^*$. Write $z = (z_0, z_1, \cdots, z_n)$, and its equivalence class as $[z]$. Then the open set

$$U_k = \{[z] \mid z_k \neq 0\}$$

for $k = 0, 1, \cdots, n$, cover $\mathbb{C}P^n$, and the coordinate charts $\phi_k : U_k \longrightarrow \mathbb{R}^{2n}$ given by $\phi_k([z]) = (x_0, x_1, \cdots, x_{n-1}, y_0, y_1, \cdots, y_{n-1})$, where

$$\begin{aligned} \frac{z_j}{z_k} &= x_j + \sqrt{-1}\, y_j,\ j = 0, \ldots, k-1, \\ \frac{z_{j+1}}{z_k} &= x_j + \sqrt{-1}\, y_j,\ j = k, \ldots, n-1 \end{aligned}$$

constitute a smooth atlas of $\mathbb{C}P^n$. Note that the origin $(0, \ldots, 0)$ in $\phi_k(U_k)$ corresponds to the point $[e_k]$ in U_k, where $e_0, e_1, \ldots, e_n$ are the elements of the standard basis of $\mathbb{C}^{n+1}$.

For $n+1$ arbitrarily fixed real numbers $a_0 < a_1 < \cdots < a_n$, define a function $f : \mathbb{C}P^n \longrightarrow \mathbb{R}$ by

$$f([z]) \;=\; \frac{a_0|z_0|^2 + a_1|z_1|^2 + \cdots + a_n|z_n|^2}{|z_0|^2 + |z_1|^2 + \cdots + |z_n|^2}.$$

The local representation of f in terms of the coordinates

$$(x_0, \ldots, x_{n-1}, y_0, \ldots, y_{n-1})$$

in U_k is given by $f \circ \phi_k^{-1}(x_0, \ldots, x_{n-1}, y_0, \ldots, y_{n-1}) =$

$$\frac{\sum_{j=0}^{k-1} a_j(x_j^2 + y_j^2) + a_k + \sum_{j=k}^{n-1} a_{j+1}(x_j^2 + y_j^2)}{\sum_{j=0}^{n-1}(x_j^2 + y_j^2) + 1}.$$

A simple computation shows that the origin $(0, \cdots, 0)$ is the only critical point of this function, and its Hessian at the origin is

$$\begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix},$$

where A is the diagonal matrix

$$\mathrm{diag(2(a_0 - a_k), \ldots, 2(a_{k-1} - a_k), 2(a_{k+1} - a_k), \ldots, 2(a_n - a_k))}.$$

Since $a_0 < a_1 < \cdots < a_n$, the matrix A has k negative eigenvalues. Thus the function f has $n+1$ non-degenerate critical points at the points $[e_0], [e_1], \ldots, [e_n]$ with critical values $a_0, a_1, \ldots, a_n$ respectively, and the k-th critical point has index $2k$. Therefore

$$\mathcal{M}_t(f) = 1 + t^2 + t^4 + \cdots + t^{2n},$$

and the lacunary principle implies that this expression is also equal to $\mathcal{P}_t(\mathbb{C}P^n)$.

Thus all the odd dimensional homology group of $\mathbb{C}P^n$ are zero, and any even dimensional homology group is isomorphic to $\mathbb{R}$.

9.5. Triangulations of manifolds

In this section we shall prove using Morse theory that every compact manifold has a simplicial structure. Much stronger results, which are due to Whitehead [**55**], deal with the existence and uniqueness of a smooth triangulation of any manifold, and require quite different proofs. We shall not go into the proofs of these results, because we do not need to use them. We shall consider only the following topological result.

Theorem 9.5.1. *A compact manifold is a polyhedron.*

Suppose that M is a compact n-manifold, and it is embedded in some bigger Euclidean space $\mathbb{R}^m$. By Morse theory, M may be built up from a disk D^n by attaching successively a finite number of handles of dimension n and of various indices. Since each handle is a polyhedron in $\mathbb{R}^m$, the proof of the theorem will follow trivially after we prove that two finite polyhedra in an Euclidean space may be subdivided into simplicial complexes in such a way that their intersection is a subcomplex of each of them. In other words, it is required to prove the following theorem.

Theorem 9.5.2. *If K_1 and K_2 are finite simplicial complexes in $\mathbb{R}^m$, then there exist simplicial subdivisions K_1' and K_2' of K_1 and K_2 respectively such that $K_1' \cup K_2'$ is a simplicial complex.*

The proof of the theorem is a consequence of some elementary properties of linear cell complexes.

Recall that points $x_0, x_1, \ldots, x_k$ of $\mathbb{R}^m$ are independent (in the affine sense) if the points $x_1 - x_0, x_2 - x_0, \ldots, x_k - x_0$ are linearly independent. This is equivalent to saying that for real numbers $\lambda_0, \lambda_1, \ldots, \lambda_k$

$$\sum_{i=0}^{k} \lambda_i x_i = 0 \text{ and } \sum_{i=0}^{k} \lambda_i = 0 \text{ imply each } \lambda_i = 0.$$

A convex combination of the points $x_0, x_1, \ldots, x_k$ is an expression of the form $\sum_{i=0}^{k} \lambda_i x_i$, where λ_i are real numbers ≥ 0 with $\sum_{i=0}^{k} \lambda_i = 1$. If S is a subset of $\mathbb{R}^m$, then the convex hull (or span) of S is the set of all finite convex combinations of elements of S.

If $x_0, x_1, \cdots, x_k$ are independent points in $\mathbb{R}^m$, then the simplex A spanned by them is the convex hull of the set

$$\{x_0, x_1, \ldots, x_k\}.$$

The dimension of A is k, which is the dimension of the smallest plane containing A, and A is called a k-simplex. If x_0 is 0 (the origin), and x_i is the i-th standard unit vector of $\mathbb{R}^m$, $i = 1, \ldots, k$, then A becomes the standard k-simplex E^k of $\mathbb{R}^m$. The simplex spanned by a subset of $r+1$ points of $\{x_0, x_1, \ldots, x_k\}$, $r = 0, \ldots, k-1$, is a face of dimension r, or an r-face of A. A 0-face is a vertex, and the k-face is the simplex A itself. A simplex may also be defined by using "join of sets". If A and B are two subsets of $\mathbb{R}^m$, then the join $A * B$ of A and B is the union of all closed line segments from a point of A to a point of B. Then the simplex A spanned by an independent set of points $x_0, x_1, \ldots, x_k$ of $\mathbb{R}^m$ is given by $A = x_0 * x_1 * \cdots * x_k$.

A linear cell C in $\mathbb{R}^m$ is a compact subset of $\mathbb{R}^m$ consisting of points $x = (x_1, \ldots, x_m)$ which satisfy a finite number of linear equations $f_i(x) = a_i$, and linear inequalities $g_i(x) \geq b_i$, where f_i and g_i are functions on $\mathbb{R}^m$ of the form

$$(x_1, \ldots, x_m) \mapsto \lambda_1 x_1 + \cdots + \lambda_m x_m, \ \lambda_i \in \mathbb{R}.$$

We say that C is an k-cell (or a cell of dimension k) if C contains $(k+1)$ independent points, but no more; in other words, k is the smallest of the dimensions of the planes which contain C. A face of C is a cell obtained by replacing some of the inequalities $g_i(x) \geq b_i$ by equalities $g_i(x) = b_i$. A vertex of C is a face consisting of only one point.

For example, The standard k-simplex E^k is given by the following set of inequalities

$$x_1 \geq 0, \dots, x_k \geq 0, \text{ and } x_1 + x_2 + \cdots + x_k \leq 1.$$

Replacing one of the inequalities by equality, we may get a $(k-1)$-face of E^k.

The following facts may be proved easily.

- A cell is a convex set[2]. Moreover, it is the convex hull of its vertices.
- The intersection, and the Cartesian product of cells is a cell (identifying $\mathbb{R}^p \times \mathbb{R}^q \equiv \mathbb{R}^{p+q}$).
- If $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is an affine linear map[3], and A is a cell of $\mathbb{R}^n$, then $f(A)$ is a cell of $\mathbb{R}^m$.
- Any simplex is a cell.

The faces of C are determined by C itself, and not by any particular choice of equations and inequalities defining C. Indeed, one can show that if C and D are cells, and P is the plane spanned by D, then D is a face of C if and only if

(1) $P \cap C = D$, and
(2) no point of P lies between two points of the set $C - D$.

The following lemma describes a special case of this result.

Lemma 9.5.3. *If C is an k-cell, then its topological boundary is the union of a finite number of $(k-1)$-cells, which are faces uniquely determined by C.*

PROOF. Let P be the smallest plane of dimension k containing C, and C be given by a minimal set of inequalities $g_i(x) \geq b_i$, $i = 1, \dots, r$, besides some equalities for points of P. Let N be the set of points of P for which all these inequalities are strict. We shall show that $N = \text{Int } C$. Clearly, $N \subseteq \text{Int } C$. For, if x is a point of P for which $g_j(x) > b_j$, then, by continuity, there is an open neighbourhood U of x in P such that for any $x' \in U$ we have $g_j(x') > b_j$. The reverse inclusion is true, because if x is a point of P for which $g_j(x) = b_j$, then x cannot be a point of Int C. This follows because in this case P does not lie wholly in the region $g_j(x) \geq b_j$, otherwise any point of C would satisfy this inequality automatically, and so the inequality could be discarded without disturbing the set C. This means that any open neighbourhood of x contains points y for which $g_j(y) < b_j$ (by continuity again), and so x cannot be an interior point of C. Thus $N = \text{Int } C$.

Let, for $i = 1, \dots r$, C_i be the set of points x of C for which $g_i(x) = b_i$. Then C_i is a cell, and the boundary of C is contained in the union $C_1 \cup \cdots \cup C_r$. Moreover, the intersection of the plane $g_i(x) = b_i$ with P is a $(k-1)$-plane P_i so that $C \cap P_i = C_i$.

[2]A convex set is a set which contains the line segment joining two of its points.

[3]An affine linear map is a linear map composed with a translation.

We shall show that the dimension of C_i is $k-1$. Let Q be the subset of P consisting of points x for which $g_j(x) > b_j$ for all $j \neq i$. Then Q is a convex set containing Int C. In particular, it contains a point x for which $g_i(x) > b_i$. It also contains a point y for which $g_i(y) < b_i$, otherwise $\overline{Q}$ would lie entirely in the region $g_i(x) \geq b_i$, and so the inequality $g_i(x) \geq b_i$ might be discarded from the set of inequalities for C without changing the set C. Therefore, by convexity of Q, there is a point $z \in Q$ such that $g_i(z) = b_i$. This means that $P_i \cap Q \neq \emptyset$. Since Q is open in P, $P_i \cap Q$ is open in P_i. Since $P_i \cap Q \subset C_i$, the cell C_i must have dimension $k-1$.

For the uniqueness of the cells C_i, the following arguments may be employed. Let B denote the boundary of C, and L be any plane of dimension $k-1$ which is different from $P_1, \ldots, P_r$. Since B is contained in the union of the planes P_i, each of which has dimension $k-1$, the intersection $B \cap L$ must have dimension less than $k-1$. This completes the proof. □

Definition 9.5.4. A linear cell complex K in $\mathbb{R}^m$ is a finite set of cells in $\mathbb{R}^m$ such that

(1) if $C \in K$, then every face of C is in K,

(2) if C and D are in K, then either $C \cap D = \emptyset$ or $C \cap D$ is a common face of C and D.

If K is a cell complex, then we denote by $|K|$ the union of all cells in K; $|K|$ is called the underlying polyhedron of K.

A simplicial complex is a cell complex all of whose cells are simplexes

Definition 9.5.5. If K and L are cell complexes, then L is called a subdivision of K, if

(1) $|L| = |K|$,

(2) every cell of L is a subset of some cell of K, each cell of K is a union of cells of L.

A subdivision L of K is called simplicial if L is a simplicial complex.

Lemma 9.5.6. *If K_1 and K_2 are cell complexes such that $|K_1| = |K_2|$, then K_1 and K_2 have a common subdivision.*

Proof. Let L be the collection of all intersections $C_1 \cap C_2$, where C_1 is a cell of K_1 and C_2 is a cell of K_2. Then L is a cell complex, since the intersection of two cells is again a cell, and any face of a cell $C_1 \cap C_2$ is of the form $D_1 \cap D_2$, where D_1 is a face of C_1, and D_2 is a face of C_2. Clearly, we have $|L| = |K_1| = |K_2|$, and so L is the required common subdivision. □

Lemma 9.5.7. *Any cell complex K has a simplicial subdivision.*

Proof. The proof is by induction on the dimension[4] n of K. If $n = 0, or 1$, then K is already a simplicial complex. Suppose that the lemma has been proved for cell complexes of dimension $n-1$. Let L be a simplicial subdivision of the $(n-1)$-skeleton [5] K^{n-1} of K. Let C be an n-cell of K, and $A_1, \ldots, A_k$ be the simplexes of L lying in the boundary of C. Let x be at a point lying in the Int C. Adjoin to the collection of simplexes of L the simplexes $A_1 * x, \cdots A_k * x$, where $A_i * x$ is the join of A_i and x. Performing this construction for each n-cell of K, we obtain a simplicial complex K' which is a subdivision of K.

We can make K' canonically defined by taking x to be the centroid c of C. $\square$

Corollary 9.5.8. *If K_1 and K_2 are simplicial complexes such that $|K_1| = |K_2|$, then $|K_1|$ and $|K_2|$ have a common simplicial subdivision.*

Definition 9.5.9. A rectilinear triangulation of $\mathbb{R}^m$ is a simplicial complex L such that $|L| = \mathbb{R}^m$.

Lemma 9.5.10. *If K is a finite simplicial complex, then there is a subdivision K' of K which is a subcomplex of a rectilinear triangulation L of $\mathbb{R}^m$.*

Proof. Let $E_1, \ldots, E_k$ be simplexes of K. For each simplex E_i, $i = 1, \ldots, k$, we choose a rectilinear triangulation L_i of $\mathbb{R}^m$ which contains E_i in the following way. First take any rectilinear triangulation T_i of $\mathbb{R}^m$, and then find a non-singular affine linear map $f_i : \mathbb{R}^m \longrightarrow \mathbb{R}^m$ such that f_i maps one of the simplexes of T_i onto E_i. Then $f_i(T_i)$ is the required rectilinear triangulation L_i of $\mathbb{R}^m$.

Let L be a common simplicial subdivision of $L_1, \ldots, L_k$ (Corollary 9.5.8). Then $|K|$ is the polyhedron of a subcomplex K' of L, and this subcomplex K' is a subdivision of K. $\square$

Proof of Theorem 9.5.2. Using the above lemma, we find, for $i = 1, 2$, a rectilinear triangulation L_i of $\mathbb{R}^m$ containing some subdivision K_i' of K_i as a subcomplex. Then if L is a common subdivision of L_1 and L_2, K_1' and K_2' will be subcomplexes of L such that $K_1' \cap K_2'$ is also a subcomplex of L. This means that $K_1' \cup K_2'$ is a simplicial complex. $\square$

[4]The dimension of a cell complex K is the maximum of the dimensions of the cells in K.

[5]The p-skeleton K^p of a cell complex K is the set of all cells of K having dimension $\leq p$. It is a subcomplex of K.

CHAPTER 10

THEORY OF HANDLE PRESENTATIONS

All manifolds that will appear in this chapter will be compact unless stated otherwise. Our first result is the handle presentation theorem which says any cobordism M admits a Morse function f, and a decomposition, which is obtained by successively attaching handles corresponding to the critical points of f. In fact, there are two such decompositions, one using f and the other using $-f$. These decompositions give the Poincaré duality theorem. The arguments are essentially the same as those used in algebraic topology using dual cell complexes. We next turn to the fundamental problem of simplifying the decomposition as far as possible, subject to some constraints imposed by the homology groups. The main result of Smale asserts that under some conditions one can modify f so as to obtain a Morse function with fewer critical points. A special case of Smale's result is when M has the homology (or homotopy) of the n-sphere S^n. In this case one can obtained a Morse function with two critical points from which it follows immediately that M is homeomorphic to S^n, $n \geq 6$ (the result is true for $n = 5$ also, but its proof is little more complicated). This is the generalised Poincaré's conjecture. A variant for manifolds with boundary gives the h-cobordism theorem. However, in our exposition, which owes much to the writings of J. Milnor, and also to C.T.C. Wall, we first prove the h-cobordism theorem and then obtain generalised Poincaré's conjecture as an application. We also discuss in this chapter as an interlude the classification of closed surfaces.

The original conjecture was formulated by a French mathematician Henri Poincaré in 1904 at the end of a sixty-five-page research paper. The version stated by Poincaré is equivalent to the following: any closed simply connected 3-manifold is diffeomorphic to the standard 3-dimensional sphere S^3. Note that a manifold is closed if it is compact and without boundary, and that a manifold is simply connected if it is path-wise connected and has trivial fundamental group. Generalised Poincaré conjecture says that if an n-manifold is a homotopy n-sphere, then it is homeomorphic to the n-sphere. The smooth version of this says that under the above conditions M is diffeomorphic to S^n. In 1961 Stephen Smale proved the generalised Poincaré conjecture in dimensions ≥ 5. In 1982 Michael Freedman proved this in dimension 4. In 2002-2003 Grigori Perelman proved the conjecture in dimension 3 in three papers using Richard Hamilton's Ricci flow with surgery. The smooth conjecture is true in dimensions 1, 2, 3, 5 and 6. The case of dimension 4 remains unsolved as of now. Milnor's exotic

A. Mukherjee, *Differential Topology*, DOI 10.1007/978-3-319-19045-7_10

spheres show that the smooth conjecture is false in dimension 7. We shall indicate the proof of the smooth conjecture in dimensions 5 and 6 at the end of this chapter.

10.1. Existence of handle presentation

Throughout I will denote the unit interval $[0,1]$.

Definition 10.1.1. If M is an n-manifold and V_0 is a component of ∂M, then a **handle presentation** of M on V_0 is a sequence of n-manifolds

$$M_0 \subset M_1 \subset \cdots \subset M_m = M,$$

where $M_0 = V_0 \times I$, and for each $j = 1, \ldots, m$, M_j is obtained from M_{j-1} by attaching an r_j- handle $\mathcal{H}_{r_j} = D^{r_j} \times D^{n-r_j}$ by means of an embedding

$$f_j : S^{r_j-1} \times D^{n-r_j} \longrightarrow \partial_+ M_{j-1},$$

where $\partial_+ M_0 = V_0 \times 1$, and for $1 < j \leq m$

$$\partial_+ M_{j-1} = [\partial M_{j-1} - \text{Int}\,(\text{Image}\, f_{j-1})] \cup [D^{r_j} \times S^{n-r_j-1}].$$

The images of the embeddings f_j are required to have disjoint interiors.

Thus $M = (V_0 \times I) \cup_{f_1} \mathcal{H}_{r_1} \cup_{f_2} \cdots \cup_{f_m} \mathcal{H}_{r_m}$.

A handle presentation of M on V_0 is denoted by

$$\chi(M; V_0, f_1, \ldots, f_m).$$

A handle presentation on a disk D^{n-1} is called a **handlebody**.

Definition 10.1.2. Two handle presentations of M on V_0

$$\chi(M; V_0, f_1, \ldots, f_m) \text{ and } \chi'(M; V_0, f'_1, \ldots, f'_m),$$

where f_j and f'_j are embeddings $S^{r_j-1} \times D^{n-r_j} \longrightarrow \partial_+ M_{j-1}$, are **equivalent** if any one of the following conditions is satisfied:

(1) The f'_j are the permutations of the f_j.

(2) There exist diffeomorphisms $h_j : M_{j-1} \longrightarrow M_{j-1}$ such that $f'_j = h_j \circ f_j$.

(3) There exist diffeomorphisms $h_j : D^{r_j} \times D^{n-r_j} \longrightarrow D^{r_j} \times D^{n-r_j}$, which maps $S^{r_j-1} \times D^{n-r_j}$ onto itself, such that $f'_j = f_j \circ h_j$.

The definition is justified by Corollary 7.6.7. In general, if each f_j is isotopic to f'_j, then the presentations are equivalent by Corollary 7.6.8.

The handle presentation theorem says that any manifold can be obtained by attaching handles successively. The theorem of Smale and Wallace describes this fact in terms of cobordism which we want to consider now.

Definition 10.1.3. A **cobordism** is a manifold M whose boundary ∂M is the disjoint union of two open and closed submanifolds V_0 and V_1 called respectively **left-hand** and **right-hand boundary of** M. Sometimes the triple $(M; V_0, V_1)$ is called a cobordism from V_0 to V_1, and we write $V_0 = \partial_- M$, $V_1 = \partial_+ M$.

The simplest example is provided by a triple $(V \times I, V \times 0, V \times 1)$, where V is without boundary. This, or any manifold diffeomorphic to this, is called a **trivial cobordism**. Other cobordisms are obtained from this by adding handles. For example, the manifold $M = V \times I \cup_f \mathcal{H}_r$, where $f : S^{r-1} \times D^{n-r} \longrightarrow V \times 1$, $n - 1 = \dim V$, is a cobordism with

$$\partial_- M = V \times 0, \text{ and } \partial_+ M = (V \times 1 - \text{Int }(\text{Image} f)) \cup (D^r \times S^{n-r-1}).$$

This is called a **cobordism of index** r. A cobordism of index 0 is a disjoint union $V \times I \cup D^n$ whose left-hand boundary is $V \times 0$, and right-hand boundary is the disjoint union $V \times 1 \cup S^{n-1}$. A trivial cobordism may be called a cobordism of index -1.

Two manifolds without boundary V_0 and V_1 are called **cobordant** if there is a manifold M such that ∂M is the disjoint union of V_0 and V_1.

Lemma 10.1.4. *If (M, V_0, V_1) is a cobordism, and $f : M \longrightarrow [0, 1]$ is a Morse function with $f^{-1}(0) = V_0$ and $f^{-1}(1) = V_1$, and c is a point in $(0, 1)$ which is not a critical value of f, then $f^{-1}([0, c])$ and $f^{-1}([c, 1])$ are manifolds with boundary.*

Proof. This follows from Lemma 6.2.7. The point is that there is a coordinate system $(x_1, \ldots, x_n)$ about $p \in f^{-1}(c)$ in which f looks like the projection $(x_1, \ldots, x_n) \mapsto x_n$. □

Lemma 10.1.5. *If $(M; V_0, V_1)$ is a cobordism, then there is a Morse function $f : M \longrightarrow \mathbb{R}$ having distinct critical values such that f has no critical points in a neighbourhood of ∂M, and $f^{-1}(0) = V_0$ and $f^{-1}(1) = V_1$.*

Proof. Take collar neighbourhoods $V_0 \times [0, 1]$ and $V_1 \times [0, 1]$ of V_0 and V_1 in M, which are disjoint. Let $K = M - (V_0 \times [0, 1) \cup V_1 \times [0, 1))$. Take a continuous function $g : K \longrightarrow [0, 1]$ which takes the value 0 on $V_0 \times 1$, and the value 1 on $V_1 \times 1$ (this is possible, by Urysohn's lemma, since K is normal). Define a continuous function $h : M \longrightarrow [-1, 2]$ by

$$\begin{aligned} h(x) &= t - 1 \quad \text{if } x = (u, t),\ u \in V_0,\ 0 \le t \le 1, \\ &= 2 - t \quad \text{if } x = (u, t),\ u \in V_1,\ 0 \le t \le 1, \\ &= g(x) \quad \text{if } x \in K. \end{aligned}$$

Approximate h by a smooth function h_1 agreeing with h on a neighbourhood of ∂M (Theorem 2.2.3). Next, approximate h_1 by a Morse function f with distinct critical values agreeing with h_1 (and therefore with h) near ∂M (Corollary 9.1.5). Note that h and h_1 have no critical point in a neighbourhood of ∂M. This completes the proof. □

Definition 10.1.6. The **Morse number** of a cobordism (M, V_0, V_1) is the minimum among all Morse functions $f : M \longrightarrow [0,1]$ of the number of critical points of f

A cobordism of Morse number one is called an **elementary cobordism**, its index is the index of the critical point.

Proposition 10.1.7. *If an interval $[a,b]$ contains no critical value of a Morse function $f : M \longrightarrow \mathbb{R}$, and V is the level surface $f^{-1}(b)$, then there is a diffeomorphism of $W = f^{-1}([a,b])$ onto $V \times [a,b]$ under which the restriction $f|W$ corresponds to the projection $V \times [a,b] \longrightarrow [a,b]$.*

In particular, a cobordism of Morse number zero is a product cobordism.

Proof. The situation is the same as that of Theorem 9.2.10. We have constructed in the proof of that theorem a vector field Y on M, which is orthogonal to the level surface of f through any point of $[a,b]$, such that if $\phi_t : M \longrightarrow M$ is the flow generated by Y, then

$$f(\phi_t(x)) = f(x) - t, \quad x \in W.$$

In the following Figure 10.1, the horizontal curves denote the level surfaces of f, and the vertical curves represent the flowlines of the vector field Y.

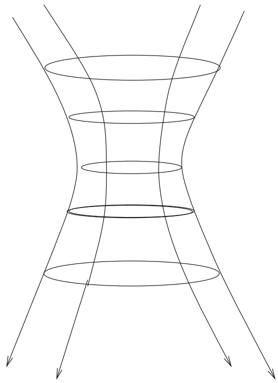


Figure 10.1

Define a map $h : V \times [a,b] \longrightarrow W$ by $h(x,t) = \phi_{b-t}(x)$. Then

$$f(h(x,t)) = f(\phi_{b-t}(x)) = f(x) - b + t = t,$$

as $f(x) = b$, and so $h_t = \phi_{b-t}$ maps V diffeomorphically onto the level surface $V' = f^{-1}(t)$. Also, if $v \in \tau(W)_x$, then

$$dh(v, d/dt) = dh_t(v) - d(\phi_t(x))(d/dt) = dh_t(v) - Y.$$

Since

$$\tau(W)_{\phi_t(x)} = \tau(V')_{\phi_t(x)} \oplus < Y_{\phi_t(x)} >,$$

it follows from the inverse function theorem that h is a diffeomorphism. $\square$

Given two corbordisms $C = (M, V_0, V_1)$, $C' = (M', V_0', V_1')$, and a diffeomorphism $h : V_1 \longrightarrow V_0'$, we may join V_1 and V_0' by Theorem 7.6.2, and get a cobordism $C \cup C' = (M \cup_h M', V_0, V_1')$, which is unique up to diffeomorphisms of M and M', but depends on the isotopy class of h. The cobordism $C \cup C'$ is called the **composition** of C and C'. We say that $C = C'$ if there is a diffeomorphism $h : M \longrightarrow M'$ carrying V_0 to V_0' and V_1 to V_1'. A cobordism (M, V_0, V_1) can be expressed as the composition of two cobordisms, one from V_0 to $f^{-1}(c)$ and the other from $f^{-1}(c)$ to V_1, where f is a Morse function provided by Lemma 10.1.5, and c is a regular value of f.

$\Diamond$ **Exercise 10.1.** Show that if C is a cobordism (M, V_0, V_1), and C' is a trivial cobordism $V \times I$, where V_1 is diffeomorphic to V, then $C \cup C' = C$.

Hint. Identify C' with a collar neighbourhood of V_1 in M.

Theorem 10.1.8 (Smale and Wallace). *Any nontrivial cobordism C can be expressed as a composition*

$$C = C_1 \cup C_2 \cup \cdots \cup C_k,$$

where C_i is an elementary cobordism of index r_i. Moreover, one may assume that $r_i \leq r_j$ for $i < j$.

PROOF. Take a Morse function f provided by Lemma 10.1.5 with distinct critical values $c_1, \ldots, c_m$ in $(0, 1)$; their number is finite, since M is compact. Take $(m + 1)$ numbers $a_0, a_1, \ldots, a_m$ such that $a_0 = 0$, $a_m = 1$, and $c_i < a_i < c_{i+1}$ for $i = 1, \ldots, m - 1$. Then

$$M = \cup_{i=0}^{m-1} f^{-1}[a_i, a_{i+1}].$$

By Theorem 9.2.18, each cobordism $f^{-1}[a_i, a_{i+1}]$ is a composition of elementary cobordisms.

The second part will follow from Proposition 10.4.1 below which says that the handles may be arranged in increasing order of indices. $\square$

Corollary 10.1.9. *If $(M, \partial_- M, \partial_+ M)$ is a cobordism, then there is a handle presentation of M on $\partial_- M$*

$$M_{-1} \subset M_0 \subset M_1 \subset \cdots M_n = M,$$

where $M_{-1} = \partial_- M \times I$, and, for each $r = 0, \ldots, n$, M_r is obtained from M_{r-1} by attaching a number of r-handles to the right-hand boundary of M_{r-1} so that

$$M_r = M_{r-1} \cup (\partial_+ M_{r-1} \times I \cup \mathcal{H}_r).$$

PROOF. Denote the cobordism by C, and consider its expression as a composition of elementary cobordisms C_i, given by the theorem. Let M_{r-1} be the union of all cobordisms of indices $\leq r - 1$ in this expression. Then M_r is obtained from M_{r-1} by attaching a number of r-handles successively. Since the index of the handles are the same, they can be attached simultaneously to the right hand boundary $\partial_+ M_{r-1}$. $\square$

Proposition 10.1.10. *If M has a handle presentation on $\partial_- M$ with α_r r-handles for $0 \leq r \leq n$, where $n = \dim M$, then it has also a handle presentation on $\partial_+ M$ with α_r $(n-r)$-handles.*

PROOF. If f is a Morse function on M, then $-f$ is also a Morse function on M with the same critical points as those of f. However, if $p \in M$ is a critical point of f of index r, then p is a critical point of $-f$ of index $n-r$. This follows by multiplying the local representation of f

$$f(x) = f(p) - x_1^2 - \cdots - x_r^2 + x_{r+1}^2 + \cdots + x_n^2$$

by -1. The proof of the proposition now follows from the correspondence between Morse functions and handle presentations. $\square$

Of the two handle presentations of M corresponding to Morse functions f and $-f$, one is called **dual presentation** of the other. One is obtained by turning the other "upside down" so that an r-handle of one becomes an $(n-r)$-handle of the other. We shall consider these presentations in Theorem 10.2.8 below.

We shall end this section by describing briefly the effect of attaching a handle to a manifold M on its boundary ∂M.

Recall from Definition 7.6.12 the operation of surgery of type $(n, n-r)$. If V be a manifold of dimension $(n-1)$ without boundary, and $f : S^{r-1} \times D^{n-r} \longrightarrow V$ is an embedding, then a spherical modification of V of type $(n, n-r)$ determined by f is the manifold

$$V' = (V - \text{Int}\ (\text{Image}\, f)) \cup_{f'} (D^r \times S^{n-r-1}),$$

where $f' = f|(S^{r-1} \times S^{n-r-1})$.

The following assertions may be seen easily from our previous discussions.

(1) The effect of surgery is determined by the isotopy class of f.
(2) If $M = V \times I \cup_f \mathcal{H}_r$, then $\partial_- M$ is $V \times 0$, and its spherical modification V' is $\partial_+ M$, so that $(M, \partial_- M, \partial_+ M)$ is a cobordism. The manifold M is called the **supporting manifold** of the surgery.
(3) If V' is obtained from V by a surgery of type $(n, n-r)$, then there is an elementary cobordism (M, V, V'), and a Morse function $M \longrightarrow \mathbb{R}$ with exactly one critical point of index r.
(4) If $M = V \times I \cup_f \mathcal{H}_r$ with V' as the spherical modification of V, and $X = V \cap V'$, then up to homotopy
    (a) M is obtained from V by attaching an r-cell.
    (b) M is obtained from V' by attaching an $(n-r)$-cell.
    (c) V is obtained from X by attaching an $(n-r)$-cell, and an $(n-1)$-cell.
    (d) V' is obtained from X by attaching an r-cell and an $(n-1)$-cell.

We may get homotopy and homology of V and V' from the standard properties of cell complexes.

A further remark is that if V' is obtained from V by a surgery of type $(n, n-r)$, then V can be obtained from V' by a surgery of type $(n-r, n)$; both the modifications have the same supporting manifold. This will be clear from our discussion of duality in the next section.

$\Diamond$ **Exercise 10.2.** Show that two $(n-1)$-manifolds are cobordant if and only if one can be obtained from the other by a series of spherical modifications.

10.2. Duality theorems

In this section we shall deduce two duality theorems, which give relations between homology and cohomology of a manifold, from the existence of a handle presentation.

We begin with some preparatory lemmas.

Lemma 10.2.1. *Let K and M be two transverse submanifolds of an m-manifold N, where $\dim K = r$, $\dim M = n$, and $r+n=m$. Let p be a point in $K \cap M$. Then there is in N a coordinate chart $\phi : U \longrightarrow \mathbb{R}^m = \mathbb{R}^r \times \mathbb{R}^n$ about p such that*

$$\phi(U \cap K) = \mathbb{R}^r \times 0, \text{ and } \phi(U \cap M) = 0 \times \mathbb{R}^n.$$

Proof. There is a coordinate chart (V, ψ) about p in N such that $\psi(V) = \mathbb{R}^m = \mathbb{R}^r \times \mathbb{R}^n$, $\psi(p) = 0$, and $\psi^{-1}(\mathbb{R}^r \times 0) = V \cap K$. Then the manifold $V \cap M$ can be realised by an embedding $f : \mathbb{R}^n \longrightarrow V \cap M$ with $f(0) = p$, that is, by the embedding $g = \psi \circ f : \mathbb{R}^n \longrightarrow \mathbb{R}^r \times \mathbb{R}^n$ with $g(0) = 0$, which is of the form $g(v) = (\alpha(v), \beta(v))$, where $\alpha : \mathbb{R}^n \longrightarrow \mathbb{R}^r$ and $\beta : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ are the components of g with $\alpha(0) = \beta(0) = 0$. The transversality condition means that the map β is a submersion, i.e. β has rank n at 0. Then the map $h : \mathbb{R}^r \times \mathbb{R}^n \longrightarrow \mathbb{R}^r \times \mathbb{R}^n$ given by $h(u, v) = (u + \alpha(v), \beta(v))$ has rank $r+n$ at 0, and $h(0, v) = g(v)$, $h(u, 0) = (u, 0)$. Therefore we get by the inverse function theorem the required coordinate chart (U, ϕ) about p in N, where U is an open neighbourhood of p in V, and $\phi = h \circ \psi$. $\square$

Lemma 10.2.2. *Let K, M, and M' be submanifolds of an m-manifold N, where $\dim K = r$, $\dim M = \dim M' = n$, and $r+n=m$. Let M and M' intersect K transversely at a point $p \in M \cap M'$. Then there is a diffeotopy of N which makes M' coincide with M in a neighbourhood of p, keeping K fixed.*

Proof. By the last lemma, there is a coordinate neighbourhood U of p in N, which we may identify as $U = \mathbb{R}^r \times \mathbb{R}^n$ with p as the origin such that $U \cap K = \mathbb{R}^r \times 0$ and $U \cap M = 0 \times \mathbb{R}^n$. Then a small coordinate neighbourhood V of p in M' is represented inside U as the image of an embedding $\lambda : \mathbb{R}^n \longrightarrow \mathbb{R}^r \times \mathbb{R}^n$ transverse to $\mathbb{R}^r \times 0$ and intersecting it at the origin.

The embedding λ is isotopic to a linear embedding $\mathbb{R}^n \longrightarrow \mathbb{R}^r \times \mathbb{R}^n$. This follows because a tubular neighbourhood of λ and a tubular neighbourhood $\mathbb{R}^n$ in $\mathbb{R}^r \times \mathbb{R}^n$ are equivalent (see Exercise 7.2, p.214). Therefore we may suppose that the image of λ is a linear subspace of $\mathbb{R}^r \times \mathbb{R}^n$ intersecting $\mathbb{R}^r \times 0$ transversely. Then we can make Image $\lambda = V$ coincident with $0 \times \mathbb{R}^n$ by an isotopy from V into N keeping K fixed. By the isotopy extension theorem, the restriction of this isotopy to a disk D^n about p in V extends to a diffeotopy of N which sends the disk D^n in M' to a disk about p in M, keeping K fixed. $\square$

Throughout this chapter, the coefficient system for homology (and cohomology) will be $\mathbb{Z}$.

Theorem 10.2.3. *If V is a manifold of dimension n, and M' is a connected submanifold of it of dimension $n-r$, then there is a natural isomorphism*

$$\psi : H_0(M') \longrightarrow H_r(V, V - M')$$

(*'natural' means 'natural equivalence' between homology functors*).

We shall call the isomorphism ψ **Thom isomorphism** (actually, ψ^{-1} is referred to as Thom isomorphism in the literature).

Proof. A general version of the theorem says

$$H_i(M') \cong H_{i+r}(V, V - M'),\ i \geq 0,$$

and this may be accomplished by using the Leray-Hirsch theorem.

The statement of a special case of the Leray-Hirsch theorem is as follows. Suppose that $E \longrightarrow X$ is a vector bundle of fibre dimension r over a connected space X. Let F be the fibre, and $i : X \longrightarrow E$ the zero-section. Let $E_0 = E - i(X)$, and $F_0 = F \cap E_0$. Then there is a natural isomorphism of degree 0 of graded rings

$$H_*(X) \otimes H_*(F, F_0) \longrightarrow H_*(E, E_0).$$

This is defined using the cohomology extension of the fibre and cap product.

The theorem is proved first for a trivial bundle. Then the general result is obtained by using Meyer-Vietoris arguments and direct limits. The details are in Spanier [**43**], p. 258.

Since the homology $H_i(F, F_0) = H_i(\mathbb{R}^r, \mathbb{R}^r - 0)$ is $\mathbb{Z}$ if $i = r$, and 0 otherwise. the above isomorphism reduces to the isomorphism

$$H_i(X) \longrightarrow H_{i+r}(E, E_0).$$

In the special case when $X = M'$ and E is the normal bundle of M' in V, the isomorphism is

$$H_i(M') \longrightarrow H_{i+r}(E, E_0).$$

By the tubular neighbourhood theorem, there is a neighbourhood U of M' in V, and a diffeomorphism $(E, E - i(M')) \longrightarrow (U, U - M')$. Then the isomorphism induced on homology by this diffeomorphism and the excision isomorphism (exc) give

$$H_{i+r}(E, E - i(M')) \simeq H_{i+r}(U, U - M') \overset{exc}{\simeq} H_{i+r}(V, V - M').$$

The excision isomorphism is justified by the facts that the union of the open sets U and $V - M'$ is V, and their intersection is $U - M'$.

This completes the proof. $\square$

Now suppose that V is a connected oriented manifold of dimension n, and M, M' are connected oriented submanifolds of V such that $\dim M = r$, $\dim M' = s = n - r$, and they intersect transversely in points $p_1, \ldots, p_k$ of V. Then the normal space $\nu(M')_{p_i}$ can be oriented in two ways. Firstly, $\nu(M')_{p_i}$ has a natural orientation compatible with the given orientations of M' and V so that a positively oriented basis of $\tau(M')_{p_i}$ followed by a positively oriented basis of $\nu(M')_{p_i}$ is a positively oriented basis of $\tau(V)_{p_i}$. Secondly, if $(X_1, \ldots, X_r)$ is a positively oriented basis of the tangent space $\tau(M)_{p_i}$, then it also represents a basis of the normal space $\nu(M')_{p_i}$ (since the intersection is transverse at p_i), and so gives an orientation to $\nu(M')_{p_i}$. We say that the intersection number ϵ_i of M and M' at p_i is $+1$ or -1 according as two orientations of $\nu(M')_{p_i}$ agree or not. The intersection number $I(M, M')$ of M and M' is the sum of the intersection numbers ϵ_i at the points p_i. Clearly, we have

$$I(M, M') = (-1)^{rs} \cdot I(M', M).$$

It will follow from the following lemma that the number $I(M, M')$ does not change under an isotopy of M or M'.

Lemma 10.2.4. *Let V, M, and M' be as above. Let α be the canonical generator of $H_0(M') \cong \mathbb{Z}$, and $[M] \in H_r(M)$ the orientation generator. Then for the sequence of homomorphisms*

$$H_r(M) \xrightarrow{g} H_r(V) \xrightarrow{g'} H_r(V, V - M'),$$

which are induced by inclusion maps, we have $g' \circ g([M]) = I(M, M') \cdot \psi(\alpha)$, where ψ is the Thom isomorphism $H_0(M') \longrightarrow H_r(V, V - M')$.

Proof. Let $U_1, \ldots, U_k$ be disjoint open r-disks in M containing $p_1, \ldots, p_k$ respectively. Then the orientation of M at p_i is represented by a generator γ_i of $H_r(U_i, U_i - p_i)$ (see §5.7), and this may be obtained from the canonical generator of $H_0(M)$ by Thom isomorphism ψ followed by excision isomorphism (exc)

$$H_0(M) \xrightarrow{\psi} H_r(V, V - M) \xrightarrow{exc} H_r(U, U - p_i).$$

On the other hand, the orientation class of the fibre F of the normal bundle $E = \nu(M')$ at p_i is a generator ω_i of $H_r(F, F_0) = \mathbb{Z}$, and this gives a well-defined generator $\epsilon_i \psi(\alpha)$ of $H_r(V, V - M')$ ($\epsilon_i = \pm 1$ is the intersection number $I(M, M')$ of M and M' at p_i) under the sequence of isomorphisms

$$H_r(F, F_0) \longrightarrow H_r(E, E_0) \longrightarrow H_r(U, U - M') \xrightarrow{exc} H_r(V, V - M'),$$

where the first one corresponds to the Leray-Hirsch isomorphism

$$H_0(M') \otimes H_r(F, F_0) \longrightarrow H_r(E, E_0),$$

and the second one comes from the tubular neighbourhood theorem and the excision theorem.

Now the homomorphism $\lambda_i : H_r(U_i, U_i - p_i) \longrightarrow H_r(V, V - M')$ induced by inclusion is an isomorphism, by the naturality of Thom isomorphism.

$$\begin{tikzcd}
H_0(p_i) \arrow[r, "\psi"] \arrow[d, "\cong"'] & H_r(U_i, U_i - p_i) \arrow[d, "\lambda_i"] \\
H_0(M') \arrow[r, "\psi"'] & H_r(V, V - M')
\end{tikzcd}$$

Therefore $\lambda_i(\gamma_i) = I(M, M') \cdot \psi(\alpha)$.

Then the proof of the lemma follows from the following commutative diagram, where the lower horizontal arrow is from excision, and other arrows are induced by inclusions.

$$\begin{tikzcd}
H_r(M) \arrow[r, "g"] \arrow[d] & H_r(V) \arrow[r, "g'"] & H_r(V, V - M') \\
H_r(M, M - M \cap M') \arrow[rr, "\cong"] & & \oplus_{i=1}^k H_r(U_i, U_i - p_i) \arrow[u, "\sum \lambda_i"']
\end{tikzcd}$$

□

Consider now a presentation with a single handle $M \cup_f \mathcal{H}_r$ obtained by attaching to M a handle $\mathcal{H}_r = D^r \times D^{n-r}$ by an embedding $f : S^{r-1} \times D^{n-r} \longrightarrow \partial M$. As we know from the isotopy extension theorem, the manifold $M \cup_f \mathcal{H}_r$ is determined up to diffeomorphism by the isotopy class of f. Any diffeomorphism g of M induces a diffeomorphism $M \cup_f \mathcal{H}_r \longrightarrow M \cup_{gf} \mathcal{H}_r$. It is also determined uniquely by the isotopy class of $f \mid S^{r-1} \times 0$ and a homotopy class of normal framings of $f(S^{r-1} \times 0)$ in ∂M (see the tubular neighbourhood theorem).

Let $\partial_+(M \cup_f \mathcal{H}_r)$ be the part of $\partial(M \cup_f \mathcal{H}_r)$ defined by

$$\partial_+(M \cup_f \mathcal{H}_r) = (\partial M - \text{ Int }(\text{Image} f)) \cup (D^r \times S^{n-r-1}).$$

Definition 10.2.5. The a**-sphere** (or attaching sphere) of the handle $\mathcal{H}_r$ is the sphere $f(S^{r-1} \times 0)$ in ∂M. The b**-sphere** (or belt sphere) of $\mathcal{H}_r$ is the sphere $0 \times S^{n-r-1}$ in $\partial_+(M \cup_f \mathcal{H}_r)$. The **core** of the handle is the disk $D^r \times 0$.

The a**-disk** (resp. b**-disk**) of $\mathcal{H}_r$ is the disk whose boundary sphere is the a-sphere (resp. b-sphere) of $\mathcal{H}_r$.

The handle $\mathcal{H}_r$ corresponds to a critical point p of a Morse function which lies in the intersection of a-disk and b-disk of $\mathcal{H}_r$.

Consider a handle presentation of W on V consisting of r- and $(r+1)$-handles

$$W = W' \cup ((r+1)\text{-handles}), \qquad W' = V \cup (r\text{-handles}),$$

where the number of r-handles is k and they are denoted by $\mathcal{H}_1, \dots, \mathcal{H}_k$, and the number of $(r+1)$-handles is ℓ and they are $\mathcal{H}'_1, \dots, \mathcal{H}'_\ell$. They arise from a Morse function on W with critical points $p_1, \dots, p_k$ of index r all belonging to the same level, and critical points $p'_1, \dots, p'_\ell$ belonging to another level so that $\mathcal{H}_i$ corresponds to p_i, and $\mathcal{H}'_j$ corresponds to p'_j. The manifold V is a non-critical level of the Morse function. We introduce further notations.

D_{ai} and D_{bi} are a-disk and b-disk of the handle $\mathcal{H}_i$, $i = 1, \dots, k$,

S_{ai} and S_{bi} are a-sphere and b-sphere of the handle $\mathcal{H}_i$, $i = 1, \dots, k$,

D'_{aj} and D'_{bj} are a-disk and b-disk of the handle $\mathcal{H}'_j$, $j = 1, \dots, \ell$,

S'_{aj} and S'_{bj} are a-sphere and b-sphere of the handle $\mathcal{H}'_j$, $j = 1, \dots, \ell$.

We orient the normal bundle $\nu(D_{bi})$ of D_{bi} in W' so that the intersection number of D_{ai} and D_{bi} is $+1$ at the point p_i. This orientation of $\nu(D_{bi})$ determines the orientation of the normal bundle $\nu(S_{bi})$ of S_{bi} in $\partial_+ W'$, because $\nu(S_{bi})$ is naturally isomorphic to the restriction of $\nu(D_{bi})$ to S_{bi}. By the definition of the intersection number, the orientations of a-disks in W and W' determine in a natural way the orientation of S_{ai} in $\partial_+ W'$ and $\nu(S_{bi})$ in $\partial_+ W'$. Therefore the intersection numbers $I(S_{bi}, S'_{aj})$ are well-defined in $\partial_+ W'$.

The homology groups $H_r(W', V)$ and $H_{r+1}(W, W')$ are free abelian groups on generators

$$[D_{a1}], \dots, [D_{ak}] \text{ and } [D'_{a1}], \dots, [D'_{a\ell}]$$

respectively, represented by oriented a-disks (Theorem 9.3.7).

Lemma 10.2.6. *Let M be an oriented boundaryless submanifold of dimension r of $\partial_+ W'$ with orientation generator $[M] \in H_r(M)$. Let $h : H_r(M) \longrightarrow H_r(W', V)$ be the homomorphism induced by inclusion. Then*

$$h([M]) = I(S_{b1}, M) \cdot [D_{a1}] + \cdots + I(S_{bk}, M) \cdot [D_{ak}],$$

where the intersection numbers $I(S_{bi}, M)$ are taken in $\partial_+ W'$.

PROOF. We shall prove the lemma for $k = 1$. The proof of the general case is similar. We write $p = p_1$, $D_a = D_{a1}$, and $S_b = S_{b1}$. Then consider the following diagram:

$$\begin{tikzcd}
H_r(W', W' - D_b) \arrow[r, "h_2"] & H_r(V \cup D_a, V \cup (D_a - p)) \arrow[d, "h_3"] \\
H_r(\partial_+ W', \partial_+ W' - S_b) \arrow[u, "h_1"] & H_r(V \cup D_a, V) \arrow[d, "h_4"] \\
H_r(M) \arrow[u, "h_0"] \arrow[r, "h"] \arrow[d] & H_r(W', V) \\
H_r(\partial_+ W') \arrow[r] & H_r(W') \arrow[u]
\end{tikzcd}$$

where all the homomorphisms, except h_2 and h_3, are induced by obvious inclusion maps. The homomorphism h_2 is induced by the deformation retraction $W' \longrightarrow V \cup D_a$ defined in the proof of Theorem 9.3.4, which sends $W' - D_b$ to $V \cup (D_a - p)$. The homomorphism h_3 is induced by the obvious retraction $V \cup (D_a - p) \longrightarrow V$.

Each of the homomorphisms h_2, h_3, and h_4 are isomorphisms with inverses i_*, j_*, and k_* respectively, where i_* and j_* are induced by the inclusions

$$i : (V \cup D_a, V \cup (D_a - p)) \longrightarrow (W', W' - D_b),$$

$$j : (V \cup D_a, V) \longrightarrow (V \cup D_a, V \cup (D_a - p)),$$

and k_* is induced by the retraction $k : (W', V) \longrightarrow (V \cup D_a, V)$. The above diagram is commutative, because if $\ell : H_r(M) \longrightarrow H_r(V \cup D_\alpha, V)$ is homomorphism induced by inclusion, then $h_1 \circ h_0 = i_* \circ j_* \circ \ell$, and therefore $h_4 \circ h_3 \circ h_2 \circ h_1 \circ h_0 = h$.

By Lemma 10.2.4, we have $h_0([M]) = I(S_b, M) \cdot \psi(\alpha)$, where $\alpha \in H_0(S_b)$ is the canonical generator, and $\psi : H_0(S_b) \longrightarrow H_r(\partial_+ W', \partial_+ W' - S_b)$ is the Thom isomorphism. Therefore in order to prove $h([M]) = I(S_b, M) \cdot [D_a]$, it is sufficient to prove that $h_4 \circ h_3 \circ h_2 \circ h_1(\psi(\alpha)) = [D_a]$.

The class $\psi(\alpha)$ is represented by any disk $D^r \subset \partial_+ W'$ which intersects S_b transversely in one point with intersection number $I(S_b, D^r) = +1$ at that point. Taking $D^r = D_a$, it follows by our choice of orientation for $\nu(S_b)$ that $I(D_b, D_a) = I(S_b, D_a) = +1$. Thus D_a represents $\psi(\alpha)$ in $H_r(\partial_+ W', \partial_+ W' - S_b)$, and also $\psi'(\alpha)$ in $H_r(W', W' - D_b)$, where $\psi' : H_0(D_b) \longrightarrow H_r(W', W' - D_b)$ is another Thom isomorphism. By naturality of Thom isomorphism, we have $h_1 \psi(\alpha) = \psi'(\alpha)$. Also D_a represents the generator $[D_a]$ in $H_r(W', V)$, and by commutativity, D_a represents $\psi'(\alpha) = h_1 \psi(\alpha)$. This means

$$(h_2)^{-1}(h_3)^{-1}(h_4)^{-1}([D_a]) = h_1 \psi(\alpha),$$

or $h_4 \circ h_3 \circ h_2 \circ h_1(\psi(\alpha)) = [D_a]$ as required. $\square$

Corollary 10.2.7. *With respect to the basis represented by the oriented a-disks, the boundary homomorphism*

$$\overline{\partial} : H_{r+1}(W, W') \longrightarrow H_r(W', V)$$

for the triple $V \subset W' \subset W$ is given by the matrix (a_{ij}) of the intersection numbers $a_{ij} = I(S_{bi}, S'_{aj})$ in $\partial_+ W'$.

Proof. Consider the following homomorphisms induced by inclusions

$$j : H_r(\partial_+ W') \longrightarrow H_r(W'), \quad k : H_r(W') \longrightarrow H_r(W', V).$$

Let $\lambda^{-1} : H_{r+1}(W - W', \partial_+ W') \longrightarrow H_{r+1}(W, W')$ be the excision isomorphism, and $\partial : H_{r+1}(W, W') \longrightarrow H_r(W')$ be the boundary homomorphism for the pair (W, W'). Then the boundary homomorphism for the triple $V \subset W' \subset W$ is defined by $\overline{\partial} = k \circ \partial$. By naturality of ∂, we may replace it by $j \circ \partial \circ \lambda$

$$H_{r+1}(W, W') \xrightarrow{\lambda} H_{r+1}(W - W', \partial_+ W') \xrightarrow{\partial} H_r(\partial_+ W') \xrightarrow{j} H_r(W').$$

Then $\overline{\partial} = k \circ j \circ \partial \circ \lambda$.

The fundamental class of $H_{r+1}(W - W', \partial_+ W')$ corresponds to the fundamental class $[D'_{aj}] \in H_{r+1}(W, W')$ via the isomorphism λ^{-1}. Also the fundamental class of $H_r(\partial_+ W')$ is $\mu([S'_{aj}])$, where $\mu : H_r(S'_{aj}) \longrightarrow H_r(\partial_+ W')$ is induced by the inclusion map. Then by a property of the boundary homomorphism ∂, as stated in Exercise 5.14 (p.163), we have

$$\partial \circ \lambda([D'_{aj}]) = \mu([S'_{aj}]).$$

The result then follows by taking $M = S'_{aj}$ and $h = k \circ j$ in Lemma 10.2.6. $\square$

Theorem 10.2.8 (Duality Theorem). *If $(W, \partial_- W, \partial_+ W)$ is a cobordism where W is orientable,then*

$$H_r(W, \partial_- W) \cong H^{n-r}(W, \partial_+ W)$$

for $0 \leq r \leq n$.

Proof. By Corollary 10.1.9 we have a handle presentation of W using a Morse function $f : M \longrightarrow \mathbb{R}$

$$\partial_- W \times I = W_{-1} \subset W_0 \subset W_1 \subset \cdots \subset W_n = W,$$

where W_r is W_{r-1} plus a number of r-handles attached. Let α_r be the number of r-handles. By Theorem 9.3.4, the homology $H_r(W_r, W_{r-1})$ is a free abelian group generated by α_r r-cells, which are the a-disks of the r-handles. Thus we have a chain complex $C_* = (C_r, \partial)$, where $C_r = H_r(W_r, W_{r-1})$, and ∂ is the boundary homomorphism of the exact homology sequence of the triple (W_r, W_{r-1}, W_{r-2}). As shown in Corollary 10.2.7, ∂ is given by the matrix of the intersection numbers $a_{ij} = I(S_{bi}, S'_{aj})$. Then, it follows from Dold [**7**], Chap. V, §1 (p. 85), that $H_*(C_*) \cong H_*(W, \partial_- W)$.

Using the Morse function $-f$ we get the dual handle presentation of W

$$\partial_+ W \times I = W'_{-1} \subset W'_0 \subset W'_1 \subset \cdots \subset W'_n = W,$$

where $W'_r = W_{n-r} \cup W_{n-r+1} \cup \cdots \cup W_n$, $0 \le r \le n$. We may say that W'_r is W'_{r-1} plus a number of r-handles attached, where the r-handles correspond to the $(n-r)$-handles of the presentation given by f. Let $C'_r = H_r(W'_r, W'_{r-1})$. and $\partial' : C'_r \longrightarrow C'_{r-1}$ be given as in Corollary 10.2.7. Since the a-spheres and b-spheres are interchanged, ∂' is given by a matrix (a'_{ij}), where $a'_{ij} = (-1)^{r(n-r-1)} a_{ji}$. Let $C'^* = (C'^r, \delta)$ be the cochain complex dual to the chain complex $C'_* = (C'_r, \partial')$. Then $C'^r = \mathrm{Hom}(\mathrm{C}'_r, \mathbb{Z})$ and the matrix (b_{ij}) of the coboundary δ is the transpose of the matrix (a'_{ij}). Thus $b_{ij} = (-1)^{r(n-r-1)} a_{ij}$ and δ is the same as ∂ up to sign. The isomorphisms of chain groups induce isomorphism $H_r(C_*) \cong H^{n-r}(C'^*)$ for each value of r. □

Corollary 10.2.9 (Poincaré Duality). *If $\partial W = \emptyset$, then*

$$H_r(W) \cong H^{n-r}(W).$$

PROOF. Consider the cobordism $(W, \emptyset, \emptyset)$. □

Corollary 10.2.10 (Lefschetz Duality). *There are two types of isomorphisms*

$$(1) \quad H_r(W) \cong H^{n-r}(W, \partial W), \ \textit{and}$$

$$(2) \quad H^r(W) \cong H_{n-r}(W, \partial W).$$

PROOF. Consider the cobordism $(W, \emptyset, \partial W)$. □

10.3. Normalisation of presentation

Consider now a presentation with two handles of consecutive indices

$$M \cup_f \mathcal{H}_r \cup_g \mathcal{H}_{r+1}.$$

Let us fix some notations.

$$\mathcal{H}_r = D_1^r \times D_1^{n-r}, \ \mathcal{H}_{r+1} = D_2^{r+1} \times D_2^{n-r-1},$$

$$f : S_1^{r-1} \times D_1^{n-r} \longrightarrow \partial M, \ g : S_2^r \times D_2^{n-r-1} \longrightarrow \partial_+(M \cup_f \mathcal{H}_r).$$

The suffices 1 and 2 are put to distinguish the handles, although the disks and spheres are the standard ones:

$$D^{r+1} = \{(x_0, \cdots, x_r) \mid \|x\| \le 1\}, \quad S^r = \{(x_0, \cdots, x_r) \mid \|x\| = 1\}.$$

We shall denote the upper and lower hemisphere of S^r by D_+^r and D_-^r. They are given by

$$D_+^r = \{(x_0, \cdots, x_r) \in S^r \mid x_r \ge 0\}, \ D_-^r = \{(x_0, \cdots, x_r) \in S^r \mid x_r \le 0\}.$$

We have the identification $P_r^{-1} : D_+^r \longrightarrow D^r$, where P_r is the stereographic projection from the south pole of S^r.

Definition 10.3.1. The handles of the presentation $M \cup_f \mathcal{H}_r \cup_g \mathcal{H}_{r+1}$ are said to be in a **normal position** if the attaching map g can be isotoped to an embedding $S_2^r \times D_2^{n-r-1} \longrightarrow \partial_+(M \cup_f \mathcal{H}_r)$ which is given on $D_+^r \times D_2^{n-r-1}$ as a product of two embeddings $\alpha : D_+^r \longrightarrow D_1^r$ and $\beta : D_2^{n-r-1} \longrightarrow S_1^{n-r-1}$

$$D_+^r \times D_2^{n-r-1} \xrightarrow{\alpha\times\beta} D_1^r \times S_1^{n-r-1} \subset M \cup_f (D_1^r \times D_1^{n-r}),$$

and which maps $D_-^r \times D_2^{n-r-1}$ into M.

In the situation of a normal position the a-sphere of $\mathcal{H}_{r+1}$ intersects the b-sphere $\mathcal{H}_r$ in one and only one point, unless α fails to cover the origin of D_1^r. But this possibility of α is not serious at all, because we may replace α and β by any two embeddings. By the Disk Theorem (7.4.10), the new product will be isotopic to $\alpha \times \beta$. In particular, we may replace α by the identity map, or by the stereographic projection P_r^{-1}.

The manifold $\partial_+(M \cup_f \mathcal{H}_r)$ contains the a-sphere $g(S_2^r \times 0)$ of $\mathcal{H}_{r+1}$ and the b-sphere $0 \times S_1^{n-r-1}$ of $\mathcal{H}_r$. The embedding

$$\overline{g} = g|(S_2^r \times 0) : S_2^r \longrightarrow \partial_+(M \cup_f \mathcal{H}_r)$$

may be approximated by a smooth map transverse to the b-sphere, and if the approximation is sufficient close we may suppose that the approximated map is still an embedding isotopic to the original map. Since $\dim S_2^r = \operatorname{codim} S_1^{n-r-1}$, the intersection $\overline{g}(S_2^r) \cap S_1^{n-r-1}$ consists of isolated points (Theorem 6.2.5). Since S_2^r is compact, there are only finitely many such points.

Proposition 10.3.2. *If in a handle presentation*

$$M \cup_f \mathcal{H}_r \cup_g \mathcal{H}_{r+1},$$

with two handles of consecutive indices, the a-sphere of $\mathcal{H}_{r+1}$ intersects the b-sphere of $\mathcal{H}_r$ transversely at a single point, then the handles can be brought into a normal position by means of suitable isotopies.

The proof will follow after the following lemma.

Let $D^n(\epsilon)$ denote the disk of radius ϵ in $\mathbb{R}^n$, $D^n(1) = D^n$. Let $D_+^n(\epsilon)$ be the polar cap in $D_+^n \subset S^n$ given by $D_+^n(\epsilon) = P_n(D^n(\epsilon)),\ 0 < \epsilon \leq 1,\ D_+^n(1) = D_+^n$.

Lemma 10.3.3. *The hemisphere D_+^n can be shrunk to a polar cap $D_+^n(\epsilon)$ by an isotopy $h_t : D_+^n \longrightarrow D_+^n$ with $h_0 = \mathrm{Id}$.*

Proof. Define h_t by

$$h_t(x) = P_n \circ [(1 - t + t\epsilon) \cdot P_n^{-1}(x)].$$

Then $h_0 = \mathrm{Id}$ and $h_1(D_+^n) = D_+^n(\epsilon)$. □

Proof of Proposition 10.3.2. Let $p \in \overline{g}(S_2^r) \cap S_1^{n-r-1}$. Consider p as a point $(0, q) \in D_1^r \times S_1^{n-r-1}$. Then the submanifolds $\overline{g}(S_2^r)$ and $D_1^r \times q$ of $D_1^r \times S_1^{n-r-1}$ intersect $0 \times S_1^{n-r-1}$ transversely at $(0, q)$, and therefore, by

Lemma 10.2.2, a portion of $\overline{g}(S_2^r)$ near $(0, q)$ may be made to lie in a small disk D_0^r in $D_1^r \times q$ centred at $(0, q)$ by a diffeotopy of $D_1^r \times S_1^{n-r-1}$, which is stationary on $0 \times S_1^{n-r-1}$.

We may suppose without loss of generality that $(\overline{g})^{-1}(p)$ is the north pole of S_2^r, and $(\overline{g})^{-1}(D_0^r)$ is a polar cap $D_+^r(\epsilon)$ in S_2^r. Then by Lemma 10.3.3, $\overline{g}|D_+^r$ is isotopic to an embedding $\overline{g}_1$ that maps the upper hemisphere D_+^r into

$$D_1^r \times q \subset D_1^r \times S_1^{n-r-1}.$$

The tubular neighbourhood $D_1^r \times S_1^{n-r-1}$ of $0 \times S_1^{n-r-1}$ in $\partial_+(M \cup_f \mathcal{H}_r)$ may be shrunk by a diffeotopy to a smaller concentric tubular neighbourhood which intersects $D_1^r \times q$ in Image $\overline{g}_1$. The inverse of this diffeotopy extends to a diffeotopy of

$$\partial_+(M \cup_f \mathcal{H}_r),$$

which stretches the image of $\overline{g}_1$ to the whole of $D_1^r \times q$, making $\overline{g}_1$ an embedding from D_+^r onto $D_1^r \times q$.

By Proposition 7.3.5, $\overline{g}_1$ extends to an embedding

$$g_1 : S_2^r \times D_2^{n-r-1} \longrightarrow \partial_+(M \cup_f \mathcal{H}_r)$$

which is isotopic to g. Therefore g_1 is $\overline{g}_1$ on D_+^r and an embedding on the second factor, and so the handles are in a normal position. □

10.4. Cancellation of handles

In this section we shall show that a pair of handles of consecutive indices $(\mathcal{H}_r, \mathcal{H}_{r+1})$ satisfying the condition of the last proposition may be cancelled.

Proposition 10.4.1. *In a handle presentation*

$$(M \cup_f \mathcal{H}_r) \cup_g \mathcal{H}_s,$$

where $r > s$, the handles may be attached simultaneously, or in the reverse order.

Proof. Let $\dim M = n$, and $V = M \cup_f \mathcal{H}_r$. By the definition of handle presentation, the interiors of the images of the embeddings f and g are disjoint, and we have embedded in $\partial_+ V$ the a-sphere $g(S^{s-1})$ of $\mathcal{H}_s$ and the b-sphere S^{n-r-1} of $\mathcal{H}_r$. Since $r > s$, the sum of their dimensions is less than the dimension of $\partial_+ V$, which is $n-1$. Therefore $g \mid S^{s-1}$ may be approximated by an embedding of S^{s-1} in $\partial_+ V$ not intersecting S^{n-r-1}, and isotopic to $g \mid S^{s-1}$ (supposing the approximation is close enough). By using isotopy extension theorem, and the fact that a tubular neighbourhood may be shrunk by isotopy, we may assume that $g(S^{s-1})$ does not intersect the tubular neighbourhood $D^r \times S^{n-r-1}$ of S^{n-r-1}. Next the tubular neighbourhood $S^{s-1} \times D^{n-s}$ may be shrunk so that $g(S^{s-1} \times D^{n-s})$ also does not intersect $D^r \times S^{n-r-1}$. Now the attaching part $\mathcal{H}_s$ is disjoint from the handle $\mathcal{H}_r$, and it lies in $\partial M - \text{Image}\, f$. Therefore the handles can be attached in any order. □

Proposition 10.4.2. *Let $P(= P_{n-r-1}) : D^{n-r-1} \longrightarrow D^{n-r}$ be the the stereographic projection of the equatorial disk D^{n-r-1} of D^{n-r} from the south pole $(0, 0, \ldots, 0, -1)$ of ∂D^{n-r} onto the northern hemisphere D_+^{n-r} of ∂D^{n-r} corresponding to non-negative last coordinate. Then there is a diffeomorphism*

$$S^r \times D^{n-r} \cup_{\mathrm{Id}\times P} \mathcal{H}_{r+1} \cong D^n,$$

where $\mathcal{H}_{r+1}$ is the handle $D^{r+1} \times D^{n-r-1}$, and $\mathrm{Id} \times P$ *is the attaching map*

$$S^r \times D^{n-r-1} \longrightarrow S^r \times D^{n-r}.$$

Proof. In the special case $n = 2$, $r = 0$, two copies of D^2, namely $S^0 \times D^2$, can be joined by a handle $D^1 \times D^1$ using the embedding $\mathrm{Id} \times P$ to obtain D^2 again.

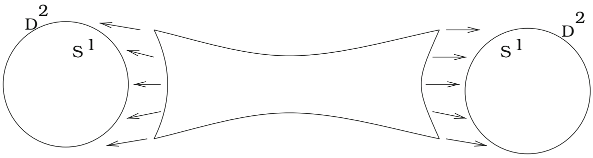


Figure 10.2

Let us consider this special case in another way, with a view to generalisation in higher dimension. Let E be the ellipse $f(x, y) = \frac{1}{2}x^2 + y^2 = 1$, and H be the hyperbola $g(x, y) = 2x^2 - 2y^2 = 1$. Let A and B denote the following regions of the plane

$$A = \mathrm{Int}\ E \cap \mathrm{Int}\ H = \{(x, y) \in \mathbb{R}^2 \mid f(x, y) \leq 1, g(x, y) \geq 1\},$$
$$B = \mathrm{Int}\ E \cap \mathrm{Ext} H = \{(x, y) \in \mathbb{R}^2 \mid f(x, y) \leq 1, g(x, y) \leq 1\}.$$

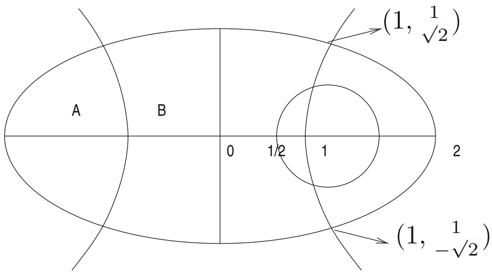


Figure 10.3

The component A_+ of A which lies in $\mathbb{R}^2_+$ (where $x \geq 0$) is a manifold with two corner points in $E \cap H$. After straightening these by the method of Example 7.5.4, we get a smooth manifold which is diffeomorphic to the region

$(x-1)^2+y^2 \leq \frac{1}{4}$, as may be seen by moving its boundary along the directions of the vector field given by the rays emanating from the focus $(1,0)$. For example, the corner points $(1, \pm\frac{1}{\sqrt{2}}) \in E \cap H$ correspond to the points $(1, \pm\frac{1}{2})$ which are the opposite ends of a diameter of the circle $(x-1)^2+y^2 = \frac{1}{4}$. We may say that the region $(x-1)^2+y^2 \leq \frac{1}{4}$ is obtained by straightening the corner of the component A_+. Note that the rays are everywhere transverse to the boundary ∂A_+, and they form a fibre bundle over ∂A_+, of which the circle $(x-1)^2+y^2 = \frac{1}{4}$ is a smooth section.

The region B is diffeomorphic to the handle or rectangle

$$\{(u,v) \in \mathbb{R}^2 \mid |u| \leq \frac{1}{\sqrt{2}}, |v| \leq 1\}.$$

To see this, note that a point (x,y) in B with $(x,y) \neq (0,0)$ lies on a hyperbola

$$\frac{x^2}{1-\lambda_1} - \frac{y^2}{\lambda_1} = 1, \quad \frac{1}{2} \leq \lambda_1 < 1,$$

and an ellipse

$$\frac{x^2}{1+\lambda_2} + \frac{y^2}{\lambda_2} = 1, \quad 0 < \lambda_2 \leq 1.$$

These intersect in four points $(\pm\sqrt{1-\lambda_1}\sqrt{1+\lambda_2}, \pm\sqrt{\lambda_1\lambda_2})$. Therefore, writing $u^2 = 1-\lambda_1$, $v^2 = \lambda_2$, we may set the transformation as

$$x = u\sqrt{1+v^2}, \quad y = v\sqrt{1-u^2}, \quad |u| \leq \frac{1}{\sqrt{2}}, |v| \leq 1,$$

(only positive square roots are considered) with the origin remaining fixed. The Jacobian determinant of the transformation is

$$\frac{1-u^2+v^2}{\sqrt{1-u^2}\sqrt{1+v^2}},$$

so it is non-zero for $|u| \leq \frac{1}{\sqrt{2}}, |v| \leq 1$.

For the general case, write

$$x = (x_1, \ldots, x_{r+1}), \quad y = (y_1, \ldots, y_{n-r-1}),$$
$$u = (u_1, \ldots, u_{r+1}), \quad v = (v_1, \ldots, v_{n-r-1}).$$
$$A = \{(x,y) \in \mathbb{R}^{r+1} \times \mathbb{R}^{n-r-1} \mid f(\|x\|, \|y\|) \leq 1, g(\|x\|, \|y\|) \geq 1\},$$
$$B = \{(x,y) \in \mathbb{R}^{r+1} \times \mathbb{R}^{n-r-1} \mid f(\|x\|, \|y\|) \leq 1, g(\|x\|, \|y\|) \leq 1\}.$$

Then $A \cup B \cong D^n$, and the diffeomorphism of $D^{r+1}(\frac{1}{\sqrt{2}}) \times D^{n-r-1}$ onto the region B is given by

$$x_i = u_i\sqrt{1+\|v\|^2}, \qquad y_i = v_i\sqrt{1-\|u\|^2}.$$

Note that the corner along $\|x\| = 1$, $\|y\| = \frac{1}{\sqrt{2}}$ in $\partial A \cap \partial B$ correspond to $S^r(\frac{1}{\sqrt{2}}) \times S^{n-r-2}$.

As in the case $n = 2$, we round off A along $\|x\| = 1$, $\|y\| = \frac{1}{\sqrt{2}}$, and then deform it onto the manifold $M : (\|x\|-1)^2 + \|y\|^2 \leq \frac{1}{4}$ by moving along the trajectories of the vector field formed by the rays from the centre of the r-sphere $\|x\| = 1, y = 0$. Note that the vector field is transverse to the boundary of M, and that of the region A (after rounding the corner along $\|x\| = 1, \|y\| = \frac{1}{\sqrt{2}}$).

The manifold M is diffeomorphic to $S^r \times D^{n-r}$. To see this consider $S^r \times D^{n-r}$ as a subset of $\mathbb{R}^{r+1} \times \mathbb{R}^{n-r-1} \times \mathbb{R}$ with coordinates (a, b, c) where $\|a\| = 1$ and $\|b\|^2 + c^2 \leq 1$. Define diffeomorphism $S^r \times D^{n-r} \longrightarrow M$ by sending (a, b, c) to (x, y), where

$$x = \left(1 + \frac{1}{2}c\right) \cdot a, \qquad y = \frac{1}{2}b.$$

Its inverse is

$$a = \frac{x}{\|x\|}, \quad b = 2y, \quad c = 2(\|x\| - 1)$$

(note that the inverse is smooth, as $\|x\|$ is nowhere zero).

Finally, let us see how the a-sphere $S^r \times 0$ of the $(r + 1)$-handle $D^{r+1} \times D^{n-r-1}$ is embedded in $S^r \times D^{n-r}$. The points $(u, v) \in S^r \times 0$ with $\|u\| = 1$, $v = 0 \in D^{n-r-1}$ correspond to the points $(x, y) \in S^r(\frac{1}{\sqrt{2}}) \times 0$ with $\|x\| = \frac{1}{\sqrt{2}}$, $y = 0 \in D^{n-r-1}$, by the diffeomorphism $x = \frac{u}{\sqrt{2}}$, $y = v$. These points (x, y) belong to ∂B, since $f(\|x\|, \|y\|) < 1$ and $g(\|x\|, \|y\|) = 1$. Also the points (x, y) become the points $(a, b, c) \in S^r \times D^{n-r}$, where

$$a = \frac{x}{\|x\|} = \frac{u}{\|u\|} = u, \quad b = 0, \quad c = -1.$$

Thus the attaching map on $S^r \times 0$ is the identity map. Therefore the full attaching map $S^r \times D^{n-r-1} \longrightarrow S^r \times D^{n-r}$ (which is a tubular neighbourhood of the embedding $S^r \times 0 \longrightarrow S^r \times D^{n-r}$) is isotopic to the embedding $\mathrm{Id} \times P$, by the disk theorem. $\square$

Proposition 10.4.3. *If in a handle presentation $D^n \cup_f \mathcal{H}_r \cup_g \mathcal{H}_{r+1}$, the a-sphere of $\mathcal{H}_{r+1}$ meets the b-sphere of $\mathcal{H}_r$ transversely at a single point, then*

(a) *the embedding f can be chosen so that $D^n \cup_f \mathcal{H}_r \cong S^r \times D^{n-r}$,*

(b) *the embedding g can be chosen so that $D^n \cup_f \mathcal{H}_r \cup_g \mathcal{H}_{r+1} \cong D^n$.*

Proof. First note that if two copies of $D^r \times D^{n-r}$ are glued together along $S^{r-1} \times D^{n-r}$ by using the identity map, then the resulting manifold is $S^r \times D^{n-r}$. We regard the first copy of $D^r \times D^{n-r}$ as D^n by rounding off the corner, and the second copy as an r-handle $\mathcal{H}_r$ (see Figure 10.4 below).

Therefore the proof of the proposition may be completed, by showing that the attaching map f is isotopic to the inclusion map of $S^{r-1} \times D^{n-r}$ into $D^r \times D^{n-r}$. and that g is isotopic to the map $\mathrm{Id} \times P$ of Proposition 10.4.2. We shall show these by bringing g into a normal position.

The upper and lower hemisphere of S^r are D^r_+ and D^r_-. By Proposition 10.3.2, we may suppose that $g|D^r_+ \times D^{n-r-1}$ is the product map

$$P_r^{-1} \times P_{n-r-1} : D^r_+ \times D^{n-r-1} \longrightarrow D^r \times S^{n-r-1} \subset \partial_+(D^n \cup_f \mathcal{H}_r),$$

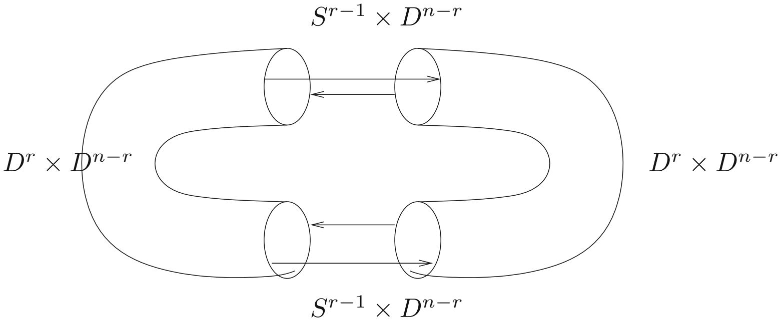


FIGURE 10.4

where the P's are the stereographic projections. Note that we may suppose further that $g|D^r_+ \times D^{n-r-1}$ is isotopic to any embedding with the same orientation, by the Disk Theorem. Thus, $\overline{g} = g|S^{r-1} \times S^{n-r-2}$ is essentially the identity map. Since D^n is convex, the map $\overline{f} = f|S^{r-1} \times S^{n-r-2}$ and the map $\overline{g}$ are homotopic. The continuous homotopy can be approximated by a smooth homotopy, and if the approximation is small enough, we may suppose that $\overline{f}$ is isotopic to $\overline{g}$. By Proposition 7.3.5, there is an embedding $f' : S^{r-1} \times D^{n-r} \to D^n$ which is isotopic to $f : S^{r-1} \times D^{n-r} \to D^n$ such that $f'|S^{r-1} \times S^{n-r-2} = \overline{g}$. Therefore $f'|S^{r-1} = \mathrm{Id}$, and f' is also a product map. Thus f may be taken as the inclusion map, and g as the product map $\mathrm{Id} \times P_{n-r-1}$. This completes the proof. □

Theorem 10.4.4 (**Cancellation of handles**). *The handles may be cancelled from the presentation $M \cup_f \mathcal{H}_r \cup_g \mathcal{H}_{r+1}$, where the a-sphere of $\mathcal{H}_{r+1}$ meets the b-sphere of $\mathcal{H}_r$ transversely at a single point, in the following sense*

$$M \cup_f \mathcal{H}_r \cup_g \mathcal{H}_{r+1} \cong M.$$

PROOF. We may suppose that $\partial_+ M$ contains a disk D^{n-1} to which the handles are attached. Then, we may write M as the connected sum $M \cong M + D^n$ so that the handles are added to $D^{n-1} \subset D^n$. Therefore we have from the last proposition

$$M \cup_f \mathcal{H}_r \cup_g \mathcal{H}_{r+1} \cong M + (D^n \cup_f \mathcal{H}_r \cup_g \mathcal{H}_{r+1}) \cong M + D^n \cong M.$$

□

Definition 10.4.5. A pair $(\mathcal{H}_r, \mathcal{H}_{r+1})$ of handles of consecutive indices is called a **complementary pair** if the a-sphere $\mathcal{H}_{r+1}$ intersects the b-sphere of $\mathcal{H}_r$ transversely at a single point.

Theorem 10.4.6. *At any point of a handle presentation of a manifold M, a pair of complementary handles may be added.*

PROOF. The proof is now easy. First write M as $M + D^n$, and then add a pair of complementary handles to D^n, and hence to M. □

Proposition 10.4.7. *A cobordism* (W, V_0, V_1)*, where* W *is connected, admits a handle presentation of the following kind:*

(a) *if* $V_0 \neq \emptyset$*, there are no* 0*-handles,*

(b) *if* $V_0 = \emptyset$*, there is one* 0*-handle.*

PROOF. Attaching 0-handles is the same as taking disjoint union of $V_0 \times I$ and some n-disks D^n. If V_0 is connected, we may join with $V_0 \times I$ the 0-handles one by one (the order is immaterial) by 1-handles so that at any stage the resulting manifold remains connected. These bridging operations will not change the diffeomorphism type of the manifold. Because, in a handle presentation

$$M = V_0 \times I \cup D^0 \times D^n \cup D^1 \times D^{n-1}$$

the a-sphere ∂D^1 (consisting of two points) of the 1-handle and the b-sphere S^{n-1} of the 0-handle intersect transversely at a single point, and therefore M is diffeomorphic to $V_0 \times I$ by the cancellation theorem (10.4.4). Therefore, if V_0 is connected, there are no 0-handles. If V_0 has components, then also there are no 0-handles.

If $V_0 = \emptyset$, then all the 0-handles can be absorbed in one 0-handle. □

10.5. Classification of closed surfaces

A compact 2-manifold without boundary is called a closed surface. Here are some examples.

- $\Sigma_0 =$ sphere S^2,
- $\Sigma_g =$ the connected sum $T \# \cdots \# T$ of k copies of the torus T,
- $\Omega_k =$ the connected sum $P^2 \# \cdots \# P^2$ of k copies of the projective plane P^2.

Theorem 10.5.1. (1) *A closed connected orientable surface is diffeomorphic to one and only one of surfaces* Σ_g*,* $g = 0, 1, \ldots$.

(2) *A closed connected non-orientable surface is diffeomorphic to one and only one of surfaces* Ω_k*,* $k = 1, 2, \ldots$.

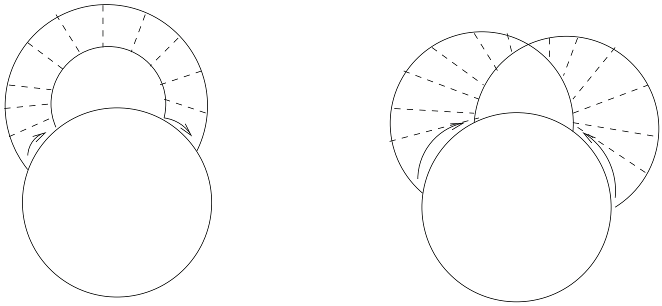
PROOF. A closed surface M admits a Morse function with non-degenerate critical points of indices 0, 1, and 2. It has therefore a handle presentation with handles of indices 0, 1, and 2. Since the surface is connected, it has a handle presentation with only one 0-handle, only one 2-handle, and k number of 1-handles where k may be any integer $= 0, 1, 2, \ldots$ (Proposition 10.4.7).

If $k = 0$, the Morse function f on M has one maximum point p, and one minimum point q. The point p has a Morse coordinate neighbourhood U with coordinates x, y in which $f(x, y) = P - x^2 - y^2$, where P is the maximum value of f. Let

$$D_p = \{(x, y) \in U \mid P - \epsilon < f(x, y) \leq P\}$$

for some $\epsilon > 0$. Then D_p is a 2-disk. Similarly, we have a 2-disk neighbourhood D_q around q. Then $M_0 = M - \mathrm{Int}(\mathrm{D_p} \cup \mathrm{D_q})$ contains no critical point of f, and therefore M_0 is a cylinder with boundary $\partial D_p \cup \partial D_q$. Then $M = D_p \cup M_0 \cup D_q$, and it is diffeomorphic to the sphere S^2.

If $k = 1$, then the surface M has one 1-handle attached to a 0-handle D^2. Now a 1-handle can be attached to a disk in two ways: (a) in an orientation preserving way (attach an ordinary band), or (b) in an orientation reversing way (attach a Möbius band $\mathcal{M}$), see Figures 10.5 and 10.6.



(a) (b)

FIGURE 10.5. Attaching to a 2-disk
(a) an ordinary 1-handle (b) a twisted 1-handle

In case (a), the resulting surface has two disconnected circle boundaries, and they cannot be eliminated by attaching a 2-handle D^2 in order to obtain a closed surface.

In case (b), the resulting surface has a circle boundary, and attaching a 2-handle to this we obtain a projective plane P^2. Thus in this case we have only

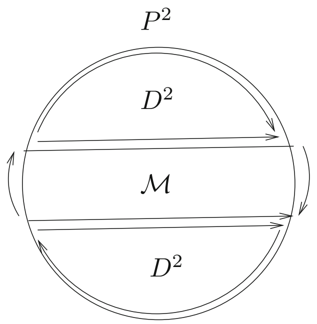


FIGURE 10.6. Projective plane P^2 =Möbius band $\mathcal{M}$ + Disk D^2

one non-orientable surface P^2, since a non-orientable Möbius band is embedded in it (see also Example 5.5.9).

For $k = 2$ either (1) both the 1-handles $\mathcal{H}^1$ and $\mathcal{H}^2$ are ordinary or (2) both of them are twisted. In each of the two cases, we have two possibilities before us: either (a) $\mathcal{H}^1 \cap \mathcal{H}^2 = \emptyset$ or (b) $\mathcal{H}^1 \cap \mathcal{H}^2 \neq \emptyset$. The case 1(a) is ruled out, because the boundary of $D^2 \cup \mathcal{H}^1 \cup \mathcal{H}^2$ is disconnected. In case 1(b), the surface $D^2 \cup \mathcal{H}^1 \cup \mathcal{H}^2$ becomes a torus with a 2-disk removed from it (see Example 9.2.19). Thus the attachment of the final 2-handle to $D^2 \cup \mathcal{H}^1 \cup \mathcal{H}^2$ will produce the full torus T^2. In case 2(a), we get after joining a 2-handle to $D^2 \cup \mathcal{H}^1 \cup \mathcal{H}^2$ the connected sum of two projective planes $P^2 \# P^2$. In case 2(b), we slide the foot of the handle $\mathcal{H}^2$ over the circle boundary of $D^2 \cup \mathcal{H}^1$ until the handles become disjoint. Thus case 2(b) reduces to case 2(a) and we get after joining a 2-handle to $D^2 \cup \mathcal{H}^1 \cup \mathcal{H}^2$ the connected sum $P^2 \# P^2$. The sliding operation means that we are changing the attaching map of the handle $\mathcal{H}^2$ by an isotopy of the boundary $\partial(D^2 \cup \mathcal{H}^1 \cup \mathcal{H}^2)$, and therefore, as we know, the operation will not change the diffeomorphism type of the manifold $M = D^2 \cup \mathcal{H}^1 \cup \mathcal{H}^2 \cup D^2$. Thus we either get the non-orientable surface $P^2 \# P^2$, or the orientable surface T^2.

For $k = 3$ at least one of the three 1-handles must be twisted. If all of them are disjoint, then we get the connected sum $P^2 \# P^2 \# P^2$. If all three or two of them are intersecting, then also we get $P^2 \# P^2 \# P^2$ by applying the sliding operation. If one of the three 1-handles is twisted, then the other two must be ordinary and intersecting. It follows that in this case the surface M is diffeomorphic to

$$P^2 \# T^2 \cong P^2 \# P^2 \# P^2.$$

We may now apply induction on k to finish the proof of the theorem. □

10.6. Removal of intersection points

For a handle presentation $W = V_0 \cup \mathcal{H}_r \cup \mathcal{H}_{r+1}$, where W is an n-manifold, the $(n-1)$-submanifold $V = \partial_+(V_0 \cup \mathcal{H}_r)$ contains the a-sphere of $\mathcal{H}_{r+1}$ and the b-sphere of $\mathcal{H}_r$. We denote them by M and M' respectively. Thus $M = S^r$ and $M' = S^{n-r-1}$. For simplicity, we write the dimension of M' as s. Then $s = n - r - 1$.

Let M and M' intersect transversely in points $p_1, \dots, p_k$ of V, and $M \cap M'$ contain no other points. We select two points p and q from $M \cap M'$, say $p = p_1$ and $q = p_2$. Let C (resp. C') be a smoothly embedded curve in M (resp. M') from p to q such that both C and C' do not intersect $M \cap M' - \{p, q\}$.

Lemma 10.6.1. (1) *There is a Riemannian metric on V with respect to which M and M' are totally geodesic submanifolds.*

(2) *There are coordinates neighbourhoods U_p about p and U_q about q in which the metric in* (1) *is Euclidean such that the portions of C and C' in U_p, and in U_q, are straight line segments.*

PROOF. We find a finite covering of $M \cup M'$ by coordinate neighbourhoods U_i in V with diffeomorphisms $\phi_i : U_i \longrightarrow \mathbb{R}^{r+s}$, $i = 1, \ldots, m$, and disjoint coordinate neighbourhoods $V_1, \ldots, V_k$ such that

(i) $p_i \in V_i \subset \overline{V}_i \subset U_i$ and $V_i \cap U_j = \emptyset$ for $1 \leq i \leq k$ and $k+1 \leq j \leq m$,

(ii) $\phi_i(U_i \cap M) \subset \mathbb{R}^r \times 0$ and $\phi_i(U_i \cap M') \subset 0 \times \mathbb{R}^s$ for $1 \leq i \leq k$,

(iii) $\phi_i(U_i \cap C)$ and $\phi_i(U_i \cap C')$ are straight line segments in $\mathbb{R}^{r+s}$ for $i = 1, 2$.

We construct a Riemannian metric $\langle\, , \rangle$ on the open set

$$U_0 = U_1 \cup \cdots \cup U_m$$

by gluing together the metrics on U_i induced from $\mathbb{R}^{r+s}$ by ϕ_i, by means of a partition of unity. Then, by (i), the metric is Euclidean on V_i for $i = 1, \ldots, k$. Using this metric, we construct open tubular neighbourhoods [1] $\mathcal{T}$ and $\mathcal{T}'$ of M and M' in U_0 respectively, which are thin enough so that $\mathcal{T} \cap \mathcal{T}' \subset V_1 \cup \cdots \cup V_k$, and

$$\phi_i(\mathcal{T} \cap \mathcal{T}' \cap V_i) \subset \mathrm{Int}\ D^r(\epsilon) \times \mathrm{Int}\ D^s(\epsilon') \subset \mathbb{R}^r \times \mathbb{R}^s = \mathbb{R}^{r+s},\ 1 \leq i \leq k,$$

for some $\epsilon,\ \epsilon' > 0$ depending on i.

Let $\alpha : \mathcal{T} \longrightarrow \mathcal{T}$ be the map which is antipodal map $(v \mapsto -v)$ on each fibre of $\mathcal{T}$. Then α is smooth, $\alpha^2 = \mathrm{Id}$, and M is the fixed point set of α. Define a new metric $\langle\, , \rangle_\alpha$ on $\mathcal{T}$ by

$$\langle v, w \rangle_\alpha = \frac{1}{2}[\langle v, w \rangle + \langle d\alpha(v), d\alpha(w) \rangle].$$

Then M is totally geodesic with respect to this new metric. This follows, because α is an isometry of $\mathcal{T}$ in the new metric, and $d\alpha$ fixes vectors in $\tau(M)$, and so if σ is a geodesic in $\mathcal{T}$ tangent to M at a point $x \in M$, then $\alpha(\sigma)$ is also a geodesic in $\mathcal{T}$ with the same tangent vector at $\alpha(x) = x$. By the uniqueness of geodesic, α is identity on σ, and therefore Image $\sigma \subset M$, showing that M is totally geodesic. Similarly, M' is totally geodesic with respect to a new metric $\langle\, , \rangle_{\alpha'}$ on $\mathcal{T}'$.

It follows from (ii), and the form of $\mathcal{T} \cap \mathcal{T}'$ that these two new metrics coincide with the old metric on $\mathcal{T} \cap \mathcal{T}'$. Therefore they fit together to define a metric on $\mathcal{T} \cup \mathcal{T}'$. Restricting this metric to $\overline{U}$, where U is an open set such that $M \cup M' \subset U \subset \overline{U} \subset \mathcal{T} \cup \mathcal{T}'$, and then extending it to the whole of V, we get a metric on V satisfying the required conditions. The extension follows from Lemma 5.3.3, considering a metric as a section of the vector bundle of bilinear forms (Example 5.2.4(5)). $\square$

Let B be a 2-submanifold of V containing $C \cup C'$ that is diffeomorphic to a 2-dimensional ball. For simplicity, we shall write in the following lemma C, C' in place of $B \cap C,\ B \cap C'$ respectively.

[1] An open tubular neighbourhood concerns a bundle whose fibre is an open disk.

Lemma 10.6.2. *Let V be given the Riemannian structure of Lemma 10.6.1. Suppose that the intersection number of M and M' at p and q are* $+1$ *and* -1 *respectively, and* $\dim M' = s \geq 3$*. Then there exist smooth vector fields $X_1, \ldots, X_{r-1}$, $Y_1, \ldots, Y_{s-1}$ on B such that*

(1) they are orthonormal, and orthogonal to B,

(2) $X_1, \ldots, X_{r-1}$ are tangent to M along C,

(3) $Y_1, \ldots, Y_{s-1}$ are tangent to M' along C'.

Proof. Let T and T' be unit tangent vector fields along C and C' respectively. Let N be the field of unit normal vectors along C' which are orthogonal to T' pointing inwards, and tangent to B. Then $N(p) = T(p)$ and $N(q) = -T(q)$.

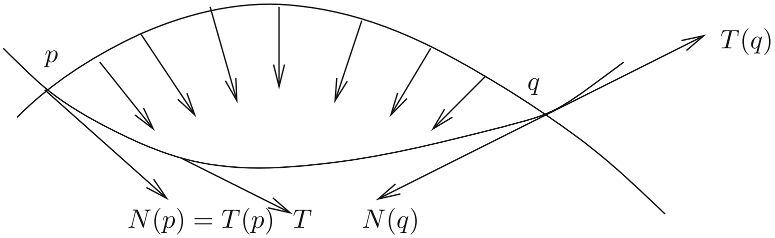


Figure 10.7

The construction of N is possible, since the normal bundle of M' over the contractible space C' is trivial.

We can find $r-1$ vectors $X_1(p), \ldots, X_{r-1}(p)$ that are tangent to M at p, and orthogonal to B such that the ordered set of vectors

$$(T(p), X_1(p), \ldots, X_{r-1}(p))$$

gives the positive orientation in $\tau(M)_p$. By parallel translation along C, these vectors give $r-1$ smooth vector fields $X_1, \ldots, X_{r-1}$ along C. These vector fields satisfy (1) because parallel translation preserves inner product, and they satisfy (2) because parallel translation along a curve in the totally geodesic submanifold M transports tangent vectors to M into tangent vectors to M (Theorem 4.6.4). Then, by continuity, the field of r-frames $(T, X_1, \ldots, X_{r-1})$ is positively oriented in $\tau(M)$ at every point of C.

Next, we apply parallel translation to the vectors $X_1(p), \ldots, X_{r-1}(p)$ along $U_p \cap C'$, and to the vectors $X_1(q), \ldots, X_{r-1}(q)$ along $U_q \cap C'$, where U_p and U_q are coordinate neighbourhoods on p and q provided by Lemma 10.6.1(2). By assumption about the intersection numbers,

$$(T(p), X_1(p), \ldots, X_{r-1}(p))$$

is positively oriented in the normal space $\nu(M')_p$, and

$$(T(q), X_1(q), \ldots, X_{r-1}(q))$$

is negatively oriented in $\nu(M')_q$. Since $N(p) = T(p)$ and $N(q) = -T(q)$, it follows that $N, X_1, \ldots, X_{r-1}$ is positively oriented in $\nu(M')$ at all points of the subspace $A = (U_p \cap C') \cup (U_q \cap C')$. Let us denote the frame field $(X_1, \ldots, X_{r-1})$ by ξ. Then ξ is a section over A of the bundle over C' of $(r-1)$-frames $(X_1, \ldots, X_{r-1})$ orthogonal to M' and to B such that $(N, X_1, \ldots, X_{r-1})$ is positively oriented in $\nu(M')$. The fibre of this bundle is $SO(r-1)$ which is connected, and C' is obtained from A by attaching a 1-cell. Therefore, by an extension theorem, we can extend ξ to a continuous, and hence smooth, field of $(r-1)$-frames on C', and hence on $C \cup C'$, satisfying (1) and (2). The extension theorem says that if (X, A) is a relative CW-complex with $\dim(X) = r$, and $E \longrightarrow X$ is a fibre bundle with $(r-1)$-connected fibre, then any section of $E|A$ extends to a section of E (Husemollar [**18**], Theorem 1(7.1), p. 21).

We repeat the above arguments for the bundle over B of orthonormal $(r-1)$-frames orthogonal to B with fibre the Stiefel manifold of $(r-1)$-frames in $\mathbb{R}^{r+s-2}$. The above frame field ξ, which we have constructed, constitutes a section of this bundle over $C \cup C'$. The pair $(B, C \cup C')$ is a relative CW-complex with $\dim B = 2$, and the Stiefel manifold $V_{r-1}(\mathbb{R}^{r+s-2})$ is simply connected, since $s \geq 3$ (see Husemoller [**18**], Theorem 5.1 in p. 83). Therefore we may extend $\xi = (X_1, \ldots, X_{r-1})$ to a smooth section over all of B satisfying (1) and (2),using the extension theorem again.

To construct the remaining fields over B, note that the bundle over B of orthonormal $(s-1)$ frames $Y_1, \ldots, Y_{s-1}$ in $\tau(V)$ such that each Y_i is orthogonal to B and to $X_1, \ldots, X_{r-1}$ is a trivial bundle, since B is contractible. Then we take the required field of frames $(Y_1, \ldots, Y_{s-1})$ on B to be a smooth section of this bundle. Then $X_1, \ldots, X_{r-1}, Y_1, \ldots, Y_{s-1}$ satisfy (1). Also $Y_1, \ldots, Y_{s-1}$ satisfy (3), because $X_1, \ldots, X_{r-1}$ are orthogonal to M' along C'. This completes the proof. $\square$

We now consider the following plane model of the above situation.

Let C_0 and C_0' be smoothly embedded open curves in $\mathbb{R}^2$ intersecting transversely in points a and b, and enclosing a disk D (with corners at a and b rounded off). We choose an embedding $\phi_0 : C_0 \cup C_0' \longrightarrow M \cup M'$ such that

$$\phi_0(C_0) = C, \quad \phi_0(C_0') = C', \quad \phi_0(a) = p, \quad \phi_0(b) = q.$$

See Figure 10.8 below.

Lemma 10.6.3. *Suppose that ϕ_0 extends to an embedding*

$$\phi_1 : B_0 \longrightarrow V,$$

where B_0 is some open ball neighbourhood of D in $\mathbb{R}^2$ (this we shall construct in course of our arguments in the next theorem (Theorem 10.6.5)). Then ϕ_1 extends further to an embedding $\phi : B_0 \times \mathbb{R}^{r-1} \times \mathbb{R}^{s-1} \longrightarrow V$ such that

$$\phi^{-1}(M) = (B_0 \cap C_0) \times \mathbb{R}^{r-1} \times 0, \qquad \phi^{-1}(M') = (B_0 \cap C_0') \times 0 \times \mathbb{R}^{s-1}.$$

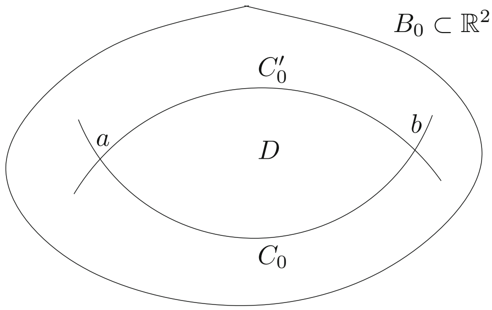


FIGURE 10.8

PROOF. Define a map $B_0 \times \mathbb{R}^{r-1} \times \mathbb{R}^{s-1} \longrightarrow V$ by

$$(b, x_1, \ldots, x_{r-1}, y_1, \ldots, y_{s-1}) \mapsto \exp\Big[\sum_{i=1}^{r-1} x_i X_i(\phi_1(b)) + \sum_{j=1}^{s-1} y_j Y_j(\phi_1(b))\Big],$$

where X_i, Y_j are the vector fields on $\phi_1(B_0)$, constructed in Lemma 10.6.2 (note that $\dim M' = s \geq 3$ by our assumption). This map is a local diffeomorphism. Therefore, by Lemma 6.1.3, there is an open ϵ-ball neighbourhood V_ϵ of the origin in $\mathbb{R}^{r-1} \times \mathbb{R}^{s-1} = \mathbb{R}^{r+s-2}$ such that the restriction of this map to $B_0 \times V_\epsilon$ is an embedding $\psi : B_0 \times V_\epsilon \longrightarrow V$. Then define an embedding

$$\phi : B_0 \times \mathbb{R}^{r-1} \times \mathbb{R}^{s-1} \longrightarrow V$$

by

$$\phi(b, z) = \psi\Big(b, \frac{\epsilon z}{\sqrt{1 + \|z\|^2}}\Big).$$

Then $\phi((B_0 \cap C_0) \times \mathbb{R}^{r-1} \times 0) \subset M$ and $\phi((B_0 \cap C_0') \times 0 \times \mathbb{R}^{s-1}) \subset M'$, because M and M' are totally geodesic submanifolds of V. Let us write $B = \phi(B_0 \times 0) = \phi_1(B_0)$. Then B intersects M and M' in C and C' transversely. Therefore, for sufficiently small $\epsilon > 0$, Image ϕ intersects M and M' precisely in the above product neighbourhoods of C and C'; explicitly,

$$\phi^{-1}(M) = B_0 \cap C_0 \times \mathbb{R}^{r-1} \times 0, \qquad \phi^{-1}(M') = B_0 \cap C_0' \times 0 \times \mathbb{R}^{s-1}.$$

This finishes the proof. $\square$

Lemma 10.6.4. *Suppose that P is a manifold of dimension ≥ 5, and Q is a submanifold of codimension ≥ 3. Suppose that P and Q are without boundary, and Q is a closed subset of P. Then any loop in $P - Q$ that is contractible in P is also contractible in $P - Q$.*

PROOF. Let $f : S^1 \longrightarrow P - Q \subset P$ be a loop so that $f(S^1)$ is contractible in P. Then f extends to a continuous map $D^2 \longrightarrow P$, which we still denote by f. (This is a basic fact of homotopy theory: a continuous map $f : S^1 \longrightarrow X$

extends to a continuous map $D^2 \longrightarrow X$ if and only if $f \simeq 0$ in X (see, for example, Spanier [**43**], p. 27).) We have then a map of the pairs

$$f : (D^2, S^1) \longrightarrow (P, P - Q).$$

Since $\dim(P) \geq 5$, f can be approximated by an embedding $g : D^2 \longrightarrow P$ such that $g \simeq f$, and $g|S^1 = f|S^1$ by Complement 6.1.8. Since $g(D^2)$ is contractible in P, the normal bundle of $g(D^2)$ in P is trivial. Therefore there is an embedding $h : D^2 \times \mathbb{R}^{r-2} \longrightarrow P$ such that $h(x, 0) = g(x)$, and so $h|(S^1 \times 0) \subset P - Q$. Choose an $\epsilon > 0$ so that $\|v\| \leq \epsilon$, $v \in \mathbb{R}^{r-2}$, implies $h(S^1 \times v) \subset P - Q$. The map $h|(D^2 \times v)$ is homotopic to a map $k : D^2 \times v \longrightarrow P$ that is transverse to Q (Theorem 6.2.11). Since codim $(Q) \geq 3$, the transversality means $k(D^2 \times v) \cap Q = \emptyset$, or $k(D^2 \times v) \subset P - Q$. This implies $k|(S^1 \times v) \simeq 0$. Thus in $P - Q$ we have

$$f|S^1 = g|S^1 = h|(S^1 \times 0) \simeq h|(S^1 \times v) \simeq k|(S^1 \times v) \simeq 0.$$

This completes the proof. $\square$

Theorem 10.6.5. *Recall our basic data : M and M' are compact oriented transversely intersecting submanifolds without boundary of dimensions r and s in an $(r + s)$-manifold V without boundary.*

Suppose that $r + s \geq 5$, $s \geq 3$, and, when $r = 1$ or $r = 2$, the inclusion map induces monomorphism of fundamental groups

$$\pi_1(V - M') \longrightarrow \pi_1(V).$$

Suppose further that an orientation has been chosen for M, and for the normal bundle $\nu(M')$ of M' in V.

Let p and q be points in $M \cap M'$ where the intersection numbers are $+1$ and -1 respectively. Let C and C' be a smoothly embedded curve in M and M' from p to q such that C and C' do not intersect $M \cap M' - \{p, q\}$, and such that the loop L formed by C followed by the reverse of C' is contractible in V.

Under these conditions there is an isotopy $h_t : V \longrightarrow V$ such that

(1) $h_0 = \mathrm{Id}$,

(2) $h_t = \mathrm{Id}$ in a neighbourhood of $M \cap M' - \{p, q\}$ for all $0 \leq t \leq 1$,

(3) $h_1(M) \cap M' = M \cap M' - \{p, q\}$.

Proof. We endow V with a Riemannian structure as given by Lemma 10.6.1. Orient the curves C and C' from p to q. Let $T(p)$, $T(q)$ and $T'(p)$, $T'(q)$ be the unit tangent vectors to the curves C and C' at points p and q respectively. Since C is contractible, the bundle $\nu(M)|C$ is trivial. Therefore, we can construct a unit vector field along C orthogonal to M and equal to the parallel translates of $T'(p)$, and of $-T'(q)$, along $U_p \cap C$ and $U_q \cap C$ respectively, where U_p and U_q are coordinate neighbourhoods about p and q. Thus we get a subbundle ξ of $\nu(M)|C$ with the unit 1-disk as fibre.

Consider the plane model of Figure 10.8 (p. 325), where we have two smooth curves C_0 and C_0' intersecting transversely in points a and b and enclosing a disk D. There we have chosen an embedding

$$\phi_0 : C_0 \cup C_0' \longrightarrow M \cup M'$$

mapping C_0, C_0' onto C, C', and a, b onto p, q respectively. Then the pull-back $\phi_0^*\xi$ is the normal 1-disk bundle of the embedding $\phi_0 : C_0 \longrightarrow C \subset V$. By Proposition 7.1.3 we get an embedding of a neighbourhood of C_0 in $\mathbb{R}^2$ into V which extends $\phi_0|C_0$. Similarly, $\phi_0|C_0'$ extends to an embedding of a neighbourhood of C_0', using a field of unit vectors along C' orthogonal to M' which along $U_p \cap C'$ and $U_q \cap C'$ consists of parallel translates of $T(p)$ and $-T(q)$ respectively. The constructions are also possible when $r = 1$, because the intersection numbers at p and q are opposite.

By Lemma 10.6.1(2), the two embeddings agree in a neighbourhood of $C_0 \cup C_0'$ and thus define an embedding $\phi_1' : K \longrightarrow V$, where K is a closed annular neighbourhood of ∂D such that

$$\phi_1'^{-1}(M) = K \cap C_0, \text{ and } \phi_1'^{-1}(M') = K \cap C_0'.$$

Let S denote the inner boundary of K, and $D_0 \subset D$ be the disk bounded by S in $\mathbb{R}^2$.

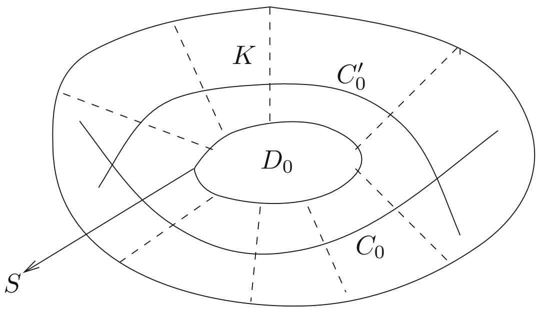


FIGURE 10.9

Since the given contractible loop L is homotopic to the loop $\phi_1'(S)$, $\phi_1'(S)$ is contractible in V, and hence in $V - (M \cup M')$. To see this, first note that $\phi_1'(S)$ contractible in V implies it is contractible in $V - M'$ for any value of r. The assertion is also true when $r = 1$ or 2, because we have by hypothesis

$$\pi_1(V - M') \longrightarrow \pi_1(V)$$

is injective. If $r \geq 3$ the assertion is true by Lemma 10.6.4. Therefore, since codim $M = s \geq 3$, and $\dim(V - M') \geq 5$ ($V - M'$ being open in V), $\phi_1'(S)$ is contractible in

$$(V - M') - M = V - (M \cup M'),$$

by Lemma 10.6.4 again.

We may therefore choose a continuous extension of ϕ_1' to $B_0 = K \cup D_0$

$$\phi_1^{''} : B_0 \longrightarrow, V$$

which maps Int D into $V - (M \cup M')$. Applying Theorems 2.2.3 and 8.9.5 to $\phi_1^{''}|(B_0 - \text{Int } D)$, we get a smooth embedding $\phi_1 : B_0 \longrightarrow V$ coinciding with ϕ_1' on a neighbourhood of $B_0 - \text{Int } D$, and such that $\phi_1(b) \notin M \cup M'$ if $b \notin C_0 \cup C_0'$.

Choose an isotopy $G_t : B_0 \longrightarrow B_0$ so that

(i) $G_0 = \text{Id}$,
(ii) $G_t = \text{Id}$ in a neighbourhood of the boundary $\overline{B}_0 - B_0$ of B_0, $0 \leq t \leq 1$,
(iii) $G_1(B_0 \cap C_0) \cap C_0' = \emptyset$,

(see Figure 10.10 below).

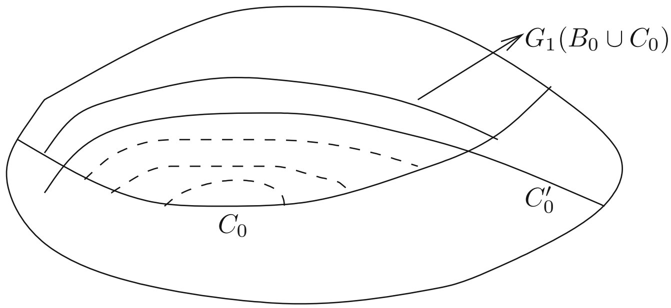


Figure 10.10

Let $\rho : \mathbb{R}^{r-1} \times \mathbb{R}^{s-1} \longrightarrow [0, 1]$ denote the function $\mathcal{B}(2 - \|x\|^2 - \|y\|^2)$, where $\mathcal{B}$ is a bump function. Then

$$\rho(x, y) = \begin{cases} 1, & \text{if} \quad \|x\|^2 + \|y\|^2 \leq 1 \\ 0, & \text{if} \quad \|x\|^2 + \|y\|^2 \geq 2. \end{cases}$$

Define an isotopy $H_t : B_0 \times \mathbb{R}^{r-1} \times \mathbb{R}^{s-1} \longrightarrow B_0 \times \mathbb{R}^{r-1} \times \mathbb{R}^{s-1}$ by

$$H_t(u, x, y) = \Big(G_{t\rho(x,y)}(u), x, y\Big).$$

Let $\phi : B_0 \times \mathbb{R}^{r-1} \times \mathbb{R}^{s-1} \longrightarrow V$ be the embedding of Lemma 10.6.3 which extends $\phi_1 : B_0 \longrightarrow V$. Then define an isotopy $h_t : V \longrightarrow V$ as follows

$$h_t(v) = \begin{cases} v, & \text{if} \quad v \in V - \text{Image } \phi, \\ \phi \circ H_t \circ \phi^{-1}(v), & \text{if} \quad v \in V. \end{cases}$$

Then $h_0 = \text{Id}$, $h_1(M) \cap M' = M \cap M' - \{p, q\}$, and $h_t = \text{Id}$ outside Image ϕ for $0 \leq t \leq 1$. This completes the proof. $\square$

Theorem 10.6.6. *Suppose that $W = V_0 \times I \cup \mathcal{H}_r \cup \mathcal{H}_{r+1}$, where $\dim W = n$ and $r \geq 2$, $r + 1 \leq n - 3$, and W, V_0 are simply connected. Suppose that an orientation has been chosen for the a-sphere S^r of $\mathcal{H}_{r+1}$, and for the normal bundle of the b-sphere S^{n-r-1} of $\mathcal{H}_r$. If the intersection number $I(S^r, S^{n-r-1}) = \pm 1$, then S^r and S^{n-r-1} intersect transversely in just one single point.*

PROOF. Let $M = S^r$ and $M' = S^{n-r-1}$. They are embedded in $V = \partial_+(V_0 \cup \mathcal{H}_r)$ whose dimension is $n-1 \geq 5$. We may suppose that they intersect transversely. If $M \cap M'$ is not a single point, then $I(M, M') = \pm 1$ implies that there exists a pair of points p', q' in $M \cap M'$ with opposite intersection numbers. Then, by the arguments of Theorem 10.6.5, we can adjust the things so that M and M' have two fewer intersection points. Repeating the process finitely many times we see that ultimately M intersects M' transversely in a single point.

To complete the proof, we must verify the following two conditions of Theorem 10.6.5 in the present context.

(1) The loop L is contractible in V.

(2) For $r = 2$, the homomorphism $\pi_1(V - M') \longrightarrow \pi_1(V)$ is injective.

(We need not worry about $r = 1$, since this case has been excluded by the theorem.)

First note that the conditions $r \geq 2$, $n - r > 3$, and W simply connected imply V is simply connected. This may be seen as follows. If $W_1 = V_0 \cup \mathcal{H}_r$, then, up to homotopy, W is W_1 plus an $(r+1)$-cell D^{r+1} attached. Therefore the pair (W, W_1) is r-connected so that $\pi_i(W, W_1) = 0$ for $i \leq r$, and hence the homomorphism induced by inclusion $\pi_1(W_1) \longrightarrow \pi_1(W)$ is an isomorphism since $r \geq 2$, by the homotopy exact sequence of the pair (W, W_1). Similarly, up to homotopy, W_1 is V with an $(n-r)$-cell D^{n-r} attached. So (W_1, V) is $(n-r-1)$-connected, and $\pi_1(V) \simeq \pi_1(W_1)$ since $n - r \geq 3$. Therefore V is simply connected.

Next note that $M - S$ and $M' - S$, where $S = M \cap M' - \{p, q\}$, are complete Riemannian manifolds. Then Hopf-Rinow theorem (Theorem 4.5.4(v)) ensures that there is a smoothly embedded curve from p to q in M, and a smoothly embedded curve from q to p in M', which do not intersect S. Then the loop L formed by these curves is contractible in V, since V is simply connected. This verifies (1).

Turning now to (2), note that if $r = 2$, then the b-sphere of $\mathcal{H}_2$ is $M' = S^{n-3}$, and its a-sphere is S^1, and

$$V = \partial_+(V_0 \cup \mathcal{H}_2) = [V_0 - \text{Int } (S^1 \times D^{n-2})] \cup (D^2 \times S^{n-3}).$$

It is easy to construct homotopy equivalences

$$V - S^{n-3} \simeq V_0 - \text{Int } (S^1 \times D^{n-2}) \simeq V_0 - S^1,$$

(see Figure 10.11 below). Therefore $\pi_1(V_0 - S^1) \simeq \pi_1(V - M')$.

Let $\mathcal{T} = S^1 \times D^{n-2}$ be a tubular neighbourhood of S^1 in V_0. Then

$$\mathcal{T} \cup (V_0 - S^1) = V_0, \ \mathcal{T} \cap (V_0 - S^1) = \mathcal{T} - S^1, \ \text{ and } \pi_1(\mathcal{T} - S^1) = \pi_1(\mathcal{T}) = \mathbb{Z}.$$

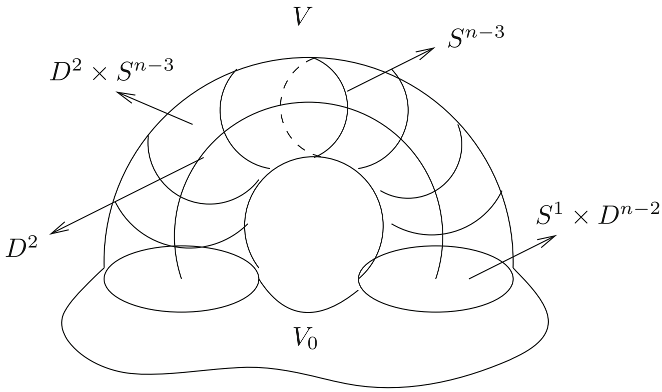


FIGURE 10.11

Also we have $\pi_1(V_0) = 1$, by hypothesis. Therefore, by Van Kampen's theorem ([**49**], p. 115)[2]

$$1 = \pi_1(V_0) \simeq \pi_1(\mathcal{T}) *_{\pi_1(\mathcal{T}-S^1)} \pi_1(V_0 - S^1) \simeq \pi_1(V_0 - S^1) \simeq \pi_1(V - M').$$

This verifies (2), and completes the proof of the theorem. □

10.7. Addition of handles

In this section we shall discuss addition of handles in terms of homotopy.

If $f : S^{r-1} \times D^{n-r} \longrightarrow \partial_+ W$ is the attaching embedding of a handle $\mathcal{H}_r$, then its a-sphere $f(S^{r-1} \times 0)$ does not represent a well-determined homotopy class, unless $\partial_+ W$ is simply connected. This ambiguity may be resolved by introducing as further structure a base point $x_0 \in \partial_+ W$, and for each r-handle with attaching map f a path ξ in $\partial_+ W$ from x_0 to $f(1 \times 0)$. If

$$\xi_* : \pi_{r-1}(\partial_+ W, f(1 \times 0)) \longrightarrow \pi_{r-1}(\partial_+ W, x_0)$$

is the isomorphism obtained by moving the base point along ξ, then $\alpha = \xi_*([f|(S^{r-1} \times 0)])$ is a well-defined homotopy class. We shall say that f represents $\alpha \in \pi_{r-1}(\partial_+ W)$ via the a-sphere $f(S^{r-1} \times 0)$.

Note that the isotopy class of the attaching embedding f determines α to within the action of $\pi_1(\partial_+ W)$ on $\pi_{r-1}(\partial_+ W)$. To see this, let

$$f_0, f_1 : S^{r-1} \times D^{n-r} \longrightarrow \partial_+ W$$

[2] Van Kampen's theorem says that if X is a topological space with pathwise connected subspaces X_0, X_1, X_2 and inclusion maps $j_i : X_0 \longrightarrow X_i$, $k_i : X_i \longrightarrow X$, $i = 1, 2$, such that $X = X_1 \cup X_2 = \text{Int } X_1 \cup \text{Int } X_2$, and $X_0 = X_1 \cap X_2$, then the homomorphisms k_{i*}, $i = 1, 2$ extend to a homomorphism of the free product $h : \pi_1(X_1) \star \pi_1(X_2) \longrightarrow \pi_1(X)$, with some base point in X_0, which is surjective, and $\ker h$ is the normal subgroup generated by the words of the form $j_{1*}(\alpha) j_{2*}(\alpha)^{-1}$ for $\alpha \in \pi_1(X_0)$. Thus $\pi_1(X)$ is isomorphic to $\pi_1(X_1) \star \pi_1(X_2)/\ker h$. This is called the amalgamated free product, and is denoted by $\pi_1(X_1) \star_{\pi_1(X_0)} \pi_1(X_2)$.

be isotopic attaching embeddings representing $\alpha_0, \alpha_1 \in \pi_{r-1}(\partial_+ W, x_0)$ respectively. Let f_t be an isotopy from f_0 to f_1, and ζ be the path in $\partial_+ W$ from $f_0(1 \times 0)$ to $f_1(1 \times 0)$ given by $\zeta(t) = f_t(1 \times 0)$. Then it can be seen easily that

$$\zeta_* : \pi_{r-1}(\partial_+ W, \zeta(1)) \longrightarrow \pi_{r-1}(\partial_+ W, \zeta(0)$$

maps $[f_1|(S^{r-1} \times 0)]$ on to $[f_0|(S^{r-1} \times 0)]$. Therefore there is an element $\sigma \in \pi_1(\partial_+ W, x_0)$ such that $\sigma_*(\alpha_1) = \alpha_0$. In fact, σ is represented by the loop $\xi_0 \cdot \zeta \cdot \xi_1^{-1}$.

Theorem 10.7.1 (Handle addition theorem). *Let W be an n-manifold with $\partial_+ W$ connected. Let*

$$W' = W \cup_f \mathcal{H}_r^1 \cup_g \mathcal{H}_r^2, \quad 1 < r < n-1,$$

where $f,\ g : \partial D^r \times D^{n-r} \longrightarrow \partial_+ W$ are embeddings with disjoint images representing α and $\beta \in \pi_{r-1}(\partial_+ W)$ respectively, via the respective a-spheres. Let $\sigma \in \pi_1(\partial_+ W)$. Then

$$W' \cong W \cup_f \mathcal{H}_r^1 \cup_{h_\pm} \mathcal{H}_r^2,$$

where $h_+, h_- : \partial D^r \times D^{n-r} \longrightarrow \partial_+ W$ are embeddings disjoint from f representing $\beta + \sigma_*(\alpha)$, $\beta - \sigma_*(\alpha)$ in $\pi_{r-1}(\partial_+ W)$ respectively.

Proof. We shall show that the sphere $g(S^{r-1} \times 0)$ is isotopic to a sphere which is $f(S^{r-1} \times 0)$ joined by a tube to $g(S^{r-1} \times 0)$.

Consider the point $1 \in \partial D^{n-r}$, and let D denote the r-disk $D^r \times 1$ in $\mathcal{H}_r^1$ whose boundary ∂D is the sphere $f(\partial D^r \times 1)$ in $\mathcal{H}_r^1$. Let

$$\psi : f(\partial D^r \times \partial D^{n-r}) \times I \longrightarrow \partial_+ W$$

be a collar neighbourhood of $f(\partial D^r \times \partial D^{n-r})$ in $\partial_+ W$, chosen so that Image ψ does not intersect the handle $\mathcal{H}_r^2$. Let

$$D' = D \cup \psi(f(\partial D^r \times 1) \times I) = D \cup \psi(\partial D \times I).$$

Then $\partial D'$ is diffeomorphic to $f(\partial D^r \times 1)$.

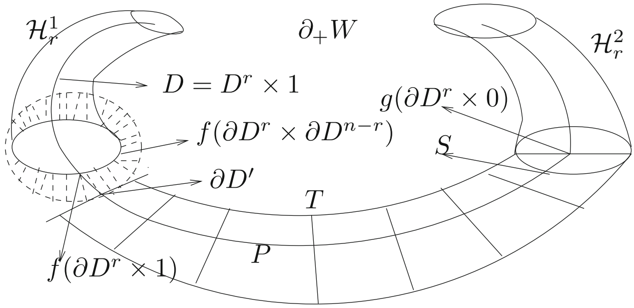


Figure 10.12

Let S denote the sphere $g(\partial D^r \times 1)$ in $\mathcal{H}_r^2$. Since $\partial_+ W$ is connected, there is a path P in $\partial_+ W$ from the point $a = f(1 \times 1) \in \partial D'$ to the point

$b = g(1 \times 1) \in S$ such that P hits $\partial D'$ and S in the normal directions of the spheres, and $P \cap \mathcal{H}_r^1 = \emptyset$. Note that we may take P in any homotopy class, and that the free choice of the homotopy class of P is equivalent to the free choice of the class $\sigma \in \pi_1(\partial_+ W)$. We may also suppose by the general position arguments that P is a smoothly embedded path with $(\mathrm{Int}\ P) \cap D' = \emptyset$, and $(\mathrm{Int}\ P) \cap S = \emptyset$, and that P lies near its end points a and b along the normal directions of the spheres $\partial D'$ and S. Choose a tubular neighbourhood T of P in $\partial_+ W$ with fibre the $(n-2)$-disk D^{n-2} so that $T \cap \partial D'$ and $T \cap S$ are $(r-1)$-disks (the boundary of each of which is a great $(r-2)$-sphere of the $(n-3)$-sphere in the fibre). Note that $r \leq n-2$, and the fibre T_a (resp. T_b) is tangent to $\partial D'$ (resp. S) so that $T_a \cap \tau(\partial D')_a = D^{n-2} \cap \mathbb{R}^{r-1} = D^{r-1}$ (resp. $T_b \cap \tau(S)_b = D^{r-1}$). We can therefore choose a normal $(n-2)$-framing for P so that the first $(r-1)$ of them give the standard orientation of S, and one of the two possible orientations of $\partial D'$. We can choose a Riemannian metric on $\partial_+ W$ in which $\partial D'$ and S are totally geodesic, that is, geodesics in $\partial_+ W$ for directions tangent to $\partial D'$ (resp. S) lie entirely in $\partial D'$ (resp. S). Then the restriction of the exponential map on the normal $(r-1)$-vectors to P gives an embedding

$$\phi : [-1, 0] \times D^{r-1} \longrightarrow \partial_+ W$$

with $\phi(-1 \times D^{r-1}) \subset \partial D'$ and $\phi(0 \times D^{r-1}) \subset S$

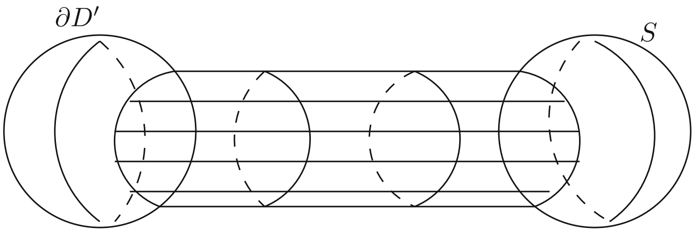


FIGURE 10.13

We extend P by a diameter of $D^r \times 1$ in $\partial(\partial_+ W \cup_f \mathcal{H}_r^1)$. Let

$$\phi' : [-1, 1] \times D^{r-1} \longrightarrow \partial(\partial_+ W \cup_f \mathcal{H}_r^1)$$

be the corresponding embedding. Define an isotopy

$$\lambda_t : \partial(\partial_+ W \cup_f \mathcal{H}_r^1) \longrightarrow \partial(\partial_+ W \cup_f \mathcal{H}_r^1)$$

by

$$\begin{array}{lll} \lambda_t(x) & = x & \text{if} \quad x \notin \phi'(0 \times D^{r-1}), \\ \lambda_t(\phi'(0, v)) & = \phi'(2t\mathcal{B}(1 - \|v\|), v) & \text{otherwise.} \end{array}$$

The definitions fit together smoothly by the properties of the bump function. The isotopy pulls the cell $\phi'(0 \times D^{r-1}) \subset g(S^{r-1} \times 1)$ across part of the disk $D^r \times 1$ covering the central point. Thus we have an isotopy of $g|(S^{r-1} \times 1)$, and hence an isotopy of $g|(S^{r-1} \times 0)$, since $g(S^{r-1} \times 0)$ is diffeotopic to $g(S^{r-1} \times 1)$.

Let the isotopy be from the embedding $f|S^{r-1}\times 0$ to an embedding $\overline{h}: S^{r-1}\times 0 \longrightarrow \partial_+ W$. Clearly, $\overline{h}$ represents the class β plus or minus $\sigma_*(\alpha)$, the sign depending on an orientation chosen earlier (where P is determined by σ).

By Proposition 7.3.5, $\overline{h}$ extends to an embedding $h: S^{r-1}\times D^{n-r}\longrightarrow \partial_+ W$ which is isotopic to g. By the isotopic extension theorem (7.3.3), there is a diffeomorphism ϕ of $V = W\cup_h \mathcal{H}_r^1$ such that $\phi\circ g = h$. Finally, by Corollary 7.6.8,

$$V\cup_g \mathcal{H}_r^2 \cong V\cup_h \mathcal{H}_r^2.$$

This completes the proof. □

Let (W, V_0, V_1) be a cobordism. Consider a handle presentation of W on V_0 consisting of k r-handles

$$W = V_0\times I\cup_{f_1}\mathcal{H}_r^1\cup_{f_2}\cup\cdots\cup_{f_k}\mathcal{H}_r^k.$$

We shall denote this presentation of W by the symbol

$$\sigma = (W, V_0, f_1, f_2, \ldots, f_k; r)$$

The group $H_r(W, V_0)$ is free on k generators $(g_1,\ldots,g_k)$, which are the oriented a-disks of the k r-handles. Define a homomorphism

$$F_\sigma : H_r(W,V_0)\longrightarrow \pi_{r-1}(V_0)$$

by $F_\sigma(g_i) = \phi_i$ where $\phi_i\in\pi_{r-1}(V_0)$ is the homotopy class of

$$\overline{f}_i : \partial D^r\times 0\longrightarrow V_0\times 1,$$

which is the restriction of $f_i : \partial D^r\times D^{n-r}\longrightarrow V_0\times 1$. Note that F_σ is well-defined by our convention of base point.

We say that the homomorphism F_σ is induced by the presentation σ. If $F: H_r(W,V_0)\longrightarrow \pi_{r-1}(V_0)$ is a homomorphism, then we say that W **realises** F if some presentation of W induces F. Note that manifolds realising a given homomorphism are not necessarily unique.

Theorem 10.7.2. *Let $\sigma = (W, V_0, f_1, f_2, \ldots, f_k; r)$ be a presentation of W inducing a homomorphism $F_\sigma : H_r(W,V_0)\longrightarrow \pi_{r-1}(V_0)$. Then, for any automorphism $\lambda : H_r(W,V_0)\longrightarrow H_r(W,V_0)$, W realises the homomorphism $F_\sigma\circ\lambda$.*

The proof will appear after the following two lemmas.

Lemma 10.7.3. *Let $f_1 : \partial D^r\times D^{n-r}\longrightarrow V_0$ and $\overline{f}_2 : \partial D^r\times 0\longrightarrow V_0$ be embeddings so that $\overline{f}_2$ is smoothly isotopic in V_0 to the restriction $\overline{f}_1$ of f_1 to $\partial D^r\times 0$. Then there is an embedding $f_2 : \partial D^r\times D^{n-r}\longrightarrow V_0$ extending $\overline{f}_2$ and a diffeomorphism $h: W\longrightarrow W$ such that $h\circ f_2 = f_1$.*

Proof. The proof follows from Proposition 7.3.5, and the isotopy extension theorem. □

Lemma 10.7.4. *We have a diffeomorphism*

$$(V_0 \times I) \cup_{f_1} \mathcal{H}_r \longrightarrow (V_0 \times I) \cup_{f_2} \mathcal{H}_r.$$

PROOF. The proof is immediate from Corollary 7.6.8. $\square$

Proof of Theorem 10.7.2. A standard results of group theory says that if G is a free abelian group on generators $g_1, \ldots, g_k$, then the group of automorphisms of G is generated by the following automorphism (called elementary automorphisms) ([**25**], Theorem N1, p. 163):

$R_i: \ g_1 \longrightarrow g_i, \ g_i \longrightarrow g_1, \ g_j \longrightarrow g_j$ for $j \neq 1, j \neq i, 2 \leq i \leq k$,

$S: \ g_1 \longrightarrow g_1^{-1}, \ g_i \longrightarrow g_i$ for $i > 1$,

$T: \ g_1 \longrightarrow g_1 g_2, \ g_i \longrightarrow g_i$ for $i > 1$.

Therefore it is sufficient to check the theorem for elementary automorphisms. The case $\lambda = R_i$ follows immediately from Definition 10.1.2(1). For the case $\lambda = S$, define a diffeomorphism $h : D^r \times D^{n-r} \longrightarrow D^r \times D^{n-r}$ by $h(x, y) = (\rho(x), y)$, where $\rho : D^r \longrightarrow D^r$ is a reflection through an equatorial $(r-1)$-plane. Let $f_i' = f_i \circ h$, and $\sigma' = (W, V_0, f_1', f_2, \ldots, f_k; r)$. Then the presentation

$$W = V_0 \times I \cup_{f_1'} \mathcal{H}_r^1 \cup_{f_2} \cup \cdots \cup_{f_k} \mathcal{H}_r^k$$

realises $F_{\sigma'} = F_\sigma \circ S$, and it is diffeomorphic to W, by (10.1.2)(2). Finally, the case $\lambda = T$ follows from Theorem 10.7.1. This completes the proof.

10.8. Simplification of handle presentations

Theorem 10.8.1. *Let* $W = M \times I \cup \mathcal{H}_r \cup k\mathcal{H}_{r+1}$*, where* $M = \partial_- W$*,* $\dim W = n$*,* $n \geq 2r+3$*. Then*

$$W \cong M \times I \cup k\mathcal{H}_{r+1} \cup \mathcal{H}_{r+2}.$$

PROOF. The maps that will appear in the proof will be assumed to be embeddings, after close enough approximations. Also, we shall write down the proof without mentioning the attaching embeddings of the handles. Thus the a-sphere of the handle $\mathcal{H}_{r+1} = D^{r+1} \times D^{n-r-1}$ will be denoted simply by $S^r \times 0$. We may suppose that this a-sphere intersects the disk $D^r \times 1 \subset \mathcal{H}_r$ transversely. Since $n > 2r$, this means $S^r \times 0$ does not intersect $D^r \times 1$.

The disk $D^r \times 1$ determines an elements of $\pi_r(W, M)$, which is zero. Therefore it is homotopic in W (rel its boundary) to a disk in M. This homotopy is a map $F : D^{r+1} \longrightarrow W$ which maps the upper hemisphere of S^r into $D^r \times 1$, and the lower hemisphere into M. Again, since $n \geq 2r+3$, we may suppose that $F(D^{r+1})$ is disjoint from the cores $D^r \times 0$ and $D^{r+1} \times 0$ of $\mathcal{H}_r$ and $\mathcal{H}_{r+1}$. Therefore we can deform F so that its image is disjoint from the tubular neighbourhoods of these cores, and lies in M. We may suppose that $F|S^r : S^r \longrightarrow M$ is an embedding. This is null homotopic, and so homotopic to a constant map. Take a point $P \in S^r$, and a constant map f of S^r onto a point q belonging to the b-sphere $0 \times S^{n-r-2}$ of $\mathcal{H}_{r+1}$. Let N_q be the normal space to $0 \times S^{n-r-2}$ at q. Then f can be approximated by a smooth map $g : S^r \times 0 \longrightarrow N_q \subset M$ such that $g(p) = q$, and g is homotopic to f (rel p) (Theorem 6.1.5). We may assume that g is an embedding if the approximation is small. Since $g(S^r)$ intersects the b-sphere at a single point, we can take S^r as the a-sphere of the first member of a pair of complementary handles $(\mathcal{H}^1_{r+1}, \mathcal{H}^2_{r+2})$, where $\mathcal{H}^1_{r+1}$ is disjoint from $\mathcal{H}_{r+1}$. Thus we have two pairs of complementary handles $(\mathcal{H}_r, \mathcal{H}^1_{r+1})$ and $(\mathcal{H}^1_{r+1}, \mathcal{H}^2_{r+2})$. Therefore

$$\begin{aligned} W &\cong M \times I \cup \mathcal{H}_r \cup k\mathcal{H}_{r+1} \cup (\mathcal{H}^1_{r+1} \cup \mathcal{H}^2_{r+2}) \\ &\cong M \times I \cup (\mathcal{H}_r \cup \mathcal{H}^1_{r+1}) \cup k\mathcal{H}_{r+1} \cup \mathcal{H}^2_{r+2} \\ &\cong M \times I \cup k\mathcal{H}_{r+1} \cup \mathcal{H}^2_{r+2}, \end{aligned}$$

where in the first line the complementary pair $(\mathcal{H}^1_{r+1}, \mathcal{H}^2_{r+2})$ has been introduced (10.4.6), in the second line the $(r+1)$-handles have been rearranged (10.4.1), and in the last line a pair of complementary handles has been cancelled (10.4.4). This completes the proof. $\square$

We may replace the trivial cobordism $M \times I$ in the above theorem by a non-trivial cobordism M with $\partial_- M = M_0$ and $\partial_+ M = M_1$, provided an extra condition is introduced.

Corollary 10.8.2. *Let M be a cobordism (M, M_0, M_1) so that the inclusion $M_1 \hookrightarrow M$ induces isomorphism $\pi_1(M_1) \cong \pi_1(M)$, and*

$$W = M \cup \mathcal{H}_r \cup k\mathcal{H}_{r+1},$$

where $\dim M = n$*, $n \geq 2r+3$, and $\pi_r(W, M) = 0$. Then*

$$W \cong M \cup k\mathcal{H}_{r+1} \cup \mathcal{H}_{r+2}.$$

Proof. First note that M is diffeomorphic to $M \cup M_1 \times I$ so that

$$W = M \cup M_1 \times I \cup \mathcal{H}_r \cup k\mathcal{H}_{r+1}.$$

Let $W_1 = M_1 \times I \cup \mathcal{H}_r \cup k\mathcal{H}_{r+1}$.

The proof will follow from the last theorem if we show that

$$\pi_r(W_1, M_1) = 0.$$

This is true for $r = 1$. Because, by Van Kampen's Theorem,

$$\pi_1(W) \cong \pi_1(M) *_{\pi_1(M_1)} \pi_1(W_1) \cong \pi_1(W_1),$$

since $\pi_1(M_1) \cong \pi_1(M)$, and therefore $\pi_1(W_1, M_1) \cong \pi_1(W, M) = 0$, by the five lemma applied to the map between homotopy exact sequences (note that all the spaces are 0-connected).

Next, suppose that $r \geq 2$, and M, and hence M_1, is 1-connected. Since there is no handle, and hence no cell, of dim $\leq r-1$, the CW-complex pair (W_1, M_1) is $(r-1)$-connected (i.e. $\pi_q(W_1, M_1) = 0$ for $0 \leq q \leq r-1$). Therefore by Relative Hurewicz Isomorphism Theorem (Spanier [**43**], Theorem 4, p. 397),

$\pi_r(W_1, M_1) \cong H_r(W_1, M_1)$ (because M_1 is 1-connected and (W_1, M_1) $(r-1)$-connected, $r \geq 2$). Similarly, we have $\pi_r(W, M) \cong H_r(W, M)$. Now, by homology excision theorem, $H_r(W_1, M_1) \cong H_r(W, M)$[3]. Therefore

$$\pi_r(W, M) = 0 \text{ implies } \pi_r(W_1, M_1) = 0.$$

If $r \geq 2$, but M and M_1 are not 1-connected, then we will have to pass to the universal covering spaces. Let $\widetilde{W}$ be the universal covering space of W with covering map $p : \widetilde{W} \longrightarrow W$. Then

$$\widetilde{W_1} = p^{-1}(W_1),\ \ \widetilde{M_1} = p^{-1}(M_1), \ \text{ and } \widetilde{M} = p^{-1}(M)$$

are the universal covering spaces of W_1, M_1, and M respectively, because the fundamental groups are isomorphic (note that the pair (W, M) is at least 2-connected, since $\pi_r(W, M) = 0$ so that the inclusion $M \hookrightarrow W$ induces isomorphism

$$\pi_1(M) \longrightarrow \pi_1(W)).$$

Then, by the theory of covering spaces, we have isomorphisms

$$p_* : \pi_r(\widetilde{W_1}, \widetilde{M_1}) \cong \pi_r(W_1, M_1) \text{ and } p_* : \pi_r(\widetilde{W}, \widetilde{M}) \cong \pi_r(W, M),$$

and therefore by Hurewicz theorem

$$\pi_r(W_1, M_1) \cong H_r(\widetilde{W_1}, \widetilde{M_1}) \text{ and } \pi_r(W, M) \cong H_r(\widetilde{W}, \widetilde{M}).$$

We may now use excision theorem as before, and get the result in this case also. □

Corollary 10.8.3. *If* $W = M \cup \ell\mathcal{H}_r \cup k\mathcal{H}_{r+1}$ *with* $\dim W = n$, $n \geq 2r+3$, $\pi_r(W, M) = 0$, *and* $\pi_1(\partial_+ M) \cong \pi_1(M)$, *then*

$$W \cong M \cup k\mathcal{H}_{r+1} \cup \ell\mathcal{H}_{r+2}.$$

Proof. Write $W = W_1 \cup \mathcal{H}_r \cup k\mathcal{H}_{r+1}$, where

$$W_1 = M \cup (\ell-1)\mathcal{H}_r \cong M \cup M_1 \times I \cup (\ell-1)\mathcal{H}_r.$$

Let $W_2 = \overline{W_1 - M}$.

Then $W_2 = M_1 \times I \cup (\ell - 1)\mathcal{H}_r$ and $\partial_+ W_2 = \partial_+ W_1$. This has a dual presentation

$$W_2 = \partial_+ W_2 \times I \cup (\ell-1)\mathcal{H}_{n-r}.$$

Since $n - r \geq 3$, the pair $(W_2, \partial_+ W_2)$ is 2-connected, and so $\pi_1(W_2) \cong \pi_1(\partial_+ W_2) = \pi_1(\partial_+ W_1)$. Then, by Van Kampen's Theorem applied to the original presentation $W_1 = M \cup M_1 \times I \cup (\ell-1)\mathcal{H}_r$, we have

$$\pi_1(W_1) \cong \pi_1(M) *_{\pi_1(M_1)} \pi_1(W_2) \cong \pi_1(W_2) \cong \pi_1(\partial_+ W_1).$$

[3]To get this result, we have to eliminate the left-hand boundary M_0 by attaching an open collar $M_0 \times [0,1)$ getting $W' = W \cup M_0 \times [0,1)$ and $M' = M \cup M_0 \times [0,1)$ which have the homotopy type of W and M respectively, and then excise the open set Int M' from W' and M'.

Again, by the homotopy exact sequence of the triple $M \subset W_1 \subset W$

$$\longrightarrow \pi_r(W, M) \longrightarrow \pi_r(W, W_1) \longrightarrow \pi_{r-1}(W_1, M) \longrightarrow \pi_{r-1}(W, M) \longrightarrow,$$

$\pi_r(W, M) = 0$ and $\pi_{r-1}(W_1, M) = 0$ imply $\pi_r(W, W_1) = 0$. We are therefore in a position to apply Corollary 10.8.2 to the presentation $W = W_1 \cup \mathcal{H}_r \cup k\mathcal{H}_{r+1}$ to conclude that $W \cong W_1 \cup k\mathcal{H}_{r+1} \cup \mathcal{H}_{r+2}$. The proof may now be completed by induction on ℓ. □

Corollary 10.8.4. *If* (W, M) *is* r*-connected, where* $\dim W = n$, $n \geq 2r+3$, *and*

$$\pi_1(\partial_+ M) \cong \pi_1(M),$$

then W *has a handle presentation on* M *with no* i*-handle for* $0 \leq i \leq r$.

Proof. We use Corollary 10.8.3 repeatedly to replace i-handles by $(i+2)$-handles for $i = 0, 1, \ldots, r$. □

Theorem 10.8.5 (Smale). *Let* $W = M \cup \mathcal{H}_r \cup k\mathcal{H}_{r+1}$, *where* $\dim W = n$, $r > 1$ *and* $n - r \geq 4$. *Let* W *and* M *be simply connected, and* $H_r(W, M) = 0$. *Then*

$$W \cong M \cup (k-1)\mathcal{H}_{r+1}.$$

Proof. Let $n_1, \ldots, n_k$ denote the intersection numbers of the a-spheres of $\mathcal{H}_{r+1}$ with the b-sphere of $\mathcal{H}_r$. As we can add and subtract the $(r+1)$-handles by Theorem 10.7.2, we may replace n_i by $n_i \pm n_j$ if n_i and n_j are non-zero. Therefore, by induction on $|n_1| + \cdots + |n_k|$, we can replace each of the n_i by zero, except for one, say n_1. As there is only one r-handle and no $(r-1)$-handle, the hypothesis $H_r(W, M) = 0$ implies that the boundary homomorphism $\partial_{r+1} : C_{r+1}(W, M) \longrightarrow \mathbb{Z}$ is surjective, and so the incidence number n_1 must be $+1$ or -1.

Now we normalise the handles in the presentation $W = M \cup \mathcal{H}_r \cup k\mathcal{H}_{r+1}$. The a-sphere of one of the $\mathcal{H}_{r+1}$ and the b-sphere of $\mathcal{H}_r$ meet transversely and have intersection number ± 1 in $\partial_+(M \cup \mathcal{H}_r)$. Then, since W and M are simply connected, we may apply Theorem 10.6.6 to reduce the number of intersections to one. But then $\mathcal{H}_r$ and $\mathcal{H}_{r+1}$ are complementary handles, so can be cancelled by Theorem 10.4.4. □

Remark 10.8.6. The theorem is also true without the simple-connectivity condition. Its proof is more technical, and uses the simple homotopy theory.

10.9. h-cobordism theorem and generalised Poincaré conjecture

Definition 10.9.1. Let W be a cobordism with $\partial_- W = V_0$ and $\partial_+ W = V_1$. Then W is called an **h-cobordism** if the inclusions of V_0 and V_1 in W are homotopy equivalences. In this case V_0 is said to be **h-cobordant** to V_1.

Theorem 10.9.2 (h-cobordism theorem). *If* W *is a simply connected* h*-cobordism with* $\dim W = n \geq 6$, *then* W *is diffeomorphic to* $V_0 \times I$.

PROOF. The inclusion $V_0 \hookrightarrow W$ is a homotopy equivalence implies that

$$H_i(W, V_0) = 0 \text{ and } \pi_i(W, V_0) = 0,$$

for all i, by the exact homology and homotopy sequences of the pair (W, V_0). In the same way, $H_i(W, V_1) = \pi_i(W, V_1) = 0$ for all i. Then, by Corollary 10.8.4 for $r = 1$, W has a handle presentation on V_1 with no 0- and 1-handles. Thus the dual presentation on V_0 has no n- and $(n-1)$-handles. Moreover, we can cancel from this presentation on V_0 the r-handles for $2 \leq r \leq n-4$ using Theorem 10.8.5, since W and V_0 are simply connected. Thus we are left only with the cases $r = n-3$ and $n-2$, or with the cases $r = 2$ and 3 for the presentation on V_1. Since W and V_1 are simply connected, we can remove the 2-handles by using Theorem 10.8.5 for $r = 2$. Now there are as many 3-handles as there are 2-handles, since the chain complex of the pair (W, V_1) consists of a single isomorphism $\partial : C_3(W, V_1) \longrightarrow C_2(W, V_1)$. Thus there are no 3-handles either. Since all the handles may be cancelled, W is diffeomorphic to $V_0 \times I$, by Proposition 10.1.7. $\square$

Corollary 10.9.3. *If two simply connected boundaryless manifolds of dimension ≥ 5 are h-cobordant, then they are diffeomorphic.*

We now present another form of the h-cobordism theorem.

Theorem 10.9.4. *If for a cobordism (W, V_0, V_1) with $\dim W \geq 6$, the manifolds W, V_0, and V_1 are simply connected, and $H_*(W, V_0) = 0$, then W is diffeomorphic to $V_0 \times [0, 1]$.*

PROOF. In fact, the given conditions imply the conditions of the h-cobordism theorem. To see this, note that $\pi_1(V_0) = 0$, $\pi_1(W, V_0) = 0$, and $H_*(W, V_0) = 0$ imply $\pi_i(W, V_0) = 0$ for all i , by relative Hurewicz theorem. Therefore $\pi_i(V_0) \simeq \pi_i(W)$ for all i, by the homotopy exact sequence. Then the inclusion map $V_0 \hookrightarrow W$ is a homotopy equivalence, by Whitehead theorem (see Spanier [**43**], Corollary 24, p. 405). Similarly the other inclusion map $V_1 \hookrightarrow W$ is a homotopy equivalence (note that $H_*(W, V_0) = 0$ implies $H^*(W, V_0) = 0$, by the universal coefficient theorem for cohomology, and $H^*(W, V_0) = 0$ implies $H_*(W, V_1) = 0$, by the duality theorem (10.2.8)). Thus W is an h-cobordism. $\square$

We now consider the characterisation of n-disk D^n for $n \geq 6$.

Theorem 10.9.5. *If W is a compact simply connected n-manifold with a simply connected boundary, and with $n \geq 6$, then W is diffeomorphic to D^n if and only if W has the integral homology of a point.*

PROOF. We prove the 'if' part. The only if' part is obvious. So suppose that W has the homology of a point. Let D be a submanifold in Int W which is diffeomorphic to an n-disk, and $N = W - \text{Int } D$. Then the cobordism

$(N, \partial D, \partial W)$ satisfies the conditions of Theorem 10.9.4. Indeed, we have, by Van Kampen's theorem,

$$\pi_1(W) \simeq \pi_1(N) *_{\pi_1(\partial D)} \pi_1(D) \simeq \pi_1(N),$$

and therefore N is simply connected; also by the excision theorem and exactness theorem of homology theory,

$$H_*(W - \text{Int } D, \partial D) \simeq H_*(W, D) = 0.$$

Therefore $(N, \partial D, \partial W)$ is a product cobordism. Since $(W^n, \emptyset, \partial W)$ is a composition of the cobordisms $(D, \emptyset, \partial D)$ and $(W - \text{Int } D, \partial D, \partial W)$, it follows that W is diffeomorphic to D (see Exercise 10.1 in p. 303). □

We may now prove the generalised Poincaré conjecture in dimensions ≥ 6.

Theorem 10.9.6 (Smale). *If M is a closed simply connected n-manifold with $n \geq 6$ and M has the integral homology of the n-sphere S^n, then M is homeomorphic to S^n.*

Proof. Let $D \subset M$ be a closed n-disk, and $N = \overline{M - \text{Int } D}$. Then for the cobordism $(N, \emptyset, \partial D)$, N is simply connected by Van Kampen's theorem, and it has the integral homology of a point, because by Poincaré duality, excision, and exactness

$$H_i(N, \emptyset) \simeq H^{n-i}(N, \partial D) \simeq H^{n-i}(M, D) \simeq \begin{cases} \mathbb{Z} & \text{if } i = 0, \\ 0 & \text{if } i > 0. \end{cases}$$

Therefore, by Theorem 10.9.5, N is diffeomorphic to D, and so M is diffeomorphic to a union of two copies D_1^n and D_2^n of an n-disk with the boundaries glued together by a diffeomorphism $\lambda : S_1^{n-1} \longrightarrow S_2^{n-1}$, that is, $M \simeq D_1^n \cup_\lambda D_2^n$.

Construct a homeomorphism $\mu : D_1^n \cup_\lambda D_2^n \longrightarrow S^n$ in the following way. Let D_-^n and D_+^n be the lower and the upper hemisphere of S^n corresponding to the north pole $e_{n+1} = (0, \ldots, 0, 1)$. Take a diffeomorphism $\mu_1 : D_1^n \longrightarrow D_-^n$ which maps the boundary onto the boundary. Then define $\mu_2 : D_2^n \longrightarrow D_+^n$ by

$$\mu_2(tv) = \sin\left(\frac{\pi t}{2}\right) \cdot \mu_1(\lambda^{-1}(v)) + \cos\left(\frac{\pi t}{2}\right) \cdot e_{n+1},$$

where tv denotes a point of D_2^n, $v \in S_2^{n-1}$, $0 \leq t \leq 1$. These give a map $\mu : D_1^n \cup_\lambda D_2^n \longrightarrow S^n$, which is well-defined, since $\mu_2(S_2^{n-1}) = \mu_1(S_1^{n-1})$. Clearly μ is a continuous bijection, and hence a homeomorphism. This completes the proof. □

Corollary 10.9.7. *If $n \geq 6$, and M is a closed n-manifold having the homotopy type of the n-sphere S^n, then M is homeomorphic to S^n.*

Theorem 10.9.8. *If M is an n-manifold as in Theorem 10.9.6, where $n = 5$ or 6, then M is diffeomorphic to S^n.*

Proof. The proof follows immediately from a result of Kervaire and Milnor [**20**], and of also Wall [**51**] which states that if $n = 4, 5$, or 6, and M is a closed simply connected smooth n-manifold having the homology of S^n, then M is the boundary of a smooth compact contractible manifold W. Then for $n = 5$ or 6, $\dim W \geq 6$, and W has the integral homology of a point. Therefore by Theorem 10.9.5, W is diffeomorphic to D^n, and hence M is diffeomorphic to S^n. $\square$

Bibliography

[1] E.Artin, *Geometric algebra*, Interscience, New York, 1957.
[2] R.L. Bishop and R.J. Crittenden, *Geometry of Manifolds*, Academic Press, New York, 1964.
[3] G.E. Bredon, *Topology and geometry*, Springer-Verlag, 1993.
[4] F. Brickell and R.S. Clark, *Differentiable Manifolds : An Introduction*, Van Nostrand, 1970
[5] C. Chevalley, *Theory of Lie Groups*, Princeton Univ. Press, Princeton, 1946.
[6] J. Dieudonné, *Foundations of modern analysis*, Academic Press, New York and London, 1969.
[7] A. Dold, *Lectures on algebraic topology*, Springer-Verlag, Berlin Heidelberg New York, 1972.
[8] A. du Plessis, Homotopy classification of regular sections, *Compositio Math.* **32**(1976), 301-333.
[9] J. Eells and K.D. Elworthy, On the differential topology of Hilbert manifolds in "Global Analysis" (S.S.Chern and S. Smale eds.), Part II (*Proc. Symp. Pure Mathematics*, Vol. 15). Amer. Math. Soc., Providence, Rhode Island, 1970.
[10] R. Godement, *Topologie Algèbrique et Théorie des Faisceaux*, Hermann,1958.
[11] M. Golubitsky and V. Guillemin, *Stable Mappings and their Singularities*, Second Edition, Springer-Verlag, Berlin Heidelberg New York, 1980.
[12] V. Guillemin and V.A. Pollack, *Differential Topology*, Prentice-Hall, Englewook Cliffs, NJ, 1974.
[13] V. Guillemin and S. Sternberg, *Symplectic Techniques in Physics*, Cambridge University Press, Cambridge, 1984.
[14] M.W. Hirsch, *Differential Topology*, Springer-Verlag Berlin Heidelberg New York, 1976.
[15] J.G. Hocking and G.S.Young, *Topology*, Addison-Wesley Publ. Co. Inc., 1961.
[16] L. Hörmander, *An Introduction to Complex Analysis in Several Variables*, North-Holland Publishing Co., 1973.
[17] W. Hurewicz, *Lectures on Ordinary Differential Equations*, The MIT Press, Cambridge, Mass., 1958.
[18] D. Husemoller, *Fibre Bundles*, McGraw-Hill, New York, 1975.
[19] J.L. Kelley, *General Topology*, Nostrand; Springer-Verlag, New york, 1975.
[20] M. Kervaire and J. Milnor, Groups of homotopy spheres 1, *Ann. of Math.* **77**(1963), 504-537.
[21] H. Kneser, Analytische Struktur und Abzählbarkeit *Ann. Acad. Sci. Fenn.* Ser. A. I., **251/5**(1958).
[22] N. Kuiper, The homotopy-type of the unitary group of Hilbert space, *Topology* **3**(1965), 19-30.
[23] S. Lang, *Introduction to Differentiable Manifolds*, Wiley Interscience, New York, 1962.
[24] S. Lang, *Algebra*, Addison-Wesley Publ. Co., 1969.
[25] W. Magnus, A. Karrass, and D. Solitar, *Combinatorial group theory*, Interscience Publishers, New York, London, Sydney, 1966.

A. Mukherjee, *Differential Topology*, DOI 10.1007/978-3-319-19045-7

[26] B. Mazur, Differential topology from the point of view of simple homotopy theory, *Publ. Math. Institut des hautes études scientifiques*, No. 15, 1963.

[27] D. McDuff and Lafontaine *Introduction to Symplectic Topology*, Clarendon Press, Oxford, 1995.

[28] J.Milnor, On manifolds homeomorphic to the 7-sphere, *Ann. of Math.* **64**(1956), 399-405.

[29] J. Milnor, *Morse Theory*, Annals Study 51, Princeton Univ. Press, Princeton, 1963.

[30] J. Milnor, *Lectures on the h-cobordism theorem*, Princeton Mathematical Series, Princeton Univ. Press, Princeton, New Jersey, 1965.

[31] J. Milnor and D.W. Weaver, *Topology from the Differentiable Viewpoint*, Univ. of Virginia Press, Charlottesville 1969.

[32] J. Milnor and J. Stascheff, *Characteristic classes*, Princeton Math Series, Princeton Univ. Press, 1974.

[33] E.E.Moise, Affine structure in 3-manifolds I, II, III, IV, V, *Ann.of Math.* **54**(1951), 506-533; **55**(1952), 172-176; **55**(1952), 203-214; **55**(1952), 215-222; **56**(1952), 96-114.

[34] M. Morse, Relations between the Critical Points of a Real Analytic Function of n Independent Variables, *Trans. Amer. Math. Soc.*, **27**(1925), 345-396.

[35] J.R. Munkres, Obstructions to the smoothing of piecewise differentiable homeomorphisms, *Ann.of Math.* **72**(1960), 521-554.

[36] J.R. Munkres, *Elementary differential topology*, Princeton Univ. Press, Princeton, 1963.

[37] R.S. Palais, Morse theory on Hilbert manifolds, *Topology* 2 (1963), 299-340.

[38] R.S. Palais and C-L. Terng, *Critical Points Theory and Submanifold Geometry*, Lecture Notes in Math. No. 1353, Springer-Verlag, 1988.

[39] T, Rado, Über den Begriff der Riemannanschen Fläche, *Acta Litt. Sci. Univ. Szegd.* **2** (1925), 101-121.

[40] J.G. Ratcliffe, *Foundations of hyperbolic manifolds*, Springer-Verlag, 1994.

[41] S. Smale, Generalized Poincaré conjecture in dimensions greater than four, *Ann. of Math.* **74**(1961), 391-406.

[42] S. Smale, On the structure of manifolds, *Amer. J. Math.* **84** (1962), 387-399.

[43] E. Spanier, *Algebraic Topology*, McGraw-Hill, New York, 1966.

[44] M. Spivak, *A comprehensive introduction to differential geometry*, Publish or Perish, Inc., Boston, Mass.

[45] J. Stallings, The piecewise-linear structure of Euclidean space, *Proc. Camb. Philos. Soc.* **58**(1962), 481-488.

[46] N. Steenrod, *The topology of fibre bundles*, Princeton Univ. Press, Princeton, 1951.

[47] S. Sternberg, *Lectures on differential geometry*, 2nd Ed. Chelsea, New York 1983.

[48] R. Thom, Quelques propriétés globales des variétés différentiables, *Comment. Math. Helv.* **29**(1954), 17-85.

[49] J.W. Vick, *Homology Theory: An Introduction to Algebraic Topology*, Springer-Verlag, 1973

[50] C.T.C. Wall,*Lectures on Differentiable Manifolds*, unpublished lecture notes, Liverpool Univ.

[51] C.T.C. Wall, Killing the middle homotopy groups of odd dimensional manifolds, *Trans. Amer. Math. Soc.* **103**(1962), 421-433. .

[52] A. Wallace, Modifications and cobounding manifolds, *Canadian J. Math.* **12**(1960), 503-528.

[53] A. Weinstein, *Lectures on symplectic manifolds*, CBMS Conference Series, 29, Amer. Math. Soc., Providence, Rhode Island.

[54] G.W. Whitehead, *Elements of homotopy theory*, Springer-Verlag, New York-Heidelberg-Berlin, 1978.

[55] J.H.C. Whitehead, On C^1-complexes, *Ann. of Math.* **41**(1940), 809-824.

[56] J.H.C. Whitehead, Sphere spaces, *Proc. Nat. Acad. of Sci. USA* **21**(1935), 464-468.

[57] H. Whitney, A function not constant on a connected set of critical points, *Duke Math. J.* **1**(1935), 514-517.

[58] H. Whitney, Analytic extensions of differentiable functions defined in closed sets, *Trans. Amer. Math. Soc.* **36**(1936), 63-89.

[59] H. Whitney, Differentiable manifolds, *Ann. of Math.* **37**(1936), 645-680.

[60] H. Whitney, The self-intersections of a smooth manifold in $2n$-space, *Ann. of Math.* **45**(1944), 220-246.

Index

A. Mukherjee, *Differential Topology*, DOI 10.1007/978-3-319-19045-7