Classical and Modern Morse Theory with Applications

Francesco Mercuri

Paolo Piccione

Daniel Victor Tausk

with an Appendix by Claudio Gorodski

Contents

List of Figures

Preface

This is the textbook for a short course given by the authors during the 23rd Brazilian Colloquium of Mathematics, held at the *Instituto de Matemática Pura e Aplicada* (IMPA), Rio de Janeiro, in July 2001. The purpose of the course is to introduce the reader, supposedly a second or third year graduate student in Mathematics, to the main ideas and techniques of Morse Theory, as well as some of its most well known applications in Geometry and Analysis.

Even though the lectures of the course are planned to be given in portuguese, the choice of english for this textbook is due to the hope that these notes may serve for a wider purpose and that they could be used elsewhere. In its current form, the presentation of the material is very far from being optimal, due partly to the short amount of time in which the book had to be written.

The central idea of Morse Theory is to describe the relationship between the topology of a differentiable manifold and the structure of critical points of a real valued differentiable function defined on it. The choice of this subject for a course was based on two main reasons. First and foremost, Morse Theory is both an *elegant* and a *powerful* theory; such aspects of the theory could not be described better than it was done by the words of Richard Palais (see [**119**]):

> The essence of Morse Theory is a collection of theorems describing the intimate relationship between the topology of a manifold and the critical point structure of real valued functions on the manifold. This body of theorems has over and over proved itself to be one of the most powerful and far-reaching tools available for advancing our understanding of differential topology and analysis. But a good mathematical theory is more than *just* a collection of theorems; in addition it consists of a tool box of related conceptualizations and techniques that have been gradually been built up to help understand some circle of mathematical problems. Morse Theory is no exception, and its basic concepts and constructions have an unusual appeal derived from an underlying geometric naturality, simplicity and elegance.

The second reason that motivated the choice of this subject for our short course is the fact that Morse Theory is a truly *interdisciplinary* issue. As it will be evident to the reader of this booklet, an enormous amount of different results from a wide variety of areas of Mathematics play some role in the theory: General and Algebraic Topology, Homological Algebra, finite and infinite dimensional Differential Geometry, Real and Functional Analysis as well as some theory of ODE's participate in the construction of this magnificent "tool box". As a result, the theory offers many different aspects and it can be employed under several viewpoints to obtain results of interest by mathematicians working in different areas. For instance, typically an analyst would use Morse Theory to determine existence and multiplicity results for solutions of ordinary or partial differential equations satisfying suitable boundary conditions, and that can be described as solutions of variational problems. Under a different perspective, a typical geometer's approach to Morse Theory is to use the property of some well known functions to obtain results concerning the topology and the geometry of the underlying manifold.

Morse Theory can be thought as one of the keystones of Critical Point Theory, which, very roughly speaking, is a theory devoted to finding topological invariants for the critical points of a smooth map and to developing techniques for estimating the value of such invariants. The nice feature here is that most results of Critical Point Theory have an analytical statement and a geometrical counterpart. Just to mention a very elementary example of what a Critical Point Theorem looks like, one can think of the following statement: "if M is a compact manifold and $f : M \to \mathbb{R}$ is a smooth map, then f has at least two critical points"; namely, the maximum and the minimum. A geometric counterpart of the above statement is the following: "Assume that M is a manifold such that every smooth $f : M \to \mathbb{R}$ has at least two critical points. Then M is compact." In a sense, the topological property of compactness for the manifold produces the invariant "two" for the number of critical point of just about *any* smooth map, and vice versa.

The very basic idea of Morse theory is the following.

Given a smooth map $f : M \to \mathbb{R}$, with M, say, a compact manifold, then the sublevels $f^a = f^{-1}\big(\,]-\infty, a]\,\big)$ and $f^b = f^{-1}\big(\,]-\infty, b]\,\big)$ are homeomorphic if there is no critical value of f in the interval $[a, b]$; on the other hand, if there is a critical value $c \in \,]a, b[$, then f^b is obtained, as a topological space, by "attaching" to f^a one cell for each critical point in $f^{-1}(c)$, whose dimension equals the index of such critical point. Since the operation of attaching a cell by its boundary produces an effect in the homology of a topological space, the presence and the quantity of critical points having a given index can be measured by looking at the homology groups of M.

In this book we have made an effort to offer a presentation of the different aspects of the subject, both in the general theory and in the choice of its applications. We hope we have been able to pass to the reader at least the flavor of all the ingredients of the theory, whichever his/her personal tastes might be. We will consider a major accomplishment of our efforts if we knew that this book has managed to motivate an analyst to learn about the elegant constructions of Riemannian Geometry and Algebraic Topology, or to convince a geometer or a topologist to get involved into the delicate estimates that produce powerful techniques in Real and Functional Analysis. We must confess that, during the writing of the book we have often opted for mathematical statements or arguments that could have a stronger impact on the reader's curiosity, rather than following the most direct path to the desired conclusion.

The book was written with a purely didactical purpose, at least its first four chapters where the classical theory and some well known applications are discussed. Keeping in mind a typical student's exigences, we have made our best to make a self-contained text and to provide *many* technical details of pretty much all the statements and claims made. In order to get the reader more directly involved into the development of the theory (and also in order to remove excessive burden from some technical proofs) we propose a series of exercises at the end of each one of the first four chapters. The results obtained in the exercises are in general secondary to the main development of the theory; however, in the course of some proofs we have made explicit use of results mentioned among the exercises. Usually, in these circumstances we have presented the exercise with a suggestion of consecutive steps to be followed to get to its solution.

Caring about the *visual aspect* of the material of the book was one of our original goals, which ended up suffering very much from time limitation. Many figures that ought to have been inserted to visualize some technical proofs are still missing, and in some parts of the text we may even have forgotten to remove references to some figure which in fact never came to life. We very regretfully apologize with the reader for such failure. On the other hand, we have made a massive use of diagrams to visualize compositions of maps or even association of concepts, as customary in modern Algebraic Topology textbooks. Some boring "formula hunting" sort of proof has been occasionally replaced by a more appealing "diagram chasing" procedure, and in some parts of the text we have made of "diagram commutativity" a real philosophy of life; the choice of this language is merely a matter of personal taste.

The material of the book is organized according to the following outline. Chapter 1 contains everything the reader needs to know concerning the algebraic topological notions involved in Morse Theory: starting from the very basic definitions and properties of singular homology, relative and local homology, orientation on topological manifolds to the theory of CW-complexes and their homology. The results that are more relevant in the context of Morse Theory are contained in Section 1.16, where we prove the relations between the Betti numbers of a CW-complex and its cellular structure. Propositions 1.16.19, 1.16.20, 1.16.21 and 1.16.22 constitute the body of what could be called a "topological Morse Theory".

The basic notions of Morse Theory for real valued maps on compact manifolds are discussed in Chapter 2. After a brief review of differential and Riemannian geometry, as well as some basic notions concerning measures and densities on manifolds, we study the local and global properties of the so called *Morse functions*, which are smooth maps whose critical points are nondegenerate. The kernel of Morse Theory consists of the so called *deformation Lemmas* (Sections **??** and 2.5), that tell us how the topology of the sublevels of a Morse function changes when passing through a non critical interval and through a critical value. The general theory is introduced by a simple and instructive example, given by the height function on a torus (Section 2.2). As observed in [**119**], this is everyone's favorite example, because it has the nice features of being non trivial, easy to understand

and sufficiently general to describe satisfactorily all the distinctive characteristics of the theory.

In Chapter 3 we discuss in some details three classical applications of Morse Theory in Submanifolds Theory: a generalized Gauss-Bonnet theorem for even dimensional compact manifolds (Corollary 3.3.3), the theorem of Chern and Lashof (Theorem 3.4.1), and a characterization of Riemannian immersions with non negative isotropic curvature (Theorem 3.6.14).

In Chapter 4 we discuss the generalization of Morse Theory for smooth maps defined on non compact manifolds. Such generalized theory holds for maps that are bounded from below (or from above) and that satisfy a suitable technical condition, known as the Condition (C) of Palais and Smale. Moreover, in order to avoid trivial results, the critical points of the map under consideration should have finite index[1] (or co-index). It is a surprising fact that, once these assumptions are established, Morse theory is extended at once from the case of compact manifolds to the case of infinite dimensional[2] Hilbert manifolds. Adapting the proofs of all the results of Morse Theory for this general situation is a matter of minor changes, mostly simple restatements of results in a form which makes sense in an infinite dimensional Hilbertian setting. It would be a legitimate doubt to ask oneself why bothering about Morse Theory in compact manifolds, which causes an unnecessary duplication of results (compare the statements of the results in Section **??** with those in Sections **??** and **??**!) when a full extension of the theory can be presented by such minor adaptations. Our decision of splitting the theory in a seemingly irrational way was based on two considerations. In first place, the compact case can be handled with relatively elementary notions of differential geometry and topology, without assuming knowledge of sophisticated techniques from Hilbert space and Hilbert manifold theory. Observe that using the Morse Theory for smooth maps on compact manifolds one is able to obtain deep and non trivial results (as for instance the theorem of Reeb, Theorem 2.3.13) in a form which is accessible to a wider audience. Secondly, given the didactical purpose of the book, we felt that treating the compact case first and leaving the non compact case to a later stage would lead the student to comprehension by a gentler approach.

In Chapter 4 we also discuss one of the most well known applications of infinite dimensional Morse Theory, which is the study of the Riemannian energy functional in the space of curves joining two fixed points in a finite dimensional Riemannian manifold. It is well known that the critical points of this functional are precisely the geodesics joining the two points, and Morse theory in this case gives highly non trivial global results in Riemannian geometry. Given the importance of this example, and considering also a certain lack of rigorosity in the classical literature, we have treated the subject with a very special care of all technical details. We give a somewhat original approach to the study of the manifold structure for the space

[1]this is due to the disturbing occurrence that infinite dimensional Hilbertian balls *are* indeed retractible onto their boundaries (Proposition **??**), and therefore the operation of attaching an infinite dimensional cell produces no effect in the homology of the sublevels of the map.

[2]Observe that such manifolds are not even *locally compact*!!

of curves in a differentiable manifold satisfying suitable regularity assumptions by introducing the notion of *one-parameter family of charts* (Definition 5.1.7). We have preferred this approach which seems more suited for a didactical presentation than the classical *Vector Bundle Neighborhood* approach of Palais (see [**116, 119**]); however, it must be observed that the two methods are essentially equivalent. We then prove the details of the smoothness of the energy functional and the Palais–Smale condition, obtaining the desired results.

In Chapter **??** we give a short and informal presentation of some recent results obtained by the authors and some collaborators concerning the Morse Theory for geodesics in manifolds endowed with a non positive definite metric. The idea here is simply to show how the theory can be used successfully also in circumstances when the most crucial assumptions of the infinite dimensional Morse theory do not quite “fit as a glove” in the variational setup. For instance, in the case of non positive definite metrics, the energy functional does *not* satisfy the Palais–Smale condition, it is *not* bounded from below, and it is *strongly indefinite*, i.e., all its critical points have infinite index. In the case of *partially definite* positive metric tensors (the so-called *sub-Riemannian metrics*), the main problem in applying techniques from Morse Theory is given by the fact that, in general, the space of trial paths for the variational problem has singularities. Chapter **??** is written under a totally different perspective from the previous chapters, and most of the proofs are either simply sketched or totally omitted. The reader should also be warned of the fact that some minor discrepancies between the notation used in this chapter and that used in the previous chapters may occur occasionally.

Appendix E was written by Claudio Gorodski; the author gives a survey and detailed bibliography on some developments of the theory of the so-called “tight” and “taut” immersions in Riemannian manifolds. These immersions are characterized by the property of “minimizing” (in a suitable sense) the number of critical points respectively for the height and the distance functions that are Morse functions. This is a very active field of research today and it should attract the attention of graduate students and researchers who work in the area of Differential Geometry.

The subject of Morse Theory is far from being exhaustively treated in this book; many important aspects of the theory have not even been mentioned in these pages. Most notably, we have not touched at all the issue of *equivariant* Morse Theory, which studies situations where functional is invariant by a group of transformations of the underlying manifold. Applications of such theory are ubiquitous both in Analysis and in Geometry; as an example, we mention here the celebrated result of Gromoll and Meyer on the existence of infinitely many closed geodesics in a broad class of compact Riemannian manifolds ([**63**]). We have also omitted mentions to several applications of Morse Theory to the theory of Hamiltonian systems and Symplectic Geometry, particularly with the works of C. Conley, E. Zehnder, H. Hofer, D. Salamon, I. Ekeland, A. Floer, M. Struwe, Y. Long and many others. Extensions of Morse Theory to the case of Finsler (Banach) manifolds have not been touched upon in this book. It should also be given a mention to the existence of an alternative approach to Morse Theory, not discussed in this book, which is known as *Morse homology*. The Morse homology approach consists in studying the gradient flow lines connecting the critical points of a smooth functional f on a Riemannian manifold M: under generic assumptions, they constitute a manifold whose dimension equals the difference between the Morse indexes of the two critical points, and their combinatorics can be used to build a complex whose homology coincides with the homology of M. Most of the interest of such an approach relies in its infinite dimensional generalizations: in some situations the spaces of gradient flow lines connecting two critical points are finite dimensional, also if the critical points have infinite Morse index, so this approach can be used in cases where standard deformation arguments are not applicable.

In spite of these omissions, we have nevertheless tried to provide a sufficiently general bibliography, in which the interested reader may find suggestions for further reading on these subjects. Hopefully, future versions of this book will reduce the amount of such regretful gaps by including a discussion of some of the above mentioned topics.

Thanks are due to many friends and colleagues who have given support to us during the writing of these notes.

Our colleague Fabio Giannoni has offered mathematical support during different stages of the writing, since the beginning until the very end. He is a deep *connaisseur* of Morse Theory and very many of its modern applications in Mathematical Analysis, and he is probably the main responsible for making two of us addicted to the beauties of the theory. Particularly, Fabio's constant support and encouragement to the second author went *way* beyond the call of duty. We thank him a lot for doing so.

Our colleague Claudio Gorodski has written a beautiful appendix (Appendix E) about the so called “tight” and “taut” immersions in Riemannian manifolds. His contribution is extremely valuable and it gives the book a distinctive touch of sophistication we are so proud of.

Our old friend Antonella Marquez has helped us finding the correct text of the Kant's citation which was used as *overture* of the book; we appreciated very much her enthusiastic contribution to the work. In the cited words, the philosopher indicate the two things that cause him a profound admiration: a starry sky above him and the moral law inside him. We like to share his point of view.

Our own institutions, the *Universidade de São Paulo* and the *Universidade Estadual de Campinas* provide the most adequate environment to do, read and write Mathematics. All the people of the Differential Geometry group at USP and Unicamp have surrounded us with constant support and appreciation for our work; we wish to thank each and everyone of them. Our state and federal agencies that support the scientific research, FAPESP, CNPq and CAPES, have provided funds and equipment to carry on our research. We gratefully acknowledge their generosity.

We wish to take the opportunity for expressing our gratitude to the organizers of the 23rd Brazilian Colloquium of Mathematics who gave us the opportunity of teaching the course and publishing these notes. Particularly, we want to thank them for their patience and willingness to tolerate the great delay with which the final manuscript of these notes was delivered.

Finally, the second author wishes to express his deep gratitude to his son Pietro and his wife Diacuy for constantly reminding him that there is life beyond Morse Theory.

The authors
São Paulo, June 23rd 2001.

... der bestirnte Himmel über mir
und das moralische Gesetz in mir.
(I. Kant)

CHAPTER 1

Singular Homology and CW-complexes

1.1. Short Review of Abelian Groups

A few generalities about abelian groups should be studied before we enter the territory of homology theory. General group theory is not relevant here; only *abelian groups* are needed. In fact, when studying homology theory and homological algebra, we rarely think of abelian groups as *groups* but rather as $\mathbb{Z}$*-modules*; more explicitly, the binary operation of an abelian group is always denoted by a *plus* sign, the neutral element is always denoted by *zero* and the inverse element of g is denoted by $-g$. If g is an element of an abelian group G and $n \in \mathbb{Z}$ is an integer we can as usual define the product $ng \in G$ (see Exercise 1.1). A *group homomorphism* (or, more simply, a *homomorphism*) between abelian groups G, H is a map $f : G \to H$ such that $f(g_1 + g_2) = f(g_1) + f(g_2)$ for all $g_1, g_2 \in G$; group homomorphisms are automatically $\mathbb{Z}$*-linear*, i.e., $f(ng) = nf(g)$ for all $n \in \mathbb{Z}$, $g \in G$. We repeat below, in the context of abelian groups, some definitions that are probably better known by students in the context of linear algebra. By G we will denote an arbitrary abelian group.

1.1.1. DEFINITION. A family $(n_i)_{i\in I}$ of integer numbers is called *essentially zero* if the set $\{i \in I : n_i \neq 0\}$ is finite. Given an essentially zero family $(n_i)_{i\in I}$ of integer numbers and a family $(g_i)_{i\in I}$ of elements of G then the sum $\sum_{i\in I} n_i g_i$ is well-defined in the obvious way and it is called a *linear combination* of the family $(g_i)_{i\in I}$. The family[1] $(g_i)_{i\in I}$ is called *linearly independent* if $\sum_{i\in I} n_i g_i = 0$ implies $n_i = 0$ for all $i \in I$; a family that is not linearly independent is called *linearly dependent*. The family $(g_i)_{i\in I}$ is called *generating* for G if every element of G is a linear combination of the family (see also Remark 1.1.2 below). A family $(g_i)_{i\in I}$ that is both generating for G and linearly independent is called a *basis* for G; in this case the abelian group G is also said to be *free* over $(g_i)_{i\in I}$. An abelian group that admits a basis is called a *free abelian group*.

[1]We in principle distinguish the *family* $(g_i)_{i\in I}$ from the *set* $\{g_i\}_{i\in I} = \{g_i : i \in I\}$ in the sense that the family keeps some extra information, namely, the indexing map $i \mapsto g_i$. It is quite possible for instance that $g_i = g_j$ for $i \neq j$, i.e., that a family contains "repeated elements", while such repetition is obviously lost when we look at the corresponding set. On some occasions, however, we will (with a little abuse) use in the context of sets, notions that were in principle defined only for families (and vice-versa). Observe, for instance, that if a family $(g_i)_{i\in I}$ is linearly independent then the map $i \mapsto g_i$ is indeed injective, so that ignoring the difference between the family $(g_i)_{i\in I}$ and the set $\{g_i\}_{i\in I}$ is not so bad.

In linear algebra the term "free" is rarely used since any vector space admits a basis. It is quite easy though to give examples of abelian groups that are not free (see Exercise 1.9).

1.1.2. REMARK. One usually defines the *span* of a subset of an abelian group G to be the smallest subgroup of G containing such set. A family $(g_i)_{i\in I}$ is thus generating for G when the span of the set $\{g_i : i \in I\}$ equals G (see Exercise 1.5).

Keeping in mind the analogy with linear algebra, the following should be no surprise:

1.1.3. PROPOSITION. *Let G, H be abelian groups. Given a basis $(g_i)_{i\in I}$ for G and an arbitrary family $(h_i)_{i\in I}$ of elements of H then there exists a unique homomorphism $f : G \to H$ such that $f(g_i) = h_i$ for all $i \in I$.*

PROOF. Define f by $f\big(\sum_{i\in I} n_i g_i\big) = \sum_{i\in I} n_i h_i$. The proof that f is well-defined and is indeed a homomorphism is straightforward. $\square$

The proposition above can be nicely pictured in the language of commutative diagrams as follows. Let G, H be abelian groups and $(g_i)_{i\in I}$ a basis for G. Denoting by A the set $\{g_i\}_{i\in I}$ then Proposition 1.1.3 says that given an *arbitrary* map $f_0 : A \to H$, there exists a unique homomorphism $f : G \to H$ that extends f_0, i.e., that fills in the place of the dotted arrow in the commutative diagram:

$$\begin{tikzcd} G \arrow[r, dotted, "f"] & H \\ A \arrow[u, "\text{inclusion}"] \arrow[ur, "f_0"'] & \end{tikzcd}$$

Conversely, the validity of the property above implies that $(g_i)_{i\in I}$ is a basis for A (see Exercise 1.11).

A basic notion needed in the construction of singular homology theory is that of a free abelian group spanned by a set. The idea is the following: one starts with an arbitrary set A where no operations are defined. Then, one wants to create an environment where expressions like $a + b$ with $a, b \in A$ make sense; more specifically, one wants an abelian group G that contains the set A in such a way that A form a basis for G. This is achieved in the following:

1.1.4. DEFINITION. Let A be a set. The *free abelian group spanned by* A is the group $\mathrm{Free}[A]$ consisting of all maps $f : A \to \mathbb{Z}$ that are *essentially zero*, i.e., such that the set $\{a \in A : f(a) \neq 0\}$ is finite. The group operation in $\mathrm{Free}[A]$ is *pointwise addition*, i.e., $(f_1 + f_2)(a) = f_1(a) + f_2(a)$ for all $a \in A$ and all $f_1, f_2 \in \mathrm{Free}[A]$. We think of A as a subset of $\mathrm{Free}[A]$ in the following way: each $a \in A$ is identified with the map that carries $A \setminus \{a\}$ to zero and a to $1 \in \mathbb{Z}$.

By considering A as a subset of $\mathrm{Free}[A]$ in the way explained above, an element $f \in \mathrm{Free}[A]$ is written as:

$$f = \sum_{a\in A} f(a)a;$$

it is thus very easy to see that A is indeed a basis for $\mathrm{Free}[A]$. Typically, the elements of $\mathrm{Free}[A]$ are thought of as finite linear combinations of elements of A and not as $\mathbb{Z}$-valued maps on A. Actually, the only important thing to be kept in mind about $\mathrm{Free}[A]$ is that it is a free abelian group over A. The use of $\mathbb{Z}$-valued maps on A is just a "trick" to produce a set-theoretic rigorous construction for $\mathrm{Free}[A]$. In Exercise 1.12, the reader is asked to show that $\mathrm{Free}[A]$ is actually the *unique* free abelian group over A, up to isomorphisms (in a suitable sense).

When studying algebraic topology and homological algebra one has frequently to deal with lots of abelian groups and homomorphisms and several relations between such objects. The use of commutative diagrams to visualize such relations is unavoidable. The terminology we introduce below is another important tool to describe relations between groups and homomorphisms.

1.1.5. DEFINITION. If $G_1, \ldots, G_n$ are abelian groups and $f_i : G_i \to G_{i+1}$, $i = 1, \ldots, n-1$ are homomorphisms then we say that the sequence:

$$G_1 \xrightarrow{f_1} G_2 \xrightarrow{f_2} \cdots \xrightarrow{f_{i-1}} G_i \xrightarrow{f_i} G_{i+1} \xrightarrow{f_{i+1}} \cdots \xrightarrow{f_{n-1}} G_n$$

is *exact* at G_i ($i = 2, \ldots, n-1$) if $\mathrm{Ker}(f_i) = \mathrm{Im}(f_{i-1})$. The sequence above is called *exact* if it is exact at every G_i, $i = 2, \ldots, n-1$.

A particularly important type of exact sequences are the short exact sequences; a *short exact sequence* is an exact sequence of the form:

$$0 \longrightarrow G_1 \xrightarrow{f_1} G_2 \xrightarrow{f_2} G_3 \longrightarrow 0 \tag{1.1.1}$$

where G_1, G_2, G_3 are abelian groups, f_1, f_2 are homomorphisms and 0 denotes the one element group $\{0\}$ (obviously there is no need to explain who the unlabelled arrows are!). Exactness of (1.1.1) means that f_1 is injective, f_2 is surjective and that $\mathrm{Ker}(f_2) = \mathrm{Im}(f_1)$. If G is an abelian group and $H \subset G$ is a subgroup of G then:

$$0 \longrightarrow H \xrightarrow{i} G \xrightarrow{q} G/H \longrightarrow 0 \tag{1.1.2}$$

is a short exact sequence, where $i : H \to G$ denotes inclusion and $q : G \to G/H$ denotes the canonical quotient map. The short exact sequence (1.1.2) is essentially the most general type of short exact sequence, in the following sense: if (1.1.1) is exact, $G = G_2$ and $H = \mathrm{Im}(f_1)$ then H is a subgroup of G, f_1 gives an isomorphism between G_1 and H and, since $\mathrm{Ker}(f_2) = H$, f_2 induces an isomorphism from G/H onto G_3. The mere existence of isomorphisms $G_1 \cong H$, $G_2 \cong G$ and $G_3 \cong G/H$ doesn't quite say that (1.1.1) and (1.1.2) are "equivalent" sequences for the existence of such isomorphisms does not relate the maps appearing in both sequences. Observe though that the particular isomorphisms we have described give us a commutative diagram:

$$\begin{tikzcd}
0 \arrow[r] & G_1 \arrow[r, "f_1"] \arrow[d, "\cong"'] & G_2 \arrow[r, "f_2"] \arrow[d, dashed, "\cong"'] & G_3 \arrow[r] \arrow[d, "\cong"] & 0 \\
0 \arrow[r] & H \arrow[r, "i"'] & G \arrow[r, "q"'] & G/H \arrow[r] & 0
\end{tikzcd}$$

the existence of such diagram is the correct notion of equivalence for exact sequences!

It is well-known from linear algebra that a subspace of a vector space is always a direct summand, i.e., admits a complementary subspace. The corresponding statement in the theory of abelian groups is not true: if G is an abelian group and $H \subset G$ is a subgroup then there may not exist a subgroup $K \subset G$ with $G = H \oplus K$; in fact, it is even possible that G be not isomorphic to the direct sum $H \oplus (G/H)$ (take for instance $G = \mathbb{Z}$ and $H = 2\,\mathbb{Z}$). These considerations imply that it is not in general possible to determine the middle group G_2 of a short exact sequence (1.1.1) (up to isomorphism) only from the knowledge of G_1 and G_3. In some situations though, one has an extra bit of information about (1.1.1) that implies that $G_2 \cong G_1 \oplus G_3$; this is the subject of the following:

1.1.6. DEFINITION. We say that a short exact sequence (1.1.1) *splits* if one of the following equivalent conditions hold:

- f_1 has a left inverse that is a homomorphism, i.e., there exists a homomorphism $\phi : G_2 \to G_1$ with $\phi \circ f_1 = \mathrm{Id}_{G_1}$;
- $\mathrm{Im}(f_1) = \mathrm{Ker}(f_2)$ is a direct summand of G, i.e., there exists a subgroup $K \subset G$ with $G = \mathrm{Im}(f_1) \oplus K = \mathrm{Ker}(f_2) \oplus K$;
- f_2 has a right inverse that is a homomorphism, i.e., there exists a homomorphism $\psi : G_3 \to G_2$ with $f_2 \circ \psi = \mathrm{Id}_{G_3}$.

The equivalence between the three conditions in the definition above follows directly from the result of Exercise 1.16.

The following result gives a sufficient condition for a short exact sequence to split:

1.1.7. PROPOSITION. *Any short exact sequence* (1.1.1) *with G_3 free splits.*

PROOF. Follows directly from the result of Exercise 1.17. □

An abelian group G is said to be *finitely generated* if it admits a finite generating family. We recall the following theorem from elementary courses of group theory:

1.1.8. THEOREM (classification of finitely generated abelian groups). *Every finitely generated abelian group G is isomorphic to a direct sum of the form:*

$$\mathbb{Z} \oplus \mathbb{Z} \oplus \cdots \oplus \mathbb{Z} \oplus \mathbb{Z}_{p_1^{\alpha_1}} \oplus \mathbb{Z}_{p_2^{\alpha_2}} \oplus \cdots \oplus \mathbb{Z}_{p_k^{\alpha_k}},$$

where $p_i \geq 2$ is a prime and α_i is a positive integer for $i = 1, \ldots, k$. Moreover, the number of summands $\mathbb{Z}$ and the number of summands $\mathbb{Z}_{p^\alpha}$ (for a fixed prime $p \geq 2$ and a fixed positive integer α) is uniquely determined by G.

PROOF. See [**75**, Section 3.8]. □

1.1.9. DEFINITION. The *Betti number* of a finitely generated abelian group G is defined to be the number of summands $\mathbb{Z}$ in the decomposition given by Theorem 1.1.8. If G is not finitely generated, we define the Betti number of G to be $+\infty$.

We finish this section with a short exposition of the notion of the tensor product of two abelian groups. We remark that this material won't be used until we study homology with arbitrary coefficients in Section 1.13.

Most students begin to learn the concept of tensor product and tensors in linear algebra courses, in the context of finite dimensional real (or complex) vector spaces. In such context, tensors are often identified with multi-linear maps or (when basis are chosen) with multi-indexed matrices. Such simplifications already fail when one studies tensor products of infinite dimensional vector spaces and they are far away from reality in the context of (possibly not free) modules over arbitrary rings. We will be concerned below with the case of tensor products of $\mathbb{Z}$-modules, i.e., abelian groups; we develop a bit of the general theory of tensor products of general modules in Exercises 1.21 and 1.22.

Although the definition of tensor product may seem a bit awkward when studied for the first time, the use of tensor products is spread around several fields of mathematics. For now, let us just say that the basic motivation for the definition of a tensor product is the possibility of identifying bilinear maps with linear maps. In Section 1.13, we will make use of tensor products to give an algebraically elegant treatment of the theory of singular homology with arbitrary coefficients.

If G_1, G_2 and H are abelian groups, we recall that a map $B : G_1 \times G_2 \to H$ is called *bilinear* (or *$\mathbb{Z}$-bilinear*) if for every $g_1 \in G_1$, $g_2 \in G_2$ the maps $B(g_1, \cdot) : G_2 \to H$ and $B(\cdot, g_2) : G_1 \to H$ are group homomorphisms.

1.1.10. DEFINITION. Let G_1, G_2 be abelian groups. A *tensor product* of G_1 and G_2 is a pair (T, b) where T is an abelian group and $b : G_1 \times G_2 \to T$ is a bilinear map such that the following property holds: given an arbitrary abelian group H and an arbitrary bilinear map $B : G_1 \times G_2 \to H$ then there exists a unique homomorphism $\overline{B} : T \to H$ such that the diagram:

$$\begin{tikzcd} G_1 \times G_2 \arrow[r, "B"] \arrow[d, "b"'] & H \\ T \arrow[ur, dotted, "\overline{B}"'] & \end{tikzcd}$$

commutes.

1.1.11. LEMMA (uniqueness of tensor product). *If (T, b) and (T', b') are both tensor products of two abelian groups G_1 and G_2 then there exists a unique isomorphism $\phi : T \to T'$ such that the diagram:*

$$\begin{tikzcd} & G_1 \times G_2 \arrow[dl, "b"'] \arrow[dr, "b'"] & \\ T \arrow[rr, "\cong", "\phi"'] & & T' \end{tikzcd} \tag{1.1.3}$$

commutes.

PROOF. Since b' is bilinear and (T, b) is a tensor product of G_1 and G_2, there exists a unique homomorphism ϕ for which (1.1.3) commutes; our problem is to prove that ϕ is actually an isomorphism. Since b is bilinear and (T', b') is a tensor product of G_1 and G_2, there exists a unique homomorphism ψ for which the diagram:

(1.1.4)
$$\begin{tikzcd} & G_1 \times G_2 \arrow[dl, "b"'] \arrow[dr, "b'"] & \\ T & & T' \arrow[ll, "\psi"] \end{tikzcd}$$

commutes. We show that ϕ and ψ are mutually inverse. The commutativity of (1.1.3) and (1.1.4) imply easily that also the diagram:

$$\begin{tikzcd} & G_1 \times G_2 \arrow[dl, "b"'] \arrow[dr, "b"] & \\ T \arrow[rr, dashed, "\psi\circ\phi"'] & & T \end{tikzcd}$$

commutes. The diagram above still commutes if we replace $\psi \circ \phi$ in the dashed arrow by the identity of T; but by the definition of tensor product, the map on the dashed arrow is supposed to be unique and so $\psi \circ \phi = \mathrm{Id}$. An analogous argument shows that $\phi \circ \psi = \mathrm{Id}$, concluding the proof. □

We now give an explicit construction of a tensor product of two abelian groups G_1 and G_2. Let $F = \mathrm{Free}[G_1 \times G_2]$ be the free abelian group spanned by the set $G_1 \times G_2$; let $R \subset F$ be the subgroup spanned by the elements of the form $(g_1 + g_1', g_2) - (g_1, g_2) - (g_1', g_2)$ and $(g_1, g_2 + g_2') - (g_1, g_2) - (g_1, g_2')$ with $g_1, g_1' \in G_1$ and $g_2, g_2' \in G_2$. We set $G_1 \otimes G_2 = F/R$ and we define $b : G_1 \times G_2 \to G_1 \otimes G_2$ by $b(g_1, g_2) = g_1 \otimes g_2$, where $g_1 \otimes g_2$ denotes the coset in F/R of the element $(g_1, g_2) \in F$; below we will denote by the symbol $\otimes$ the map b. We have the following:

1.1.12. LEMMA (existence of tensor product). *The pair* $(G_1 \otimes G_2, \otimes)$ *is a tensor product of* G_1 *and* G_2.

PROOF. Observe first that the map $(g_1, g_2) \mapsto g_1 \otimes g_2$ is indeed bilinear. Let then $B : G_1 \times G_2 \to H$ be a bilinear map, where H is an arbitrary abelian group; we have to show that there exists a unique homomorphism $\overline{B} : G_1 \otimes G_2 \to H$ such that

(1.1.5)
$$\overline{B}(g_1 \otimes g_2) = B(g_1, g_2),$$

for all $g_1 \in G_1$, $g_2 \in G_2$. To prove the uniqueness of $\overline{B}$, observe that (1.1.5) determines $\overline{B}$ on the set $\{g_1 \otimes g_2 : g_1 \in G_1,\ g_2 \in G_2\}$ which is a generating set for G (because q is surjective and the pairs (g_1, g_2) generate $F = \mathrm{Free}[G_1 \times G_2]$). Finally, to prove the existence of $\overline{B}$, observe first that B extends uniquely to a homomorphism $\widetilde{B} : F \to H$; the bilinearity of B implies easily that $\widetilde{B}$ vanishes on the subgroup R of F and therefore defines a homomorphism $\overline{B}$ on the quotient $F/R = G_1 \otimes G_2$. Obviously $\overline{B}$ satisfies (1.1.5). □

From now on we will call $G_1 \otimes G_2$ *the* tensor product of G_1 and G_2; we may think of the construction given above as the definition of tensor product, but as Lemma 1.1.11 shows, any other construction of a tensor product would yield essentially the same object. The important property of the tensor product is the one given in Definition 1.1.10; namely, that given an abelian group H and a bilinear map $B : G_1 \times G_2 \to H$ then there exists a unique homomorphism $\overline{B} : G_1 \otimes G_2 \to H$ such that (1.1.5) holds for all $g_1 \in G_1$, $g_2 \in G_2$.

1.1.13. REMARK. As it was observed in the proof of Lemma 1.1.12, elements of the form $g_1 \otimes g_2$ with $g_1 \in G_1$, $g_2 \in G_2$ form a generating set for the tensor product $G_1 \otimes G_2$. It is not in general true however that all the elements of $G_1 \otimes G_2$ are of the form $g_1 \otimes g_2$ (see Exercise 1.19). Elements of the form $g_1 \otimes g_2$ are usually called *simple tensors*. It is also not true in general that $g_1 \otimes g_2 = g_1' \otimes g_2'$ implies $g_1 = g_1'$ and $g_2 = g_2'$ (for instance, $0 \otimes g_2 = g_1 \otimes 0 = 0$ for all $g_1 \in G_1$, $g_2 \in G_2$). In fact, there is no general simple algorithm to decide whether two linear combinations of simple tensors are equal (unless G_1 and G_2 are free — see Exercise 1.20)

1.1.14. EXAMPLE. The tensor product $G \otimes \mathbb{Z}$ can be naturally identified with G; more explicitly, the homomorphism:

$$G \ni g \longmapsto g \otimes 1 \in G \otimes \mathbb{Z} \tag{1.1.6}$$

is an isomorphism. To prove that, we construct explicitly an inverse for (1.1.6). Let $\overline{B} : G \otimes \mathbb{Z} \to G$ be the unique homomorphism that corresponds to the bilinear map $B(g, n) = ng$, i.e., $\overline{B}$ has the property that $\overline{B}(g \otimes n) = ng$ for all $g \in G$, $n \in \mathbb{Z}$. It is obvious that (1.1.6) is a right inverse for $\overline{B}$. We now have to check that $\overline{B}(u) \otimes 1 = u$ for all $u \in G \otimes \mathbb{Z}$; but this is obvious when u is a simple tensor. Since simple tensors form a set of generators for $G \otimes \mathbb{Z}$ the conclusion follows.

Homomorphisms between abelian groups induce homomorphisms between the corresponding tensor products. More explicitly, we have the following:

1.1.15. DEFINITION. Let $f_1 : G_1 \to G_1'$ and $f_2 : G_2 \to G_2'$ be homomorphisms. The *tensor product* of f_1 and f_2 is the unique homomorphism $f_1 \otimes f_2 : G_1 \otimes G_2 \to G_1' \otimes G_2'$ such that $(f_1 \otimes f_2)(g_1 \otimes g_2) = f_1(g_1) \otimes f_2(g_2)$ for all $g_1 \in G_1$, $g_2 \in G_2$.

The fact that $f_1 \otimes f_2$ is well-defined follows from the observation that the map:

$$G_1 \times G_2 \ni (g_1, g_2) \longmapsto f_1(g_1) \otimes f_2(g_2) \in G_1' \otimes G_2'$$

is bilinear.

1.2. Singular Homology

The basic idea of algebraic topology is to associate algebraic structures to topological spaces in such a way that homeomorphic spaces correspond to isomorphic algebraic structures. The algebraic structure (like groups, modules, vector spaces) should capture part of the geometric properties of the original topological space. In general, information will be lost during the passage from the topological space to the algebraic structure and that's pretty much the key aspect of the theory; namely, since the algebraic structures are in general simpler than the original topological spaces, it should be easier checking that the algebraic structures are not isomorphic than checking that topological spaces are not homeomorphic. Thus algebraic topology provides help in the (huge) problem of classifying topological spaces up to homeomorphisms. It can also be used to produce "clean" definitions for notions like intersection numbers of manifolds and degrees of maps; such notions are usually defined in differential topology courses by somewhat messier techniques involving, for instance, approximation theory.

The program of associating algebraic structures to topological spaces goes through two separate routes: *homotopy* theory and *homology* theory. The objective of this section is to define the singular homology groups of a topological space. There are actually several ways of associating abelian groups to topological spaces that go by the name of "homology theories"; such "homology theories" are supposed to agree on "nice" spaces. Singular homology is probably the easiest to define and it has the advantage of making sense for *all* topological spaces; it is also very easy to prove its invariance under homeomorphisms (*simplicial homology*, for instance, is defined only for polyhedrons and its invariance under homeomorphisms is quite hard to be proven). The definition of singular homology though, may look a little obscure in principle and it takes some work before we can actually compute it, even for very simple spaces. Enough being said for motivation, let's dive into the technical stuff.

For every integer $p \geq 0$ we denote by $(e_i)_{i=1}^p$ the canonical basis of $\mathbb{R}^p$ and by e_0 the origin of $\mathbb{R}^p$. The convex hull of the set $\{e_i\}_{i=0}^p$ is denoted by Δ_p and is called the *standard p-simplex*; more explicitly:

$$\Delta_p = \Big\{ \sum\nolimits_{i=0}^p t_i e_i : \sum\nolimits_{i=0}^p t_i = 1,\ t_i \geq 0,\ i = 0, \dots, p \Big\}.$$

Observe that Δ_0 consists of the single point $e_0 = 0$, Δ_1 is the unit interval $[0, 1]$ in $\mathbb{R}$, Δ_2 is a triangle in $\mathbb{R}^2$ and Δ_3 is a tetrahedron in $\mathbb{R}^3$; obviously, one should picture Δ_p for $p \geq 4$ as a "hyper-tetrahedron" in $\mathbb{R}^p$. It should be pointed out that the notation just introduced has a harmless ambiguity: for $q > p \geq i \geq 0$, e_i denotes a point of both $\mathbb{R}^p$ and $\mathbb{R}^q$. If one is bothered by that, simply identify $\mathbb{R}^p$ with a subspace of $\mathbb{R}^q$ (and both of them with a subspace of $\mathbb{R}^\infty$).

We will also take this opportunity to introduce some general notation that will be used throughout the book concerning balls and spheres. For $p \geq 0$ we set:

$$\overline{\mathrm{B}}^p = \Big\{ x \in \mathbb{R}^p : \sum\nolimits_{i=1}^p x_i^2 \leq 1 \Big\}, \quad \mathrm{B}^p = \Big\{ x \in \mathbb{R}^p : \sum\nolimits_{i=1}^p x_i^2 < 1 \Big\},$$
$$S^p = \Big\{ x \in \mathbb{R}^{p+1} : \sum\nolimits_{i=1}^{p+1} x_i^2 = 1 \Big\},$$

i.e., $\overline{\mathrm{B}}^p$ is the *closed unit ball* of $\mathbb{R}^p$, B^p is the *open unit ball* of $\mathbb{R}^p$ and S^p is the *unit sphere* of $\mathbb{R}^{p+1}$. Observe that $\overline{\mathrm{B}}^0 = \mathrm{B}^0 = \{0\}$, $S^0 = \{-1, 1\}$; it will also be convenient to make the convention that S^{-1} (the "-1-th dimensional sphere") equals the empty set. In some occasions we may also have to talk about balls and spheres of arbitrary radii and centers; to fix the notation once and for all, consider an arbitrary metric space (M, dist) and set:

$$\mathrm{B}[x_0; r] = \big\{x \in M : \mathrm{dist}(x, x_0) \le r\big\}, \ \ \mathrm{B}(x_0; r) = \big\{x \in M : \mathrm{dist}(x, x_0) < r\big\},$$
$$S(x_0; r) = \big\{x \in M : \mathrm{dist}(x, x_0) = r\big\},$$

for every $x_0 \in M$, $r > 0$.

Let X be a topological space. Our goal now is to define the singular homology groups of X. This will take quite a few preliminary definitions. We start with the following:

1.2.1. DEFINITION. A *singular p-simplex* in the space X ($p \ge 0$) is a continuous map $T : \Delta_p \to X$ from the standard p-simplex Δ_p to X. We denote by $\mathfrak{S}_p(X)$ the free abelian group spanned by the set of all singular p-simplexes in X. The elements of $\mathfrak{S}_p(X)$ will be called *singular p-chains* in X.

The reader may draw a mental picture of a singular p-simplex in X as a sort of a "curved tetrahedron" embedded in X. Of course, this may not in general be a very good picture. For instance, a constant map $T : \Delta_p \to X$ is a singular[2] p-simplex; moreover, a singular simplex is a *mapping* T and not just the set $\mathrm{Im}(T)$. A singular p-chain is a finite formal linear combination of singular p-simplexes with integer coefficients; typically, the whole group $\mathfrak{S}_p(X)$ is huge and very hard to picture.

1.2.2. REMARK. We will not distinguish between a singular 0-simplex $T : \Delta_0 \to X$ and the unique point $x \in X$ such that $\mathrm{Im}(T) = \{x\}$; the group $\mathfrak{S}_0(X)$ of singular 0-chains is therefore identified with the free abelian group spanned by the set X itself. We remark also that a singular 1-simplex in X is simply a continuous curve $T : [0, 1] \to X$.

The next ingredient we need is the notion of the boundary of a singular simplex. Roughly speaking, the boundary of a singular p-simplex T should be a singular $(p-1)$-chain which is a linear combination of the faces of T. We thus need first to introduce formally the notion of a face of a singular simplex. Of course, for $i = 0, \ldots, p$, the i-th face of $T : \Delta_p \to X$ should be defined as the restriction of T to the i-th face of the standard simplex Δ_p, i.e., the convex hull of the vertices e_j, $j = 0, \ldots, p$ with $j \ne i$. We want however the faces of T to be singular $(p-1)$-simplexes and so the definition just proposed doesn't work (for $i \ne p$); we need an auxiliary map which performs the identification between the i-th face of Δ_p and the standard $(p-1)$-simplex Δ_{p-1}. With that purpose in mind we give the following:

1.2.3. DEFINITION. Given points v_i, $i = 0, \ldots, p$ in a real vector space V we denote by $\ell(v_0, \ldots, v_p) : \Delta_p \to V$ the restriction to Δ_p of the affine map from $\mathbb{R}^p$ to V which takes e_i to v_i, $i = 0, \ldots, p$; more explicitly:

$$\ell(v_0, \ldots, v_p) : \Delta_p \ni \sum_{i=0}^{p} t_i e_i \longmapsto \sum_{i=0}^{p} t_i v_i \in V,$$

[2]Actually, the name *singular* comes from the fact that the image of a singular p-simplex can be a "degenerate" version of Δ_p.

for every $(p+1)$-tuple $(t_i)_{i=0}^p$ of non negative real numbers with $\sum_{i=0}^p t_i = 1$.

The image of $\ell(v_0, \ldots, v_p)$ is the convex hull of the set $\{v_i\}_{i=0}^p$. Obviously if V has a topology for which the vector space operations are continuous (i.e., if V is a *topological vector space*) then $\ell(v_0, \ldots, v_p)$ is a singular p-simplex in V; typically, V will be finite dimensional.

We can now formally define the boundary of a singular simplex.

1.2.4. DEFINITION. For $p \geq 1$ and $i = 0, \ldots, p$, the *i-th face* of a singular p-simplex $T : \Delta_p \to X$ is the singular $(p-1)$-simplex $T \circ \ell(e_0, \ldots, \hat{e}_i, \ldots, e_p)$, where the hat means that the term has been omitted from the sequence. The *boundary* of T is the singular $(p-1)$-chain defined by:

$$\partial_p T = \sum_{i=0}^{p} (-1)^i \, T \circ \ell(e_0, \ldots, \hat{e}_i, \ldots, e_p) \in \mathfrak{S}_{p-1}(X). \tag{1.2.1}$$

The *p-th boundary homomorphism*:

$$\partial_p : \mathfrak{S}_p(X) \longrightarrow \mathfrak{S}_{p-1}(X)$$

is defined as the unique group homomorphism satisfying (1.2.1) for every singular p-simplex $T : \Delta_p \to X$. For every singular p-chain $c \in \mathfrak{S}_p(X)$ we call $\partial_p c$ the *boundary* of c.

One should note that $\ell(e_0, \ldots, \hat{e}_i, \ldots, e_p)$ is an affine linear homeomorphism from Δ_{p-1} onto the i-th face of Δ_p. At this point one could be curious about the sign $(-1)^i$ in the definition of the boundary of T; the signs in formula (1.2.1) are motivated by the following:

1.2.5. LEMMA. *For every $p \geq 2$ we have $\partial_{p-1} \circ \partial_p = 0$.*

PROOF. It suffices to show that $\partial_{p-1} \circ \partial_p$ vanishes on singular p-simplexes. We compute:

$$\begin{aligned} \partial_{p-1} \partial_p T &= \sum_{j<i} (-1)^{i+j} \, T \circ \ell(e_0, \ldots, \hat{e}_j, \ldots, \hat{e}_i, \ldots, e_p) \\ &\quad + \sum_{j>i} (-1)^{i+j-1} \, T \circ \ell(e_0, \ldots, \hat{e}_i, \ldots, \hat{e}_j, \ldots, e_p) = 0. \end{aligned}$$ □

The property $\partial_{p-1} \circ \partial_p = 0$ is in the very heart[3] of homology theory. This will be made clear in the following:

1.2.6. DEFINITION. For $p \geq 1$ we set $Z_p(X) = \mathrm{Ker}(\partial_p)$ and for $p \geq 0$ we set $B_p(X) = \mathrm{Im}(\partial_{p+1})$; we also set $Z_0(X) = \mathfrak{S}_0(X)$. We call $Z_p(X)$ the *p-th singular cycle group* (or *p-th dimensional cycle group*) of the space X and $B_p(X)$ the *p-th singular boundary group* (or *p-th dimensional boundary group*) of X; the elements of $Z_p(X)$ and $B_p(X)$ are respectively called *singular p-cycles*

[3]The reader which is familiarized with *de Rham cohomology theory* for differentiable manifolds should keep in mind the analogy between Lemma 1.2.5 and the fact that the iterated exterior differentiation of differential forms is zero. One should observe though that exterior differentiation *increases* degree while the boundary operator decreases dimension.

(or *cycles of dimension* p) and *singular* p*-boundaries* (or *boundaries of dimension* p). By Lemma 1.2.5 we obviously have $B_p(X) \subset Z_p(X)$ and the quotient group $H_p(X) = Z_p(X)/B_p(X)$ is called the p*-th singular homology group* (or p*-th dimensional singular homology group*) of the topological space X. If $c \in Z_p(X)$ is a singular p-cycle then the equivalence class $c + B_p(X)$ determined by c in $H_p(X)$ is called the *singular homology class* determined by c. If $c_1, c_2 \in Z_p(X)$ determine the same singular homology class, i.e., if $c_1 - c_2 \in B_p(X)$ then we say that the singular p-cycles c_1 and c_2 are *homologous*.

While the singular cycle, boundary and chain groups are usually unreasonably large, it is an amazing fact that the singular homology groups can be explicitly computed when the space is not too complicated. Of course, one should not expect to compute singular homology groups directly from Definition 1.2.6, except for very simple cases; in Exercise 1.23 the reader is asked to compute the singular homology of the empty space $X = \emptyset$ and of a one-point space. The examples below should illustrate the geometrical ideas behind Definition 1.2.6.

1.2.7. EXAMPLE. Consider a singular 1-simplex T in X, i.e., a continuous curve $T : [0,1] \to X$. Its boundary $\partial_1 T$ is the formal difference $T(1) - T(0)$ (recall Remark 1.2.2). It follows that T is a cycle iff T is a closed curve. Assume that $X \subset \mathbb{R}^2$ denotes the boundary of the square $[0,1]^2$, i.e.:

$$X = \big(\{0,1\} \times [0,1]\big) \cup \big([0,1] \times \{0,1\}\big);$$

if $v_0 = (0,0)$, $v_1 = (1,0)$, $v_2 = (1,1)$, $v_3 = (0,1)$ denote the vertices of X then the singular 1-chain

$$c = \ell(v_0, v_1) + \ell(v_1, v_2) + \ell(v_2, v_3) - \ell(v_0, v_3) \in \mathfrak{S}_1(X)$$

is a cycle. If we had taken X to be the whole square $[0,1]^2$ then c would be the boundary of the 2-chain $\ell(v_0, v_1, v_2) + \ell(v_0, v_2, v_3)$, so that $c \in B_1(X)$. But if X is just the boundary of the square $[0,1]^2$ then c is *not* a singular boundary in X; this should be geometrically plausible and it is indeed true, although not easy to be proven rigorously with the theory developed so far. The 1-cycle c thus determines a *non trivial singular homology class* in the space X.

1.2.8. EXAMPLE. Consider the unit closed sphere $\overline{\mathrm{B}}^3$ and its boundary S^2. Choose a homeomorphism $h : \Delta_3 \to \overline{\mathrm{B}}^3$ that carries[4] the boundary of Δ_3 (i.e., the union of the 4 faces of Δ_3) onto S^2. The singular 2-chain

$$c = h \circ \ell(e_1, e_2, e_3) - h \circ \ell(e_0, e_2, e_3) + h \circ \ell(e_0, e_1, e_3) - h \circ \ell(e_0, e_1, e_2)$$

is a cycle in both the space $X = S^2$ and in the space $X = \overline{\mathrm{B}}^3$; in the latter, c is the boundary of $h \in \mathfrak{S}_3\big(\overline{\mathrm{B}}^3\big)$. It can be shown that c is not a singular boundary in $X = S^2$.

[4]Actually every homeomorphism from Δ_3 to $\overline{\mathrm{B}}^3$ carries the boundary of Δ_3 onto the boundary of $\overline{\mathrm{B}}^3$. The proof of this fact depends on the *theorem of the invariance of the boundary* (see Exercise 1.62) whose proof depends on the theory of *local homology*, developed in Section 1.9.

1.2.9. EXAMPLE (zero-dimensional homology). By convention, every singular 0-chain is a cycle. If the space X is path connected then two points $x_0, x_1 \in X$ always determine homologous 0-cycles; namely, a continuous curve $T : [0,1] \to X$ from x_0 to x_1 is a singular 1-chain whose boundary is $x_1 - x_0$. It follows that the 0-th singular homology group $H_0(X)$ is cyclic and spanned by the homology class of any point $x \in X$. We will prove now that the homology class of a point $x \in X$ is a *free* generator of $H_0(X)$, i.e., that if $n \in \mathbb{Z}$ is a non zero integer then $n \cdot x$ is not a boundary. It will then follow that

$$H_0(X) \cong \mathbb{Z}$$

whenever X is a non empty path connected space. The trick to prove that $n \cdot x$ can not be a boundary is the introduction of the homomorphism

$$\epsilon : \mathfrak{S}_0(X) \longrightarrow \mathbb{Z} \tag{1.2.2}$$

characterized by the property that ϵ maps every point of X to $1 \in \mathbb{Z}$. The homomorphism (1.2.2) is known as the *augmentation map*. We have the identity

$$\epsilon \circ \partial_1 = 0; \tag{1.2.3}$$

to check (1.2.3) one simply computes its lefthand side on any singular 1-simplex $T : [0,1] \to X$. It follows that the augmentation map vanishes on 0-boundaries and since $\epsilon(n \cdot x) = n$, the 0-chain $n \cdot x$ is a boundary iff $n = 0$.

1.2.10. EXAMPLE. Let V be a real topological vector space (for instance, a finite dimensional real vector space endowed with the topology induced by any norm) and let $X \subset V$ be a subset that is *star-shaped* around a point $x_0 \in X$, i.e., for every $x \in X$ the line segment $[x_0, x] = \mathrm{Im}\big(\ell(x_0, x)\big)$ is contained in X. It follows from Example 1.2.9 that $H_0(X) \cong \mathbb{Z}$. We will now show that $H_p(X) = 0$ for every $p \geq 1$. To this aim, we introduce the following notation. If $T : \Delta_p \to X$ is a singular p-simplex then we denote by $[x_0, T] : \Delta_{p+1} \to X$ the singular $(p+1)$-simplex that coincides with T on $\Delta_p \subset \Delta_{p+1}$, maps the vertex $e_{p+1} \in \Delta_{p+1}$ to the point $x_0 \in X$ and is affine on each line segment of the form $[u, e_{p+1}]$ with $u \in \Delta_p$; more explicitly:

(1.2.4)

$$[x_0, T] : \Delta_{p+1} \ni (1-t)u + te_{p+1} \longmapsto (1-t)T(u) + tx_0 \in X, \quad u \in \Delta_p,\ t \in [0,1].$$

The reader is asked to show in Exercise 1.24 that (1.2.4) indeed defines a continuous map in Δ_{p+1}. The operation $T \mapsto [x_0, T]$ extends uniquely to a group homomorphism

$$[x_0, \cdot] : \mathfrak{S}_p(X) \longrightarrow \mathfrak{S}_{p+1}(X).$$

For a singular p-chain $c \in \mathfrak{S}_p(X)$ the boundary of $[x_0, c] \in \mathfrak{S}_{p+1}(X)$ is given by:

$$\partial_{p+1}[x_0, c] = \begin{cases} c - [x_0, \partial_p c], & p \geq 1, \\ c - \epsilon(c)x_0, & p = 0, \end{cases} \tag{1.2.5}$$

where ϵ is the augmentation map. Namely, since the three expressions in (1.2.5) define group homomorphisms when seen as functions of $c \in \mathfrak{S}_p(X)$, in order to check (1.2.5) it suffices to consider the case when c is a singular p-simplex T; this is easily done using the definition of the boundary of a singular simplex. It follows directly from (1.2.5) that if c is a singular p-cycle with $p \geq 1$ then c is the boundary of $[x_0, c]$; hence $H_p(X) = 0$ for all $p \geq 1$.

Motivated by the constructions above we now give the following abstract definition:

1.2.11. DEFINITION. A *chain complex* is a pair $(\mathfrak{C}, \partial)$ where $\mathfrak{C} = (\mathfrak{C}_p)_{p\in\mathbb{Z}}$ is a family of abelian groups and $\partial = (\partial_p)_{p\in\mathbb{Z}}$ is a family of group homomorphisms $\partial_p : \mathfrak{C}_p \to \mathfrak{C}_{p-1}$ satisfying $\partial_{p-1} \circ \partial_p = 0$ for all $p \in \mathbb{Z}$. We call $\mathfrak{C}_p$ the *p-dimensional group* of the chain complex $(\mathfrak{C}, \partial)$ and ∂_p the *p-th boundary homomorphism* of $(\mathfrak{C}, \partial)$. The group $Z_p(\mathfrak{C}) = \mathrm{Ker}(\partial_p)$ is called the *p-th cycle group* (or *p-th dimensional cycle group*) of $\mathfrak{C}$ and the group $B_p(\mathfrak{C}) = \mathrm{Im}(\partial_{p+1})$ is called the *p-th boundary group* (or *p-th dimensional boundary group*) of $\mathfrak{C}$. Obviously $B_p(\mathfrak{C})$ is contained in $Z_p(\mathfrak{C})$ and the quotient $H_p(\mathfrak{C}) = Z_p(\mathfrak{C})/B_p(\mathfrak{C})$ is called the *p-th homology group* (or *p-th dimensional homology group*) of the chain complex $(\mathfrak{C}, \partial)$.

1.2.12. EXAMPLE. Let X be a topological space. For every integer $p \geq 0$ denote by $\mathfrak{S}_p(X)$ the group of singular p-chains in X and for every $p \geq 1$ let ∂_p denote the p-th boundary homomorphism introduced in Definition 1.2.4. If we set $\mathfrak{S}_p(X) = 0$ for $p < 0$ and $\partial_p = 0$ for $p \leq 0$ then, by Lemma 1.2.5, $\big(\mathfrak{S}(X), \partial\big)$ becomes a chain complex called the *singular chain complex* of the topological space X. Another important example of a chain complex associated to a topological space X is the *augmented singular chain complex* $\big(\widetilde{\mathfrak{S}}(X), \partial\big)$ defined as follows: for $p \neq -1$ we set $\widetilde{\mathfrak{S}}(X) = \mathfrak{S}_p(X)$ and for $p \neq 0$ we take the p-th boundary homomorphism of $\big(\widetilde{\mathfrak{S}}(X), \partial\big)$ to be equal to the p-th boundary homomorphism of the singular chain complex $\big(\mathfrak{S}(X), \partial\big)$; moreover, we set $\widetilde{\mathfrak{S}}_{-1}(X) = \mathbb{Z}$ and we take the 0-th boundary homomorphism of $\big(\widetilde{\mathfrak{S}}(X), \partial\big)$ to be the augmentation map introduced in Example 1.2.9; equality (1.2.3) implies that $\big(\widetilde{\mathfrak{S}}(X), \partial\big)$ is indeed a chain complex. The p-th homology group of the chain complex $\big(\widetilde{\mathfrak{S}}(X), \partial\big)$ is called the *p-th reduced singular homology group* of the topological space X and is denoted by $\widetilde{H}_p(X)$. We obviously have:

$$\widetilde{H}_p(X) = H_p(X),$$

for all $p \notin \{0, -1\}$. If X is empty then $\widetilde{H}_{-1}(X) = \mathbb{Z}$ and $H_{-1}(X) = 0$; otherwise, if X is not empty, the augmentation map is surjective so that $\widetilde{H}_{-1}(X) = H_{-1}(X) = 0$. The reader is asked to determine the relation between $H_0(X)$ and $\widetilde{H}_0(X)$ in Exercise 1.28.

Our next task is to show how a continuous map between topological spaces induces a homomorphism between their singular homology groups. We start with the algebraic part of the task.

1.2.13. DEFINITION. Given chain complexes $\mathfrak{C}$ and $\mathfrak{D}$, then a *chain map* from $\mathfrak{C}$ to $\mathfrak{D}$ is a family $\phi_p : \mathfrak{C}_p \to \mathfrak{D}_p$, $p \in \mathbb{Z}$, of homomorphisms such that:

$$\partial_p \circ \phi_p = \phi_{p-1} \circ \partial_p, \tag{1.2.6}$$

for all $p \in \mathbb{Z}$.

Condition (1.2.6) can be pictured in the following commutative diagram:

$$\begin{tikzcd}
\cdots \arrow[r, "\partial_{p+2}"] & \mathfrak{C}_{p+1} \arrow[r, "\partial_{p+1}"] \arrow[d, "\phi_{p+1}"'] & \mathfrak{C}_p \arrow[r, "\partial_p"] \arrow[d, "\phi_p"] & \mathfrak{C}_{p-1} \arrow[r, "\partial_{p-1}"] \arrow[d, "\phi_{p-1}"] & \cdots \\
\cdots \arrow[r, "\partial_{p+2}"'] & \mathfrak{D}_{p+1} \arrow[r, "\partial_{p+1}"'] & \mathfrak{D}_p \arrow[r, "\partial_p"'] & \mathfrak{D}_{p-1} \arrow[r, "\partial_{p-1}"'] & \cdots
\end{tikzcd}$$

Chain maps induce homology homomorphisms in a natural way:

1.2.14. PROPOSITION. *If $\phi : \mathfrak{C} \to \mathfrak{D}$ is a chain map then ϕ maps $Z_p(\mathfrak{C})$ to $Z_p(\mathfrak{D})$ and $B_p(\mathfrak{C})$ to $B_p(\mathfrak{D})$. In particular, ϕ induces a homomorphism $\phi_* : H_p(\mathfrak{C}) \to H_p(\mathfrak{D})$ so that the diagram:*

$$\begin{tikzcd}
Z_p(\mathfrak{C}) \arrow[r, "\phi|_{Z_p(\mathfrak{C})}"] \arrow[d, "\text{quotient map}"'] & Z_p(\mathfrak{D}) \arrow[d, "\text{quotient map}"] \\
H_p(\mathfrak{C}) \arrow[r, "\phi_*"'] & H_p(\mathfrak{D})
\end{tikzcd}$$

commutes.

PROOF. If $\partial_p c = 0$ then $\partial_p \phi_p(c) = \phi_{p-1}\big(\partial_p c\big) = 0$, which means that ϕ_p maps $Z_p(\mathfrak{C})$ to $Z_p(\mathfrak{D})$. Similarly, if $c = \partial_{p+1}(d)$ then $\phi_p(c) = \partial_{p+1} \phi_{p+1}(d)$, which means that ϕ_p maps $B_p(\mathfrak{C})$ to $B_p(\mathfrak{D})$. $\square$

Observe now that if X, Y are topological spaces and $f : X \to Y$ is a continuous map then, for each $p \geq 0$, we can define a homomorphism:

$$(f_\#)_p : \mathfrak{S}_p(X) \longrightarrow \mathfrak{S}_p(Y)$$

by requiring that $(f_\#)_p(T) = f \circ T$ for every singular p-simplex $T : \Delta_p \to X$. Moreover, if we set $(f_\#)_p = 0$ for $p < 0$ then it is easy to see that $f_\# : \mathfrak{S}(X) \to \mathfrak{S}(Y)$ becomes a chain map (see Exercise 1.31). Observe that $\epsilon \circ (f_\#)_0 = \epsilon$ so that we also obtain a chain map $f_\# : \widetilde{\mathfrak{S}}(X) \to \widetilde{\mathfrak{S}}(Y)$ by setting $(f_\#)_{-1} = \mathrm{Id} : \mathbb{Z} \to \mathbb{Z}$. Keeping in mind Proposition 1.2.14 we can now give the following:

1.2.15. DEFINITION. Given topological spaces X, Y and a continuous map $f : X \to Y$ then we denote by:

$$f_* : H_p(X) \longrightarrow H_p(Y) \tag{1.2.7}$$

the homomorphism induced in homology by the chain map $f_\# : \mathfrak{S}(X) \to \mathfrak{S}(Y)$. We also denote by:

$$f_* : \widetilde{H}_p(X) \longrightarrow \widetilde{H}_p(Y) \tag{1.2.8}$$

the homomorphism induced in homology by the chain map $f_\# : \widetilde{\mathfrak{S}}(X) \to \widetilde{\mathfrak{S}}(Y)$.

It is easy to see that the homomorphisms (1.2.7) and (1.2.8) are equal for $p > 0$; if $p = 0$ then (1.2.8) is simply a restriction of (1.2.7).

1.3. Relative Homology

The examples discussed in Section 1.2 should give the reader a vague geometrical idea of what a non zero homology class is. One can picture it as a linear combination of singular simplexes in X that "surrounds a hole". The methods used so far to compute singular homology are very primitive. In order to increase our computational power of singular homology, given a complicated topological space X, we will employ a certain strategy that can be roughly described as follows:

Step 1. identify a subspace $A \subset X$ that contains the complicated part of X, i.e., such that the space obtained by collapsing A to a point is topologically simple;

Step 2. define a notion of homology of X modulo A;

Step 3. prove a general principle that allows to remove the complicated part of A and X when computing the homology of X modulo A;

Step 4. establish nice algebraic relations between the homology of X, the homology of A and the homology of X modulo A;

Step 5. use the relations established in Step 4 to "determine"[5] the homology of X in terms of the homology of A and the homology of X modulo A.

Observe that the machinery above can be applied recursively to determine the homology of A; hopefully one gets the homology of X after a finite number of reductions. If a finite number of reductions is not sufficient, some limit processes can be used (see Exercises 1.32, 1.33 and 1.34).

In practical computations of homology, the scheme above is rarely used directly because such scheme was condensed in a more systematic method of computation of homology which is known as cellular homology (the understanding of such method is actually the final goal of this whole chapter). The reader should think of the scheme above as a motivation for all the technical definitions and results presented below.

Let X be a topological space and $A \subset X$ a subspace; we then say that (X, A) is a *pair of topological spaces*. By a *map of pairs* $f : (X, A) \to (Y, B)$ we mean a continuous map $f : X \to Y$ which carries A into B, i.e., such that $f(A) \subset B$.

Singular p-simplexes in A can obviously be seen as singular p-simplexes in X and therefore $\mathfrak{S}_p(A)$ can be seen as the subgroup of $\mathfrak{S}_p(X)$ which has as a basis

[5] The reader should be warned that the homology of A and the homology of X modulo A do not literally determine the homology of X. The precise meaning of Step 5 will be stated in Corollary 1.3.7 below.

the singular p-simplexes $T : \Delta_p \to X$ with image in A. Moreover, the boundary homomorphism $\partial_p : \mathfrak{S}_p(X) \to \mathfrak{S}_{p-1}(X)$ of $\mathfrak{S}(X)$ restricts to the boundary homomorphism of $\mathfrak{S}(A)$; we thus say that $\mathfrak{S}(A)$ is a *chain subcomplex* of $\mathfrak{S}(X)$ (see Exercise 1.36). Obviously the chain map $i_\# : \mathfrak{S}(A) \to \mathfrak{S}(X)$ induced by the inclusion $i : A \to X$ is simply the inclusion of $\mathfrak{S}(A)$ in $\mathfrak{S}(X)$. For each p we set:

$$\mathfrak{S}_p(X, A) = \mathfrak{S}_p(X)/\mathfrak{S}_p(A),$$

and we consider the homomorphism $\bar{\partial}_p : \mathfrak{S}_p(X, A) \to \mathfrak{S}_{p-1}(X, A)$ induced by $\partial_p : \mathfrak{S}_p(X) \to \mathfrak{S}_{p-1}(X)$ on the quotient. Clearly we obtain a chain complex $(\mathfrak{S}(X, A), \bar{\partial})$.

1.3.1. DEFINITION. The chain complex $\mathfrak{S}(X, A) = \mathfrak{S}(X)/\mathfrak{S}(A)$ is called the *singular chain complex of the pair* (X, A). The homology groups of $\mathfrak{S}(X, A)$ are called the *singular relative homology groups of the pair* (X, A) and are denoted by $H_p(X, A)$.

The p-th cycle and p-th boundary groups of $\mathfrak{S}(X, A)$ are subgroups of the quotient $\mathfrak{S}_p(X)/\mathfrak{S}_p(A)$ and hence the relative homology group $H_p(X, A)$ is by definition a quotient of quotients. As usual (see Exercise 1.15), we have a "cancellation rule" $\frac{G_1/H}{G_2/H} \cong G_1/G_2$ that says that a quotient of subgroups G_1/H, G_2/H of a quotient group G/H of an abelian group G can be naturally identified with a quotient of subgroups G_1, G_2 of G. Keeping this in mind, for a pair of spaces (X, A) we define the following subgroups of $\mathfrak{S}_p(X)$:

$$\begin{aligned} Z_p(X, A) &= \big\{c \in \mathfrak{S}_p(X) : \partial_p\, c \in \mathfrak{S}_{p-1}(A)\big\} = \partial_p^{-1}\big(\mathfrak{S}_{p-1}(A)\big), \\ B_p(X, A) &= \big\{\partial_{p+1}\, c + d : c \in \mathfrak{S}_{p+1}(X),\ d \in \mathfrak{S}_p(A)\big\} = B_p(X) + \mathfrak{S}_p(A); \end{aligned}$$

we call $Z_p(X, A)$ the group of *relative p-cycles* and $B_p(X, A)$ the group of *relative p-boundaries* of the pair (X, A). Observe that $Z_p(X, A)$ (respectively, $B_p(X, A)$) equals the inverse image of the cycle group $Z_p\big(\mathfrak{S}(X, A)\big)$ (respectively, of the boundary group $B_p\big(\mathfrak{S}(X, A)\big)$) by the quotient map $\mathfrak{S}_p(X) \to \mathfrak{S}_p(X, A)$. The relative homology group can therefore be identified with the quotient:

$$H_p(X, A) \cong Z_p(X, A)/B_p(X, A),$$

for all $p \in \mathbb{Z}$.

1.3.2. EXAMPLE. If A is empty then the subcomplex $\mathfrak{S}(A)$ of $\mathfrak{S}(X)$ is identically zero, so that $\mathfrak{S}(X, A)$ is just the singular chain complex $\mathfrak{S}(X)$ of X. Therefore, the relative homology groups $H_p(X, \emptyset)$ are simply equal to the absolute homology groups $H_p(X)$.

As in the case of absolute singular homology, continuous maps between pairs of spaces induce chain maps (and therefore homology homomorphisms).

1.3.3. DEFINITION. Assume that (X, A), (Y, B) are pairs of spaces and that $f : (X, A) \to (Y, B)$ is a map of pairs. The chain map $f_\# : \mathfrak{S}(X) \to \mathfrak{S}(Y)$ takes $\mathfrak{S}(A)$ to $\mathfrak{S}(B)$ and therefore induces a chain map $f_\# : \mathfrak{S}(X, A) \to \mathfrak{S}(Y, B)$; the latter is called the *chain map induced by the map of pairs* f. The chain map

$f_\# : \mathfrak{S}(X,A) \to \mathfrak{S}(Y,B)$ induces a homomorphism from $H_p(X,A)$ to $H_p(Y,B)$ that will be denoted by f_*.

In order to relate the homologies of A, X and (X,A), we start by observing that for every $p \in \mathbb{Z}$ we have a short exact sequence:

$$0 \longrightarrow \mathfrak{S}_p(A) \xrightarrow{i_\#} \mathfrak{S}_p(X) \xrightarrow{\mathfrak{q}} \mathfrak{S}_p(X,A) \longrightarrow 0 \tag{1.3.1}$$

of abelian groups and homomorphisms, where $i : A \to X$ is the inclusion and $\mathfrak{q} : \mathfrak{S}_p(X) \to \mathfrak{S}_p(X,A)$ is the quotient map. Observe also that the quotient map $\mathfrak{q} : \mathfrak{S}(X) \to \mathfrak{S}(X,A)$ can also be thought as the chain map induced by the *inclusion* $j : (X,\emptyset) \to (X,A)$ (j is just the identity of X seen as a map of pairs).

In general, a sequence of chain complexes and chain maps will be called *exact* if it is exact at every dimension. Thus (1.3.1) can actually be seen as a *short exact sequence of chain complexes and chain maps*:

$$0 \longrightarrow \mathfrak{S}(A) \xrightarrow{i_\#} \mathfrak{S}(X) \xrightarrow{\mathfrak{q}} \mathfrak{S}(X,A) \longrightarrow 0 \tag{1.3.2}$$

where 0 denotes the zero complex. The following algebraic result will take care of Step 4 of our program:

1.3.4. LEMMA (zig-zag). *Consider a short exact sequence*

$$0 \longrightarrow \mathfrak{C} \xrightarrow{f} \mathfrak{D} \xrightarrow{g} \mathcal{E} \longrightarrow 0 \tag{1.3.3}$$

of chain complexes and chain maps. One has a long exact homology sequence:

$$\cdots \xrightarrow{\partial_*} H_p(\mathfrak{C}) \xrightarrow{f_*} H_p(\mathfrak{D}) \xrightarrow{g_*} H_p(\mathcal{E}) \xrightarrow{\partial_*} H_{p-1}(\mathfrak{C}) \xrightarrow{f_*} \cdots \tag{1.3.4}$$

where the connecting homomorphism $\partial_* : H_p(\mathcal{E}) \to H_{p-1}(\mathfrak{C})$ *is defined by:*

$$\partial_*\big(e + B_p(\mathcal{E})\big) = c + B_{p-1}(\mathfrak{C}),$$

for all $e \in Z_p(\mathcal{E})$*, where* $c \in \mathfrak{C}_{p-1}$ *is chosen so that* $f(c) = \partial_p\, d$ *and* $d \in \mathfrak{D}_p$ *is chosen with* $g(d) = e$*. The definition of* ∂_* *does not depend on the various choices involved. The long exact homology sequence is* natural *in the sense that given a commutative diagram of chain complexes and chain maps*

$$\begin{tikzcd} 0 \arrow[r] & \mathfrak{C} \arrow[r, "f"] \arrow[d, "\phi"'] & \mathfrak{D} \arrow[r, "g"] \arrow[d, dashed, "\psi"] & \mathcal{E} \arrow[r] \arrow[d, "\tau"] & 0 \\ 0 \arrow[r] & \mathfrak{C}' \arrow[r, "f'"'] & \mathfrak{D}' \arrow[r, "g'"'] & \mathcal{E}' \arrow[r] & 0 \end{tikzcd} \tag{1.3.5}$$

with exact rows then the diagram

(1.3.6)

$$\begin{tikzcd} \cdots \arrow[r, "\partial_*"] & H_p(\mathfrak{C}) \arrow[r, "f_*"] \arrow[d, "\phi_*"'] & H_p(\mathfrak{D}) \arrow[r, "g_*"] \arrow[d, "\psi_*"'] & H_p(\mathcal{E}) \arrow[r, "\partial_*"] \arrow[d, "\tau_*"] & H_{p-1}(\mathfrak{C}) \arrow[r, "f_*"] \arrow[d, "\phi_*"] & \cdots \\ \cdots \arrow[r, "\partial_*"'] & H_p(\mathfrak{C}') \arrow[r, "f'_*"'] & H_p(\mathfrak{D}') \arrow[r, "g'_*"'] & H_p(\mathcal{E}') \arrow[r, "\partial_*"'] & H_{p-1}(\mathfrak{C}') \arrow[r, "f'_*"'] & \cdots \end{tikzcd}$$

commutes, where the rows in (1.3.6) *are the long exact homology sequences corresponding to the rows of* (1.3.5).

PROOF. See [**107**, §24]. □

As a corollary, we obtain the nice algebraic relations between the homologies of A, X and (X,A) that were mentioned in Step 4:

1.3.5. COROLLARY (long exact homology sequence of a pair). *Given a pair of spaces (X,A) there exists a long exact sequence:*

(1.3.7)
$$\cdots \xrightarrow{\partial_*} H_p(A) \xrightarrow{i_*} H_p(X) \xrightarrow{j_*} H_p(X,A) \xrightarrow{\partial_*} H_{p-1}(A) \xrightarrow{i_*} \cdots$$

where $i : A \to X$ and $j : (X,\emptyset) \to (X,A)$ are inclusions and the connecting homomorphism $\partial_* : H_p(X,A) \to H_{p-1}(A)$ *is defined by:*

$$\partial_*\big(c + B_p(X,A)\big) = \partial_p\, c + B_{p-1}(A),$$

for every relative cycle $c \in Z_p(X,A)$. The long exact homology sequence of a pair is natural *in the sense that given a map of pairs $f : (X,A) \to (Y,B)$ then the diagram*

(1.3.8)
$$\begin{tikzcd}
\cdots \arrow[r, "\partial_*"] & H_p(A) \arrow[r, "i_*"] \arrow[d, "(f|_A)_*"'] & H_p(X) \arrow[r, "j_*"] \arrow[d, "f_*"'] & H_p(X,A) \arrow[r, "\partial_*"] \arrow[d, "f_*"] & H_{p-1}(A) \arrow[r, "i_*"] \arrow[d, "(f|_A)_*"] & \cdots \\
\cdots \arrow[r, "\partial_*"'] & H_p(B) \arrow[r, "i_*"'] & H_p(Y) \arrow[r, "j_*"'] & H_p(Y,B) \arrow[r, "\partial_*"'] & H_{p-1}(B) \arrow[r, "i_*"'] & \cdots
\end{tikzcd}$$

commutes, where the rows in (1.3.8) *are the long exact homology sequences of the pairs (X,A) and (Y,B). There is also a long exact sequence in reduced homology*

$$\cdots \xrightarrow{\partial_*} \widetilde{H}_p(A) \xrightarrow{i_*} \widetilde{H}_p(X) \xrightarrow{j_*} H_p(X,A) \xrightarrow{\partial_*} \widetilde{H}_{p-1}(A) \xrightarrow{i_*} \cdots$$

where the connecting homomorphism ∂_* is defined as before. The long exact sequence in reduced homology is also natural with respect to maps of pairs $f : (X,A) \to (Y,B)$ in the sense above.

PROOF. The long exact sequence (1.3.7) follows by applying the Zig-Zag Lemma to the short exact sequence (1.3.2). The long exact sequence in reduced homology is obtained by applying the Zig-Zag Lemma to the short exact sequence:

$$0 \longrightarrow \widetilde{\mathfrak{S}}(A) \xrightarrow{i_\#} \widetilde{\mathfrak{S}}(X) \xrightarrow{\mathfrak{q}} \mathfrak{S}(X,A) \longrightarrow 0$$

of augmented singular chain complexes (note that we do not need an augmented version of $\mathfrak{S}(X,A)$!). □

The following technical algebraic lemma (and its corollary) will take care of Step 5 of our program.

1.3.6. LEMMA (Steenrod's five lemma). *Consider a commutative diagram of abelian groups and homomorphisms:*

$$
\begin{tikzcd}
A_1 \arrow[r] \arrow[d, "f_1"'] & A_2 \arrow[r] \arrow[d, "f_2"'] & A_3 \arrow[r] \arrow[d, "f_3"] & A_4 \arrow[r] \arrow[d, "f_4"] & A_5 \arrow[d, "f_5"] \\
B_1 \arrow[r] & B_2 \arrow[r] & B_3 \arrow[r] & B_4 \arrow[r] & B_5
\end{tikzcd}
\tag{1.3.9}
$$

If the rows in (1.3.9) *are exact and* f_1, f_2, f_4, f_5 *are isomorphisms then also* f_3 *is an isomorphism.*

PROOF. See [**107**, Lemma 24.3]. □

1.3.7. COROLLARY. *Let* $f : (X, A) \to (Y, B)$ *be a map of pairs. If any two of the maps* $f : X \to Y$, $f|_A : A \to B$, $f : (X, A) \to (Y, B)$ *induce homology isomorphisms (in all dimensions) then also the third one does.*

PROOF. Follows by applying the five lemma to a suitable portion of the diagram (1.3.8). □

1.3.8. EXAMPLE. A space X will be called *acyclic* if $\widetilde{H}_p(X) = 0$ for all $p \in \mathbb{Z}$. It follows directly from the long exact sequence of the pair (X, A) in reduced homology that if A is acyclic then $H_p(X, A) \cong \widetilde{H}_p(X)$; the isomorphism is induced by the inclusion $(X, \emptyset) \to (X, A)$. Similarly, if X is acyclic then $H_p(X, A) \cong \widetilde{H}_{p-1}(A)$; in this case, the isomorphism is induced by the connecting homomorphism ∂_*.

1.4. Excision in Singular Homology

In Section 1.2 we have shown how one can associate a chain complex $\mathfrak{S}(X)$ (an algebraic entity) to a topological space X (a geometric entity). If one modifies a topological space X or, more generally, if one combines several topological spaces by means of a geometric construction (like unions, intersections and products) then it is natural to expect that such geometric constructions will have algebraic analogues in the world of chain complexes. In some situations the relation between the geometric and the algebraic construction is quite direct, while in others it requires more care and some additional technical assumptions on the topological spaces involved. It is an amazing fact that algebraic constructions in the theory of chain complexes that were originally inspired by geometric constructions turn out to be fundamental tools in abstract algebra; this phenomenon gave birth to the field of homological algebra.

In order to motivate the material in this section we will examine below the algebraic analogues of the geometric operations of union and intersection. We remark that when dealing with other homology theories (like simplicial homology) the link between the algebraic and the geometric constructions is more concrete. Singular homology has the advantage of being defined for arbitrary topological spaces (simplicial homology is defined only for triangulable spaces); the harder understanding of the relations between algebraic and geometric constructions is a price to be paid for such generality.

1.4.1. EXAMPLE. Given subspaces X_1, X_2 of a topological space X then a singular p-simplex $T : \Delta_p \to X$ has image in X_1 and in X_2 iff its image is in the intersection $X_1 \cap X_2$. It follows easily that (see Exercise 1.13):

$$\mathfrak{S}_p(X_1) \cap \mathfrak{S}_p(X_2) = \mathfrak{S}_p(X_1 \cap X_2).$$

The equality above means that the geometric operation of intersection of topological subspaces corresponds directly to the algebraic operation of intersection of chain subcomplexes (in Exercise 1.40, the reader is asked to generalize this result to an arbitrary family of subspaces of X).

1.4.2. EXAMPLE. Motivated by Example 1.4.1, one could guess that the algebraic operation of sum of subcomplexes should somehow correspond to the geometric operation of union of subspaces of a topological space. This is indeed the case, but the relation is not direct: given subspaces $X_1, X_2 \subset X$ then obviously it is not true in general that $\mathfrak{S}_p(X_1 \cup X_2) = \mathfrak{S}_p(X_1) + \mathfrak{S}_p(X_2)$, since it is quite possible that the image of a singular p-simplex $T : \Delta_p \to X_1 \cup X_2$ is not entirely contained either in X_1 or in X_2. The equality $\mathfrak{S}_p(X_1 \cup X_2) = \mathfrak{S}_p(X_1) + \mathfrak{S}_p(X_2)$ holds only under quite restrictive hypotheses (when X_1 and X_2 are arc-connected components of X, for instance). The theory presented in this section will show that not everything is lost, though.

Although the union of spaces does not usually correspond to the sum of complexes *at the chain level*, such correspondence holds in many important situations *at the homology level*. Let us give the following:

1.4.3. DEFINITION. A pair $\{X_1, X_2\}$ of subspaces of X is called *excisive* if the inclusion $\mathfrak{S}(X_1) + \mathfrak{S}(X_2) \to \mathfrak{S}(X_1 \cup X_2)$ induces an isomorphism in homology.

The lemma below (or its corollary) gives a sufficient condition for a pair to be excisive.

1.4.4. LEMMA (small simplices). *Let X be a topological space and let $\mathfrak{A}$ be a collection of subsets of X whose interiors cover X. Setting $\mathfrak{S}(X;\mathfrak{A}) = \sum_{A\in\mathfrak{A}} \mathfrak{S}(A)$ then the inclusion $\mathfrak{S}(X;\mathfrak{A}) \to \mathfrak{S}(X)$ induces an isomorphism in homology.*

PROOF. See [**107**, Theorem 31.5]. □

1.4.5. COROLLARY. *Given a pair $\{X_1, X_2\}$ of subspaces of X, if the interiors of X_1 and X_2 cover X then $\{X_1, X_2\}$ is an excisive pair.* □

1.4.6. REMARK. The condition that $\{X_1, X_2\}$ is an excisive pair depends only on the spaces X_1, X_2 and $X_1 \cup X_2$, not on the environment space X. For this reason, we will usually work with the hypothesis $X = X_1 \cup X_2$ for simplicity. Observe for instance that the hypotheses of Corollary 1.4.5 imply $X = X_1 \cup X_2$. One can generalize Corollary 1.4.5 in the following way: if the interiors of X_1 and X_2 *with respect to* $X_1 \cup X_2$ cover $X_1 \cup X_2$ then $\{X_1, X_2\}$ is an excisive pair. The proof is obtained by taking $X = X_1 \cup X_2$ in Corollary 1.4.5.

We want to show now how one can rephrase the condition that a pair $\{X_1, X_2\}$ is excisive in terms of certain relative homology groups. Observe that from Example 1.4.1 and basic group theory (see Exercise 1.14) we conclude that there is a chain isomorphism

$$\frac{\mathfrak{S}(X_1)}{\mathfrak{S}(X_1 \cap X_2)} \xrightarrow{\cong} \frac{\mathfrak{S}(X_1) + \mathfrak{S}(X_2)}{\mathfrak{S}(X_2)} \tag{1.4.1}$$

induced by inclusion. Consider the following chain maps

$$\frac{\mathfrak{S}(X_1)}{\mathfrak{S}(X_1 \cap X_2)} \longrightarrow \frac{\mathfrak{S}(X_1 \cup X_2)}{\mathfrak{S}(X_2)}, \tag{1.4.2}$$

$$\frac{\mathfrak{S}(X_1) + \mathfrak{S}(X_2)}{\mathfrak{S}(X_2)} \longrightarrow \frac{\mathfrak{S}(X_1 \cup X_2)}{\mathfrak{S}(X_2)}, \tag{1.4.3}$$

both induced by inclusion. It follows from (1.4.1) that (1.4.2) induces a homology isomorphism iff (1.4.3) does. Using Exercise 1.38 we conclude that (1.4.3) induces a homology isomorphism iff the inclusion $\mathfrak{S}(X_1) + \mathfrak{S}(X_2) \to \mathfrak{S}(X_1 \cup X_2)$ does. We have proven the following:

1.4.7. LEMMA. *A pair $\{X_1, X_2\}$ is excisive iff the inclusion*

$$(X_1, X_1 \cap X_2) \longrightarrow (X_1 \cup X_2, X_2)$$

induces a homology isomorphism. □

The following corollary takes care of Step 3 of the program presented in the beginning of Section 1.3.

1.4.8. COROLLARY (excision). *Let (X, A) be a pair of spaces and let $U \subset A$ be a subset whose closure is contained in the interior of A. Then for every $p \in \mathbb{Z}$ we have an isomorphism*

$$H_p(X \setminus U, A \setminus U) \xrightarrow{\cong} H_p(X, A)$$

induced by inclusion.

PROOF. The hypothesis $\overline{U} \subset \mathrm{int}(A)$ implies that the interiors of the subsets $X \setminus U, A \subset X$ cover X; from Lemma 1.4.4, we get that $\{X \setminus U, A\}$ is an excisive pair. The conclusion follow from Lemma 1.4.7. □

1.5. The Mayer-Vietoris Sequence

Given a chain complex $\mathfrak{C}$ and chain subcomplexes $\mathfrak{C}_1$, $\mathfrak{C}_2$ with $\mathfrak{C} = \mathfrak{C}_1 + \mathfrak{C}_2$ then we have the following short exact sequence of chain complexes and chain maps:

$$0 \longrightarrow \mathfrak{C}_1 \cap \mathfrak{C}_2 \xrightarrow{(i,-j)} \mathfrak{C}_1 \oplus \mathfrak{C}_2 \xrightarrow{k+l} \mathfrak{C} \longrightarrow 0 \tag{1.5.1}$$

where $\mathfrak{C}_1 \oplus \mathfrak{C}_2$ denotes the (external) direct sum of $\mathfrak{C}_1$ and $\mathfrak{C}_2$ and i, j, k, l denote inclusion maps. An application of the Zig-Zag Lemma 1.3.4 to (1.5.1) yields the following long exact homology sequence (see also Exercise 1.25):

$$\cdots \longrightarrow H_p(\mathfrak{C}_1 \cap \mathfrak{C}_2) \xrightarrow{(i_*,-j_*)} H_p(\mathfrak{C}_1) \oplus H_p(\mathfrak{C}_2) \xrightarrow{k_*+l_*} H_p(\mathfrak{C}) \xrightarrow{\partial_*} H_{p-1}(\mathfrak{C}_1 \cap \mathfrak{C}_2) \longrightarrow \cdots$$

We consider now the case where $\mathfrak{C}_1$ and $\mathfrak{C}_2$ are the singular chain complexes $\mathfrak{S}(X_1)$ and $\mathfrak{S}(X_2)$ of subspaces X_1, X_2 of a topological space X. By Example 1.4.1, we have $\mathfrak{C}_1 \cap \mathfrak{C}_2 = \mathfrak{S}(X_1 \cap X_2)$. If the pair $\{X_1, X_2\}$ is excisive then $\mathfrak{C} = \mathfrak{C}_1 + \mathfrak{C}_2$ and $\mathfrak{S}(X_1 \cup X_2)$ may not in general be the same chain complexes but they play the same role at the homology level. We have proven the following:

1.5.1. PROPOSITION (Mayer-Vietoris sequence). *If $\{X_1, X_2\}$ is an excisive pair then we have a long exact homology sequence:*

$$\cdots \longrightarrow H_p(X_1 \cap X_2) \xrightarrow{(i_*,-j_*)} H_p(X_1) \oplus H_p(X_2) \xrightarrow{k_*+l_*} H_p(X_1 \cup X_2) \xrightarrow{\partial_*} H_{p-1}(X_1 \cap X_2) \longrightarrow \cdots$$

where $i : X_1 \cap X_2 \to X_1$, $j : X_1 \cap X_2 \to X_2$, $k : X_1 \to X_1 \cup X_2$ and $l : X_2 \to X_1 \cup X_2$ denote inclusion maps. □

1.6. Rudiments of Homotopy Theory

In this section we simply recall a few basic definitions from homotopy theory. The relations between homotopy and homology will be explored in Section 1.7 below.

1.6.1. DEFINITION. Given topological spaces X, Y then a *homotopy of maps from X to Y* is a continuous map $H : X \times [0,1] \to Y$; for each $t \in [0,1]$ we write $H_t : X \to Y$ for the map $x \mapsto H(x,t)$. If (X, A) and (Y, B) are pairs of spaces then we say that H is a *homotopy of maps from (X, A) to (Y, B)* if in addition H is a map of pairs $H : \big(X \times [0,1], A \times [0,1]\big) \to (Y, B)$, i.e., if $H_t(A) \subset B$ for all $t \in [0,1]$. A homotopy H is called *relative* to a subset S of X if the map $t \mapsto H(x,t)$ is constant on $[0,1]$ for all $x \in X$, i.e., if $H(x,0) = H(x,t)$ for all $x \in S$, $t \in [0,1]$. If H is a homotopy relative to X, i.e., if $H_t = H_0$ for all $t \in [0,1]$ then we call H a *constant homotopy*. If $H_0 = f$ and $H_1 = g$ then we call H a *homotopy from f to g* and we write $H : f \cong g$. When a homotopy from f to g exists we say that f and g are *homotopic*.

We define two basic operations between homotopies. If $H : X \times [0,1] \to Y$ and $K : X \times [0,1] \to Y$ are homotopies with $H_1 = K_0$ then the *concatenation* of H and K is the homotopy $(H \cdot K) : X \times [0,1] \to Y$ defined by:

$$(H \cdot K)(x,t) = \begin{cases} H(x,2t), & t \in \big[0, \frac{1}{2}\big], \\ K(x,2t-1), & t \in \big[\frac{1}{2}, 1\big], \end{cases}$$

for all $x \in X$. The *inverse* of the homotopy H is the homotopy $H^{-1} : X \times [0,1] \to Y$ defined by:

$$H^{-1}(x,t) = H(x, 1-t),$$

for all $x \in X$, $t \in [0,1]$. It is easy to see that the formulas above do define continuous maps $H \cdot K$ and H^{-1}; it follows in particular that "f is homotopic to g" is an equivalence relation in the set of continuous maps from X to Y. In Exercise 1.57 the reader is asked to show a few basic properties of the homotopy operations defined above.

We recall below another couple of definitions from homotopy theory.

1.6.2. DEFINITION. Two continuous maps $f : X \to Y$ and $g : Y \to X$ are called *homotopy inverses* if $g \circ f : X \to X$ is homotopic to the identity of X and $f \circ g : Y \to Y$ is homotopic to the identity of Y. If $f : X \to Y$ admits a homotopy inverse $g : Y \to X$ then we call f a *homotopy equivalence*; if a homotopy equivalence $f : X \to Y$ exists we say that X and Y have *the same homotopy type*.

1.6.3. DEFINITION. If A is a subspace of X then a continuous map $r : X \to A$ such that $r|_A = \mathrm{Id}_A$ is called a *retraction* from X to A; if a retraction $r : X \to A$ exists we call A a *retract* of X. We say that A is a *deformation retract* of X if there exists a retraction $r : X \to A$ that is homotopic to the identity of X when considered as a map $r : X \to X$. If there exists a homotopy relative to A between $r : X \to X$ and the identity of X then we call A a *strong deformation retract* of X.

Observe that if $A \subset X$ is a deformation retract of X then the inclusion $i : A \to X$ is a homotopy equivalence.

1.7. Homotopy Invariance of Singular Homology

The goal of this section is to show that homotopic maps induce the same homomorphisms in homology and that spaces having the same homotopy type have isomorphic homology. We start by introducing an algebraic version of the notion of homotopy.

1.7.1. DEFINITION. Given chain maps $f, g : \mathfrak{C} \to \mathfrak{D}$ then a *chain homotopy* from f to g is a family $D_p : \mathfrak{C}_p \to \mathfrak{D}_{p+1}$, $p \in \mathbb{Z}$, of group homomorphisms such that

$$f_p - g_p = \partial_{p+1} \circ D_p + D_{p-1} \circ \partial_p, \tag{1.7.1}$$

for all $p \in \mathbb{Z}$. We write $D : f \cong g$. If there exists a chain homotopy from f to g then we say that f and g are *chain homotopic*.

1.7.2. PROPOSITION. *If $f, g : \mathfrak{C} \to \mathfrak{D}$ are chain homotopic then f and g induce the same homomorphisms in homology.*

PROOF. If $c \in \mathfrak{C}_p$ is a cycle then (1.7.1) implies that $f_p(c_p) - g_p(c_p) = \partial_{p+1} D_p(c_p)$; in particular, $f_p(c_p)$ and $g_p(c_p)$ are homologous and hence $f_* = g_*$. □

The lemma below relates the algebraic and geometric notions of homotopy.

1.7.3. LEMMA. *Let X be a topological space. The continuous maps*

$$i_X, j_X : X \longrightarrow X \times [0,1]$$

defined by $i_X(x) = (x,0)$, $j_X(x) = (x,1)$ induce chain homotopic maps $(i_X)_\#$, $(j_X)_\#$. More specifically, one can associate a chain homotopy $D_X : (i_X)_\# \cong (j_X)_\#$ to every topological space X in a natural *way, i.e., in such a way that for every continuous map $f : X \to Y$ between topological spaces X, Y the diagram:*

$$\begin{tikzcd} \mathfrak{S}_p(X) \arrow[r, "D_X"] \arrow[d, "f_\#"'] & \mathfrak{S}_{p+1}\big(X \times [0,1]\big) \arrow[d, "(f\times\mathrm{Id})_\#"] \\ \mathfrak{S}_p(Y) \arrow[r, "D_Y"'] & \mathfrak{S}_{p+1}\big(Y \times [0,1]\big) \end{tikzcd}$$

commutes for every $p \in \mathbb{Z}$.

PROOF. See [**107**, Lemma 30.6]. □

The homotopy invariance of singular homology now follows simply by putting all the pieces together.

1.7.4. THEOREM (homotopy invariance). *If $f, g : (X, A) \to (Y, B)$ are homotopic maps of pairs then the induced homomorphisms $f_*, g_* : H_p(X, A) \to H_p(Y, B)$ are equal for every $p \in \mathbb{Z}$.*

PROOF. Let $H : \big(X \times [0,1], A \times [0,1]\big) \to (Y, B)$ be a homotopy from f to g. By Lemma 1.7.3, there exists a chain homotopy $D_X : i_X \cong j_X$; the naturality of this chain homotopy implies that D_X carries $\mathfrak{S}_p(A)$ to $\mathfrak{S}_{p+1}\big(A \times [0,1]\big)$ for all $p \in \mathbb{Z}$ (observe indeed that D_X restricts to D_A). It follows that D_X induces a chain homotopy from

$$(i_X)_\# : \mathfrak{S}(X, A) \longrightarrow \mathfrak{S}\big(X \times [0,1], A \times [0,1]\big)$$

to

$$(j_X)_\# : \mathfrak{S}(X, A) \longrightarrow \mathfrak{S}\big(X \times [0,1], A \times [0,1]\big).$$

Since $f = H \circ i_X$, $g = H \circ j_X$, it follows from Exercise 1.41 that $f_\#$ is chain homotopic to $g_\#$. The conclusion is now obtained from Proposition 1.7.2. □

1.7.5. COROLLARY. *If $f : X \to Y$ is a homotopy equivalence then $f_* : H_p(X) \to H_p(Y)$ is an isomorphism for all $p \in \mathbb{Z}$.*

PROOF. It follows from Theorem 1.7.4 that if g is a homotopy inverse for f then f_* and g_* are mutually inverse homology homomorphisms. □

1.7.6. COROLLARY. *If X is contractible then X is acyclic, i.e., $\widetilde{H}_p(X) = 0$ for all $p \in \mathbb{Z}$.*

PROOF. Observe that a contractible space has the same homotopy type of a one point space (see Exercise 1.58). □

1.7.7. COROLLARY. *If $f : (X, A) \to (Y, B)$ is a map of pairs and if both $f : X \to Y$ and $f|_A : A \to B$ are homotopy equivalences then $f_* : H_p(X, A) \to H_p(Y, B)$ is an isomorphism for all $p \in \mathbb{Z}$.*

PROOF. Follows from Corollaries 1.7.5 and 1.3.7. $\square$

1.8. A Computation of the Singular Homology of Spheres and the Torus

In this section we show how the techniques developed in the previous sections can be used to compute the singular homology of some simple spaces.

1.8.1. EXAMPLE. Denote by $\mathfrak{n}$, $\mathfrak{s}$ respectively the north and the south pole of S^n and identify S^{n-1} with the equator of S^n, i.e., $\mathfrak{n}$ is the $(n+1)$-th vector of the canonical basis of $\mathbb{R}^{n+1}$, $\mathfrak{s} = -\mathfrak{n}$ and S^{n-1} is identified with the intersection of S^n with the hyperplane $x_{n+1} = 0$. Since $S^n \setminus \{\mathfrak{s}\}$ is contractible, it follows from Example 1.3.8 and Corollary 1.7.6 that we have an isomorphism:

$$\widetilde{H}_p(S^n) \xrightarrow[\cong]{\text{induced by inclusion}} H_p\big(S^n, S^n \setminus \{\mathfrak{s}\}\big). \tag{1.8.1}$$

By excision of the subset $U = \{\mathfrak{n}\}$ (see Corollary 1.4.8), we have an isomorphism:

$$H_p\big(S^n \setminus \{\mathfrak{n}\}, S^n \setminus \{\mathfrak{n}, \mathfrak{s}\}\big) \xrightarrow[\cong]{\text{induced by inclusion}} H_p\big(S^n, S^n \setminus \{\mathfrak{s}\}\big). \tag{1.8.2}$$

Since $S^n \setminus \{\mathfrak{n}\}$ is contractible, Example 1.3.8 and Corollary 1.7.6 give us again an isomorphism:

$$H_p\big(S^n \setminus \{\mathfrak{n}\}, S^n \setminus \{\mathfrak{n}, \mathfrak{s}\}\big) \xrightarrow[\cong]{\partial_*} \widetilde{H}_{p-1}\big(S^n \setminus \{\mathfrak{n}, \mathfrak{s}\}\big). \tag{1.8.3}$$

Finally, we observe that the equator S^{n-1} is a deformation retract of $S^n \setminus \{\mathfrak{n}, \mathfrak{s}\}$ and therefore, by Corollary 1.7.5, we have an isomorphism:

$$\widetilde{H}_{p-1}(S^{n-1}) \xrightarrow[\cong]{\text{induced by inclusion}} \widetilde{H}_{p-1}\big(S^n \setminus \{\mathfrak{n}, \mathfrak{s}\}\big). \tag{1.8.4}$$

We have proven that:

$$\widetilde{H}_p(S^n) \cong \widetilde{H}_{p-1}(S^{n-1}).$$

By induction it follows that:

$$\widetilde{H}_p(S^n) \cong \widetilde{H}_{p-n}(S^0).$$

Since S^0 consists of two points, it follows that $\widetilde{H}_0(S^0) \cong \mathbb{Z}$ and $\widetilde{H}_q(S^0) = 0$ for $q \neq 0$, from which we obtain:

$$\widetilde{H}_p(S^n) \cong \begin{cases} \mathbb{Z}, & p = n, \\ 0, & p \neq n. \end{cases}$$

For later use we will compute in the example below the homomorphism induced on $\widetilde{H}_n(S^n)$ by the reflection on the equator.

1.8.2. EXAMPLE. Let $R : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ denote the reflection map:

$$R(x_1, x_2, \ldots, x_{n+1}) = (-x_1, x_2, \ldots, x_{n+1}). \tag{1.8.5}$$

We will show by induction on n that the homeomorphism $R|_{S^n} : S^n \to S^n$ induces minus the identity map on $\widetilde{H}_n(S^n)$. For $n = 0$ the result is obvious, since $R|_{S^0} : S^0 \to S^0$ simply exchanges the two points of S^0. For $n \geq 1$, consider the isomorphism:

$$\phi : \widetilde{H}_n(S^n) \longrightarrow \widetilde{H}_{n-1}(S^{n-1})$$

obtained by the composition of the isomorphisms (1.8.1)—(1.8.4) (with $p = n$). The validity of the induction step will follow if we can show that the diagram:

$$\begin{tikzcd}
\widetilde{H}_n(S^n) \arrow[r, "(R|_{S^n})_*"] \arrow[d, "\phi"', "\cong"] & \widetilde{H}_n(S^n) \arrow[d, "\cong"', "\phi"] \\
\widetilde{H}_{n-1}(S^{n-1}) \arrow[r, "(R|_{S^{n-1}})_*"'] & \widetilde{H}_{n-1}(S^{n-1})
\end{tikzcd}$$

commutes. To this aim we argue as follows. The map R preserves the spaces S^n, $S^n\setminus\{\mathfrak{n}\}$, $S^n\setminus\{\mathfrak{s}\}$, $S^n\setminus\{\mathfrak{n},\mathfrak{s}\}$ and S^{n-1}. Moreover, the homology homomorphisms induced by R "intertwine" with all the isomorphisms (1.8.1)—(1.8.4); for example, we have a commutative diagram:

$$\begin{tikzcd}
\widetilde{H}_n(S^n) \arrow[r, "(R|_{S^n})_*"] \arrow[d, "\text{the isomorphism (1.8.1)}"'] & \widetilde{H}_n(S^n) \arrow[d, "\text{the isomorphism (1.8.1)}"] \\
H_n\big(S^n, S^n \setminus \{\mathfrak{s}\}\big) \arrow[r, "(R|_{S^n})_*"'] & H_n\big(S^n, S^n \setminus \{\mathfrak{s}\}\big)
\end{tikzcd}$$

and one can construct similar commutative diagrams corresponding to the isomorphisms (1.8.2)—(1.8.4). This proves that $(R|_{S^n})_*$ equals minus the identity of $\widetilde{H}_n(S^n)$ for all $n \geq 0$.

In the example below we compute the singular homology of the two-dimensional torus $\mathbb{T} = S^1 \times S^1$. In spite of the large amount of technology developed so far, the reader will see that such computation is rather involved. This serves as a motivation for developing a more systematic method of computation for singular homology.

1.8.3. EXAMPLE. In what follows we will think of S^1 as the unit circle in the plane $\mathbb{R}^2$; let's start by naming the auxiliary spaces and maps that will be used in the computation.

- $\mathbb{T} = S^1 \times S^1$ denotes the torus;
- $x_0, x_1 \in S^1$ are two antipodal points;
- $\alpha \subset S^1$ is a closed arc that is symmetric around x_0;
- x_2 and x_3 are the two endpoints of α;

- $A = S^1 \times \big(S^1 \setminus \{x_0\}\big)$;
- $C = S^1 \times \alpha$;
- $D_i = S^1 \times \{x_i\}$, $i = 0, 1, 2, 3$ and $D = D_2 \cup D_3$;
- $r : S^1 \to S^1$ is the reflection with respect to the line x_0x_1;
- $R : \mathbb{T} \to \mathbb{T}$ is the map $R = \mathrm{Id} \times r$.

Clearly the circles D_2 and D_3 are (strong) deformation retracts of both A and C; moreover, D_0 is a (strong) deformation retract of C, D_1 is a (strong) deformation retract of A and D is a (strong) deformation retract of $C \setminus D_0$. Consider the pairs (C, D), $(\mathbb{T}, A)$ and $(C, C \setminus D_0)$; we have the following commutative diagram whose arrows are induced by inclusions:

$$\begin{tikzcd}
 & H_p(C, C \setminus D_0) \arrow[dr, "\cong", "\substack{\text{isomorphism}\\ \text{by excision}}"'] & \\
H_p(C, D) \arrow[ur, "\substack{\text{isomorphism}\\ \text{by homotopy invariance}}", "\cong"'] \arrow[rr, dotted] & & H_p(\mathbb{T}, A)
\end{tikzcd}$$

It follows that the dotted arrow above is an isomorphism.

We can now prove easily that $H_p(\mathbb{T}) = 0$ for $p \geq 3$. We have the following exact sequences:

$$H_p(A) \longrightarrow H_p(\mathbb{T}) \longrightarrow H_p(\mathbb{T}, A) \tag{1.8.6}$$

$$H_p(C) \longrightarrow H_p(C, D) \longrightarrow H_{p-1}(D) \tag{1.8.7}$$

extracted from the long exact homology sequences of the pairs $(\mathbb{T}, A)$ and (C, D). Since A and C have the homotopy type of a circle and D has two arc-connected components homeomorphic to circles, by Example 1.8.1, we have:

$$H_p(A) = H_p(C) = H_{p-1}(D) = 0,$$

for $p \geq 3$. The exactness of (1.8.7) implies $H_p(C, D) = 0$ and since $H_p(\mathbb{T}, A) \cong H_p(C, D)$, the exactness of (1.8.6) gives us:

$$H_p(\mathbb{T}) = 0, \quad \text{for all } p \geq 3.$$

Let us now compute the remaining homology groups of the torus. Using the long exact sequence in reduced homology (and its naturality) we get a commutative diagram with exact lines:

(1.8.8)

$$\begin{tikzcd}
H_2(D) \arrow[r] \arrow[d] & H_2(C) \arrow[r] \arrow[d] & H_2(C, D) \arrow[r] \arrow[d, "\cong"'] & H_1(D) \arrow[r] \arrow[d] & H_1(C) \arrow[r] \arrow[d] & {} \\
H_2(A) \arrow[r] & H_2(\mathbb{T}) \arrow[r] & H_2(\mathbb{T}, A) \arrow[r] & H_1(A) \arrow[r] & H_1(\mathbb{T}) \arrow[r] & {}
\end{tikzcd}$$

$$\begin{tikzcd}
{} \arrow[r] & H_1(C, D) \arrow[r] \arrow[d, "\cong"] & \widetilde{H}_0(D) \arrow[r] \arrow[d] & \widetilde{H}_0(C) \arrow[d] \\
{} \arrow[r] & H_1(\mathbb{T}, A) \arrow[r] & \widetilde{H}_0(A) \arrow[r] & \widetilde{H}_0(\mathbb{T})
\end{tikzcd}$$

We now try to identify explicitly some of the groups and homomorphisms appearing in diagram (1.8.8). We will see that all the groups appearing in diagram (1.8.8) are free of finite rank and therefore we will be able to identify them as a finite direct sum of copies of $\mathbb{Z}$. Nevertheless, we emphasize that in order to describe the homomorphisms in diagram (1.8.8) we will have to make choices of generators for some of these groups.

- The groups $H_2(A)$ and $H_2(C)$ are zero because A and C have the same homotopy type of a circle (see Example 1.8.1); moreover, $H_2(D) = H_2(D_2) \oplus H_2(D_3) = 0$.

- For the reasons above $H_1(A) \cong \mathbb{Z}$, $H_1(C) \cong \mathbb{Z}$ and $H_1(D) \cong \mathbb{Z} \oplus \mathbb{Z}$.

- $\widetilde{H}_0(C) = \widetilde{H}_0(A) = \widetilde{H}_0(\mathbb{T}) = 0$ because C, A and T are arc-connected and $\widetilde{H}_0(D) \cong \mathbb{Z}$ because D has two arc-connected components.

- Denote by $i^\nu : D_\nu \to C$, $j^\nu : D_\nu \to A$, $\nu = 2, 3$, the inclusion maps and let γ_2 denote a generator of $H_1(D_2) \cong \mathbb{Z}$. Since D_2 is a deformation retract of C and R maps D_2 homeomorphically onto D_3, it follows that $c = i^2_*(\gamma_2)$ is a generator of $H_1(C) \cong \mathbb{Z}$ and that $\gamma_3 = R_*(\gamma_2)$ is a generator of $H_1(D_3)$. Similarly, $a = j^2_*(\gamma_2)$ is a generator of $H_1(A) \cong \mathbb{Z}$. The homomorphism $H_1(D) \to H_1(C)$ in diagram (1.8.8) is (i^2_*, i^3_*) and the homomorphism $H_1(D) \to H_1(A)$ is (j^2_*, j^3_*). We know that i^3_* maps the generator γ_3 of $H_1(D_3)$ to *some* generator of $H_1(C)$, i.e., $i^3_*(\gamma_3) = \pm c$. It will be proven below that $i^3_*(\gamma_3) = c$. Similarly, we know that $j^3_*(\gamma_3) = \pm a$ and we will see below that indeed $j^3_*(\gamma_3) = a$. It follows that both the homomorphisms $H_1(D) \to H_1(C)$ and $H_1(D) \to H_1(A)$ in (1.8.8) are given by *sum* : $\mathbb{Z} \oplus \mathbb{Z} \to \mathbb{Z}$ and in particular they have the same kernel. By the exactness of the first line of (1.8.8), the composition of $H_2(C, D) \to H_1(D)$ and $H_1(D) \to H_1(A)$ is zero in (1.8.8); the commutativity of the third square implies that the homomorphism $H_2(\mathbb{T}, A) \to H_1(A)$ in (1.8.8) is zero.

- Since $H_2(C) = 0$, the homomorphism $H_2(C, D) \to H_1(D)$ in (1.8.8) is injective and hence $H_2(C, D)$ is isomorphic to the kernel of the homomorphism

$$H_1(D) \cong \mathbb{Z} \oplus \mathbb{Z} \xrightarrow{\text{sum}} \mathbb{Z} \cong H_1(C). \tag{1.8.9}$$

  Then $H_2(C, D) \cong \mathbb{Z}$ and also $H_2(\mathbb{T}, A) \cong \mathbb{Z}$.

- Since (1.8.9) is surjective, the map $H_1(C) \to H_1(C, D)$ in (1.8.8) is zero. Therefore, $H_1(C, D) \to \widetilde{H}_0(D)$ is an isomorphism and

$$H_1(C, D) \cong H_1(\mathbb{T}, A) \cong \mathbb{Z}.$$

We now rewrite the diagram (1.8.8) using all the information we obtained in the items above:
(1.8.10)

$$\begin{tikzcd}
0 \arrow[r] \arrow[d] & 0 \arrow[r] \arrow[d] & \mathbb{Z} \arrow[r, "(\mathrm{Id},-\mathrm{Id})"] \arrow[d, "\cong"'] & \mathbb{Z}\oplus\mathbb{Z} \arrow[r, "\mathrm{sum}"] \arrow[d, "\mathrm{sum}"'] & \mathbb{Z} \arrow[r, "0"] \arrow[d] & \mathbb{Z} \arrow[r, "\cong"] \arrow[d, "\cong"] & \mathbb{Z} \arrow[r] \arrow[d] & 0 \arrow[d] \\
0 \arrow[r] & H_2(\mathbb{T}) \arrow[r] & \mathbb{Z} \arrow[r, "0"'] & \mathbb{Z} \arrow[r] & H_1(\mathbb{T}) \arrow[r] & \mathbb{Z} \arrow[r] & 0 \arrow[r] & 0
\end{tikzcd}$$

Exactness of the second line in (1.8.10) implies immediately that:

$$H_2(\mathbb{T}) \cong \mathbb{Z};$$

moreover, we get a short exact sequence:

$$0 \longrightarrow \mathbb{Z} \longrightarrow H_1(\mathbb{T}) \longrightarrow \mathbb{Z} \longrightarrow 0$$

that splits, because $\mathbb{Z}$ is free. Hence:

$$H_1(\mathbb{T}) \cong \mathbb{Z} \oplus \mathbb{Z}.$$

We are left with the problem of proving $i^3_*(\gamma_3) = c$ and $j^3_*(\gamma_3) = a$. We prove $i^3_*(\gamma_3) = c$. Observe first that $R_* : H_1(C) \to H_1(C)$ is the identity; this follows from the commutative diagram[6]:

$$\begin{tikzcd}
H_1(C) \arrow[r, "R_*"] & H_1(C) \\
H_1(D_0) \arrow[r, "R_*=\mathrm{Id}"'] \arrow[u, "\cong"] & H_1(D_0) \arrow[u, "\cong"']
\end{tikzcd}$$

The equality $i^3_*(\gamma_3) = c$ now follows from the commutative diagram:

$$\begin{tikzcd}
H_1(C) \arrow[r, "R_*=\mathrm{Id}"] & H_1(C) \\
H_1(D_2) \arrow[r, "R_*"'] \arrow[u, "i^2_*"] & H_1(D_3) \arrow[u, "i^3_*"']
\end{tikzcd}$$

The proof of the equality $j^3_*(\gamma_3) = a$ is carried out in a similar way, using D_0 instead of D_1.

1.9. Local Homology

Homology groups are a global topological invariant of topological spaces: if one establishes that the p-th dimensional homology group of X is not isomorphic to the p-th dimensional homology group of Y then X cannot be homeomorphic to (or even have the same homotopy type of) Y. But what about if one wants to decide whether some small portion of X is homeomorphic to a small portion of Y? For instance, it is quite plausible (and will be proven by the theory of this section)

[6] Observe that $R_* : H_1(D_0) \to H_1(D_0)$ is the identity because $R|_{D_0} : D_0 \to D_0$ is the identity.

that a non empty open subset of $\mathbb{R}^m$ cannot be homeomorphic to an open subset of $\mathbb{R}^n$ if $m \neq n$. There is a special type of relative homology groups that are known as *local homology groups* that are suitable for solving this kind of problem.

For the development of the theory presented below we will have to assume that *all the topological spaces appearing in this section (*and all topological spaces in the book for which we talk about local homology*) satisfy the separation axiom* T_1. We say that a topological space X *satisfies the separation axiom* T_1 (or, more simply, that the space X is T_1) when the points of X are closed subsets of X, i.e., if given any pair of distinct points $x, y \in X$ we can find a neighborhood of x in X that does not contain y. Observe that all Hausdorff spaces are T_1 (actually, Hausdorff spaces are also called *T_2 spaces*).

1.9.1. DEFINITION. Let X be a topological space. The *local homology groups* of X with respect to a point $x_0 \in X$ are defined to be the relative homology groups $H_p\big(X, X \setminus \{x_0\}\big)$.

The name "local homology" is motivated by the following:

1.9.2. LEMMA. *If $x_0 \in X$ and V is a (not necessarily open) neighborhood of x_0 then the inclusion of $\big(V, V \setminus \{x_0\}\big)$ in $\big(X, X \setminus \{x_0\}\big)$ induces an isomorphism in homology.*

PROOF. Follows immediately from the excision principle, observing that the closure of $X \setminus V$ (i.e., the complement of the interior of V) is contained in the open set $X \setminus \{x_0\}$. $\square$

1.9.3. REMARK. In what follows we will usually not distinguish between the groups $H_n\big(X, X \setminus \{x_0\}\big)$ and $H_n\big(V, V \setminus \{x_0\}\big)$ when V is a neighborhood of x_0 in X. For example, if V is a neighborhood of x_0 in X, $h : V \to Y$ is a continuous map taking values in a topological space Y, $h(x_0) = y_0 \in Y$ and $h\big(V \setminus \{x_0\}\big) \subset Y \setminus \{y_0\}$ then we will say that h induces a homomorphism:

$$h_* : H_p\big(X, X \setminus \{x_0\}\big) \longrightarrow H_p\big(Y, Y \setminus \{y_0\}\big),$$

for every $p \in \mathbb{Z}$. More explicitly, the homomorphism above is the dashed arrow in the commutative diagram:

$$\begin{tikzcd}
H_p\big(X, X \setminus \{x_0\}\big) \arrow[dr, dashed] & \\
H_p\big(V, V \setminus \{x_0\}\big) \arrow[u, "\cong"'] \arrow[r, "h_*"'] & H_p\big(Y, Y \setminus \{y_0\}\big)
\end{tikzcd}$$

where the vertical unlabelled arrow is induced by inclusion.

1.9.4. EXAMPLE. Let's compute the local homology groups of $\mathbb{R}^n$ at an arbitrary point; we consider, for instance, the origin. By Lemma 1.9.2, the local homology groups of $\mathbb{R}^n$ at the origin are isomorphic to the relative homology groups $H_p\big(\overline{\mathrm{B}}^n, \overline{\mathrm{B}}^n_\times\big)$, where $\overline{\mathrm{B}}^n_\times$ denotes the *punctured closed ball* $\overline{\mathrm{B}}^n \setminus \{0\}$. Since the unit sphere S^{n-1} is a deformation retract of the punctured ball $\overline{\mathrm{B}}^n_\times$, it follows from the homotopy invariance of homology that the inclusion of $\big(\overline{\mathrm{B}}^n, S^{n-1}\big)$ in $\big(\overline{\mathrm{B}}^n, \overline{\mathrm{B}}^n_\times\big)$ induces an isomorphism in homology. Since $\overline{\mathrm{B}}^n$ is contractible, the long exact homology sequence of the pair $\big(\overline{\mathrm{B}}^n, S^{n-1}\big)$ implies that $H_p\big(\overline{\mathrm{B}}^n, S^{n-1}\big) \cong \widetilde{H}_{p-1}(S^{n-1})$. By Example 1.8.1, we have:

$$H_p\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big) \cong \begin{cases} \mathbb{Z}, & p = n, \\ 0, & p \neq n. \end{cases}$$

1.9.5. EXAMPLE. Denote by H^n the closed half-space $\{x \in \mathbb{R}^n : x_n \geq 0\}$ and by $\mathrm{Bd}(\mathsf{H}^n)$ the hyper-plane $\{x \in \mathbb{R}^n : x_n = 0\}$ (that we identify with $\mathbb{R}^{n-1}$). Obviously H^n is contractible because it is convex; but $\mathsf{H}^n \setminus \{0\}$ is also contractible because it is star-shaped around any point of the open half-space $\mathsf{H}^n \setminus \mathrm{Bd}(\mathsf{H}^n)$. It follows that the local homology groups of H^n at the origin (and also at any point of $\mathrm{Bd}(\mathsf{H}^n)$) are all identically zero, i.e.:

$$H_p\big(\mathsf{H}^n, \mathsf{H}^n \setminus \{0\}\big) = 0, \quad p \in \mathbb{Z}.$$

On the other hand, by Lemma 1.9.2 the local homology groups of H^n at the points of the open half-space $\mathsf{H}^n \setminus \mathrm{Bd}(\mathsf{H}^n)$ are the same as those of $\mathbb{R}^n$ (see Example 1.9.4).

The simple results obtained above have some very interesting applications that are developed in Exercises 1.61, 1.62 and 1.63. We finish the section by proving a result that will be used in Section 1.10 to relate the generators of the local homology groups of a manifold with the orientations of that manifold.

1.9.6. PROPOSITION. *Let $f : U \to \mathbb{R}^n$ be a continuous map defined on an open neighborhood U of the origin in $\mathbb{R}^n$. Assume that f is differentiable at the origin, the differential $\mathrm{d}f(0) : \mathbb{R}^n \to \mathbb{R}^n$ is an isomorphism, $f(0) = 0$ and $f\big(U \setminus \{0\}\big) \subset \mathbb{R}^n \setminus \{0\}$. Then the homomorphism:*

$$f_* : H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big) \longrightarrow H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big) \tag{1.9.1}$$

equals the identity if $\mathrm{d}f(0)$ has positive determinant and f_* equals minus the identity if $\mathrm{d}f(0)$ has negative determinant.

PROOF. Set:

$$c = \min_{\|v\|=1} \big\|\mathrm{d}f(0) \cdot v\big\| > 0.$$

Since $f(0) = 0$ and f is differentiable at the origin, it follows that:

$$\lim_{x \to 0} \frac{f(x) - \mathrm{d}f(0) \cdot x}{\|x\|} = 0;$$

in particular, we can find an open neighborhood $V \subset U$ of the origin such that:

$$\big\|f(x) - \mathrm{d}f(0) \cdot x\big\| \leq \frac{c}{2}\, \|x\|,$$

for all $x \in V$. This implies that $\big\|f(x) - \mathrm{d}f(0) \cdot x\big\| < \big\|\mathrm{d}f(0) \cdot x\big\|$ for all $x \in V \setminus \{0\}$ and therefore f is homotopic to $\mathrm{d}f(0)$ as a map from $\big(V, V \setminus \{0\}\big)$

to $\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big)$ (see Exercise 1.54). We have proven that the homomorphism (1.9.1) equals:

$$\mathrm{d}f(0)_* : H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big) \longrightarrow H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big). \tag{1.9.2}$$

If $\mathrm{d}f(0)$ has positive determinant then $\mathrm{d}f(0)$ is homotopic to the identity as a map from $\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big)$ to itself (see Exercise 1.55); therefore (1.9.2) equals the identity. On the other hand if $\mathrm{d}f(0)$ has negative determinant then $\mathrm{d}f(0)$ is homotopic to the reflection map $R : \mathbb{R}^n \to \mathbb{R}^n$ (see (1.8.5)) as a map from $\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big)$ to itself and therefore (1.9.2) equals minus the identity (see Exercise 1.59). This concludes the proof. $\square$

1.10. Orientation on Manifolds

An orientation for a differentiable manifold M is usually defined as a mapping that assigns to each point of M an orientation for the tangent space at that point; such choice of orientation should depend continuously on the point of M (such continuity can for instance be stated in terms of the existence of an atlas of positively oriented charts). In the case of topological manifolds there is no tangent space and so there is no obvious way of generalizing the notion of orientation to the topological case. The goal of this section is to show how one can use homology theory to give an elegant definition for the concept of orientation for topological manifolds.

Before one tries to find an intrinsic definition for the concept of orientation on a topological manifold, one should take a look at transition functions between charts of a topological manifold (i.e., homeomorphisms between open subsets of $\mathbb{R}^n$) and try to define a notion of orientation preserving transition function. In the differentiable case, such task is easy: a diffeomorphism between open subsets of $\mathbb{R}^n$ is called orientation preserving when its linear approximation around each point is orientation preserving, i.e., when its differential at each point has positive determinant. Once a notion of orientation preserving transition function between charts has been defined, one can proceed to give an intrinsic definition of orientation: in the differentiable case, one easily finds the idea of orienting the "linear approximations" of the differentiable manifold, i.e., its tangent spaces. Proposition 1.9.6 showed that orientation preserving diffeomorphisms between open subsets of $\mathbb{R}^n$ are precisely those that induce the identity on the local homology groups of $\mathbb{R}^n$; one now observes that this latter condition is *purely topological* and thus also makes sense for homeomorphisms. Now that a notion of positively oriented transition function has been found in the topological case, it is not so hard to guess what the intrinsic definition of orientation for topological manifolds should be; at the very least, one can guess that such definition should involve the local homology groups of the manifold.

The definition of a topological manifold (and also of a topological manifold with boundary) is recalled in Exercise 1.63. In what follows, M will always denote an n-dimensional topological manifold (without boundary).

1.10.1. DEFINITION. An *orientation* for M at a point $x \in M$ is a generator of the infinite cyclic group $H_n\big(M, M \setminus \{x\}\big)$.

The fact that $H_n\big(M, M \setminus \{x\}\big)$ is indeed infinite cyclic (i.e., isomorphic to $\mathbb{Z}$) is a rather trivial consequence of Lemma 1.9.2 and Example 1.9.4 (see also Exercise 1.63). Observe that (as it should be expected), at each point $x \in M$ there are precisely two orientations.

A global orientation for M should be defined as a continuous map that associates an orientation to each point of M; our next task is to define a notion of continuity for such maps. We denote by $\mathcal{O}(M)$ the disjoint union of the local homology groups $H_n\big(M, M \setminus \{x\}\big)$, i.e., we set:

$$\mathcal{O}(M) = \bigcup_{x \in M} \{x\} \times H_n\big(M, M \setminus \{x\}\big),$$

and we call $\mathcal{O}(M)$ the *orientation bundle* of the topological manifold M. Observe that there is a canonical projection:

$$\pi : \mathcal{O}(M) \longrightarrow M,$$

that takes $\{x\} \times H_n\big(M, M \setminus \{x\}\big)$ to x. By a *section* of $\mathcal{O}(M)$ along a subset $A \subset M$ we mean a map $\tau : A \to \mathcal{O}(M)$ such that $\pi \circ \tau : A \to M$ is the inclusion map of A in M, i.e., τ is a map that associates to each $x \in A$ an element of the infinite cyclic group $H_n\big(M, M \setminus \{x\}\big)$; when $A = M$, we say that τ is a *global section* (or simply a *section*) of the orientation bundle $\mathcal{O}(M)$. Observe that if τ is a section of $\mathcal{O}(M)$ then, for each $x \in M$, $\tau(x)$ is simply an element of $H_n\big(M, M \setminus \{x\}\big)$ and not necessarily a generator (although we will be mostly concerned with sections of $\mathcal{O}(M)$ that assign a generator of $H_n\big(M, M \setminus \{x\}\big)$ for every $x \in M$).

We now define a topology for the orientation bundle $\mathcal{O}(M)$. This will take a little work. For every pair of subsets $A, B \subset M$ with $B \subset A$, we consider the homomorphism:

$$\rho_{AB} : H_n(M, M \setminus A) \longrightarrow H_n(M, M \setminus B),$$

that is induced by the inclusion of $(M, M \setminus A)$ in $(M, M \setminus B)$; in particular, when $B = \{x\}$ consists of a single point we obtain a homomorphism:

$$\rho_{Ax} : H_n(M, M \setminus A) \longrightarrow H_n\big(M, M \setminus \{x\}\big), \tag{1.10.1}$$

taking values in the local homology group $H_n\big(M, M \setminus \{x\}\big)$ (we prefer writing ρ_{Ax} than using the awkward notation $\rho_{A\{x\}}$). When $C \subset B \subset A \subset M$ we have an obvious commutative diagram:

$$\begin{tikzcd} & H_n(M, M \setminus A) \arrow[dl, "\rho_{AB}"'] \arrow[dr, "\rho_{AC}"] & \\ H_n(M, M \setminus B) \arrow[rr, "\rho_{BC}"'] & & H_n(M, M \setminus C) \end{tikzcd} \tag{1.10.2}$$

The setup above constitutes what is usually called a *pre-sheaf of abelian groups* in M (see Exercise 1.64).

If $A \subset M$ is fixed then each homology class $\alpha \in H_n(M, M \setminus A)$ induces a section $\mathcal{O}(\alpha; A, M)$ of $\mathcal{O}(M)$ along A defined by:

$$\mathcal{O}(\alpha; A, M)(x) = \rho_{Ax}(\alpha), \tag{1.10.3}$$

for all $x \in A$. When M is fixed by the context we write simply $\mathcal{O}(\alpha; A)$ instead of $\mathcal{O}(\alpha; A, M)$.

1.10.2. REMARK. It is a simple consequence of the commutativity of the diagram (1.10.2) that if $B \subset A \subset M$ and $\alpha \in H_n(M, M \setminus A)$ then the section $\mathcal{O}\big(\rho_{AB}(\alpha); B\big)$ is simply the restriction to B of the section $\mathcal{O}(\alpha; A)$.

Our plan is to topologize the orientation bundle $\mathcal{O}(M)$ by requiring that the image of the sections $\mathcal{O}(\alpha; U)$ be a basis of open sets of $\mathcal{O}(M)$, where U runs over the *open* subsets of M and α runs over $H_n(M, M \setminus U)$. In order to make this definition valid, we have to prove a few things (see Exercise 1.65).

1.10.3. LEMMA. *Given a point $x \in M$ and a local homology class $\alpha_0 \in H_n\big(M, M \setminus \{x\}\big)$ then there exists an open neighborhood U of x and a homology class $\alpha \in H_n(M, M \setminus U)$ such that $\rho_{Ux}(\alpha) = \alpha_0$; more concisely:*

$$H_n\big(M, M \setminus \{x\}\big) = \bigcup_{\substack{U \text{ an open} \\ \text{neighborhood of } x}} \mathrm{Im}\big(\rho_{Ux}\big).$$

PROOF. It is a simple consequence of the fact that homology classes are compactly supported (see Exercise 1.47). Namely, we can find a pair (K_1, K_2) of compact topological spaces with $K_1 \subset M$, $K_2 \subset M \setminus \{x\}$ and such that α_0 belongs to the image of the homomorphism $H_n(K_1, K_2) \to H_n\big(M, M \setminus \{x\}\big)$ induced by inclusion. The conclusion is obtained by taking $U = M \setminus K_2$. $\square$

1.10.4. LEMMA. *Let subsets $A, B \subset M$ be given and choose homology classes $\alpha_1 \in H_n(M, M \setminus A)$, $\alpha_2 \in H_n(M, M \setminus B)$. Assume that for some $x \in A \cap B$ we have $\rho_{Ax}(\alpha_1) = \rho_{Bx}(\alpha_2)$, i.e., the sections $\mathcal{O}(\alpha_1; A)$ and $\mathcal{O}(\alpha_2; B)$ agree on the point x. Then $\mathcal{O}(\alpha_1; A)$ and $\mathcal{O}(\alpha_2; B)$ agree on a neighborhood of x in $A \cap B$, i.e., there exists an open neighborhood U of x in M such that $\rho_{Ay}(\alpha_1) = \rho_{By}(\alpha_2)$ for all $y \in U \cap A \cap B$.*

PROOF. Observe first that by replacing A and B with $A \cap B$ and α_1 and α_2 respectively with $\rho_{A(A\cap B)}(\alpha_1)$ and $\rho_{B(A\cap B)}(\alpha_2)$ (keeping in mind also Remark 1.10.2) one concludes that there is no loss of generality in assuming that $A = B$. Now the result is a simple consequence of the fact that homology relations are compactly supported (see Exercise 1.48). Namely, since $\alpha_1 - \alpha_2$ is mapped to zero by the homomorphism $H_n(M, M \setminus A) \to H_n\big(M, M \setminus \{x\}\big)$ induced by inclusion, we can find compact subsets $K_1 \subset M$, $K_2 \subset M \setminus \{x\}$ with $K_2 \subset K_1$ and such that $\alpha_1 - \alpha_2$ is also mapped to zero by the homomorphism $H_n(M, M \setminus A) \to H_n\big(M, (M \setminus A) \cup K_2\big)$ induced by inclusion. The conclusion is obtained by taking $U = M \setminus K_2$. $\square$

In the language of sheaf theory, Lemmas 1.10.3 and 1.10.4 above imply that the local homology group $H_n\big(M, M \setminus \{x\}\big)$ can be identified with the group of germs at x of the pre-sheaf determined by the groups $H_n(M, M \setminus U)$ and the maps ρ_{UV}. Thus, the orientation bundle $\mathcal{O}(M)$ is nothing more than the sheaf of germs corresponding to such pre-sheaf. Below we describe the topology of $\mathcal{O}(M)$ in sheaf-free language. For those who like the sheaf theory approach, take a look at Exercise 1.66.

1.10.5. PROPOSITION. *The sets $\mathrm{Im}\big[\mathcal{O}(\alpha; U)\big]$, where U runs over all open subsets of M and α runs through $H_n(M, M \setminus U)$ is a basis of open sets for a (unique) topology in $\mathcal{O}(M)$.*

PROOF. We use the criterion given in Exercise 1.65. We start by observing that Lemma 1.10.3 implies directly that the sets $\mathrm{Im}\big[\mathcal{O}(\alpha; U)\big]$ cover $\mathcal{O}(M)$. Now choose open sets U, $V \subset M$ and homology classes

$$\alpha_1 \in H_n(M, M \setminus U), \quad \alpha_2 \in H_n(M, M \setminus V);$$

assume that some α_0 belongs to the intersection $\mathrm{Im}\big[\mathcal{O}(\alpha_1; U)\big] \cap \mathrm{Im}\big[\mathcal{O}(\alpha_2; V)\big]$, i.e., $\alpha_0 = \rho_{Ux}(\alpha_1) = \rho_{Vx}(\alpha_2)$ for some $x \in U \cap V$. By Lemma 1.10.4 we can find an open neighborhood W of x (that can be assumed to be contained in $U \cap V$) such that $\mathcal{O}(\alpha_1; U)$ and $\mathcal{O}(\alpha_2; V)$ agree on W. Then (by Remark 1.10.2) $\mathrm{Im}\big[\mathcal{O}\big(\rho_{UW}(\alpha_1); W\big)\big]$ is contained in $\mathrm{Im}\big[\mathcal{O}(\alpha_1; U)\big] \cap \mathrm{Im}\big[\mathcal{O}(\alpha_2; V)\big]$. This concludes the proof. $\square$

From now on we will always assume that the orientation bundle $\mathcal{O}(M)$ is endowed with the topology defined by Proposition 1.10.5.

The following lemma gives a simple criterion for checking the continuity of sections of $\mathcal{O}(M)$.

1.10.6. LEMMA. *Let $A \subset M$ be a subset and $\tau : A \to \mathcal{O}(M)$ a section of $\mathcal{O}(M)$ along A. Then τ is continuous at a point $x \in A$ if and only if there exists an open neighborhood U of x in M and a homology class $\alpha \in H_n(M, M \setminus U)$ such that $\mathcal{O}(\alpha; U)$ equals τ on $A \cap U$.*

PROOF. Assume that τ is continuous at x. By Lemma 1.10.3 we can find an open neighborhood V of x in M and a homology class $\alpha \in H_n(M, M \setminus V)$ such that $\rho_{Vx}(\alpha) = \tau(x)$. Then $\tau(x)$ belongs to the open set $\mathrm{Im}\big[\mathcal{O}(\alpha; V)\big]$ and by the continuity of τ at x we can find an open neighborhood U of x in M such that $\tau(A \cap U) \subset \mathrm{Im}\big[\mathcal{O}(\alpha; U)\big]$. This implies that $\mathcal{O}(\alpha; U)$ equals τ on $A \cap U$. Conversely, assume that we can find an open neighborhood U of x in M and a homology class $\alpha \in H_n(M, M \setminus U)$ such that τ equals $\mathcal{O}(\alpha; U)$ on $A \cap U$. Choose a basic open set $\mathrm{Im}\big[\mathcal{O}(\beta; V)\big]$ containing $\tau(x)$, i.e., V is an open neighborhood of x in M, $\beta \in H_n(M, M \setminus V)$ and $\rho_{Vx}(\beta) = \tau(x) = \rho_{Ux}(\alpha)$. By Lemma 1.10.4 we can find an open neighborhood W of x contained in $U \cap V$ such that $\mathcal{O}(\alpha; U)$ equals $\mathcal{O}(\beta; V)$ on W. But then also τ equals $\mathcal{O}(\beta; V)$ on W and therefore $\tau(W) \subset \mathrm{Im}\big[\mathcal{O}(\beta; V)\big]$. This establishes the continuity of τ at x and concludes the proof. $\square$

1.10.7. COROLLARY. *For any subset $A \subset M$ and any homology class $\alpha \in H_n(M, M \setminus A)$ the section $\mathcal{O}(\alpha; A)$ of $\mathcal{O}(M)$ along A is continuous.* $\square$

We are now ready to give the following:

1.10.8. DEFINITION. An *orientation* for the topological manifold M is a continuous (global) section τ of $\mathcal{O}(M)$ such that $\tau(x)$ is a generator of the local homology group $H_n\big(M, M \setminus \{x\}\big)$ (i.e., $\tau(x)$ is an orientation for M at x) for every $x \in M$. If the manifold M admits an orientation then M is called *orientable*; a manifold M endowed with an orientation is called an *oriented* manifold.

If U is an open subset of M then one should expect that orientations of M can be restricted to orientations of U. In order to formalize that thought we have to relate the orientation bundles $\mathcal{O}(U)$ and $\mathcal{O}(M)$. First, for every $x \in U$ we can identify the local homology group $H_n\big(U, U \setminus \{x\}\big)$ with the local homology group $H_n\big(M, M \setminus \{x\}\big)$ via the isomorphism induced by inclusion (recall Remark 1.9.3). In particular, we can identify the orientation bundle $\mathcal{O}(U)$ with the subset of $\mathcal{O}(M)$ that projects onto U via the canonical projection $\pi : \mathcal{O}(M) \to M$. Moreover, we have the following:

1.10.9. LEMMA. *If U is open in M then $\mathcal{O}(U)$ is open in $\mathcal{O}(M)$; moreover, the topology of $\mathcal{O}(U)$ is induced from the topology of $\mathcal{O}(M)$.*

PROOF. For every subset $A \subset U$ and every homology class $\alpha \in H_n(U, U \setminus A)$, we denote by $\mathfrak{i}(\alpha) \in H_n(M, M \setminus A)$ the image of α by the homomorphism:

$$H_n(U, U \setminus A) \longrightarrow H_n(M, M \setminus A)$$

induced by inclusion. We have a commutative diagram:

$$\begin{tikzcd} & \mathcal{O}(M) \\ A \arrow[ur, "\mathcal{O}\left(\mathfrak{i}(\alpha);A,M\right)"] \arrow[r, "\mathcal{O}(\alpha;A,U)"'] & \mathcal{O}(U) \arrow[u, "\text{inclusion}"'] \end{tikzcd}$$

that implies that $\mathrm{Im}\big[\mathcal{O}(\alpha; A, U)\big] = \mathrm{Im}\big[\mathcal{O}\big(\mathfrak{i}(\alpha); A, M\big)\big]$. Let now $\mathfrak{T}$ be a subset of $\mathcal{O}(U)$. We show that $\mathfrak{T}$ is open in $\mathcal{O}(M)$ if and only if it is open in $\mathcal{O}(U)$. If $\mathfrak{T}$ is open in $\mathcal{O}(U)$ then every $\tau \in \mathfrak{T}$ belongs to some basic open set $\mathrm{Im}\big[\mathcal{O}(\alpha; V, U)\big]$, where $V \subset U$ is open and $\alpha \in H_n(U, U \setminus V)$ is a homology class; but then $\tau \in \mathrm{Im}\big[\mathcal{O}(\alpha; V, U)\big] = \mathrm{Im}\big[\mathcal{O}\big(i(\alpha); V, M\big)\big]$ and thus τ is an interior point of $\mathfrak{T}$ in $\mathcal{O}(M)$. Conversely, assume that $\mathfrak{T}$ is open in $\mathcal{O}(M)$. Then every $\tau \in \mathfrak{T}$ belongs to some basic open set $\mathrm{Im}\big[\mathcal{O}(\beta; V, M)\big]$ with $V \subset M$ open and $\beta \in H_n(M, M \setminus V)$ a homology class. We can replace V by a smaller open set such that $\overline{V} \subset U$; then, by excision the homomorphism:

$$H_n\big(U, U \setminus V\big) \longrightarrow H_n(M, M \setminus V)$$

induced by inclusion is an isomorphism. We can thus find $\alpha \in H_n(U, U \setminus V)$ with $i(\alpha) = \beta$. Then $\mathrm{Im}\big[\mathcal{O}(\alpha; V, U)\big] = \mathrm{Im}\big[\mathcal{O}(\beta; V, M)\big]$ is open in $\mathcal{O}(U)$, contains τ and is contained in $\mathfrak{T}$. This concludes the proof. $\square$

1.10.10. COROLLARY. *If $U \subset M$ is open and $\tau : M \to \mathcal{O}(M)$ is an orientation for M then $\tau|_U : U \to \mathcal{O}(U)$ is an orientation for U.* $\square$

If τ is an orientation for M then it is easy to see that $-\tau$ is also an orientation for M (see Exercise 1.67). If M is connected and orientable, we now show that M has precisely two orientations.

1.10.11. PROPOSITION. *If M is connected and τ, τ' are orientations for M then either $\tau = \tau'$ or $\tau = -\tau'$.*

PROOF. It follows easily from Lemmas 1.10.6 and 1.10.4 that the set:

$$\big\{x \in M : \tau(x) = \tau'(x)\big\}$$

is open. Similarly, its complement:

$$\big\{x \in M : \tau(x) = \tau'(x)\big\} = \big\{x \in M : \tau(x) = -\tau'(x)\big\},$$

is also open. The conclusion follows. $\square$

Homeomorphic manifolds have homeomorphic orientation bundles. More precisely, if $f : M \to N$ is a homeomorphism between topological manifolds then we can define a map:

$$\mathcal{O}(f) : \mathcal{O}(M) \longrightarrow \mathcal{O}(N),$$

by requiring that the restriction of $\mathcal{O}(f)$ to $H_n\big(M, M \setminus \{x\}\big)$ is equal to the homomorphism:

$$f_* : H_n\big(M, M \setminus \{x\}\big) \longrightarrow H_n\big(N, N \setminus \{f(x)\}\big), \tag{1.10.4}$$

for every $x \in M$. Moreover, we have the following:

1.10.12. PROPOSITION. *If $f : M \to N$ is a homeomorphism then the map $\mathcal{O}(f)$ is also a homeomorphism.*

PROOF. Since (1.10.4) is an isomorphism for every $x \in M$, it follows that $\mathcal{O}(f)$ is bijective. Moreover, for every open set $U \subset M$ and every homology class $\alpha \in H_n(M, M \setminus U)$ we have a commutative diagram:

$$\begin{tikzcd}
\mathcal{O}(M) \arrow[r, "\mathcal{O}(f)"] & \mathcal{O}(N) \\
U \arrow[u, "\mathcal{O}(\alpha;U)"] \arrow[r, "f"'] & f(U) \arrow[u, "\mathcal{O}\left(f_*(\alpha);f(U)\right)"']
\end{tikzcd}$$

that implies that $\mathcal{O}(f)$ maps the basic open set $\mathrm{Im}\big[\mathcal{O}(\alpha;U)\big] \subset \mathcal{O}(M)$ to the basic open set $\mathrm{Im}\big[\mathcal{O}\big(f_*(\alpha); f(U)\big)\big] \subset \mathcal{O}(N)$. Thus $\mathcal{O}(f)$ is an open map. But then $\mathcal{O}(f)^{-1} = \mathcal{O}\big(f^{-1}\big)$ is also an open map. This concludes the proof. $\square$

1.10.13. DEFINITION. A homeomorphism $f : M \to N$ between oriented topological manifolds (M, τ), (N, τ') is called *positively oriented* (or, more simply, *positive*) if $\mathcal{O}(f) \circ \tau = \tau'$. Similarly, we say that $f : M \to N$ is *negatively oriented* (or, more simply, *negative*) if $\mathcal{O}(f) \circ \tau = -\tau'$.

1.10.14. REMARK. If M is orientable and connected then one need not choose an orientation for M in order to talk about positivity and negativity of homeomorphisms $f : M \to M$ (or, more in general, of homeomorphisms between open subsets of M). Namely, if τ is an orientation for M then $f : (M, \tau) \to (M, \tau)$ is positively oriented (or negatively oriented) if and only if $f : (M, -\tau) \to (M, -\tau)$ is.

The following is a simple consequence of Proposition 1.10.11.

1.10.15. PROPOSITION. *Let $f : M \to N$ be a homeomorphism between oriented topological manifolds (M, τ), (N, τ'). If M is connected then f is either positively oriented or negatively oriented.*

PROOF. By Proposition 1.10.12, $\mathcal{O}(f) : \mathcal{O}(M) \to \mathcal{O}(N)$ is a homeomorphism and thus $\mathcal{O}(f)^{-1} \circ \tau' \circ f$ is an orientation for M; such orientation is either equal to τ or equal to $-\tau$, by Proposition 1.10.11. In the first case, f is positive and in the latter, negative. □

Let's now take a look at the case $M = \mathbb{R}^n$. For every $v \in \mathbb{R}^n$ we denote by $\mathfrak{t}_v$ the *translation map in the direction* v:

$$\mathfrak{t}_v : \mathbb{R}^n \ni x \longmapsto x + v \in \mathbb{R}^n.$$

Obviously $\mathfrak{t}_v$ induces an isomorphism:

$$(\mathfrak{t}_v)_* : H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big) \longrightarrow H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{v\}\big).$$

It would be natural to expect that, given a generator of $H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big)$, one can "spread around" such generator using the maps $(\mathfrak{t}_v)_*$ in order to produce an orientation for $\mathbb{R}^n$. This is indeed true, but the proof is not so straightforward as one could expect. It actually depends on the following:

1.10.16. LEMMA. *For every $n \geq 1$ and every $v \in \mathrm{B}^n$ there exists a homeomorphism $h : \mathbb{R}^n \to \mathbb{R}^n$ satisfying the following properties:*

- *h equals the identity outside B^n;*
- *h equals the translation $\mathfrak{t}_v$ in a neighborhood of the origin (in particular $h(0) = v$).*

PROOF. Let $\xi : \mathbb{R} \to \mathbb{R}$ be a differentiable map that equals 1 in a neighborhood of zero, vanishes outside $\left]-\infty, 1\right[$ and such that $\sup_{t \in \mathbb{R}} \big|\xi'(t)\big| < \frac{1}{\|v\|}$. Consider the map $h : \mathbb{R}^n \to \mathbb{R}^n$ defined by:

$$h(x) = x + \xi\big(\|x\|\big)v, \quad x \in \mathbb{R}^n;$$

using the estimate on ξ' and the mean value inequality, it is easy to see that the map $x \mapsto \xi(\|x\|)v$ is a contraction and therefore h is a global homeomorphism of $\mathbb{R}^n$ (see Exercise 1.69). Moreover, it is obvious that h satisfies the required properties. □

We can now prove the following:

1.10.17. PROPOSITION. *Choose a generator* τ_0 *of* $H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big)$. *The map* $\tau : \mathbb{R}^n \to \mathcal{O}(\mathbb{R}^n)$ *defined by:*

$$\tau(x) = (\mathfrak{t}_x)_*(\tau_0) \in H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{x\}\big), \quad x \in \mathbb{R}^n, \tag{1.10.5}$$

is an orientation for $\mathbb{R}^n$.

PROOF. Obviously each $\tau(x)$ is an orientation for the point x and thus we only have to prove the continuity of τ. We show that τ is continuous at the origin. Set $U = \mathrm{B}^n$; since $\mathbb{R}^n \setminus \mathrm{B}^n$ is a (strong) deformation retract of $\mathbb{R}^n \setminus \{0\}$ the map:

$$\rho_{U0} : H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \mathrm{B}^n\big) \longrightarrow H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big)$$

is an isomorphism (recall that ρ_{U0} is simply the homomorphism induced by inclusion). We can thus find $\alpha \in H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \mathrm{B}^n\big)$ such that $\rho_{U0}(\alpha) = \tau_0$. We claim that τ equals $\mathcal{O}(\alpha; U)$ on $U = \mathrm{B}^n$ (this will imply the continuity of τ at the origin by Lemma 1.10.6). Let $v \in \mathrm{B}^n$ be fixed and choose $h : \mathbb{R}^n \to \mathbb{R}^n$ as in Lemma 1.10.16. The commutative diagram:

$$\begin{tikzcd}
H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \mathrm{B}^n\big) \arrow[r, "h_* = \mathrm{Id}"] \arrow[d, "\rho_{U0}"'] & H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \mathrm{B}^n\big) \arrow[d, "\rho_{Uv}"] \\
H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big) \arrow[r, "h_* = (\mathfrak{t}_v)_*"'] & H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{v\}\big)
\end{tikzcd}$$

implies that $\rho_{Uv}(\alpha) = \tau(v)$, proving the claim and the continuity of τ at the origin. The continuity of τ at the other points of $\mathbb{R}^n$ can be proven in a similar way using the (obvious) version of Lemma 1.10.16 for balls with other centers. $\square$

1.10.18. COROLLARY. *For any orientation chosen on* $\mathbb{R}^n$*, the translations* $\mathfrak{t}_v$ *are positively oriented homeomorphisms.*

PROOF. Observe that Proposition 1.10.11 implies that any orientation τ for $\mathbb{R}^n$ must be of the form (1.10.5), for some generator τ_0 of $H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big)$. For such an orientation τ, the map $\mathcal{O}(\mathfrak{t}_v)$ carries $\tau(0)$ to $\tau(v)$ and hence $\mathfrak{t}_v$ must be positively oriented by Proposition 1.10.15. $\square$

1.10.19. COROLLARY. *Let* $f : U \to V$ *be a diffeomorphism between open subsets* U, $V \subset \mathbb{R}^n$. *Choose an arbitrary orientation for* $\mathbb{R}^n$ *and assume that* U *and* V *are endowed with the restriction of such orientation. Then* f *is a positively oriented homeomorphism (respectively, negatively oriented homeomorphism) if and only if* $\mathrm{d}f(x)$ *has positive determinant (respectively, negative determinant) for every* $x \in U$.

PROOF. Since translations are positively oriented, it follows that f is positively oriented (respectively, negatively oriented) if and only if the homomorphism:

$$\big(\mathfrak{t}_{-f(x)} \circ f \circ \mathfrak{t}_x\big)_* : H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big) \longrightarrow H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big)$$

equals the identity (respectively, minus the identity) for every $x \in U$. The conclusion follows from Proposition 1.9.6. $\square$

1.10.20. REMARK. Observe that during the proof of Proposition 1.10.17 we have actually shown (keeping in mind also Corollary 1.10.18) the following fact: if τ is an orientation for $\mathbb{R}^n$ and if $\alpha \in H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big)$ is mapped to $\tau(0)$ by the homomorphism:

$$H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \mathrm{B}^n\big) \longrightarrow H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big)$$

induced by inclusion then for every $v \in \mathrm{B}^n$ the homomorphism:

$$H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \mathrm{B}^n\big) \longrightarrow H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{v\}\big)$$

induced by inclusion takes α to $\tau(v)$.

1.10.21. EXAMPLE (zero-dimensional manifolds). If M is a zero-dimensional topological manifold (i.e., a discrete topological space) then the orientation bundle $\mathcal{O}(M)$ is also a discrete topological space; namely, for every $x \in M$ the set $U = \{x\}$ is open in M and for every $\alpha \in H_0(M, M \setminus U)$ the basic open set $\mathrm{Im}\big[\mathcal{O}(\alpha; U)\big]$ is the singleton $\{\alpha\}$. Thus, every section $\tau : M \to \mathcal{O}(M)$ of $\mathcal{O}(M)$ is continuous. Moreover, for every $x \in M$ the local homology group $H_0\big(M, M \setminus \{x\}\big)$ has a canonical generator, namely, the homology class of the singular 0-simplex x. If we identify the generators x and $-x$ of $H_0\big(M, M \setminus \{x\}\big)$ respectively with 1 and -1 then choosing an orientation for a zero-dimensional topological manifold M becomes the same as choosing an arbitrary map $\tau : M \to \{-1, 1\}$.

1.10.22. EXAMPLE (orientation on the sphere). For every $x \in S^n$, since the space $S^n \setminus \{x\}$ is contractible, the homomorphism $\widetilde{H}_n(S^n) \to H_n\big(S^n, S^n \setminus \{x\}\big)$ induced by inclusion is an isomorphism. In other words, the homomorphisms:

$$\rho_{S^n x} : H_n(S^n) \longrightarrow H_n\big(S^n, S^n \setminus \{x\}\big),$$

are isomorphisms for $n \ge 1$ and in the case $n = 0$, the restriction of $\rho_{S^n x}$ to $\widetilde{H}_n(S^n)$ is an isomorphism. Thus if α is a generator of the infinite cyclic group $\widetilde{H}_n(S^n)$ then $\mathcal{O}(\alpha; S^n)$ is an orientation for S^n (see Lemma 1.10.7). If $n \ge 1$ then Proposition 1.10.11 implies that we have a one-to-one correspondence:

$$\big\{\text{generators of } \widetilde{H}_n(S^n)\big\} \ni \alpha \longmapsto \mathcal{O}(\alpha; S^n) \in \big\{\text{orientations of } S^n\big\}, \tag{1.10.6}$$

between the (two element set of) generators of $\widetilde{H}_n(S_n) = H_n(S_n)$ and the set of orientations of S^n. For $n = 0$, the sphere $S^0 = \{-1, 1\}$ has actually *four* orientations (see Example 1.10.21), so that the image of the injective map (1.10.6) contains only two of them (namely, those attaching opposite signs to the two points of S^0); we will not be interested in the other two orientations of S^0. Hence, from now on, we shall identify the orientations of S^n with their corresponding generators of $\widetilde{H}_n(S^n)$ via the correspondence (1.10.6); more explicitly, if α is a generator of $\widetilde{H}_n(S^n)$ then for every $x \in S^n$ we will write $\alpha(x)$ for the image of α by the isomorphism $\widetilde{H}_n(S^n) \to H_n\big(S^n, S^n \setminus \{x\}\big)$ induced by inclusion.

1.10.23. REMARK. Regarding the convention we made in Example 1.10.22 of identifying orientations of S^n with generators of $\widetilde{H}_n(S^n)$, we observe in addition that a homeomorphism $h : S^n \to S^n$ is positively oriented (respectively, negatively oriented) if and only if the automorphism h_* of $\widetilde{H}_n(S^n)$ is the identity (respectively, minus the identity). This follows easily from the commutativity of the diagram:

$$\begin{tikzcd} \widetilde{H}_n(S^n) \arrow[r, "h_*"] \arrow[d, "\rho_{S^n x}"'] & \widetilde{H}_n(S^n) \arrow[d, "\rho_{S^n h(x)}"] \\ H_n\big(S^n, S^n \setminus \{x\}\big) \arrow[r, "h_*"'] & H_n\big(S^n, S^n \setminus \{h(x)\}\big) \end{tikzcd}$$

We now study the relations between the notion of orientation introduced in this section (let's call it *homological orientation* for the moment) and the standard notion of orientation for differentiable manifolds defined in terms of orientations for the tangent spaces (let's call it *differentiable orientation* for the moment). Most of the work is encoded in Corollary 1.10.19. A basic difficulty that appears right away when one tries to relate homological and differentiable orientation is the following: the model space for manifolds, i.e., the Euclidean space $\mathbb{R}^n$ has a canonical differentiable orientation (corresponding to the vector space orientation defined by the canonical basis) while it has in principle no obvious choice for a homological orientation. The natural way around this difficulty is to make a choice (once and for all) for an orientation on $\mathbb{R}^n$ that will be called "canonical"; we thus make the following:

1.10.24. CONVENTION. Let us choose a homological orientation $\tau^{[n]} : \mathbb{R}^n \to \mathcal{O}(\mathbb{R}^n)$ for $\mathbb{R}^n$. If $n = 0$ we orient the unique point of $\mathbb{R}^0$ with a plus sign (see Example 1.10.21), i.e., we simply take $\tau^{[0]}(0) \in H_0(\mathbb{R}^0)$ to be the homology class of the singular 0-simplex determined by the origin. Assume now that $n \geq 1$. Proposition 1.10.17 tells us that an orientation $\tau^{[n]}$ for $\mathbb{R}^n$ is obtained if one chooses a generator $\tau^{[n]}(0)$ for $H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big)$ and then set $\tau^{[n]}(v) = (\mathfrak{t}_v)_*\big(\tau^{[n]}(0)\big)$ for all $v \in \mathbb{R}^n$. Let us now choose $\tau^{[n]}(0)$. We start by fixing an isomorphism between $H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big)$ and $\widetilde{H}_{n-1}(S^{n-1})$. We choose the isomorphism used in Example 1.9.4 to compute the local homology of $\mathbb{R}^n$; namely, we consider the isomorphism given by the dotted arrow in the commutative diagram:

(1.10.7)
$$\begin{tikzcd} H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big) \arrow[dr, dotted, "\cong"] & \\ H_n\big(\overline{\mathrm{B}}^n, S^{n-1}\big) \arrow[u, "\cong"] \arrow[r, "\cong", "\partial_*"'] & \widetilde{H}_{n-1}(S^{n-1}) \end{tikzcd}$$

where the unlabelled vertical arrow is induced by inclusion. Finally, we choose a generator $\alpha^{[n]}$ for $\widetilde{H}_n(S^n)$ for every $n \geq 0$ and then, for every $n \geq 1$, we take $\tau^{[n]}(0)$ to be the inverse image of $\alpha^{[n-1]}$ by the dotted arrow in (1.10.7). We now define $\alpha^{[n]}$ recursively. We choose the generator $\alpha^{[0]}$ of $\widetilde{H}_0(S^0) = \widetilde{H}_0\big(\{-1, 1\}\big)$ by taking a plus sign on $1 \in S^0$ and a minus sign in $-1 \in S^0$. Assuming that $\alpha^{[n-1]}$ is defined for some $n \geq 1$, we take $\alpha^{[n]} \in \widetilde{H}_n(S^n)$ to be the element that is mapped to $(-1)^{n-1}\alpha^{[n-1]}$ by the isomorphism $\widetilde{H}_n(S^n) \cong \widetilde{H}_{n-1}(S^{n-1})$ defined in Example 1.8.1, i.e., the composition of the isomorphisms (1.8.1)—(1.8.4) (with $p = n$).

From now on, we will call $\tau^{[n]}$ the *canonical* homological orientation of $\mathbb{R}^n$ and $\alpha^{[n]}$ the *canonical* homological orientation of S^n (see also Example 1.10.22).

Let now M be an n-dimensional differentiable manifold with[7] $n \geq 1$. A *differentiable orientation* for M at a point $x \in M$ is by definition a vector space orientation for the tangent space T_xM.

Let $x \in M$ and let $\varphi : U \to \widetilde{U}$ be a (smooth) chart for M with U an open neighborhood of x in M and $\widetilde{U}$ an open subset of $\mathbb{R}^n$; set $\tilde{x} = \varphi(x)$. The vector space isomorphism $\mathrm{d}\varphi_x : T_xM \to \mathbb{R}^n$ induces a bijection between the (two element) set of vector space orientations of T_xM and the set of vector space orientations of $\mathbb{R}^n$. Moreover, the group isomorphism:

$$\varphi_* : H_n\big(M, M \setminus \{x\}\big) \longrightarrow H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{\tilde{x}\}\big)$$

induces a bijection between the (two element) set of generators of $H_n\big(M, M\setminus\{x\}\big)$ and the set of generators of $H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{\tilde{x}\}\big)$. We have a canonical bijection between the set of vector space orientations of $\mathbb{R}^n$ and the set of generators of $H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{\tilde{x}\}\big)$; namely, this bijection takes the orientation of $\mathbb{R}^n$ induced by the canonical basis to the generator $\tau^{[n]}(\tilde{x})$. The chart φ therefore induces a bijection between the set of vector space orientations of T_xM and the set of generators of $H_n\big(M, M \setminus \{x\}\big)$; namely, such bijection is the dotted arrow in the commutative diagram:

(1.10.8)
$$\begin{tikzcd}
\boxed{\begin{matrix}\text{Orientations}\\ \text{of } T_xM\end{matrix}} \arrow[r, dotted] \arrow[d, "\begin{matrix}\text{induced}\\ \text{by } \mathrm{d}\varphi_x\end{matrix}"'] & \boxed{\begin{matrix}\text{Generators of}\\ H_n\big(M, M \setminus \{x\}\big)\end{matrix}} \arrow[d, "\begin{matrix}\text{induced}\\ \text{by } \varphi_*\end{matrix}"] \\
\boxed{\begin{matrix}\text{Orientations}\\ \text{of } \mathbb{R}^n\end{matrix}} \arrow[r, "\text{canonical} \mapsto \tau^{[n]}(\tilde{x})"'] & \big\{\tau^{[n]}(\tilde{x}), -\tau^{[n]}(\tilde{x})\big\}
\end{tikzcd}$$

The crucial point here is that the top arrow of the diagram above *does not* depend on the choice of the chart φ. To prove that, choose another chart ψ of M around x and set $\tilde{x}' = \psi(x)$; we denote by $f = \psi \circ \varphi^{-1}$ the transition map from φ to ψ so that f is a diffeomorphism between open subsets of $\mathbb{R}^n$ and $f(\tilde{x}) = \tilde{x}'$. Consider

---

[7] If M has dimension zero, one usually takes by convention that a *differentiable orientation* on M is simply an arbitrary map $\tau : M \to \{-1, 1\}$. By Example 1.10.21 this is actually compatible with the notion of homological orientation for zero-dimensional manifolds.

the following diagram:

$$\begin{tikzcd}
& \boxed{\begin{matrix}\text{Orientations}\\ \text{of } T_xM\end{matrix}} \arrow[rr, dotted] \arrow[d, "\mathrm{d}\varphi_x"] \arrow[ddl, bend right, "\mathrm{d}\psi_x"'] & & \boxed{\begin{matrix}\text{Generators of}\\ H_n\big(M, M\setminus\{x\}\big)\end{matrix}} \arrow[d, "\varphi_*"'] \arrow[ddr, bend left, "\psi_*"] & \\
& \boxed{\begin{matrix}\text{Orientations}\\ \text{of } \mathbb{R}^n\end{matrix}} \arrow[rr, dashed] \arrow[dl, dashed, "\mathrm{d}f_{\tilde{x}}"] & & \big\{\tau^{[n]}(\tilde{x}), -\tau^{[n]}(\tilde{x})\big\} \arrow[dr, dashed, "f_*"'] & \\
\boxed{\begin{matrix}\text{Orientations}\\ \text{of } \mathbb{R}^n\end{matrix}} \arrow[rrrr, dashed] & & & & \Big\{\tau^{[n]}(\tilde{x}'), -\tau^{[n]}(\tilde{x}')\Big\}
\end{tikzcd}$$

where the dotted arrow at the top of the diagram is the bijection induced by the chart φ; by definition, the square in the back of the diagram commutes. Clearly, the triangle on the left side of the diagram commutes by the chain rule and the triangle on the right side commutes by the functoriality of singular homology. The dashed square at the bottom of the diagram also commutes by Corollary 1.10.19; it follows that the front square commutes and therefore the dotted arrow coincides with the bijection induced by ψ.

We have proven that, given a differentiable manifold M, then for every $x \in M$ there exists a *canonical* bijection between the set of orientations of M at x in the homological sense (i.e., the set of generators of $H_n\big(M, M \setminus \{x\}\big)$) and the set of orientations of M at x in the differentiable sense (i.e., the set of vector space orientations of T_xM).

Now let $\tau : M \to \mathcal{O}(M)$ be a section of the orientation bundle such that $\tau(x)$ is a generator of $H_n\big(M, M \setminus \{x\}\big)$ for all $x \in M$; let $\bar{\tau}(x)$ be the vector space orientation of T_xM that corresponds to τ_x. To finish our comparison between homological and differentiable orientation we sill have to show that τ is continuous if and only if the family $(\bar{\tau}_x)_{x\in M}$ define a differentiable orientation for M; we briefly recall below what the latter condition means.

1.10.25. DEFINITION. Let M be an n-dimensional differentiable manifold with $n \geq 1$. Assume that for every $x \in M$ one chooses a vector space orientation $\bar{\tau}_x$ for T_xM. A (smooth) chart $\varphi : U \subset M \to \widetilde{U} \subset \mathbb{R}^n$ is called *positively oriented* for the family $\bar{\tau} = (\bar{\tau}_x)_{x\in M}$ if for every $x \in U$ the isomorphism $\mathrm{d}\varphi_x : T_xM \to \mathbb{R}^n$ carries the orientation $\bar{\tau}_x$ of T_xM to the canonical orientation of $\mathbb{R}^n$. We say that the family $\bar{\tau}$ *defines an orientation* for M (in the differentiable sense) if M admits an atlas of positively oriented charts, i.e., if M can be covered by the domains U of the positively oriented charts $\varphi : U \to \widetilde{U}$.

We can now finally prove the following:

1.10.26. PROPOSITION. *Let M be an n-dimensional differentiable manifold with $n \geq 1$. Let $\tau : M \to \mathcal{O}(M)$ be a section of the orientation bundle $\mathcal{O}(M)$ such that $\tau(x)$ is a generator of $H_n\big(M, M \setminus \{x\}\big)$ for every $x \in M$; denote by $\bar{\tau}_x$ the vector space orientation of T_xM that corresponds to $\tau(x)$ by the rule explained in diagram* (1.10.8). *Then the family $\bar{\tau} = (\bar{\tau}_x)_{x\in M}$ defines an orientation for M (in* *the differentiable sense) if and only if τ is an orientation for M (in the homological sense), i.e., if and only if τ is a continuous section of $\mathcal{O}(M)$.*

PROOF. Assume that $\bar\tau$ defines an orientation for M. For every $x \in M$ we can find a positively oriented (smooth) chart $\varphi : U \to \widetilde{U}$ with $x \in U$. Then $\mathrm{d}\varphi_y$ carries the orientation $\bar\tau_y$ of T_yM to the canonical orientation of $\mathbb{R}^n$ for every $y \in U$; hence the isomorphism:

$$\varphi_* : H_n\big(M, M \setminus \{y\}\big) \longrightarrow H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{\varphi(y)\}\big) \tag{1.10.9}$$

carries $\tau(y)$ to $\tau^{[n]}\big(\varphi(y)\big)$. Therefore we have a commutative diagram:

$$\begin{tikzcd}
\mathcal{O}(M) \supset \mathcal{O}(U) \arrow[r, "\mathcal{O}(\varphi)"] & \mathcal{O}\big(\widetilde{U}\big) \subset \mathcal{O}(\mathbb{R}^n) \\
M \supset U \arrow[u, "\tau|_U"] \arrow[r, "\varphi"'] & \widetilde{U} \subset \mathbb{R}^n \arrow[u, "\tau^{[n]}|_{\widetilde{U}}"']
\end{tikzcd}$$

Since $\mathcal{O}(\varphi)$ is a homeomorphism (see Proposition 1.10.12), it follows that $\tau|_U$ is continuous; but x is arbitrary and therefore τ is continuous.

Conversely, assume that τ is continuous. Let $x \in M$ and let $\varphi : U \to \widetilde{U}$ be a (smooth) chart with U a connected open neighborhood of x. By Proposition 1.10.15, φ is either a positive or a negative homeomorphism; by composing φ on the left with a negative isomorphism of $\mathbb{R}^n$ if necessary, we can assume that φ is positive. This means that (1.10.9) carries $\tau(y)$ to $\tau^{[n]}\big(\varphi(y)\big)$ for every $y \in U$ and therefore the isomorphism $\mathrm{d}\varphi_y$ carries $\bar\tau_y$ to the canonical orientation of $\mathbb{R}^n$. Thus φ is positively oriented for $\bar\tau$. $\square$

We have completed the prove of the equivalence between the notions of homological and differentiable orientation. Actually, one should prove now (and that's very easy) that a diffeomorphism between oriented differentiable manifolds is positively oriented in the differentiable sense if and only if it is positively oriented in the homological sense (see Exercise 1.68).

In our Convention 1.10.24 we have fixed the canonical orientation $\tau^{[n]}$ for $\mathbb{R}^n$ and the canonical orientation $\alpha^{[n]}$ for the sphere S^n. But to what differentiable orientations do this conventions correspond? Well, it is pretty obvious that $\tau^{[n]}$ corresponds to the canonical differentiable orientation of $\mathbb{R}^n$, i.e., the one induced from the canonical basis. But what about $\alpha^{[n]}$? We will have to work a little to answer that. First, let's fix some terminology.

1.10.27. DEFINITION. For every $n \geq 0$ the *outward pointing orientation* on S^n is defined as follows; for $n = 0$ we simply take a plus sign for the point $1 \in S^0$ and a minus sign for the point $-1 \in S^0$. If $n \geq 1$ then for every $x \in S^n \subset \mathbb{R}^{n+1}$ we orient T_xM in such a way that $(x, b_1, \ldots, b_n)$ is a positively oriented basis of $\mathbb{R}^{n+1}$ for every positively oriented basis $(b_1, \ldots, b_n)$ of T_xM.

It is a very elementary exercise to check that the outward pointing orientation is indeed a differentiable orientation for S^n.

Now we can compare explicitly the homological and the differentiable orientations of the sphere.

1.10.28. PROPOSITION. *For every $n \ge 0$, the differentiable orientation associated to the canonical homological orientation $\alpha^{[n]}$ of S^n is the outward pointing orientation (recall Example 1.10.22, Convention 1.10.24 and Proposition 1.10.26).*

PROOF. If $n = 0$ there is nothing to do, so assume $n \ge 1$. By Proposition 1.10.11 it suffices to check that the homological orientation corresponding to the outward pointing orientation equals $\alpha^{[n]}$ at one specific point, say the south pole. We use the notation of Example 1.8.1. Let $\varphi : S^n \setminus \{\mathfrak{n}\} \to \mathbb{R}^n$ denote the stereographic projection from the north pole onto the plane containing the equator, i.e., for every $x \in S^n$, $x \ne \mathfrak{n}$, $\varphi(x)$ is the unique point of the half-line $\big\{\mathfrak{n} + t(x - \mathfrak{n}) : t \ge 0\big\}$ that belongs to the hyper-plane $\mathbb{R}^n \times \{0\} \subset \mathbb{R}^{n+1}$. If S^n is endowed with the outward pointing orientation then a straightforward computation (see Exercise 1.70) shows that $\mathrm{d}\varphi_{\mathfrak{s}} : T_{\mathfrak{s}} S^n \to \mathbb{R}^n$ is a positive isomorphism for odd n and it is a negative isomorphism for even n; hence the proof will be concluded if we can show that:

$$\varphi_*\big(\alpha^{[n]}(\mathfrak{s})\big) = (-1)^{n-1} \tau^{[n]}(0).$$

Consider the following diagram of abelian groups and isomorphisms:

$$\begin{tikzcd}
 & \widetilde{H}_{n-1}\big(S^{n-1}\big) \arrow[r, "i_*"] \arrow[dl, dotted, bend right] & \widetilde{H}_{n-1}\big(S^n \setminus \{\mathfrak{n}, \mathfrak{s}\}\big) \\
H_n(S^n) \arrow[r] \arrow[drr, dashed, bend right] & H_n(S^n, S^n \setminus \{\mathfrak{s}\}) & H_n\big(S^n \setminus \{\mathfrak{n}\}, S^n \setminus \{\mathfrak{n}, \mathfrak{s}\}\big) \arrow[l] \arrow[u, "\partial_*"] \arrow[d, "\varphi_*"] \\
 & & H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big)
\end{tikzcd}$$

where i_* and the unlabelled arrows are induced by inclusion. The dotted path in the diagram above corresponds precisely to the isomorphism between $\widetilde{H}_{n-1}(S^{n-1})$ and $H_n(S^n)$ describe in Example 1.8.1, i.e., the composition of the isomorphisms (1.8.1)—(1.8.4). Hence such dotted path carries $\alpha^{[n-1]}$ to $(-1)^{n-1}\alpha^{[n]}$. Moreover, the dashed path in the diagram carries $\alpha^{[n]}$ to $\varphi_*\big(\alpha^{[n]}\big[\mathfrak{s}]\big)$; it follows that:

$$(\varphi_* \circ \partial_*^{-1} \circ i_*)\big(\alpha^{[n-1]}\big) = (-1)^{n-1} \varphi_*\big(\alpha^{[n]}(\mathfrak{s})\big). \tag{1.10.10}$$

We will show now that the lefthand side of (1.10.10) equals $\tau^{[n]}(0)$. To this aim, consider the commutative diagram:

$$
\begin{tikzcd}
\widetilde{H}_{n-1}\big(S^n \setminus \{\mathfrak{n}, \mathfrak{s}\}\big) & H_n\big(S^n \setminus \{\mathfrak{n}\}, S^n \setminus \{\mathfrak{n}, \mathfrak{s}\}\big) \arrow[l, "\partial_*"'] \arrow[r, "\varphi_*"'] & H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big) \\
\widetilde{H}_{n-1}(S^{n-1}) \arrow[u] \arrow[d, "\varphi_* = \mathrm{Id}"'] \arrow[urr, dotted, bend left] & H_n\big(S^n \setminus \{\mathfrak{n}\}, S^{n-1}\big) \arrow[l, "\partial_*"'] \arrow[u] \arrow[d, "\varphi_*"] & \\
\widetilde{H}_{n-1}(S^{n-1}) & H_n\big(\mathbb{R}^n, S^{n-1}\big) \arrow[l, "\partial_*"'] & \\
\widetilde{H}_{n-1}(S^{n-1}) \arrow[u, "\mathrm{Id}"] \arrow[uuurr, dashed, bend right] & H_n\big(\overline{\mathrm{B}}^n, S^{n-1}\big) \arrow[l, "\partial_*"'] \arrow[u] \arrow[uuur, bend right] &
\end{tikzcd}
$$

where the unlabelled arrows are induced by inclusion. The conclusion is obtained by observing that the dotted path in the diagram above takes $\alpha^{[n-1]}$ to the lefthand side of (1.10.10) while the dashed path takes $\alpha^{[n-1]}$ to $\tau^{[n]}(0)$. $\square$

Let us now study orientations on manifolds with boundary (see Exercise 1.63 for the exact definition and the terminology we adopt). In the case of *differentiable* manifolds with boundary, there is no real additional difficulty in comparison with the case of manifolds without boundary; namely, there is a well-defined notion of tangent space also at the points of the boundary and one can consider vector space orientations on such tangent spaces. Moreover, in the differentiable case, it is well known (for instance by those who have studied Stoke's theorem on manifolds) that an orientation on a manifold with boundary induces canonically an orientation on the boundary; namely, one uses the canonical *transversal* orientation of the boundary, given by the outward pointing tangent vector. In the case of topological manifolds with boundary, there is a difficulty with the homological approach for orientation; namely, all the local homology groups vanish at the boundary points. We use the following strategy to go around this difficulty: we simply don't talk about oriented topological manifolds with boundary — we just talk about orientations for the *interior* of the manifold with boundary (which is a manifold without boundary). Nevertheless, we have to clarify how an orientation on the interior of a topological manifold with boundary induces an orientation on the boundary of the manifold. Such notion of *induced orientation on the boundary* will be achieved by an elegant trick using the connecting homomorphism ∂_* of the long exact homology sequence of a pair (keep in mind the isomorphism $\partial_* : H_n\big(\overline{\mathrm{B}}^n, S^{n-1}\big) \to \widetilde{H}_{n-1}(S^{n-1})$ as a model for the general construction we explain below).

In what follows, M will always denote an n-dimensional topological manifold with *non empty* boundary (in particular, n cannot be zero). Recall from Exercise 1.63 that if U is a non empty open subset of M then U is also an n-dimensional topological manifold with boundary and:

$$\mathrm{inter}(U) = \mathrm{inter}(M) \cap U, \quad \mathrm{Bd}(U) = \mathrm{Bd}(M) \cap U.$$

For every open subset U in M and every $x \in \mathrm{inter}(U)$ we define, in analogy with (1.10.1), a homomorphism:

$$\rho_{Ux} : H_n\big(M, M \setminus \mathrm{inter}(U)\big) \longrightarrow H_n\big(\mathrm{inter}(M), \mathrm{inter}(M) \setminus \{x\}\big)$$

by requiring the commutativity of the diagram:

$$\begin{tikzcd} H_n\big(M, M \setminus \mathrm{inter}(U)\big) \arrow[r] \arrow[dr, "\rho_{Ux}"'] & H_n\big(M, M \setminus \{x\}\big) \\ & H_n\big(\mathrm{inter}(M), \mathrm{inter}(M) \setminus \{x\}\big) \arrow[u, "\cong"'] \end{tikzcd}$$

in which the unlabelled arrows are induced by inclusion. Now, for every $\alpha \in H_n\big(M, M \setminus \mathrm{inter}(U)\big)$ we define (in analogy with (1.10.3)) a section $\mathcal{O}_{\mathrm{i}}(\alpha; U, M)$ of the orientation bundle $\mathcal{O}\big(\mathrm{inter}(M)\big)$ along the open set $\mathrm{inter}(U) \subset \mathrm{inter}(M)$ by setting:

$$\mathcal{O}_{\mathrm{i}}(\alpha; U, M)(x) = \rho_{Ux}(\alpha),$$

for all $x \in \mathrm{inter}(U)$. When M is fixed by the context we write simply $\mathcal{O}_{\mathrm{i}}(\alpha; U)$ instead of $\mathcal{O}_{\mathrm{i}}(\alpha; U, M)$.

Assume now that x belongs to the boundary of the open set $U \subset M$, i.e., $x \in \mathrm{Bd}(U)$. Observe that $\mathrm{Bd}(M)$ is a neighborhood of x in the topological space $M \setminus \mathrm{inter}(U)$; namely, $\mathrm{Bd}(U) = U \cap \big(M \setminus \mathrm{inter}(U)\big)$ is an open set in the space $M \setminus \mathrm{inter}(U)$ that contains x and is contained in $\mathrm{Bd}(M)$. It follows that the local homology groups of $M \setminus \mathrm{inter}(U)$ and of $\mathrm{Bd}(M)$ at x are isomorphic (by the usual isomorphism induced by inclusion); we can thus define a homomorphism:

$$\mathcal{J}_{Ux} : H_n\big(M, M \setminus \mathrm{inter}(U)\big) \longrightarrow H_{n-1}\big(\mathrm{Bd}(M), \mathrm{Bd}(M) \setminus \{x\}\big)$$

by requiring the commutativity of the diagram:
(1.10.11)

$$\begin{tikzcd} H_n\big(M, M \setminus \mathrm{inter}(U)\big) \arrow[r, "\partial_*"] \arrow[d, "\mathcal{J}_{Ux}"'] & H_{n-1}\big(M \setminus \mathrm{inter}(U)\big) \arrow[d] \\ H_{n-1}\big(\mathrm{Bd}(M), \mathrm{Bd}(M) \setminus \{x\}\big) \arrow[r, "\cong"] & H_{n-1}\big(M \setminus \mathrm{inter}(U), [M \setminus \mathrm{inter}(U)] \setminus \{x\}\big) \end{tikzcd}$$

in which the unlabelled arrows are induced by inclusion and the top arrow ∂_* is the connecting homomorphism of the long exact homology sequence of the pair $\big(M, M \setminus \mathrm{inter}(U)\big)$. If $\alpha \in H_n\big(M, M \setminus \mathrm{inter}(U)\big)$ is fixed then the homomorphisms $\mathcal{J}_{Ux}$ can be joined together to form a section $\mathcal{O}_{\mathrm{b}}(\alpha; U, M)$ of $\mathcal{O}\big(\mathrm{Bd}(M)\big)$ along $\mathrm{Bd}(U)$; more explicitly, we set:

$$\mathcal{O}_{\mathrm{b}}(\alpha; U, M)(x) = \mathcal{J}_{Ux}(\alpha),$$

for all $x \in \mathrm{Bd}(U)$. Again, if M is fixed by the context we write simply $\mathcal{O}_{\mathrm{b}}(\alpha; U)$.

Using the terminology introduced above, we can give the following:

1.10.29. DEFINITION. If τ is an orientation for $\mathrm{inter}(M)$ then an orientation τ^{b} for $\mathrm{Bd}(M)$ is called *induced* from τ if for every point $x \in \mathrm{Bd}(M)$ we can find an open set U in M containing x and a homology class $\alpha \in H_n\big(M, M \setminus \mathrm{inter}(U)\big)$ such that:

(1.10.12)
$$\tau|_{\mathrm{inter}(U)} = \mathcal{O}_{\mathrm{i}}(\alpha; U), \quad \tau^{\mathrm{b}}|_{\mathrm{Bd}(U)} = \mathcal{O}_{\mathrm{b}}(\alpha; U).$$

Our task now will be to prove that for every orientation τ of $\mathrm{inter}(M)$ there is a unique orientation τ^{b} on $\mathrm{Bd}(M)$ that is induced from τ; after this fact is established we shall simply say that τ^{b} is *the orientation induced by τ on the boundary of M*.

We start by stating some simple naturality results regarding the homomorphisms ρ_{Ux} and $\mathcal{J}_{Ux}$.

1.10.30. LEMMA. *If U, $V \subset M$ are open subsets with $V \subset U$ then for every $x \in \mathrm{inter}(V)$, $y \in \mathrm{Bd}(V)$ the diagrams:*

(1.10.13)
$$\begin{tikzcd}
H_n\big(M, M \setminus \mathrm{inter}(U)\big) \arrow[dd] \arrow[dr, "\rho_{Ux}"] & \\
& H_n\big(\mathrm{inter}(M), \mathrm{inter}(M) \setminus \{x\}\big) \\
H_n\big(M, M \setminus \mathrm{inter}(V)\big) \arrow[ur, "\rho_{Vx}"'] &
\end{tikzcd}$$

(1.10.14)
$$\begin{tikzcd}
H_n\big(M, M \setminus \mathrm{inter}(U)\big) \arrow[dd] \arrow[dr, "\mathcal{J}_{Uy}"] & \\
& H_{n-1}\big(\mathrm{Bd}(M), \mathrm{Bd}(M) \setminus \{y\}\big) \\
H_n\big(M, M \setminus \mathrm{inter}(V)\big) \arrow[ur, "\mathcal{J}_{Vy}"'] &
\end{tikzcd}$$

commute, where the unlabelled vertical arrows are induced by inclusion.

In particular, if $\alpha \in H_n\big(M, M \setminus \mathrm{inter}(U)\big)$ is a homology class and $\alpha' \in H_n\big(M, M \setminus \mathrm{inter}(V)\big)$ is the image of α by the homomorphism:

(1.10.15)
$$H_n\big(M, M \setminus \mathrm{inter}(U)\big) \longrightarrow H_n\big(M, M \setminus \mathrm{inter}(V)\big)$$

induced by inclusion then $\mathcal{O}_{\mathrm{i}}(\alpha'; V)$ is the restriction of $\mathcal{O}_{\mathrm{i}}(\alpha; U)$ to $\mathrm{inter}(V)$ and $\mathcal{O}_{\mathrm{b}}(\alpha'; V)$ is the restriction of $\mathcal{O}_{\mathrm{b}}(\alpha; U)$ to $\mathrm{Bd}(V)$.

PROOF. This is basically a consequence of the fact that the homomorphisms we used to assemble the ρ's and the $\mathcal{J}$'s are all natural with respect to inclusions. For example, in order to prove the commutativity of (1.10.14) one can draw a cubic diagram as follows: the bottom face of the cube is diagram (1.10.11), the top face of the cube is diagram (1.10.11) with U replaced by V; the top and the bottom faces of the cube are connected by (downward pointing) vertical arrows which are all induced by inclusion. One has now to observe that five faces of this cube are commutative and then use this fact to conclude the commutativity of the sixth face, which relates the maps $\mathcal{J}_{Ux}$ and $\mathcal{J}_{Vx}$. The proof of the commutativity of (1.10.13) is obtained in a similar way, considering a diagram having the form of a prism of triangular basis. The (boring) diagram-chase details are left to the reader. □

Observe that Lemma 1.10.30 implies that if one can find U and α that satisfy equalities (1.10.12) then for every smaller open set $V \subset U$ one can find α' (the image of α by (1.10.15)) such that (1.10.12) is satisfied with U replaced by V and α replaced by α'. In particular, we obtain the following:

1.10.31. COROLLARY. *If an orientation τ^{b} on $\mathrm{Bd}(M)$ is induced from an orientation τ on $\mathrm{inter}(M)$ then for every open set $W \subset M$ with $\mathrm{Bd}(W) \neq \emptyset$, the restriction of τ^{b} to $\mathrm{Bd}(W) = \mathrm{Bd}(M) \cap W$ is induced from the restriction of τ to $\mathrm{inter}(W) = \mathrm{inter}(M) \cap W$.*

PROOF. For every $x \in \mathrm{Bd}(W)$ one can find an open set $U \subset M$ containing x and a homology class $\alpha \in H_n\big(M, M \setminus \mathrm{inter}(U)\big)$ satisfying (1.10.12); as observed above, one can pick a smaller U such that $\overline{U} \subset W$. By excision, we know that the homomorphism:

$$H_n\big(W, W \setminus \mathrm{inter}(U)\big) \longrightarrow H_n\big(M, M \setminus \mathrm{inter}(U)\big) \tag{1.10.16}$$

induced by inclusion is an isomorphism; we can thus find a homology class $\beta \in H_n\big(W, W \setminus \mathrm{inter}(U)\big)$ that is mapped by (1.10.16) onto α. The conclusion follows from Exercise 1.72. □

1.10.32. LEMMA. *Let $h : M \to N$ be a homeomorphism between n-dimensional topological manifolds with (non empty) boundary, so that h automatically maps $\mathrm{Bd}(M)$ onto $\mathrm{Bd}(N)$ (see Exercise 1.63). For every open subset $U \subset M$ and for every $x \in \mathrm{inter}(U)$, $y \in \mathrm{Bd}(U)$ the diagrams:*

$$\begin{array}{ccc} H_n\big(M, M \setminus \mathrm{inter}(U)\big) & \xrightarrow{\rho_{Ux}} & H_n\big(\mathrm{inter}(M), \mathrm{inter}(M) \setminus \{x\}\big) \\ {\scriptstyle h_*}\big\downarrow & & \big\downarrow{\scriptstyle h_*} \\ H_n\big(N, N \setminus \mathrm{inter}(U')\big) & \xrightarrow{\rho_{U'x'}} & H_n\big(\mathrm{inter}(N), \mathrm{inter}(N) \setminus \{x'\}\big) \end{array}$$

$$\begin{array}{ccc} H_n\big(M, M \setminus \mathrm{inter}(U)\big) & \xrightarrow{\mathcal{J}_{Uy}} & H_n\big(\mathrm{Bd}(M), \mathrm{Bd}(M) \setminus \{y\}\big) \\ {\scriptstyle h_*}\big\downarrow & & \big\downarrow{\scriptstyle h_*} \\ H_n\big(N, N \setminus \mathrm{inter}(U')\big) & \xrightarrow{\mathcal{J}_{U'y'}} & H_n\big(\mathrm{Bd}(N), \mathrm{Bd}(N) \setminus \{y'\}\big) \end{array}$$

commute, where $U' = h(U) \subset N$, $x' = h(x) \in \mathrm{inter}(U')$ and $y' = h(y) \in \mathrm{Bd}(U')$.

In particular, if $\alpha \in H_n\big(M, M \setminus \mathrm{inter}(U)\big)$ is a homology class and $\alpha' \in H_n\big(N, N \setminus \mathrm{inter}(U')\big)$ is the image of α by the homomorphism:

$$h_* : H_n\big(M, M \setminus \mathrm{inter}(U)\big) \longrightarrow H_n\big(N, N \setminus \mathrm{inter}(U')\big)$$

then h also "relates" the maps $\mathcal{O}_{\mathrm{i}}(\alpha;U,M)$ and $\mathcal{O}_{\mathrm{b}}(\alpha;U,M)$ with the maps $\mathcal{O}_{\mathrm{i}}(\alpha';U',N)$ and $\mathcal{O}_{\mathrm{b}}(\alpha';U',N)$ respectively; more precisely, the diagrams:

$$\begin{tikzcd}
\mathcal{O}\big(\mathrm{inter}(M)\big) \arrow[r, "\mathcal{O}\left(h|_{\mathrm{inter}(M)}\right)"] & \mathcal{O}\big(\mathrm{inter}(N)\big) \\
\mathrm{inter}(U) \arrow[u, "\mathcal{O}_{\mathrm{i}}(\alpha;U,M)"] \arrow[r, "h|_{\mathrm{inter}(U)}"'] & \mathrm{inter}(U') \arrow[u, "\mathcal{O}_{\mathrm{i}}(\alpha';U',N)"']
\end{tikzcd}$$

$$\begin{tikzcd}
\mathcal{O}\big(\mathrm{Bd}(M)\big) \arrow[r, "\mathcal{O}\left(h|_{\mathrm{Bd}(M)}\right)"] & \mathcal{O}\big(\mathrm{Bd}(N)\big) \\
\mathrm{Bd}(U) \arrow[u, "\mathcal{O}_{\mathrm{b}}(\alpha;U,M)"] \arrow[r, "h|_{\mathrm{Bd}(U)}"'] & \mathrm{Bd}(U') \arrow[u, "\mathcal{O}_{\mathrm{b}}(\alpha';U',N)"']
\end{tikzcd}$$

commute.

PROOF. This is basically a consequence of the fact that the homomorphisms we used to assemble the ρ's and the $\mathcal{J}$'s are all natural with respect to homeomorphisms (one can also think about cubic and prismic diagrams as explained in the proof of Lemma 1.10.30). The details are left to the reader. □

1.10.33. COROLLARY. *Let $h : M \to N$ be a homeomorphism between n-dimensional topological manifolds with (non empty) boundary, so that automatically $h\big(\mathrm{Bd}(M)\big) = \mathrm{Bd}(N)$. Assume that τ, τ^{b}, σ, σ^{b} are orientations respectively for $\mathrm{inter}(M)$, $\mathrm{Bd}(M)$, $\mathrm{inter}(N)$ and $\mathrm{Bd}(N)$. If the homeomorphisms $h|_{\mathrm{inter}(M)} : \mathrm{inter}(M) \to \mathrm{inter}(N)$ and $h|_{\mathrm{Bd}(M)} : \mathrm{Bd}(M) \to \mathrm{Bd}(N)$ are positively oriented and if τ^{b} is induced from τ then also σ^{b} is induced from σ.*

PROOF. Observe that if an open set $U \subset M$ and a homology class $\alpha \in H_n\big(M, M \setminus \mathrm{inter}(U)\big)$ satisfy equalities (1.10.12) then the open set $U' = h(U)$ and the homology class $\alpha' = h_*(\alpha)$ satisfy:

$$\sigma|_{\mathrm{inter}(U')} = \mathcal{O}_{\mathrm{i}}(\alpha';U'), \quad \sigma^{\mathrm{b}}|_{\mathrm{Bd}(U')} = \mathcal{O}_{\mathrm{b}}(\alpha';U').$$

The conclusion follows. □

We now prove the uniqueness of the induced orientation on the boundary.

1.10.34. LEMMA. *If τ is an orientation for $\mathrm{inter}(M)$ then there exists at most one orientation τ^{b} for $\mathrm{Bd}(M)$ that is induced from τ.*

PROOF. Let τ_1^{b} and τ_2^{b} be both induced from τ. For any fixed $y \in \mathrm{Bd}(M)$ we will show that $\tau_1^{\mathrm{b}}(y) = \tau_2^{\mathrm{b}}(y)$. By the definition of induced orientation, we can find an open neighborhood U_i of y and a homology class $\alpha_i \in H_n\big(M, M \setminus \mathrm{inter}(U_i)\big)$ such that:

$$\tau|_{\mathrm{inter}(U_i)} = \mathcal{O}_{\mathrm{i}}(\alpha_i;U_i), \tag{1.10.17}$$

$$\tau_i^{\mathrm{b}}|_{\mathrm{Bd}(U_i)} = \mathcal{O}_{\mathrm{b}}(\alpha_i;U_i), \tag{1.10.18}$$

for $i = 1, 2$. Using a local chart around y we can find an open neighborhood U of y contained in $U_1 \cap U_2$ such that $\overline{U}$ is homeomorphic to the half closed ball $\overline{\mathrm{B}}^n \cap \mathsf{H}^n$ by a homeomorphism that carries $\mathrm{inter}(U)$ to the half open ball $\mathrm{B}^n \cap \mathrm{inter}(\mathsf{H}^n)$. Observe that for every $x \in \mathrm{inter}(U)$, the topological boundary $\partial\big[\mathrm{inter}(U)\big] = \overline{U} \setminus \mathrm{inter}(U)$ of $\mathrm{inter}(U)$ is a strong deformation retract of $\overline{U} \setminus \{x\}$; it follows that $M \setminus \mathrm{inter}(U)$ is also a strong deformation retract of $M \setminus \{x\}$ and therefore the homomorphism ρ_{Ux} is an isomorphism. Denote by α_i' the image of α_i by the homomorphism:

$$H_n\big(M, M \setminus \mathrm{inter}(U_i)\big) \longrightarrow H_n\big(M, M \setminus \mathrm{inter}(U)\big)$$

induced by inclusion. From (1.10.17) and Lemma 1.10.30 we obtain that:

$$\rho_{Ux}(\alpha_1') = \tau(x) = \rho_{Ux}(\alpha_2'),$$

for every $x \in \mathrm{inter}(U)$ and therefore $\alpha_1' = \alpha_2'$. Finally, using (1.10.18) and Lemma 1.10.30 we obtain:

$$\tau_1^{\mathrm{b}}(y) = \mathcal{J}_{Uy}(\alpha_1') = \mathcal{J}_{Uy}(\alpha_2') = \tau_2^{\mathrm{b}}(y). \qquad \square$$

Observe that we have not yet presented a single example of a situation where an orientation τ^{b} on $\mathrm{Bd}(M)$ is induced from some orientation τ on $\mathrm{inter}(M)$. A simple example is given below.

1.10.35. EXAMPLE. Let M denote the unit closed ball $\overline{\mathrm{B}}^n$ (with $n \geq 1$), so that $\mathrm{Bd}(M)$ is the sphere S^{n-1}. We claim that if τ is the orientation on $\mathrm{inter}(M) = \mathrm{B}^n$ obtained by restricting the canonical orientation $\tau^{[n]}$ of $\mathbb{R}^n$ then the canonical orientation $\tau^{\mathrm{b}} = \alpha^{[n-1]}$ of the sphere S^{n-1} is induced from τ. To prove the claim, let the open set $U \subset M$ be the whole closed ball $\overline{\mathrm{B}}^n$ and let the homology class $\alpha \in H_n\big(M, M \setminus \mathrm{inter}(U)\big) = H_n\big(\overline{\mathrm{B}}^n, S^{n-1}\big)$ be the one that is mapped to the canonical orientation $\alpha^{[n-1]} \in \widetilde{H}_{n-1}(S^{n-1})$ via the isomorphism ∂_* appearing in the long exact homology sequence of the pair $\big(\overline{\mathrm{B}}^n, S^{n-1}\big)$ (that's the horizontal arrow in diagram (1.10.7)). The claim will follow once we show that equality (1.10.12) holds. To this aim, observe first that equality $\tau^{\mathrm{b}}|_{\mathrm{Bd}(U)} = \mathcal{O}_{\mathrm{b}}(\alpha; U)$ means that $\mathcal{J}_{Ux}(\alpha) = \alpha^{[n-1]}(x)$ for every $x \in S^{n-1}$; this is a direct consequence of the definition of $\mathcal{J}_{Ux}$ and of the relation between $\alpha^{[n-1]}$ and $\alpha^{[n-1]}(x)$ (recall Example 1.10.22). Finally, the equality $\tau|_{\mathrm{inter}(U)} = \mathcal{O}_{\mathrm{i}}(\alpha; U)$ means that the homomorphism:

$$H_n\big(\overline{\mathrm{B}}^n, S^{n-1}\big) \longrightarrow H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{v\}\big) \tag{1.10.19}$$

induced by inclusion carries α to $\tau^{[n]}(v)$ for all $v \in \mathrm{B}^n$ (as usual we identify the local homology groups $H_n\big(\mathrm{B}^n, \mathrm{B}^n \setminus \{v\}\big)$ and $H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{v\}\big)$). This last assertion follows easily from Remark 1.10.20 (see Exercise 1.71) by observing that for $v = 0$ the map (1.10.19) takes α to $\tau^{[n]}(0)$ (recall Convention 1.10.24 and diagram (1.10.7)).

We can now finally prove the following:

1.10.36. PROPOSITION. *If τ is an orientation on $\mathrm{inter}(M)$ then there exists a unique orientation τ^{b} on $\mathrm{Bd}(M)$ that is induced by τ.*

PROOF. During the course of this proof we will say that a topological manifold M with non empty boundary is *acceptable* if the statement of the proposition holds for M. Our goal is to prove that all manifolds (with non empty boundary) are acceptable. Observe that the uniqueness of the orientation induced on the boundary was already proven in Lemma 1.10.34. The proof of the existence will be split in three steps.

- *If the boundary of M can be covered by a family $(M_i)_{i\in I}$ of open subsets of M, each of them acceptable, then M is acceptable*; let τ be an orientation for $\mathrm{inter}(M)$. For every $i \in I$ the orientation $\tau|_{\mathrm{inter}(M_i)}$ of $\mathrm{inter}(M_i)$ induces a orientation τ_i^{b} on $\mathrm{Bd}(M_i)$. Moreover, for $i, j \in I$, Corollary 1.10.31 implies that the orientations $\tau_i^{\mathrm{b}}|_{\mathrm{Bd}(M_i\cap M_j)}$ and $\tau_j^{\mathrm{b}}|_{\mathrm{Bd}(M_i\cap M_j)}$ are both induced from $\tau|_{\mathrm{inter}(M_i\cap M_j)}$; thus, by Lemma 1.10.34, we have that $\tau_i^{\mathrm{b}}|_{\mathrm{Bd}(M_i\cap M_j)}$ equals $\tau_j^{\mathrm{b}}|_{\mathrm{Bd}(M_i\cap M_j)}$. We can therefore define a map:
$$\tau^{\mathrm{b}} : \mathrm{Bd}(M) \longrightarrow \mathcal{O}\big(\mathrm{Bd}(M)\big),$$
by requiring that τ^{b} equals τ_i^{b} on $\mathrm{Bd}(M_i)$. It is now easy to check that τ^{b} is indeed an orientation on $\mathrm{Bd}(M)$ and that τ^{b} is induced from τ (see Exercise 1.72).
- *If M is homeomorphic to an acceptable manifold then M is also acceptable*; follows trivially from Corollary 1.10.33.
- *If $\mathrm{inter}(M)$ is connected and M is open in some acceptable orientable manifold N then M is also acceptable*; let τ be an orientation for $\mathrm{inter}(M)$. Choose an orientation τ' for N; by replacing τ' with $-\tau'$ if necessary, we can assume that τ' equals τ at some point of $\mathrm{inter}(M)$. It then follows from Proposition 1.10.11 that $\tau = \tau'|_{\mathrm{inter}(M)}$. Since N is acceptable, we can consider the orientation $(\tau')^{\mathrm{b}}$ on $\mathrm{Bd}(N)$ induced from τ'; by Corollary 1.10.31, the restriction of $(\tau')^{\mathrm{b}}$ to $\mathrm{Bd}(M)$ is induced from $\tau = \tau'|_{\mathrm{inter}(M)}$.

Finally, the thesis of the proposition (i.e., the fact that all manifolds are acceptable) follows from the fact that $\overline{\mathrm{B}}^n$ is acceptable (see Example 1.10.35) and from the fact that every n-dimensional topological manifold with boundary M can be covered by open sets that are homeomorphic to open subsets of $\overline{\mathrm{B}}^n$ having connected interior. □

1.10.37. COROLLARY. *If M is orientable then also $\mathrm{Bd}(M)$ is orientable.* □

In practical situations, how does one determine the orientation induced on the boundary? We answer this question below by given a simple interpretation for the induced orientation on the case of differentiable manifolds.

If M is a differentiable n-dimensional manifold with boundary then the tangent space T_xM (and hence the set of its vector space orientations) is well-defined for every $x \in M$ (even if $x \in \mathrm{Bd}(M)$!). One can thus adapt Definition 1.10.25

to the case of differentiable manifolds with boundary obtaining a concept of differentiable orientation for such manifolds; more explicitly, an orientation (in the differentiable sense) for M is a family $\bar{\tau} = (\bar{\tau}_x)_{x\in M}$ such that each $\bar{\tau}_x$ is a vector space orientation for the tangent space T_xM and such that M admits an atlas of charts $\varphi : U \subset M \to \widetilde{U} \subset \mathbb{R}^n$ that are positively oriented for $\bar{\tau}$ (the definition is the same as before, with the exception that now we accept that $\widetilde{U}$ may not be open in $\mathbb{R}^n$, but open on the half-space H^n).

We recall that for points $x \in \mathrm{Bd}(M)$ the tangent space $T_x\mathrm{Bd}(M)$ has a *canonical transverse orientation* on T_xM, i.e., one can distinguish canonically between the two half-spaces defined by the hyper-plane $T_x\mathrm{Bd}(M)$ in T_xM. More explicitly, one defines that a vector $v \in T_xM$ is *outward pointing* if for some chart $\varphi : U \to \widetilde{U}$, with $U \ni x$ open in M and $\widetilde{U}$ open in H^n, the vector $\mathrm{d}\varphi_x(v) \in \mathbb{R}^n$ does not belong to H^n (i.e., it has negative n-th coordinate). It is not hard to check that if such condition holds for one chart φ around x then it will hold for *every* chart φ around x. Using the notion of outward pointing vectors we can give the following:

1.10.38. DEFINITION. If M is an n-dimensional differentiable manifold with boundary ($n \geq 2$) and $\bar{\tau} = (\bar{\tau}_x)_{x\in M}$ is an orientation for M (in the differentiable sense) then the *outward pointing orientation* on $\mathrm{Bd}(M)$ is the (differentiable) orientation $\bar{\tau}^{\mathrm{b}} = \big(\bar{\tau}_x^{\mathrm{b}}\big)_{x\in\mathrm{Bd}(M)}$ for which the following property holds: if $x \in \mathrm{Bd}(M)$, $v_1 \in T_xM$ is an outward pointing vector and $(v_2, \ldots, v_n)$ is a $\bar{\tau}_x^{\mathrm{b}}$-positive basis for $T_x\mathrm{Bd}(M)$ then $(v_1, v_2, \ldots, v_n)$ is a $\bar{\tau}_x$-positive basis for T_xM.

It is well known that the property given above does define an orientation $\bar{\tau}^{\mathrm{b}}$ on $\mathrm{Bd}(M)$ (this is the orientation on $\mathrm{Bd}(M)$ used to formulate Stoke's theorem on manifolds). Observe that the outward pointing orientation for the sphere S^{n-1} is precisely the outward pointing orientation that the closed ball $\overline{\mathrm{B}}^n$ (endowed with the restriction of the canonical orientation of $\mathbb{R}^n$) induces on its boundary.

As one should be guessing by now, we have the following:

1.10.39. PROPOSITION. *Let M be an n-dimensional differentiable manifold with non empty boundary (with*[8] *$n \geq 2$). If $\bar{\tau} = (\bar{\tau}_x)_{x\in M}$ is a differentiable orientation for M and if τ is the homological orientation on $\mathrm{inter}(M)$ corresponding to $(\bar{\tau}_x)_{x\in\mathrm{inter}(M)}$ then the (homological) orientation τ^{b} on $\mathrm{Bd}(M)$ induced from τ is precisely the one that corresponds to the outward pointing orientation on $\mathrm{Bd}(M)$ induced from $\bar{\tau}$.*

PROOF. The idea of the proof is to compare M locally with the closed ball $\overline{\mathrm{B}}^n$. Let then $x_0 \in \mathrm{Bd}(M)$ be fixed and choose a diffeomorphism $\varphi : U \to V$ from an open connected neighborhood U of x_0 in M onto an open subset V of $\overline{\mathrm{B}}^n$. Assume that $\overline{\mathrm{B}}^n$ is endowed with the differentiable orientation induced from the canonical orientation of $\mathbb{R}^n$ and that S^{n-1} is endowed with the outward pointing orientation. For every $x \in U$ the isomorphisms $\mathrm{d}\varphi_x : T_xM \to T_{\varphi(x)}\overline{\mathrm{B}}^n \cong \mathbb{R}^n$ are either all positive or all negative; for definiteness, let's assume that they are all

[8] See Remark 1.10.40 below for the case $n = 1$.

positive. By Proposition 1.10.28 and Exercise 1.68 the proof will be completed once we manage to show that:

(a) $\varphi|_{\mathrm{Bd}(U)} : \mathrm{Bd}(U) \to \mathrm{Bd}(V)$ is a positively oriented diffeomorphism (in the differentiable sense) when both $\mathrm{Bd}(U)$ and $\mathrm{Bd}(V)$ are endowed with the outward pointing orientation;

(b) $\varphi|_{\mathrm{Bd}(U)} : \mathrm{Bd}(U) \to \mathrm{Bd}(V)$ is a positively oriented homeomorphism (in the homological sense) when $\mathrm{Bd}(U)$ is endowed with the restriction of τ^{b} and $\mathrm{Bd}(V)$ is endowed with the restriction of $\alpha^{[n-1]}$.

To prove (a), observe that for every $x \in \mathrm{Bd}(U)$, the isomorphism $\mathrm{d}\varphi_x$ is positive and it takes outward pointing vectors to outward pointing vectors; thus $\mathrm{d}\varphi_x$ also restricts to a positive isomorphism between the tangent spaces of the boundaries. To prove (b), observe that $\varphi|_{\mathrm{inter}(U)} : \mathrm{inter}(U) \to \mathrm{inter}(V)$ is a positively oriented homeomorphism in the homological sense and hence so is $\varphi|_{\mathrm{Bd}(U)} : \mathrm{Bd}(U) \to \mathrm{Bd}(V)$, by Corollary 1.10.33 and Example 1.10.35. This concludes the proof. □

1.10.40. REMARK (zero-dimensional boundary). Assume that M is a one-dimensional differentiable manifold with boundary, oriented in the differentiable sense. Denote by τ the homological orientation of $\mathrm{inter}(M)$ associated to such differentiable orientation and by τ^{b} the homological orientation on the zero-dimensional manifold $\mathrm{Bd}(M)$ induced from τ. By Example 1.10.21, we may identify τ^{b} with a $\{-1, 1\}$-valued map on the set $\mathrm{Bd}(M)$. We claim that for every $x \in \mathrm{Bd}(M)$, $\tau^{\mathrm{b}}(x) = 1$ (respectively $\tau^{\mathrm{b}}(x) = -1$) if and only if the outward pointing vectors at the point x define the positive orientation (respectively, the negative orientation) of the one-dimensional vector space T_xM. The claim is proved by first observing that such property holds if $M = \overline{\mathrm{B}}^1$ (recall Convention 1.10.24); for general M, one simply use diffeomorphisms to compare open subsets of M with open subsets of $\overline{\mathrm{B}}^1$ as in the proof of Proposition 1.10.39. The details are left to the reader.

1.10.41. EXAMPLE. Let $M \subset \mathbb{R}^2$ be a compact convex polygon[9]. Then M is a 2-dimensional topological manifold with boundary, because M is homeomorphic to the disc $\overline{\mathrm{B}}^2$ via radial projection from an interior point (see Exercise 1.42). The interior of M as a manifold with boundary (respectively, the boundary of M as a manifold with boundary) coincides with the topological interior (respectively, the topological boundary) of M as a subset of $\mathbb{R}^2$. Assume that $\mathrm{inter}(M) = \mathrm{int}(M)$ is endowed with the (restriction of) the canonical orientation $\tau^{[2]}$ of $\mathbb{R}^2$. Let's describe the induced orientation on the boundary of M. Let M' denote the complement in M of the (finite set) consisting of the vertices of M. Then M' is a *differentiable* manifold with boundary, with the differentiable structure that makes it embedded in $\mathbb{R}^2$. Assume that M' is endowed with the canonical differentiable

[9]We don't care much about the precise definition of polygon here because we will be using the content of this example only for the case of the *regular* n-agon (which may be defined as the convex hull of the points $e^{\frac{2k\pi i}{n}} \in \mathbb{C}$, $k = 0, \dots, n-1$).

orientation induced from $\mathbb{R}^2$, so that the corresponding homological orientation on $\mathrm{inter}(M') = \mathrm{inter}(M)$ is just (the restriction of) $\tau^{[2]}$. By Corollary 1.10.31, the orientation that M' induces on $\mathrm{Bd}(M')$ is precisely the restriction of the orientation that M induces on $\mathrm{Bd}(M)$. Therefore, if we determine the orientation that M' induces on $\mathrm{Bd}(M')$ we will have a good description of the orientation that M induces on $\mathrm{Bd}(M)$. Since M' is a differentiable manifold, we can compute the orientation induced on the boundary using Proposition 1.10.39. Let $S \subset \mathrm{Bd}(M')$ denote an open side (i.e., a side without the vertices) of the polygon M. If $x \in S$ and $v \in T_xM' \cong \mathbb{R}^2$ is an outward pointing vector (in this case, this means that $x + \varepsilon v \notin M'$ for small $\varepsilon > 0$) then a vector $w \in T_xS \subset \mathbb{R}^2$ (i.e., a vector parallel to S) defines the positive orientation for T_xS if (v, w) is a positive basis of $\mathbb{R}^2$. Hence, if x_0 and x_1 are the vertices of S and if $(v, x_1 - x_0)$ is a positive basis for $\mathbb{R}^2$ then the map:

$$]0,1[\, \ni t \longmapsto x_0 + t(x_1 - x_0) \in S,$$

is a positively oriented homeomorphism if the interval $]0,1[$ is endowed with (the restriction of) the canonical orientation $\tau^{[1]}$ of $\mathbb{R}$ and S is endowed with (the restriction of) the orientation of $\mathrm{Bd}(M)$ induced from M.

1.11. Degree Theory

The degree of a continuous map, roughly speaking, is an integer valued homotopic invariant that measures how many times a manifold is folded around another one by such map; the concept of degree generalizes the one of *winding number* of a closed curve around a point in the plane (or of a closed curve in the circle S^1). The degree of a map is also a particular case of the more general concept of *intersection number* between a map and a submanifold (the degree corresponds to the case where the submanifold reduces to a single point). When one studies integration of differential forms on differentiable manifolds, the degree of a smooth map f appears as the multiplicative factor that relates the integral of a form ω with the integral of its pull-back $f^*\omega$. The formal definition of degree can be given for instance by techniques of differential topology; one defines the degree of a smooth map f to be an algebraic count of the number of inverse images by f of a regular value. The proof that such number is independent of the regular value and the proof of the homotopy invariance takes some work (typically involving differential forms and Stoke's theorem); the generalization of the notion of degree to continuous maps is carried out using approximation theorems and the homotopy invariance. The use of techniques of algebraic topology (or, more precisely, of homology theory) provides in many cases an almost "magically" simple (although less geometric) definition for the degree of a map. The simplest case is the one concerning maps f from the sphere S^n to itself; the degree of such a map equals the multiplicative factor corresponding to the homomorphism $f_* : \widetilde{H}_n(S^n) \to \widetilde{H}_n(S^n)$ (recalling that $\widetilde{H}_n(S^n) \cong \mathbb{Z}$).

The most general definition of degree can be given for proper maps $f : M \to N$ between oriented topological manifolds of the same dimension (with N connected); such degree is invariant under proper homotopies. The amount of work required to develop degree theory in such general case is a bit extensive, so we prefer to stick with a simpler case that will be sufficient for our purposes[10]. The case we consider will be the one of a continuous map defined on an open subset of the sphere S^n taking values in an oriented n-dimensional topological manifold. In Definition 1.11.1 below, we start by introducing a notion of degree that depends on a fixed point of the counter-domain; under certain conditions, it will be possible to prove that the degree is independent of the choice of such point.

1.11.1. DEFINITION. Let $U \subset S^n$ be an open subset, M an n-dimensional topological manifold, $q \in M$ a point and τ_q an orientation for M at q, i.e., a generator of the local homology group $H_n\big(M, M \setminus \{q\}\big)$. Let $f : U \to M$ be a continuous map such that $f^{-1}(q)$ is compact; we define the *degree of f at the value q with respect to the orientation τ_q* to be the unique integer number $\deg_q(f) \in \mathbb{Z}$ for which the equality:

$$\phi\big(\alpha^{[n]}\big) = \deg_q(f)\tau_q$$

holds, where $\phi : \widetilde{H}_n(S^n) \to H_n\big(M, M \setminus \{q\}\big)$ is the homomorphism obtained by the composition of maps pictured in the diagram:

$$\begin{tikzcd}
\widetilde{H}_n(S^n) \arrow[r] \arrow[drr, bend right, "\phi"'] & H_n\big(S^n, S^n \setminus f^{-1}(q)\big) & \\
& H_n\big(U, U \setminus f^{-1}(q)\big) \arrow[u, "\cong"] \arrow[r, "f_*"] & H_n\big(M, M \setminus \{q\}\big)
\end{tikzcd} \tag{1.11.1}$$

the unlabelled arrows in the diagram above are induced by inclusion.

The fact that the vertical arrow in diagram (1.11.1) is an isomorphism follows by excision, observing that $f^{-1}(q)$ is a closed subset of the open set $U \subset S^n$.

1.11.2. REMARK. If K is any compact subset of U containing $f^{-1}(q)$ then one could replace the two occurrences of $f^{-1}(q)$ with K on diagram (1.11.1) obtaining a new commutative diagram:

$$\begin{tikzcd}
\widetilde{H}_n(S^n) \arrow[r] \arrow[drr, bend right, "\phi_K"'] & H_n\big(S^n, S^n \setminus K\big) & \\
& H_n\big(U, U \setminus K\big) \arrow[u, "\cong"] \arrow[r, "f_*"] & H_n\big(M, M \setminus \{q\}\big)
\end{tikzcd} \tag{1.11.2}$$

where (as before) the unlabelled arrows are induced by inclusion. Observe that, since K is closed in S^n, we can still use excision to conclude that the vertical arrow is an isomorphism; moreover, since K contains $f^{-1}(q)$, the map f indeed

[10]We want to use degree theory to give an explicit method of computing the cellular complex of a CW-complex. See Section 1.17.

carries $U \setminus K$ to $M \setminus \{q\}$. One can now define an integer number $\deg_q(f;K) \in \mathbb{Z}$ by the equality:

$$\phi_K\big(\alpha^{[n]}\big) = \deg_q(f;K)\tau_q.$$

What is the relation between the integers $\deg_q(f;K)$ and $\deg_q(f)$? It's simple: they are equal. Such equality follows from the fact that the homomorphisms ϕ and ϕ_K are equal; namely, we can glue diagrams (1.11.1) and (1.11.2) together obtaining the larger commutative diagram:

$$\begin{tikzcd}
\widetilde{H}_n(S^n) \arrow[d] \arrow[dr] \arrow[ddd, dotted, bend right=70] \arrow[ddd, dashed, bend left=80] & \\
H_n\big(S^n, S^n \setminus f^{-1}(q)\big) & H_n\big(S^n, S^n \setminus K\big) \arrow[l] \\
H_n\big(U, U \setminus f^{-1}(q)\big) \arrow[u, "\cong"] \arrow[d, "f_*"'] & H_n\big(U, U \setminus K\big) \arrow[l] \arrow[u, "\cong"'] \arrow[dl, "f_*"] \\
H_n\big(M, M \setminus \{q\}\big) &
\end{tikzcd}$$

where the unlabelled arrows are induced by inclusion. The dotted path in the diagram above defines ϕ, while de dashed path defines ϕ_K. The conclusion follows.

Let's now prove the basic properties of degrees.

1.11.3. PROPOSITION. *Let $f : U \subset S^n \to M$, q and τ_q be as in Definition 1.11.1. The following assertions hold.*

(1) (invariance of degree by restriction of domain) *If $V \subset U$ is an open set containing $f^{-1}(q)$ then* $\deg_q(f) = \deg_q(f|_V)$.

(2) (invariance of degree by restriction of counter-domain) *If Z is an open neighborhood of q in M containing the image of f then the degree of $f : U \to M$ at q equals the degree of $f : U \to Z$ at q (where we identify the local homology groups $H_n\big(Z, Z \setminus \{q\}\big)$ and $H_n\big(M, M \setminus \{q\}\big)$ in the usual way).*

(3) (vanishing of the degree) *If $q \notin \mathrm{Im}(f)$ then* $\deg_q(f) = 0$.

(4) (additivity of degree by disjoint union) *If U is a disjoint union $U = \bigcup_{\lambda\in L} U_\lambda$ of open subsets $U_\lambda \subset U$ then $\deg_q(f) = \sum_{\lambda\in L} \deg(f|_{U_\lambda})$ (only a finite number of terms on that sum are non zero).*

(5) (degree of a homeomorphism) *If $f : U \to M$ is a homeomorphism then $\deg_q(f) = \pm 1$; more precisely, $\deg_q(f) = 1$ (respectively, $\deg_q(f) = -1$) if the isomorphism:*

$$f_* : H_n\big(S^n, S^n \setminus \{f^{-1}(q)\}\big) \longrightarrow H_n\big(M, M \setminus \{q\}\big)$$

takes the canonical orientation $\alpha^{[n]}\big(f^{-1}(q)\big)$ to τ_q (respectively, takes the canonical orientation $\alpha^{[n]}\big(f^{-1}(q)\big)$ to $-\tau_q$).

(6) (homotopy invariance) *If* $f : U \to M$ *is homotopic to* $g : U \to M$ *by a homotopy* $H : U \times [0,1] \to M$ *for which* $H^{-1}(q)$ *is compact then* $\deg_q(f) = \deg_q(g)$.

(7) (invariance by positive homeomorphisms on the counter-domain) *If* N *is a topological manifold,* $h : M \to N$ *is a homeomorphism and:*

$$\tau'_{h(q)} = h_*(\tau_q) \in H_n\big(N, N \setminus \{h(q)\}\big),$$

then the degree of f *at* q *with respect to the orientation* τ_q *equals the degree of* $h \circ f$ *at* $h(q)$ *with respect to the orientation* $\tau'_{h(q)}$.

(8) (invariance by positive homeomorphisms on the domain) *If* $h : S^n \to S^n$ *is a positive homeomorphism then the degree of* $f \circ h : h^{-1}(U) \to M$ *at* q *equals the degree of* f *at* q.

PROOF. The proof of item (1) follows easily from the commutativity of the diagram:

$$\begin{tikzcd}
\widetilde{H}_n(S^n) \arrow[r] & H_n\big(S^n, S^n \setminus f^{-1}(q)\big) & H_n\big(V, V \setminus f^{-1}(q)\big) \arrow[l, "\cong"'] \arrow[d, "(f|_V)_*"] \\
& H_n\big(U, U \setminus f^{-1}(q)\big) \arrow[u, "\cong"] \arrow[r, "f_*"'] & H_n\big(M, M \setminus \{q\}\big)
\end{tikzcd}$$

where the unlabelled arrows are induced by inclusion. Namely, the dashed path takes the generator $\alpha^{[n]}$ of $\widetilde{H}_n(S^n)$ to $\deg_q(f)\tau_q$ and the dotted path takes $\alpha^{[n]}$ to $\deg_q(f|_V)\tau_q$.

The proof of item (2) follows from the commutativity of the diagram:

$$\begin{tikzcd}
& H_n\big(Z, Z \setminus \{q\}\big) \arrow[d] \\
H_n\big(U, U \setminus f^{-1}(q)\big) \arrow[ur, "f_*"] \arrow[r, "f_*"'] & H_n\big(M, M \setminus \{q\}\big)
\end{tikzcd}$$

by observing that the unlabelled arrow (induced by inclusion) is precisely the isomorphism we use to identify orientations of M at q and orientations of the open set Z at q.

The proof of item (3) follows from the observation that $f^{-1}(q) = \emptyset$ implies $H_n\big(U, U \setminus f^{-1}(q)\big) = H_n(U, U) = 0$.

To prove item (4), we start by observing that, since $f^{-1}(q)$ is compact, the intersection $f^{-1}(q) \cap U_\lambda$ is non empty only for a finite number of indexes $\lambda \in L$. Using items (1) and (3) we can discard the λ's for which $U_\lambda \cap f^{-1}(q) = \emptyset$ and therefore we can assume that L is finite. Consider now the commutative diagram:

$$\begin{tikzcd}
\widetilde{H}_n(S^n) \arrow[r, "\text{diagonal inclusion}"] \arrow[d] & \bigoplus_{\lambda\in L} \widetilde{H}_n(S^n) \arrow[d] \\
H_n\big(S^n, S^n \setminus f^{-1}(q)\big) \arrow[r] & \bigoplus_{\lambda\in L} H_n\big(S^n, S^n \setminus f_\lambda^{-1}(q)\big) \\
H_n\big(U, U \setminus f^{-1}(q)\big) \arrow[u, "\cong"] \arrow[d, "f_*"'] & \bigoplus_{\lambda\in L} H_n\big(U_\lambda, U_\lambda \setminus f_\lambda^{-1}(q)\big) \arrow[l, dashed, "\cong"'] \arrow[u, "\cong"'] \arrow[d, "\bigoplus_{\lambda\in L}(f_\lambda)_*"] \\
H_n\big(M, M \setminus \{q\}\big) & \bigoplus_{\lambda\in L} H_n\big(M, M \setminus \{q\}\big) \arrow[l, "\text{sum}"]
\end{tikzcd}$$

where the unlabelled arrows are induced by inclusion and $f_\lambda = f|_{U_\lambda}$. The fact that the dashed arrow is indeed an isomorphism follows from the result of Exercise 1.49. The left column of the diagram maps $\alpha^{[n]}$ to $\deg_q(f)\tau_q$ and the right column of the diagram maps the family $\big(\alpha^{[n]}\big)_{\lambda\in L}$ to the family $\big(\deg_q(f_\lambda)\tau_q\big)_{\lambda\in L}$. This proves item (4).

To prove item (5), we start by observing that since $f^{-1}(q)$ is a single point then the homomorphism $\widetilde{H}_n(S^n) \to H_n\big(S^n, S^n \setminus f^{-1}(q)\big)$ induced by inclusion is an isomorphism that maps $\alpha^{[n]}$ to $\alpha^{[n]}\big(f^{-1}(q)\big)$. The conclusion follows by observing that the homomorphism:

$$f_* : H_n\big(U, U \setminus \{f^{-1}(q)\}\big) \longrightarrow H_n\big(M, M \setminus \{q\}\big)$$

is an isomorphism and therefore it maps $\alpha^{[n]}\big(f^{-1}(q)\big)$ to either τ_q or $-\tau_q$.

To prove item (6) we argue as follows: let $K \subset U$ be the projection onto U of the compact set $H^{-1}(q) \subset U \times [0,1]$. Then K is a compact subset of U that contains both $f^{-1}(q)$ and $g^{-1}(q)$, so that (recall Remark 1.11.2):

$$\deg_q(f; K) = \deg_q(f), \quad \deg_q(g; K) = \deg_q(g).$$

The conclusion follows by observing that H actually defines a homotopy between the maps of pairs $f, g : (U, U \setminus K) \to \big(M, M \setminus \{q\}\big)$ and therefore the homomorphisms:

$$f_* : H_n(U, U \setminus K) \longrightarrow H_n\big(M, M \setminus \{q\}\big),$$
$$g_* : H_n(U, U \setminus K) \longrightarrow H_n\big(M, M \setminus \{q\}\big)$$

used to define $\deg_q(f; K)$ and $\deg_q(g; K)$ are equal.

Item (7) is trivial consequence of the equality $(h \circ f)_* = h_* \circ f_*$.

Finally, to prove item (8) consider the commutative diagram:

$$
\begin{tikzcd}
\widetilde{H}_n(S^n) \arrow[r, "h_*"] \arrow[d] \arrow[dddr, dashed, bend right=60] & \widetilde{H}_n(S^n) \arrow[d] \\
H_n\big(S^n, S^n \setminus (f\circ h)^{-1}(q)\big) \arrow[r, "h_*"] & H_n\big(S^n, S^n \setminus f^{-1}(q)\big) \\
H_n\big(h^{-1}(U), h^{-1}(U) \setminus (f\circ h)^{-1}(q)\big) \arrow[r, "h_*"'] \arrow[u, "\cong"] \arrow[dr, "(f\circ h)_*"'] & H_n\big(U, U \setminus f^{-1}(q)\big) \arrow[u, "\cong"'] \arrow[d, "f_*"] \\
 & H_n\big(M, M \setminus \{q\}\big)
\end{tikzcd}
$$

where the unlabelled arrows are induced by inclusion. The right column of the diagram takes $\alpha^{[n]}$ to $\deg_q(f)\tau_q$ and the dashed path takes $\alpha^{[n]}$ to $\deg_q(f\circ h)\tau_q$. The conclusion follows by observing that, since h is positive, the top arrow of the diagram is the identity (see Remark 1.10.23). $\square$

We now study conditions under with the degree $\deg_q(f)$ is independent of the point $q \in M$.

1.11.4. PROPOSITION. *Let $f : U \to M$ be a continuous proper map defined on an open subset $U \subset S^n$ taking values on an oriented n-dimensional connected topological manifold (M, τ). Then the integer number $\deg_q(f)$ (defined using the orientation $\tau(q)$ for M at q) is independent of $q \in M$.*

PROOF. It suffices to show that the map $q \mapsto \deg_q(f) \in \mathbb{Z}$ is locally constant. Let $q \in M$ be fixed. Since $\tau : M \to \mathcal{O}(M)$ is continuous, we can find an open neighborhood V of q in M and a homology class $\alpha \in H_n(M, M \setminus V)$ such that $\tau|_V = \mathcal{O}(\alpha; V)$ (recall Lemma 1.10.6). By passing to a smaller V (and using a local chart around q) we can assume that there exists a homeomorphism from $\overline{V}$ to $\overline{\mathrm{B}}^n$ that carries V to B^n and q to the origin; then $\overline{V} \setminus V = \partial V$ is a strong deformation retract of $\overline{V} \setminus \{q\}$ and also $M \setminus V$ is a strong deformation retract of $M \setminus \{q\}$. In particular, the homomorphism:

$$\rho_{Vq} : H_n(M, M \setminus V) \longrightarrow H_n\big(M, M \setminus \{q\}\big)$$

induced by inclusion is an isomorphism and α is a generator of $H_n(M, M\setminus V) \cong \mathbb{Z}$ (because $\rho_{Vq}(\alpha) = \tau(q)$ is a generator of $H_n\big(M, M \setminus \{q\}\big)$). Let K denote the compact set $f^{-1}\big(\overline{V}\big) \subset U$ (here we use that f is proper!) and consider the homomorphism $\lambda : \widetilde{H}_n(S^n) \to H_n(M, M \setminus V)$ defined by the commutative diagram:

$$\begin{tikzcd}
\widetilde{H}_n(S^n) \arrow[r] \arrow[drr, bend right=60, "\lambda"'] & H_n\big(S^n, S^n \setminus K\big) & \\
& H_n\big(U, U \setminus K\big) \arrow[u, "\cong"] \arrow[r, "f_*"] & H_n(M, M \setminus V)
\end{tikzcd}$$

where the unlabelled arrows are induced by inclusion. Since α is a generator of $H_n(M, M \setminus V)$ we can find an integer $d \in \mathbb{Z}$ with $\lambda\big(\alpha^{[n]}\big) = d\alpha$. But for every $q' \in V$ we have (see Remark 1.11.2):

$$\deg_{q'}(f; K)\tau(q') = \big(\rho_{Vq'} \circ \lambda\big)\big(\alpha^{[n]}\big) = d\tau(q'),$$

and therefore $\deg_{q'}(f) = \deg_{q'}(f; K) = d$. $\square$

We can finally give the following:

1.11.5. DEFINITION. If $f : U \subset S^n \to (M, \tau)$ are as in the statement of Proposition 1.11.4 then then integer number $\deg(f) = \deg_q(f) \in \mathbb{Z}$ (that does not depend on $q \in M$) is called the *degree* of the map f (with respect to the orientation τ of M).

1.11.6. EXAMPLE. If $U = S^n$ and $M = S^n$ is endowed with the canonical orientation $\alpha^{[n]}$ then the degree of a (automatically proper) continuous map f from $U = S^n$ to $M = S^n$ has a particularly simple interpretation (as mentioned in the beginning of the section). Choose any $q \in S^n$ and let $\phi : \widetilde{H}_n(S^n) \to H_n\big(S^n, S^n \setminus \{q\}\big)$ denote the homomorphism defined by diagram (1.11.1). It is easy to see that ϕ makes the following diagram:

$$\begin{tikzcd}
\widetilde{H}_n(S^n) \arrow[r, "f_*"] \arrow[dr, "\phi"'] & \widetilde{H}_n(S^n) \arrow[d, "\cong"] \\
& H_n\big(S^n, S^n \setminus \{q\}\big)
\end{tikzcd}$$

commute, where the vertical arrow is induced by inclusion. Since such vertical arrow maps $\alpha^{[n]}$ to $\alpha^{[n]}(q)$, it follows that *the degree of f equals the unique integer $d \in \mathbb{Z}$ for which the homomorphism $f_* : \widetilde{H}_n(S^n) \to \widetilde{H}_n(S^n)$ equals multiplication by d.*

Now we give a simple method for computing degrees of smooth maps.

1.11.7. PROPOSITION (differential degree). *Let $f : U \to M$ be a proper map of class C^1 defined on an open subset U of S^n (with*[11] *$n \geq 1$), taking values on an oriented connected n-dimensional differentiable manifold (M, τ). If $q \in M$*

[11] See Example 1.11.8 below for the case $n = 0$.

is a regular value of f *(i.e., if* $\mathrm{d}f_x : T_xS^n \to T_qM$ *is an isomorphism for every* $x \in f^{-1}(q)$*) then the set* $f^{-1}(q)$ *is finite and the degree of* f *is given by:*

$$\deg(f) = \sum_{x \in f^{-1}(q)} \operatorname{sign}(\mathrm{d}f_x),$$

where $\operatorname{sign}(\mathrm{d}f_x) = 1$ *if* $\mathrm{d}f_x$ *is a positively oriented isomorphism and* $\operatorname{sign}(\mathrm{d}f_x) = -1$ *if* $\mathrm{d}f_x$ *is a negatively oriented isomorphism (we consider the sphere* S^n *endowed with the outward pointing orientation).*

PROOF. It follows from the inverse function theorem that the compact set $f^{-1}(q)$ is discrete and hence finite; we write $f^{-1}(q) = \{x_1, \ldots, x_k\}$. Again by the inverse function theorem, we can choose an open neighborhood U_i of x_i in S^n such that $f(U_i)$ is open in M and $f|_{U_i} : U_i \to f(U_i)$ is a diffeomorphism; we can also assume that $U_i \cap U_j = \emptyset$ for $i \neq j$ (because S^n is Hausdorff). By item (1) of Proposition 1.11.3, the degree of f at q (which equals the degree of f, by definition) equals the degree at q of the restriction of f to the open set $\bigcup_{i=1}^k U_i$; now by item (4) of Proposition 1.11.3, we have:

$$\deg_q(f) = \sum_{i=1}^k \deg_q(f|_{U_i}).$$

Since $f|_{U_i} : U_i \to f(U_i)$ is a homeomorphism onto an open subset of M, the degree of $f|_{U_i}$ at q equals ± 1 by items (2) and (5). The sign of $\deg_q(f|_{U_i})$ depends on whether the homeomorphism $f|_{U_i}$ is positively oriented or negatively oriented. The conclusion follows from the result of Exercise 1.68 (see also Proposition 1.10.28). □

1.11.8. EXAMPLE (degree on the zero-dimensional case). Let M be a zero-dimensional topological manifold (i.e., a discrete topological space), $q \in M$ a point and τ_q an orientation for M at q. We identify τ_q with an element of $\{-1, 1\}$ as explained in Example 1.10.21. Let $U \subset S^0 = \{-1, 1\}$ be a (open) subset. Using items (2), (3), (4) and (5) of Proposition 1.11.3 and recalling Convention 1.10.24 for the definition of $\alpha^{[0]}$, one checks easily that the degree $\deg_q(f)$ of a map $f : U \to M$ is equal to:

- zero, if $f^{-1}(q)$ is either empty or equal to $S^0 = \{-1, 1\}$;
- τ_q, if $f^{-1}(q)$ contains only the "north pole" $1 \in S^0$;
- $-\tau_q$, if $f^{-1}(q)$ contains only the "south pole" $-1 \in S^0$.

1.11.9. REMARK. As mentioned in the beginning of the section, it is possible to give a notion of degree for a continuous proper map between arbitrary oriented topological manifolds of the same dimension (actually, the counter-domain should be connected to guarantee that $\deg_q(f)$ is independent of q). Let's just take a glimpse at this more general definition. First, observe that if we replace $\big(S^n, \alpha^{[n]}\big)$ with an arbitrary oriented n-dimensional topological manifold N then it would make no sense to care about maps defined on *an open subset* $U \subset N$, since such open set is again an n-dimensional topological manifold (like N). So, consider a continuous map $f : N \to M$, a point $q \in M$, a generator τ_q of $H_n\big(M, M \setminus \{q\}\big)$

and assume that $K = f^{-1}(q)$ is compact. As in Definition 1.11.1, we can consider the homomorphism:

$$f_* : H_n(N, N \setminus K) \longrightarrow H_n\big(M, M \setminus \{q\}\big);$$

the problem is: how do we choose the homology class on $H_n(N, N \setminus K)$ that is going to be pushed-forward by f_* to give us an integer multiple of τ_q? When N were an open subset of S^n then such homology class was induced from the canonical generator of $\widetilde{H}_n(S^n)$ (that defines the canonical orientation of S^n). For the general case, one has to work more to understand the structure of the homology group $H_n(N, N \setminus K)$. It can be shown that for any compact subset K of an n-dimensional topological manifold N the map:

$$H_n(N, N \setminus K) \ni \alpha \longmapsto \mathcal{O}(\alpha; K)$$

gives an isomorphism between the homology group $H_n(N, N \setminus K)$ and the abelian group of all continuous sections of the orientation bundle $\mathcal{O}(N)$ along K. Thus, the general definition of $\deg_q(f)$ can be given as follows: let $\alpha \in H_n(N, N \setminus K)$ be the unique homology class such that[12] $\mathcal{O}(\alpha; K)$ equals the restriction to K of the orientation of N; the integer $\deg_q(f) \in \mathbb{Z}$ is thus defined by the equality:

$$f_*(\alpha) = \deg_q(f)\tau_q.$$

One can now easily generalize Propositions 1.11.3, 1.11.4 and 1.11.7 to this context (for Proposition 1.11.7 one obviously has to assume that N is a differentiable manifold). See [**39**, Chapter VIII, §4] for details.

1.11.10. REMARK. In some situations we will have in hand continuous maps $f : U \to M$ defined on open subsets U of oriented n-dimensional topological manifolds X that are not *exactly* the sphere S^n but are *homeomorphic* to the sphere. In such situations, we should choose a positively oriented homeomorphism $h : \big(S^n, \alpha^{[n]}\big) \to X$ and use our degree theory on the composite map $f \circ h : h^{-1}(U) \subset S^n \to M$. Observe though that using item (8) of Proposition 1.11.3, it is easy to see that the degree of $f \circ h$ does not depend on the choice of the positively oriented homeomorphism h. We will therefore use our degree theory freely for maps defined on open subsets of oriented topological manifolds X that are homeomorphic to the sphere S^n, without making explicit references to positively oriented homeomorphisms $h : S^n \to X$.

1.12. Index of a Vector Field at an Isolated Singularity

The theory of this section will not be used elsewhere. We decided to presented this material here because it is nicely related to the notion of degree.

By a *vector field* on an open subset $U \subset \mathbb{R}^n$ we mean a continuous map $X : U \to \mathbb{R}^n$; we call a point $x_0 \in U$ a *singularity* for X if $X(x_0) = 0$.

[12] Observe that if $N = U$ is an open subset of S^n then the homology class $\alpha \in H_n(U, U \setminus K)$ obtained by pushing $\alpha^{[n]}$ forward to $H_n(S^n, S^n \setminus K)$ and then pulling it back to $H_n(U, U \setminus K)$ satisfy this condition; so we are indeed generalizing Definition 1.11.1 here.

1.12.1. DEFINITION. Let $X : U \to \mathbb{R}^n$ be a vector field and let $x_0 \in U$ be an isolated singularity of X (i.e., x_0 is a singularity of X and X has no other singularities in some neighborhood of x_0). Choose a neighborhood V of x_0 in U such that x_0 is the only singularity of $X|_V$; the dotted arrow in the commutative diagram:

$$\begin{tikzcd} H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{x_0\}\big) \arrow[dr, "(X|_V)_*"] & \\ H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big) \arrow[u, "(\mathfrak{t}_{x_0})_*", "\cong"'] \arrow[r, dotted] & H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big) \end{tikzcd}$$

defines an endomorphism of $H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big) \cong \mathbb{Z}$ that equals multiplication by an integer $\operatorname{ind}(X; x_0)$, called the *index of the vector field X at the isolated singularity* x_0.

Alternatively (recall Convention 1.10.24) one can define the index of X at the isolated singularity x_0 by the equality:

$$(X|_V)_*\big(\tau^{[n]}(x_0)\big) = \operatorname{ind}(X; x_0)\tau^{[n]}(0). \tag{1.12.1}$$

1.12.2. EXAMPLE. If x_0 is a singularity of X and if X is a homeomorphism from an open neighborhood V of x_0 onto an open neighborhood of the origin in $\mathbb{R}^n$ then $(X|_V)_*$ is an isomorphism and therefore $\operatorname{ind}(X; x_0) = \pm 1$; more precisely (see (1.12.1)), we have $\operatorname{ind}(X; x_0) = 1$ (respectively, $\operatorname{ind}(X; x_0) = -1$) if the restriction of X to V is a positively oriented homeomorphism (respectively, a negatively oriented homeomorphism) onto an open neighborhood of the origin in $\mathbb{R}^n$. In particular, by Corollary 1.10.19 and the inverse function theorem, if X is of class C^1 and if $\mathrm{d}X_{x_0} : \mathbb{R}^n \to \mathbb{R}^n$ is an isomorphism then $\operatorname{ind}(X; x_0) = \pm 1$ and $\operatorname{ind}(X; x_0)$ has the same sign of the determinant of the isomorphism $\mathrm{d}X_{x_0}$.

We now relate indexes of vector fields at isolated singularities with degrees.

Let $x_0 \in U$ be an isolated singularity of a vector field $X : U \to \mathbb{R}^n$. Let $\varepsilon > 0$ be such that the closed ball $\mathrm{B}[x_0; \varepsilon]$ is contained in U and such that X has no other singularities in $\mathrm{B}[x_0; \varepsilon]$. Consider the map $\lambda : S^{n-1} \to S(x_0; \varepsilon)$ defined by:

$$\lambda(x) = \varepsilon x + x_0$$

and denote by $r : \mathbb{R}^n \setminus \{0\} \to S^{n-1}$ the radial projection:

$$r(x) = \frac{x}{\|x\|}.$$

We can then associate to X, x_0 and ε a continuous map $f : S^{n-1} \to S^{n-1}$ defined by:

$$f = r \circ X \circ \lambda. \tag{1.12.2}$$

The index $\operatorname{ind}(X; x_0)$ can then be computed using f as shown in the following:

1.12.3. PROPOSITION. *If X, x_0 and ε are chosen as above then the index of X at the isolated singularity x_0 equals the degree of the map $f : S^{n-1} \to S^{n-1}$ defined in* (1.12.2).

PROOF. First observe that the vector field $X \circ \mathfrak{t}_{x_0}$ has an isolated singularity at the origin whose index equals $\operatorname{ind}(X; x_0)$; moreover, the continuous map f that corresponds to the singularity at the origin of $X \circ \mathfrak{t}_{x_0}$ is precisely the same as the continuous map f that corresponds to X and x_0. We may thus assume without loss of generality that $x_0 = 0$.

Consider the following commutative diagram:

$$
\begin{tikzcd}
H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big) \arrow[dr, "\substack{\text{multiplication}\\ \text{by } \operatorname{ind}(X;0)}"] \arrow[ddd, dotted, bend right] & \\
H_n\big(\mathrm{B}[0;\varepsilon], \mathrm{B}[0;\varepsilon] \setminus \{0\}\big) \arrow[r, "X_*"] \arrow[u, "\cong"'] \arrow[d, "\partial_*"', "\cong"] & H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big) \arrow[d, "\cong"', "\partial_*"] \arrow[dd, dashed, bend left] \\
\widetilde{H}_{n-1}\big(\mathrm{B}[0;\varepsilon] \setminus \{0\}\big) \arrow[r, "X_*"] & \widetilde{H}_{n-1}\big(\mathbb{R}^n \setminus \{0\}\big) \arrow[d, "\cong"', "r_*"] \\
\widetilde{H}_{n-1}(S^{n-1}) \arrow[r, "f_*"'] \arrow[u, "\lambda_*", "\cong"'] & \widetilde{H}_{n-1}(S^{n-1})
\end{tikzcd}
$$

where the unlabelled arrow is induced by inclusion. The conclusion will follow once we show that the isomorphisms defined by the dotted and the dashed paths in the diagram above are equal (recall from Example 1.11.6 that f_* equals multiplication by $\deg(f)$). But such equality can be easily established in the commutative diagram below:

$$
\begin{tikzcd}
H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big) \arrow[d, "\partial_*"', "\cong"] \arrow[dd, dashed, bend right] \arrow[ddr, dotted, bend left] & H_n\big(\mathrm{B}[0;\varepsilon], \mathrm{B}[0;\varepsilon] \setminus \{0\}\big) \arrow[l, "\cong"'] \arrow[d, "\cong"', "\partial_*"] \\
\widetilde{H}_{n-1}\big(\mathbb{R}^n \setminus \{0\}\big) \arrow[d, "r_*"', "\cong"] & \widetilde{H}_{n-1}\big(\mathrm{B}[0;\varepsilon] \setminus \{0\}\big) \arrow[l, "\cong"'] \\
\widetilde{H}_{n-1}(S^{n-1}) & \widetilde{H}_{n-1}(S^{n-1}) \arrow[l, "(r\circ\lambda)_* = \mathrm{Id}"] \arrow[u, "\cong", "\lambda_*"']
\end{tikzcd}
$$

in which the unlabelled arrows are induced by inclusion. □

1.13. Homology with General Coefficients

In Section 1.2, the singular homology of a topological space X was defined in terms of the singular chain complex $\mathfrak{S}(X)$ of X; for each $p \in \mathbb{N}$, the group $\mathfrak{S}_p(X)$ was defined as the free abelian group spanned by the set of all singular p-simplexes in X. Thus, the group $\mathfrak{S}_p(X)$ can be seen intuitively as the group of *formal linear combinations with integer coefficients* of singular p-simplexes. This is why the groups $H_p(X)$ defined in Section 1.2 are also called *singular homology groups with integer coefficients* of the topological space X. Can we use other types of coefficients? Sure we can (as you could guess from the name of this section).

Actually, one can use the elements of any fixed abelian group G as coefficients for singular chains. The group $\mathfrak{S}_p(X;G)$ of singular p-chains in X with coefficients in G could be defined as the group of all essentially zero G-valued maps defined on the set of all singular p-simplexes of X. Although it seems very reasonable to use such simple definition for the group $\mathfrak{S}_p(X;G)$, we won't do so. Alternatively, we use a more elegant algebraic technique that produces a (naturally) isomorphic object: we set $\mathfrak{S}_p(X;G) = \mathfrak{S}_p(X) \otimes G$. The advantage of doing so are the following:

- the boundary homomorphism of $\mathfrak{S}(X;G)$ can be very elegantly[13] defined by tensoring the boundary homomorphism of $\mathfrak{S}(X)$ with the identity of G.
- It becomes easier to prove theorems about singular homology with coefficients in G because we can use well-established techniques of algebra for dealing with tensor products.

Observe that the operation of "tensoring with G" makes sense not only for $\mathfrak{S}(X)$, but for arbitrary chain complexes. We will therefore consider below this more general situation.

1.13.1. DEFINITION. If $(\mathfrak{C}, \partial)$ is a chain complex and G is an arbitrary abelian group then *the tensor product of* $\mathfrak{C}$ *and* G is the chain complex $\mathfrak{C} \otimes G$ defined by setting $(\mathfrak{C} \otimes G)_p = \mathfrak{C}_p \otimes G$ for all $p \in \mathbb{Z}$; the boundary homomorphism of $\mathfrak{C} \otimes G$ is defined by $\partial_p \otimes \mathrm{Id} : \mathfrak{C}_p \otimes G \to \mathfrak{C}_{p-1} \otimes G$ for all $p \in \mathbb{Z}$, where Id stands for the identity map of G.

The fact that $\mathfrak{C} \otimes G$ is indeed a chain complex, i.e., the fact that the composition of two subsequent boundary homomorphisms is zero follows directly from the result of Exercise 1.18.

The homology of the chain complex $\mathfrak{C} \otimes G$ will be called the *homology of* $\mathfrak{C}$ *with coefficients in* G; we write:

$$H_p(\mathfrak{C}; G) = H_p(\mathfrak{C} \otimes G),$$

for all $p \in \mathbb{Z}$.

A chain map between chain complexes induces a chain map between their respective tensorizations with G; more explicitly, if $f : \mathfrak{C} \to \mathfrak{D}$ is a chain map then one obtains a chain map $f \otimes \mathrm{Id} : \mathfrak{C} \otimes G \to \mathfrak{D} \otimes G$ by setting:

$$(f \otimes \mathrm{Id})_p = f_p \otimes \mathrm{Id} : \mathfrak{C}_p \otimes G \longrightarrow \mathfrak{D}_p \otimes G,$$

for all $p \in \mathbb{Z}$, where Id stands for the identity map of G. One can also define chain maps induced from homomorphisms of coefficient groups; namely, if $\phi : G \to H$ is a homomorphism between abelian groups G and H then for every chain complex $\mathfrak{C}$ one obtains a chain map:

$$\mathrm{Id} \otimes \phi : \mathfrak{C} \otimes G \longrightarrow \mathfrak{C} \otimes H,$$

[13]Imagine how ugly the precise definition of the boundary of a singular chain with coefficients in G would be if one decided not to talk about tensor products. For instance, observe that $\mathfrak{S}(X;G)$ is not free over the set of singular p-simplexes, so it does not suffice to define the boundary homomorphism for simplexes — one has to give the formula directly for chains.

defined by $(\mathrm{Id}\otimes\phi)_p = \mathrm{Id}\otimes\phi$ for all $p\in\mathbb{Z}$, where Id stands for the identity map of $\mathfrak{C}_p$. If we are given both a chain map $f:\mathfrak{C}\to\mathfrak{D}$ and a homomorphism $\phi: G\to H$ then we can take the composition of $f\otimes\mathrm{Id}$ and $\mathrm{Id}\otimes\phi$ obtaining a chain map $f\otimes\phi$; more explicitly, we have:

$$(f\otimes\phi)_p = f_p\otimes\phi = (f_p\otimes\mathrm{Id})\circ(\mathrm{Id}\otimes\phi),$$

for all $p\in\mathbb{Z}$.

1.13.2. EXAMPLE. If X is a topological space and G is an abelian group then the *singular chain complex of X with coefficients in G* is the chain complex $\mathfrak{S}(X;G)$ defined by:

$$\mathfrak{S}(X;G) = \mathfrak{S}(X)\otimes G.$$

The homology groups of $\mathfrak{S}(X;G)$ are called the *singular homology groups of X with coefficients in G*; we write:

$$H_p(X;G) = H_p\big(\mathfrak{S}(X;G)\big) = H_p\big(\mathfrak{S}(X)\otimes G\big) = H_p\big(\mathfrak{S}(X);G\big).$$

Similarly, one can define the reduced singular homology groups of a topological space X with coefficients in G and the relative homology groups of a pair (X,A) with coefficients in G. More explicitly, we set:

$$\widetilde{\mathfrak{S}}(X;G) = \widetilde{\mathfrak{S}}(X)\otimes G,\quad \widetilde{H}_p(X;G) = H_p\big(\widetilde{\mathfrak{S}}(X);G\big),$$
$$\mathfrak{S}(X,A;G) = \mathfrak{S}(X,A)\otimes G,\quad H_p(X,A;G) = H_p\big(\mathfrak{S}(X,A);G\big),$$

for all $p\in\mathbb{Z}$. If $f:X\to Y$ is a continuous map between topological spaces X and Y then the chain map $f_\#:\mathfrak{S}(X)\to\mathfrak{S}(Y)$ induced by f induces a chain map:

$$f_\#\otimes\mathrm{Id}:\mathfrak{S}(X;G)\longrightarrow\mathfrak{S}(Y;G),$$

between the singular chain complexes with coefficients in G. Such chain map induces a homomorphism:

$$f_*: H_p(X;G)\longrightarrow H_p(Y;G)$$

between singular homology groups with coefficients in G. Moreover, if $\phi: G\to H$ is a coefficient group homomorphism then we can also look at the homomorphism:

$$\phi_*: H_p(X;G)\longrightarrow H_p(X;H),$$

induced in homology by the chain map $\mathrm{Id}\otimes\phi:\mathfrak{S}(X;G)\to\mathfrak{S}(X;H)$. Similar considerations can also be made for reduced and relative homology.

The natural question that arises now is: how do I compute homology with coefficients in G? One could imagine that $H_p(\mathfrak{C};G)$ is just $H_p(\mathfrak{C})\otimes G$; *this is not true*, as the following simple example shows.

1.13.3. EXAMPLE. Consider the chain complex $\mathfrak{C}$ defined by $\mathfrak{C}_p = \mathbb{Z}$ for $p = 0,1,2$ and $\mathfrak{C}_p = 0$ otherwise; the boundary homomorphisms are described below:

$$\cdots\longrightarrow\mathbb{Z}\xrightarrow{\text{times }2}\mathbb{Z}\xrightarrow{0}\mathbb{Z}\longrightarrow 0\longrightarrow\cdots$$

Obviously $H_2(\mathfrak{C}) = 0$, $H_1(\mathfrak{C}) = \mathbb{Z}_2$ and $H_0(\mathfrak{C}) = \mathbb{Z}$. If $G = \mathbb{Z}_2$ then the chain complex $\mathfrak{C} \otimes G$ is given by:

$$\cdots \longrightarrow \mathbb{Z}_2 \xrightarrow{\text{times } 2=0} \mathbb{Z}_2 \xrightarrow{0} \mathbb{Z}_2 \longrightarrow 0 \longrightarrow \cdots$$

and therefore $H_p(\mathfrak{C}; \mathbb{Z}_2) = \mathbb{Z}_2$ for $p = 0, 1, 2$. Observe in particular that $H_2(\mathfrak{C}; \mathbb{Z}_2)$ is not equal to $H_2(\mathfrak{C}) \otimes \mathbb{Z}_2 = 0$.

In spite of what we saw in Example 1.13.3, it is indeed true that for *free* chain complexes $\mathfrak{C}$ the homology with coefficients in G is exclusively determined by the homology of $\mathfrak{C}$; the theorem relating the homologies of $\mathfrak{C} \otimes G$ and $\mathfrak{C}$ is known as *the universal coefficient theorem for homology*. It happens that the rule relating $H_p(\mathfrak{C}; G)$ and $H_p(\mathfrak{C})$ is not as simple as $H_p(\mathfrak{C}; G) = H_p(\mathfrak{C}) \otimes G$; in fact, the group $H_p(\mathfrak{C}; G)$ depends not only on $H_p(\mathfrak{C})$ but also on $H_{p-1}(\mathfrak{C})$.

Let's first of all define a map relating $H_p(\mathfrak{C}) \otimes G$ and $H_p(\mathfrak{C} \otimes G)$. Consider a fixed chain complex $\mathfrak{C}$, an abelian group G and an integer $p \in \mathbb{Z}$. We have a canonical bilinear map:

$$\mathfrak{C}_p \times G \longrightarrow \mathfrak{C}_p \otimes G \tag{1.13.1}$$

that sends (c, g) to $c \otimes g$. Obviously, if c is a cycle (respectively, a boundary) in $\mathfrak{C}$ then also $c \otimes g$ is a cycle (respectively, a boundary) in $\mathfrak{C} \otimes G$. We can therefore restrict (1.13.1) to $Z_p(\mathfrak{C})$ and pass it to the quotient (in the first variable) obtaining a bilinear map:

$$H_p(\mathfrak{C}) \times G \longrightarrow H_p(\mathfrak{C} \otimes G) = H_p(\mathfrak{C}; G), \tag{1.13.2}$$

that takes $\big(c + B_p(\mathfrak{C}), g\big)$ to $c \otimes g + B_p(\mathfrak{C} \otimes G)$ for every $c \in Z_p(\mathfrak{C})$ and every $g \in G$. Finally, the bilinear map (1.13.2) gives rise (by the fundamental property of tensor products) to a unique homomorphism:

$$\theta_p : H_p(\mathfrak{C}) \otimes G \longrightarrow H_p(\mathfrak{C}; G). \tag{1.13.3}$$

As we already know, θ is not in general an isomorphism. Nevertheless, we have the following:

1.13.4. LEMMA. *If $\mathfrak{C}$ is free then, for every $p \in \mathbb{Z}$, the homomorphism θ_p given in* (1.13.3) *is injective and its image is a direct summand in $H_p(\mathfrak{C} \otimes G)$.*

PROOF. Since $\mathfrak{C}$ is free, the cycle group $Z_p(\mathfrak{C})$ is a direct summand in $\mathfrak{C}_p$ (see Exercise 1.39); then, the canonical projection $Z_p(\mathfrak{C}) \to H_p(\mathfrak{C})$ extends to a homomorphism:

$$\lambda_p : \mathfrak{C}_p \longrightarrow H_p(\mathfrak{C}).$$

We can regard $H(\mathfrak{C})$ as a chain complex whose p-th group is $H_p(\mathfrak{C})$ and all the boundary homomorphisms are zero; it is then easy to see that $\lambda : \mathfrak{C} \to H(\mathfrak{C})$ is a chain map. We have thus an induced chain map:

$$\lambda \otimes \mathrm{Id} : \mathfrak{C} \otimes G \longrightarrow H(\mathfrak{C}) \otimes G,$$

and a induced homomorphism in homology:

$$(\lambda \otimes \mathrm{Id})_* : H_p(\mathfrak{C} \otimes G) \longrightarrow H_p(\mathfrak{C}) \otimes G; \tag{1.13.4}$$

observe that in the complex $H(\mathfrak{C}) \otimes G$ the boundary homomorphisms are all zero and therefore its p-th homology group is indeed $H_p(\mathfrak{C}) \otimes G$. It is now easy to check that (1.13.4) is a left inverse for θ_p. The conclusion follows from the result of Exercise 1.16. □

In order to finish our task of computing the homology of $\mathfrak{C} \otimes G$ from the homology of $\mathfrak{C}$, we "only" have to determine the group $H_p(\mathfrak{C}; G)/\mathrm{Im}(\theta_p)$; but that's not so easy.

1.13.5. DEFINITION. A *short free resolution* of an abelian group G is a short exact sequence of the form:

$$0 \longrightarrow R \longrightarrow F \longrightarrow G \longrightarrow 0$$

where both R and F are free abelian groups. The *canonical free resolution* of G is defined by:

$$0 \longrightarrow \mathrm{Rel}[G] \xrightarrow{\text{inclusion}} \mathrm{Free}[G] \xrightarrow{\mathfrak{q}} G \longrightarrow 0$$

where $\mathfrak{q} : \mathrm{Free}[G] \to G$ is the homomorphism characterized by the condition $\mathfrak{q}|_G = \mathrm{Id} : G \to G$ and $\mathrm{Rel}[G] = \mathrm{Ker}(\mathfrak{q})$. The group $\mathrm{Rel}[G]$ is called the *relations group* of G.

1.13.6. DEFINITION. Let G, H be abelian groups and consider the groups $\mathrm{Rel}[G] \subset \mathrm{Free}[G]$ appearing in the canonical free resolution of G. The group $\mathrm{Tor}(G, H)$ is by definition the kernel of the homomorphism:

$$\mathrm{Rel}[G] \otimes H \xrightarrow{\text{inclusion} \otimes H} \mathrm{Free}[G]$$

obtained by tensoring with H the inclusion of $\mathrm{Rel}[G]$ in $\mathrm{Free}[G]$.

If $\phi : G_1 \to G_2$, $\psi : H_1 \to H_2$ are abelian group homomorphisms, we now want to define a homomorphism:

$$\mathrm{Tor}(\phi, \psi) : \mathrm{Tor}(G_1, H_1) \longrightarrow \mathrm{Tor}(G_2, H_2). \tag{1.13.5}$$

We proceed as follows: let $\tilde{\phi} : \mathrm{Free}[G_1] \to \mathrm{Free}[G_2]$ be the unique homomorphism that extends $\phi : G_1 \to G_2$. We have thus a commutative diagram:

$$\begin{tikzcd}
0 \arrow[d] & 0 \arrow[d] \\
\mathrm{Rel}[G_1] \arrow[r, dotted] \arrow[d, "\text{inclusion}"'] & \mathrm{Rel}[G_2] \arrow[d, "\text{inclusion}"] \\
\mathrm{Free}[G_1] \arrow[r, "\tilde{\phi}"] \arrow[d, "\mathfrak{q}"'] & \mathrm{Free}[G_2] \arrow[d, "\mathfrak{q}"] \\
G_1 \arrow[r, "\phi"'] \arrow[d] & G_2 \arrow[d] \\
0 & 0
\end{tikzcd} \tag{1.13.6}$$

with exact columns; such diagram implies that $\tilde{\phi}$ carries $\mathrm{Rel}[G_1]$ to $\mathrm{Rel}[G_2]$, i.e., we can actually put an arrow in the place where the dotted arrow is. We now take the square in (1.13.6) containing the dotted arrow and we tensor it with the commutative square

$$\begin{tikzcd} H_1 \arrow[r, "\psi"] \arrow[d, "\mathrm{Id}"'] & H_2 \arrow[d, "\mathrm{Id}"] \\ H_1 \arrow[r, "\psi"'] & H_2 \end{tikzcd}$$

As a result we get another commutative square:

(1.13.7)
$$\begin{tikzcd} \mathrm{Rel}[G_1] \otimes H_1 \arrow[r, dotted] \arrow[d] & \mathrm{Rel}[G_2] \otimes H_2 \arrow[d] \\ \mathrm{Free}[G_1] \otimes H_1 \arrow[r, "\tilde{\phi}\otimes\psi"'] & \mathrm{Free}[G_2] \otimes H_2 \end{tikzcd}$$

The dotted arrow in diagram (1.13.7) above is just the dotted arrow in (1.13.6) tensored with ψ. The commutativity of (1.13.7) implies that the dotted arrow of (1.13.7) takes the kernel of the left arrow (i.e., $\mathrm{Tor}(G_1, H_1)$) inside the kernel of the right arrow (i.e., $\mathrm{Tor}(G_2, H_2)$). We thus define (1.13.5) to be the restriction of the dotted arrow of (1.13.7).

The proposition below states the main property of Tor.

1.13.7. PROPOSITION. *To every abelian group G and every short exact sequence:*

(1.13.8)
$$0 \longrightarrow A \xrightarrow{\sigma} B \xrightarrow{\tau} C \longrightarrow 0$$

of abelian groups and homomorphisms it is possible to associate a homomorphism from $\mathrm{Tor}(C, G)$ to $A \otimes G$ (depending on G and on (1.13.8)*) that makes the following sequence:*

$$\begin{aligned} 0 \longrightarrow \mathrm{Tor}(A, G) \xrightarrow{\mathrm{Tor}(\sigma,\mathrm{Id})} \mathrm{Tor}(B, G) \xrightarrow{\mathrm{Tor}(\tau,\mathrm{Id})} \mathrm{Tor}(C, G) \longrightarrow \\ \longrightarrow A \otimes G \xrightarrow{\sigma\otimes\mathrm{Id}} B \otimes G \xrightarrow{\tau\otimes\mathrm{Id}} C \longrightarrow 0 \end{aligned}$$

exact. Let's call it the connecting *homomorphism corresponding to G and to* (1.13.8).

The connecting homomorphism $\mathrm{Tor}(C, G) \to A \otimes G$ *is* natural *in the sense that given a homomorphism $\psi : G \to G'$ and a commutative diagram:*

(1.13.9)
$$\begin{tikzcd} 0 \arrow[r] & A \arrow[r, "\sigma"] \arrow[d, "\phi_1"'] & B \arrow[r, "\tau"] \arrow[d, "\phi_2"] & C \arrow[r] \arrow[d, "\phi_3"] & 0 \\ 0 \arrow[r] & A' \arrow[r, "\sigma'"'] & B' \arrow[r, "\tau'"'] & C' \arrow[r] & 0 \end{tikzcd}$$

with exact rows then the square:

$$\begin{tikzcd} \mathrm{Tor}(C,G) \arrow[r, "\text{connecting}"] \arrow[d, "\mathrm{Tor}(\phi_3,\psi)"'] & A \otimes G \arrow[d, "\phi_1\otimes\psi"] \\ \mathrm{Tor}(C',G') \arrow[r, "\text{connecting}"'] & A' \otimes G' \end{tikzcd}$$

commutes.

PROOF. See [**107**, §54]. □

We also have the following:

1.13.8. THEOREM. *Given abelian groups G, H then $\mathrm{Tor}(G,H)$ is isomorphic to $\mathrm{Tor}(H,G)$. Such isomorphism is natural with respect to maps induced by group homomorphisms.*

PROOF. See [**107**, §54]. □

1.13.9. THEOREM. *Given abelian groups G, H, if either G or H is Torsion free then $\mathrm{Tor}(G,H) = 0$.*

PROOF. See [**107**, §54]. □

1.13.10. EXAMPLE. We have a short exact sequence:

$$0 \longrightarrow \mathbb{Z} \xrightarrow{\ n\ } \mathbb{Z} \longrightarrow \mathbb{Z}_n \longrightarrow 0$$

By tensoring it with G and using Proposition 1.13.7 and Theorem 1.13.9, it is easy to see that:

$$\mathrm{Tor}(\mathbb{Z}_n, G) \cong \mathrm{Ker}\big(G \xrightarrow{\ n\ } G\big),$$

for every abelian group G. In particular:

$$\mathrm{Tor}(\mathbb{Z}_n, \mathbb{Z}_m) \cong \mathbb{Z}_{\gcd(m,n)}, \quad \mathrm{Tor}(\mathbb{Z}, \mathbb{Z}_n) = \mathrm{Tor}(\mathbb{Z}_n, \mathbb{Z}) = 0.$$

We can now state the following:

1.13.11. THEOREM (universal coefficients theorem). *Given a free chain complex $\mathfrak{C}$ there exists a short exact sequence:*

$$0 \longrightarrow H_p(\mathfrak{C}) \otimes G \xrightarrow{\ \theta_p\ } H_p(\mathfrak{C}; G) \longrightarrow \mathrm{Tor}\big(H_{p-1}(\mathfrak{C}), G\big) \longrightarrow 0 \tag{1.13.10}$$

where θ_p is the homomorphism (1.13.3). *The sequence* (1.13.10) *splits and is* natural *with respect to chain maps $f : \mathfrak{C} \to \mathfrak{D}$ and homomorphisms $\phi : G \to H$ between coefficient groups. More precisely, if $\mathfrak{D}$ is a free chain complex, H is an abelian group, $f : \mathfrak{C} \to \mathfrak{D}$ is a chain map and $\phi : G \to H$ is a homomorphism then the diagram:*

$$\begin{tikzcd} 0 \arrow[r] & H_p(\mathfrak{C}) \otimes G \arrow[r, "\theta"] \arrow[d, "f_*\otimes\phi"'] & H_p(\mathfrak{C}; G) \arrow[r] \arrow[d, dashed, "(f\otimes\phi)_*"] & \mathrm{Tor}\big(H_{p-1}(\mathfrak{C}), G\big) \arrow[r] \arrow[d, "\mathrm{Tor}(f_*,\phi)"] & 0 \\ 0 \arrow[r] & H_p(\mathfrak{D}) \otimes H \arrow[r, "\theta"'] & H_p(\mathfrak{D}; H) \arrow[r] & \mathrm{Tor}\big(H_{p-1}(\mathfrak{D}), H\big) \arrow[r] & 0 \end{tikzcd}$$

commutes.

PROOF. . See [**107**, §55] □

1.13.12. COROLLARY. *Given free chain complexes $\mathfrak{C}$, $\mathfrak{D}$ and a chain map $f : \mathfrak{C} \to \mathfrak{D}$ that induces an isomorphism in homology (in all dimensions) then the chain map $f \otimes \mathrm{Id} : \mathfrak{C} \otimes G \to \mathfrak{D} \otimes G$ also induces an isomorphism in homology (in all dimensions) for* every *coefficient group G.*

PROOF. Follows directly from the naturality of the universal coefficient sequence (1.13.10). □

Assume now that the coefficient group G is endowed with the structure of a (left) R-module, for some ring R. Then each group $\mathfrak{C}_p \otimes G$ can be endowed with the structure of an R-module; with such R-module structure, the boundary map $\partial_p \otimes \mathrm{Id}$ becomes R-linear. We give the following:

1.13.13. DEFINITION. Given a ring R, A *chain complex of (left) R-modules* is a pair $(\mathfrak{C}, \partial)$, where $\mathfrak{C} = (\mathfrak{C}_p)_{p\in\mathbb{Z}}$ is a family of (left) R-modules and $\partial = (\partial_p)_{p\in\mathbb{Z}}$ is a family of R-linear maps $\partial_p : \mathfrak{C}_p \to \mathfrak{C}_{p-1}$ such that $\partial_{p-1} \circ \partial_p = 0$ for all $p \in \mathbb{Z}$.

Obviously a chain complex of R-modules can also be thought of as a chain complex (of abelian groups) when one forgets about the R-module structure in the chain groups $\mathfrak{C}_p$. One can therefore talk about cycles, boundaries and homology groups for chain complexes of R-modules and all the theorems presented so far can be applied in this context. We now make a few remarks to establish the compatibility of the R-module structure of the chain groups $\mathfrak{C}_p$ with some of the results we have presented so far in homology theory.

- If $\mathfrak{C}$ is a chain complex of R-modules then the cycle groups $Z_p(\mathfrak{C})$ and boundary groups $B_p(\mathfrak{C})$ are R-submodules of $\mathfrak{C}_p$ for every $p \in \mathbb{Z}$. In particular, the homology group $H_p(\mathfrak{C})$ can also be endowed with an R-module structure so that the quotient map $Z_p(\mathfrak{C}) \to H_p(\mathfrak{C})$ is R-linear.
- If $\mathfrak{C}$, $\mathfrak{D}$ are chain complexes of R-modules and if $f : \mathfrak{C} \to \mathfrak{D}$ is an *R-linear chain map*, i.e., f is a chain map and each f_p is R-linear, then the induced maps $f_* : H_p(\mathfrak{C}) \to H_p(\mathfrak{D})$ are R-linear.
- If (1.3.3) is a short exact sequence of chain complexes of R-modules and R-linear chain maps then the connecting homomorphism ∂_* in the corresponding long exact homology sequence (1.3.4) is R-linear.

We are mainly concerned with the case that the abelian group $G = R$ is a ring itself (endowed with the canonical left R-module structure) and even more specifically with the case that G is a field $\mathbb{K}$. In the latter case we talk about *chain complexes of $\mathbb{K}$-vector spaces* rather than chain complexes of $\mathbb{K}$-modules.

As observed above, if G is endowed with an R-module structure (for instance, if $G = R$ is a ring) then the homology groups $H_p(X; G)$ and $H_p(X, A; G)$ are canonically endowed with R-module structures. In particular, we can give the following:

1.13.14. DEFINITION. Given a topological space X and a field $\mathbb{K}$ then the *p-th Betti number of X* is defined as the (possibly infinite) dimension over $\mathbb{K}$ of the $\mathbb{K}$-vector space $H_p(X; \mathbb{K})$ and is denoted by $\beta_p(X; \mathbb{K})$. Similarly, given a pair (X, A)

of topological spaces one defines the Betti numbers $\beta_p(X, A; \mathbb{K})$ of the pair (X, A) with respect to the field $\mathbb{K}$.

1.13.15. LEMMA. *Consider a topological space X whose singular homology groups with integer coefficients are all finitely generated and vanish above a certain dimension. Then for every coefficient field $\mathbb{K}$ we have $\beta_p(X; \mathbb{K}) < +\infty$ for all p and $\beta_p(X; \mathbb{K}) = 0$ for p sufficiently large. Moreover, the Euler characteristic of X (recall Exercise 1.125) is given by:*

$$\chi(X) = \sum_{p\in\mathbb{Z}} (-1)^p \beta_p(X; \mathbb{K}) \in \mathbb{Z},$$

for every coefficient field $\mathbb{K}$.

PROOF. Follows from the universal coefficient theorem (Theorem 1.13.11). $\square$

1.14. Quotient Topology and Attachment Spaces

1.14.1. DEFINITION. A map $q : X \to Y$ between topological spaces is called a *quotient map* if a subset $U \subset Y$ is open if and only if $q^{-1}(U) \subset X$ is open.

Obviously, the "only if" part in the definition above means that q is continuous. Alternatively, one can define that $q : X \to Y$ is a quotient map if Y has the finest topology that makes q continuous; observe in particular that such topology is unique (see Exercise 1.81). When a topological space X and an equivalence relation $\sim$ in X is given, we will implicitly assume that the quotient set $X/\sim$ is endowed with the unique topology that makes the canonical projection $X \to X/\sim$ a quotient map.

We do not assume in principle that a quotient map $q : X \to Y$ is surjective, as some authors do; observe though that quotient maps are "almost" surjective (see Exercise 1.83).

1.14.2. PROPOSITION (universal property of quotient maps). *Given a commutative diagram*

(1.14.1)
$$\begin{tikzcd} X \arrow[d, "q"'] \arrow[dr, "f"] & \\ Y \arrow[r, dotted, "\overline{f}"'] & Z \end{tikzcd}$$

with q a quotient map and f continuous then $\overline{f}$ is also continuous.

1.14.3. DEFINITION. Given a set X and an equivalence relation $\sim$ on X then a subset $S \subset X$ is called *saturated with respect to* $\sim$ if S is a union of equivalence classes, i.e., if for $x \in X$, $x' \in S$, $x \sim x'$ implies $x \in S$. If $q : X \to Y$ is any map then q defines an equivalence relation $\sim$ in X by $x \sim_q x' \Leftrightarrow q(x) = q(x')$; we say that a subset $S \subset X$ is *saturated with respect to* q if it is saturated with respect to $\sim_q$, i.e., if $S = q^{-1}\big(q(S)\big)$.

1.14.4. LEMMA. *If $q : X \to Y$ is a quotient map and $S \subset X$ is saturated and either open or closed in X then $q|_S : S \to q(S)$ is a quotient map.*

1.14.5. PROPOSITION. *Let $q : X \to Y$ be a surjective quotient map and let Z be a locally compact Hausdorff space. Then the map $q \times \mathrm{Id} : X \times Z \to Y \times Z$ is also a quotient map.*

The following immediate corollary of Proposition 1.14.5 is a useful tool for defining homotopies by passage to the quotient.

1.14.6. COROLLARY. *Let $H : X \times [0,1] \to Y$ be a continuous map. Assume that we are given a surjective quotient map $q_1 : X \to X'$ and a continuous map $q_2 : Y \to Y'$ such that there exists a map $\overline{H} : X' \times [0,1] \to Y'$ for which the diagram*

$$\begin{tikzcd} X \times [0,1] \arrow[r, "H"] \arrow[d, "q_1 \times \mathrm{Id}"'] & Y \arrow[d, "q_2"] \\ X' \times [0,1] \arrow[r, "\overline{H}"'] & Y' \end{tikzcd}$$

commutes. Then $\overline{H}$ is continuous. $\square$

1.14.7. DEFINITION. Given a topological space X and a family $(X_i)_{i \in I}$ of subspaces of X then we say that X is the (internal) *topological sum* of the family $(X_i)_{i\in I}$ if $X = \bigcup_{i\in I} X_i$ is a disjoint union and each X_i is open in X. If $(X_i)_{i\in I}$ is a family of topological spaces then its (external) *topological sum* is defined as the disjoint union $X = \bigcup_{i\in I} \{i\} \times X_i$ topologized as follows: $U \subset X$ is open iff $U = \bigcup_{i\in I}\{i\} \times U_i$ with each U_i open in X_i.

Up to obvious identifications, the notions of internal and external topological sum are equivalent (see Exercise 1.94). If X is either the internal or the external topological sum of the spaces $(X_i)_{i\in I}$ we will write $X = \sum_{i\in I} X_i$; for finite families we may also use the simpler notation $X = X_1 + \cdots + X_n$.

Obviously, if X is a topological sum of a family of subspaces $(X_i)_{i\in I}$ then a map $f : X \to Y$ is continuous iff $f|_{X_i}$ is continuous for every $i \in I$. This property characterizes the topological sum (see Exercise 1.95).

1.14.8. DEFINITION. Given a space X and a family $(X_i)_{i\in I}$ of subspaces of X with $X = \bigcup_{i\in I} X_i$ then we say that X is the coherent union of the family $(X_i)_{i\in I}$ if the following condition holds: given a subset $U \subset X$, if $U \cap X_i$ is open in X_i for every $i \in I$ then U is open in X.

The definition above can be restated in terms of closed subsets (see Exercise 1.96).

1.14.9. PROPOSITION (universal property of coherent union). *If $X = \bigcup_{i\in I} X_i$ is a coherent union then a map $f : X \to Y$ is continuous iff $f|_{X_i}$ is continuous for every $i \in I$.*

1.14.10. DEFINITION. Let X, Y be topological spaces, $A \subset X$ a subspace and $f : A \to Y$ a continuous map. The *attachment space* $X \cup_f Y$ is the quotient space $X+Y/\sim$ of the topological sum $X+Y$, where $\sim$ is the equivalence relation in $X+Y$ spanned by $x \sim f(x)$, $x \in A$; more precisely, for $z, z' \in X+Y$ we have $z \sim z'$ iff one of the following holds:

(1) $z = z'$;
(2) $z, z' \in A$ and $f(z) = f(z')$;
(3) $z \in A$ and $z' = f(z) \in Y$;
(4) $z' \in A$ and $z = f(z') \in Y$.

1.14.11. LEMMA. *If A is closed in X then the quotient map $q : X+Y \to X \cup_f Y$ maps Y homeomorphically onto a closed set and $X \setminus A$ homeomorphically onto an open set.*

PROOF. Since $X \setminus A$ is open and saturated in $X+Y$, it follows that $q(X \setminus A)$ is open (see Exercise 1.91) and that $q|_{X \setminus A} : X \setminus A \to q(X \setminus A)$ is a quotient map (see Lemma 1.14.4); since $q|_{X \setminus A}$ is injective, then it is a homeomorphism (see Exercise 1.84). Since $q|_Y$ is injective, to prove that q maps Y homeomorphically onto a closed set, it suffices to show that $q|_Y : Y \to X \cup_f Y$ is a closed map. This follows by observing that the saturation of a closed subset $F \subset Y$ is the closed subset $F \cup f^{-1}(F)$ of $X+Y$. □

The fact that $q : X+Y \to X \cup_f Y$ maps Y homeomorphically onto a (not necessarily closed) subset of $X \cup_f Y$ is also true when A is not closed in X; the proof is a little more delicate (see Exercise 1.113).

When $f : A \subset X \to Y$ is injective, the attachment space $X \cup_f Y$ can be visualized in the following way: assume that in some environment space E there are homeomorphic copies of X and Y in a way that the copy of $x \in A \subset X$ in E coincides with the copy of $f(x) \in Y$ in E. It would be natural to expect that the subspace of E which is the union of the copies of X and Y should be homeomorphic to the attachment space $X \cup_f Y$. This holds for instance when the copies of X and Y are closed in E; more generally, we have the following:

1.14.12. LEMMA. *Let X, Y, E be topological spaces, $A \subset X$ a subspace, $f : A \to Y$ an injective continuous map and $h_1 : X \to X' \subset E$, $h_2 : Y \to Y' \subset E$ homeomorphisms. Assume that, for $x \in X$, $y \in Y$, $h_1(x) = h_2(y)$ iff $x \in A$ and $y = f(x)$. Then there exists a unique map ϕ such that the diagram*

$$\begin{tikzcd}
 & X+Y \arrow[dl, "q"'] \arrow[dr, "h"] & \\
X \cup_f Y \arrow[rr, "\phi"'] & & X' \cup Y'
\end{tikzcd}$$

commutes, where h is defined by $h|_X = h_1$ and $h|_Y = h_2$. The map ϕ is a continuous bijection; it is a homeomorphism iff $X' \cup Y'$ is a coherent union (for instance, if X' and Y' are closed in E).

PROOF. The existence and the bijectivity of ϕ is easy. The continuity of ϕ follows from the continuity of h and of the universal property of quotient maps; the continuity of h follows from the universal property of topological sums. The map ϕ is a homeomorphism iff it is a quotient map; by Exercises 1.85 and 1.86, ϕ is a quotient map iff h is a quotient map. We have a commutative diagram:

$$\begin{tikzcd} X+Y \arrow[rr, "h_1+h_2", "\cong"'] \arrow[dr, "h"'] & & X'+Y' \arrow[dl, "i"] \\ & X'\cup Y' & \end{tikzcd}$$

where i is induced by the inclusions of X' and Y' in $X' \cup Y'$. The map h is quotient iff the map i is quotient. The conclusion follows from Exercise 1.98. □

Recall that a topological space is said to be *T_4* if it is T_1 and if any two disjoint closed subsets admit disjoint open neighborhoods.

We recall the following two standard results from general topology concerning T_4 spaces.

1.14.13. THEOREM (Urysohn's lemma). *If X is a T_4 topological space then, given disjoint closed subsets $F, G \subset X$, there exists a continuous map $\phi : X \to [0,1]$ such that $\phi|_F \equiv 0$ and $\phi|_G \equiv 1$.*

PROOF. See for instance [**134**, Theorem IV.7]. □

1.14.14. THEOREM (Tietze). *If X is a T_4 topological space then every continuous map $\phi : F \to \mathbb{R}$ defined on a closed subset $F \subset X$ admits a continuous ($\mathbb{R}$-valued) extension to the whole space X.*

PROOF. See for instance [**134**, Theorem IV.11]. □

1.14.15. LEMMA. *If X, Y are T_4 topological spaces, $A \subset X$ is a closed subset and $f : A \to Y$ is a continuous map then the attachment space $X \cup_f Y$ is also T_4.*

PROOF. The fact that $X \cup_f Y$ is T_1 follows directly from the fact that A is closed in X. In order to prove that $X \cup_f Y$ is T_4, we will use the converse of Tietze's theorem (see Exercise 1.89). Let $\phi : F \to \mathbb{R}$ be a continuous map defined on a closed subset F of $X \cup_f Y$. If $q : X + Y \to X \cup_f Y$ denotes the canonical projection then $q^{-1}(F) = F_1 \cup F_2$, with F_1 closed in X, F_2 closed in Y and $F_1 \cap A = f^{-1}(F_2)$. The composite map $\phi \circ q : F_1 \cup F_2 \to \mathbb{R}$ restricts to a continuous map $\phi_1 : F_1 \to \mathbb{R}$ and to a continuous map $\phi_2 : F_2 \to \mathbb{R}$; moreover, $\phi_2 \circ f = \phi_1|_A$. Since Y is T_4, Tietze's theorem gives us a continuous extension $\tilde{\phi}_2 : Y \to \mathbb{R}$ of ϕ_2. The functions $\phi_2 \circ f : A \to \mathbb{R}$ and $\phi_1 : F_1 \to \mathbb{R}$ agree on $F_1 \cap A$ and, since F_1 and A are closed, we obtain a well-defined continuous function on the union $F_1 \cup A$. Now we can apply again Tietze's theorem on the T_4 space X to obtain a continuous function $\tilde{\phi}_1 : X \to \mathbb{R}$ that extends both ϕ_1 and $\phi_2 \circ f$. The functions $\tilde{\phi}_1 : X \to \mathbb{R}$ and $\phi_2 : Y \to \mathbb{R}$ now give us a continuous function on $X + Y$ that passes to the quotient, producing a continuous extension $\tilde{\phi} : X \cup_f Y \to \mathbb{R}$ of ϕ. □

1.14.16. LEMMA. *Let X be a topological space and assume that we have an increasing sequence $X_1 \subset X_2 \subset \cdots \subset X$ of closed subsets of X such that $X = \bigcup_{n\geq 1} X_n$ is a coherent union and such that each X_n is T_4. Then X is also T_4.*

PROOF. Obviously X is T_1. In order to prove that X is T_4 we will use again the converse of Tietze's theorem (see Exercise 1.89). Let $\phi : F \to \mathbb{R}$ be a continuous map defined on a closed subspace $F \subset X$. We will construct a sequence of continuous map $\phi_n : X_n \to \mathbb{R}$ such that ϕ_n equals ϕ on $F \cap X_n$ and such that ϕ_{n+1} extends ϕ_n for every n. Since the union $X = \bigcup_{n\geq 1} X_n$ is coherent, this will yield a continuous extension of ϕ to X. We define ϕ_n inductively using Tietze's theorem on each X_n. Start with an arbitrary continuous extension $\phi_1 : X_1 \to \mathbb{R}$ of $\phi|_{F\cap X_1}$. If $\phi_n : X_n \to \mathbb{R}$ is a continuous extension of $\phi|_{F\cap X_n}$, define a continuous map on $(F \cap X_{n+1}) \cup X_n$ by gluing $\phi|_{F\cap X_{n+1}}$ and ϕ_n; finally, let $\phi_{n+1} : X_{n+1} \to \mathbb{R}$ be an arbitrary continuous extension of such continuous map. □

1.14.17. DEFINITION. Given a topological space X and a subspace $A \subset X$ then the space X/A is defined to be the quotient space of X by the equivalence relation $\sim$ whose equivalence classes are A and the singletons $\{x\}$, $x \in X \setminus A$; more explicitly, for $x, x' \in X$ we have $x \sim x'$ iff $x = x'$ or $x, x' \in A$.

1.14.18. DEFINITION. Let $f : X \to Y$ be a continuous map. The *mapping cylinder* of f, denoted by M_f, is the attachment space $\big(X \times [0,1]\big) \cup_f Y$, where f is identified with the map $X \times \{0\} \ni (x, 0) \mapsto f(x) \in Y$. The *mapping cone* of f, denoted by C_f, is the quotient $M_f/\big(X \times \{1\}\big)$, where $X \times \{1\}$ is identified with its image under the canonical map $\big(X \times [0,1]\big) + Y \to M_f$.

1.14.19. EXAMPLE. The cylinder of the identity map $\mathrm{Id} : X \to X$ of a topological space X can be identified with the product $X \times [0,1]$ (this is sometimes called the *cylinder* of the space X). The cone of the identity map of X is identified with the quotient $(X \times [0,1])/(X \times \{1\})$ and is called the *cone* of X; it will be denoted by C_X.

1.14.20. EXAMPLE. The cone of the sphere S^{n-1} can obviously be identified with the closed ball $\overline{\mathrm{B}}^n$ via the map induced by $S^{n-1} \times [0,1] \ni (x,t) \mapsto tx \in \overline{\mathrm{B}}^n$. More generally, if $U \subset \mathbb{R}^n$ is an open bounded convex subset then the cone of the boundary of U can be identified with $\overline{U}$ (see Exercise 1.42).

The mapping cylinder and the mapping cone of a map are relatively simple notions that are useful in many situations. Our interest on these concepts come from their relation with attachment of cells:

1.14.21. LEMMA. *If Y is a topological space and $f : S^{n-1} \to Y$ is a continuous map then the mapping cone of f is homeomorphic to the attachment space $\overline{\mathrm{B}}^n \cup_f Y$ by a homeomorphism that fixes X.*

PROOF. Follows immediately from Example 1.14.20 and Exercise 1.116. □

We now want to relate the mapping cylinders and the mapping cones of homotopic maps. To this aim, given a homotopy $H : X \times [0,1] \to Y$ from f to g we define a continuous map $\phi_H : M_f \to M_g$ as follows. Let $q_f : \big(X \times [0,1]\big) + Y \to M_f$, $q_g : \big(X \times [0,1]\big) + Y \to M_g$ be the canonical quotient maps and consider the homotopies $q_g \circ H : X \times [0,1] \to M_g$ and $q_g|_{X\times[0,1]} : X \times [0,1] \to M_g$. We define the map ϕ_H by requiring that $\phi_H \circ q_f$ equals q_g on Y and equals the concatenation of homotopies $(q_g \circ H) \cdot (q_g|_{X\times[0,1]})$ on $X \times [0,1]$. More explicitly, we have:

$$
\begin{aligned}
\phi_H\big(q_f(x,t)\big) &= \begin{cases} q_g\big(H(x,2t)\big), & t \in \big[0, \frac{1}{2}\big], \\ q_g(x, 2t-1), & t \in \big[\frac{1}{2}, 1\big], \end{cases} \\
\phi_H\big(q_f(y)\big) &= q_g(y),
\end{aligned}
\tag{1.14.2}
$$

for all $x \in X$, $y \in Y$. It is easy to see that ϕ_H induces a map $\bar{\phi}_H : C_f \to C_g$ between the mapping cones of f and g.

We now present some properties of the maps ϕ_H.

1.14.22. LEMMA. *The following properties hold:*

(a) *if two homotopies $H, H' : X \times [0,1] \to Y$ from f to g are homotopic relatively to $X \times \{0,1\}$ then ϕ_H and $\phi_{H'}$ are homotopic relatively to Y;*
(b) *given homotopies $H, H' : X \times [0,1] \to Y$ with $H_1 = H'_0$ then the composite map $\phi_{H'} \circ \phi_H : M_{H_0} \to M_{H'_1}$ is homotopic relatively to Y to the map $\phi_{H\cdot H'}$ corresponding to the concatenated homotopy $H \cdot H'$;*
(c) *if $H : X \times [0,1] \to Y$ is a constant homotopy, say $H_t = f : X \to Y$ for all $t \in [0,1]$, then ϕ_H is homotopic relatively to Y to the identity map of M_f.*

Analogous statements hold by replacing ϕ_H, $\phi_{H'}$ by $\bar{\phi}_H$, $\bar{\phi}_{H'}$ and mapping cylinders by mapping cones.

PROOF. It follows easily using Exercise 1.57 and the following observations. Regarding item (b), we observe that $(\phi_{H'} \circ \phi_H) \circ q|_{X\times[0,1]}$ equals the concatenated homotopy $(q' \circ H) \cdot [(q' \circ H') \cdot (q'|_{X\times[0,1]})]$, where $q : \big(X \times [0,1]\big) + Y \to M_{H_0}$ and $q' : \big(X \times [0,1]\big) + Y \to M_{H'_1}$ are the canonical projections. Regarding item (c), we observe that $\phi_H \circ q$ equals the concatenation of the homotopy $(x,t) \mapsto q(x,0)$ and the homotopy $q|_{X\times[0,1]}$, where $q : \big(X \times [0,1]\big) + Y \to M_f$ denotes the canonical projection. $\square$

1.14.23. COROLLARY. *If $f, g : X \to Y$ are homotopic continuous maps between topological spaces X, Y then the mapping cylinders M_f and M_g (respectively, the mapping cones C_f and C_g) are homotopy equivalent. Explicitly, if $H : f \cong g$ is a homotopy then the maps $\phi_H : M_f \to M_g$ and $\phi_{H^{-1}} : M_g \to M_f$ (respectively, $\bar{\phi}_H : C_f \to C_g$ and $\bar{\phi}_{H^{-1}} : C_g \to C_f$) are mutual homotopy inverses; moreover, ϕ_H and $\phi_{H^{-1}}$ (respectively, $\bar{\phi}_H$ and $\bar{\phi}_{H^{-1}}$) restrict to the identity on Y (recall* (1.14.2)*).*

PROOF. Follows easily from Lemma 1.14.22 and Exercise 1.57. $\square$

1.14.24. LEMMA. *Let X, Y be topological spaces, $f : X \to Y$, $g : Y \to Y$ be continuous maps such that there exists a homotopy $H : g \cong \mathrm{Id}$. Denote by*

$q : X \times [0,1] + Y \to M_f$ the canonical projection and consider the continuous map $\psi : M_f \to M_f$ such that $\psi \circ q|_Y = q \circ g$ and $\psi \circ q|_{X\times[0,1]}$ equals the concatenation of the homotopies $(x,t) \mapsto q \circ H(f(x),t)$ and $q|_{X\times[0,1]}$. Then ψ is homotopic to the identity relatively to $q\big(X \times \{1\}\big)$. In particular, the map $\bar\psi : C_f \to C_f$ induced by ψ is also homotopic to the identity.

PROOF. We will explicitly exhibit a homotopy $K : \psi \cong \mathrm{Id}$ relative to $q\big(X \times \{1\}\big)$. For every $s, t \in [0,1]$, $x \in X$, $y \in Y$ we set:

$$K_s\big(q(y)\big) = q \circ H(y,s),$$

$$K_s\big(q(x,t)\big) = \begin{cases} q \circ H\big(f(x), s+2t\big), & t \in \big[0, \frac{1-s}{2}\big], \\ \\ q\big(x, \frac{2(t-1)}{1+s} + 1\big), & t \in \big[\frac{1-s}{2}, 1\big]. \end{cases}$$

The verification that K is well-defined and satisfies the required properties is left to the reader. □

1.14.25. COROLLARY. *Given topological spaces X, Y, Y', a continuous map $f : X \to Y$ and a homotopy equivalence $h : Y \to Y'$ then the map $\tilde h : M_f \to M_{h\circ f}$ induced by the identity on $X \times [0,1]$ and by h on Y is a homotopy equivalence. Moreover, $\tilde h$ induces a homotopy equivalence from C_f to $C_{h\circ f}$.*

PROOF. Let $k : Y' \to Y$ be a homotopy inverse for h and consider the map $\tilde k : M_{h\circ f} \to M_{k\circ h\circ f}$ induced by the identity on $X \times [0,1]$ and by k on Y. We claim that $\tilde k \circ \tilde h$ is a homotopy equivalence; this will imply that $\tilde h$ has a left homotopy inverse and that $\tilde k$ has a right homotopy inverse. In order to prove the claim, set $g = k \circ h$ and let $H : g \cong \mathrm{Id}$ be a homotopy from g to the identity of Y; then $(x,t) \mapsto H\big(f(x),t\big)$ defines a homotopy $\widehat H$ from $g \circ f$ to f. By Corollary 1.14.23, we have a homotopy equivalence $\phi_{\widehat H} : M_{g\circ f} \to M_f$; it is easy to see that the composite map $\phi_{\widehat H} \circ \tilde k \circ \tilde h$ is precisely the map ψ of Lemma 1.14.24. We conclude that $\phi_{\widehat H} \circ \tilde k \circ \tilde h$ and $\phi_{\widehat H}$ are both homotopy equivalences, proving the claim.

To conclude the proof of the corollary, we make the following observations: since k is a homotopy equivalence, the arguments above can also be used to show that $\tilde k$ (like $\tilde h$) has a left homotopy inverse. Since $\tilde k$ also has a right homotopy inverse, it follows that $\tilde k$ is a homotopy equivalence. Finally, since both $\tilde k$ and $\tilde k \circ \tilde h$ are homotopy equivalences, it follows that $\tilde h$ is a homotopy equivalence. The proof that the map induced by $\tilde h$ on the mapping cones is also a homotopy equivalence is totally analogous. □

1.15. CW-complexes

A CW-complex is a topological space X endowed with a special kind of decomposition that allows one to systematize the strategy described in the beginning of Section 1.3 for computing the singular homology groups of X. Such decomposition consists in fixing a partition of X into smaller subspaces that are homeomorphic to open balls; such subspaces are called the *open cells* of the decomposition.

The cells are glued together along each others boundaries to form the whole space X. The simplest example of a CW-complex is the one of a triangulable space (see Exercises 1.43, 1.44, 1.45, 1.46, 1.107 and 1.124). For instance, assume that one chooses a triangulation for the two-dimensional torus, i.e., that one identifies the torus with a polyhedron by means of a homeomorphism. Such triangulation gives a decomposition for the torus into open two-dimensional triangles (the faces of the polyhedron), that are glued together along open line segments (the edges of the polyhedron); such open line segments are glued together along isolated points (the vertices of the polyhedron). Triangulations can also be used to compute the singular homology of a space (that's what's called *simplicial* homology), but the decompositions in cells allowed for CW-complexes are usually much more economic. For instance, we will see below that it is possible to give a structure of CW-complex on the torus having only four cells; in the case of the sphere, it's possible to use only *two* cells.

In this section, we present the general theory of CW-complexes. In Section 1.16, we will show how one can compute the singular homology of a CW-complex.

We start by introducing formally the terminology of cells and open cells.

1.15.1. DEFINITION. If $p \geq 0$ is an integer then by a *cell of dimension* p (or a p-*cell*) we mean a topological space that is homeomorphic to the p-dimensional closed ball $\overline{\mathrm{B}}^p$; by an *open cell of dimension* p (or an *open* p-*cell*) we mean a topological space that is homeomorphic to the p-dimensional open ball B^p.

Observe that a 0-cell or an open 0-cell is the same thing as a topological space having only one point. Observe also that the *dimension* of a cell (or of an open cell) is well-defined, i.e., a topological space cannot be at the same time a p-cell (respectively, an open p-cell) and a q-cell (respectively, an open q-cell) for $p \neq q$ (see Exercise 1.63).

We can now give the formal definition of CW-complex. This is a very technical definition and not so easy to digest at first sight. The examples given below should be able to clarify the spirit of the definition.

1.15.2. DEFINITION. A *CW-complex* consists of a Hausdorff topological space X and a collection $\mathfrak{C}$ of subsets of X such that the following conditions hold:

(1) $X = \bigcup_{e \in \mathfrak{C}} e$ is a disjoint union;
(2) each $e \in \mathfrak{C}$ is an open cell;
(3) for every $p \geq 0$ and every open p-cell $e \in \mathfrak{C}$ there exists a continuous map $f : \overline{\mathrm{B}}^p \to X$ that restricts to a homeomorphism from B^p onto e;
(4) for every $p \geq 0$ and every open p-cell $e \in \mathfrak{C}$ the set $\dot{e}$ defined by $\dot{e} = \overline{e} \setminus e$ is contained in a finite union of open cells in $\mathfrak{C}$ of dimension less than p;
(5) the union $X = \bigcup_{e \in \mathfrak{C}} \overline{e}$ is coherent.

The collection $\mathfrak{C}$ is called a *cellular decomposition* for the topological space X.

Condition (4) above is usually called *Closure-finiteness* and condition (5) is usually called *Weak-topology* (thus the name CW-complex).

If $e \in \mathfrak{E}$ is an open p-cell then a continuous map $f : \overline{\mathrm{B}}^p \to X$ that maps B^p homeomorphically onto e is called a *characteristic map* for the cell e; thus condition (3) above says that every open cell $e \in \mathfrak{E}$ admits a characteristic map[14].

We will usually denote by $\mathfrak{E}_p$ the set of open p-cells of X, i.e., we set:

$$\mathfrak{E}_p = \big\{e \in \mathfrak{E} : \dim(e) = p\big\}.$$

The *dimension* of the CW-complex X is the (possibly infinite) natural number:

$$\dim(X) = \sup_{e \in \mathfrak{E}} \dim(e).$$

Regarding Definition 1.15.2, a few remarks are in order.

- If $f : \overline{\mathrm{B}}^p \to X$ is a characteristic map for an open p-cell $e \in \mathfrak{E}$ then the image of f equals the closure of e. Namely, since B^p is dense in $\overline{\mathrm{B}}^p$, $f\big(\mathrm{B}^p\big) = e$ is dense in $f\big(\overline{\mathrm{B}}^p\big)$; hence $f\big(\overline{\mathrm{B}}^p\big)$ is contained in the closure of e. Moreover, the set $f\big(\overline{\mathrm{B}}^p\big)$ is compact and therefore closed (X is Hausdorff!); since $f\big(\overline{\mathrm{B}}^p\big)$ contains e, it also contains the closure of e.
- If $e \in \mathfrak{E}$ is an open p-cell and f is a characteristic map for e then the set $\dot{e} = \overline{e} \setminus e$ is the image by f of the unit sphere S^{p-1} (see Exercise 1.118).
- If $e \in \mathfrak{E}$ is an open 0-cell then $\overline{e} = e$ (again we use that X is Hausdorff!) and hence $\dot{e} = \emptyset$. Property (4) is thus vacuously satisfied for 0-cells (even though there are no cells of dimension less than zero). Observe also that the existence of characteristic maps for open 0-cells is trivial.
- If $e \in \mathfrak{E}$ is an open p-cell then in general the closure $\overline{e}$ of e *is not* a p-cell and the set $\dot{e}$ is not homeomorphic to the sphere S^{p-1} (see Example 1.15.3 below).
- By Exercise 1.97, Property (5) is automatically satisfied for *finite CW-complexes*, i.e., CW-complexes having only a finite number of open cells.

Let's now give examples of cellular decompositions for some familiar topological spaces.

1.15.3. EXAMPLE (CW-complex structure for the sphere). Let's give a cellular decomposition for the p-dimensional sphere S^p. We assume $p \geq 1$ (the zero-dimensional sphere S^0 has an obvious cellular decomposition with two open 0-cells). It is not hard to see that there exists a continuous map $q : \overline{\mathrm{B}}^p \to S^p$ that is constant on $S^{p-1} \subset \overline{\mathrm{B}}^p$ and that maps the open ball B^p homeomorphically onto the complement of the point $q\big(S^{p-1}\big)$ in S^p (see Exercise 1.119). We can therefore define a cellular decomposition $\mathfrak{E} = \{e^0, e^p\}$ for S^p by taking e^0 to be the open 0-cell $q(S^{p-1})$ and e^p to be the open p-cell $S^p \setminus e^0$. Observe that the map q is a characteristic map for the open cell e^p. The sphere can thus be given the structure of a CW-complex having only two open cells.

---

[14]We observe that the characteristic maps for the open cells do not form a part of the structure of the CW-complex; only the space X and the cellular decomposition $\mathfrak{E}$ do. More precisely, a CW-complex is just a pair $(X, \mathfrak{E})$; the characteristic maps for the open cells are assumed to exist, but no particular privileged set of characteristic maps is fixed *a priori*.

1.15.4. EXAMPLE (CW-complex structure for the torus). Let R denote the square $[0,1]^2 \subset \mathbb{R}^2$ and let $\sim$ be the equivalence relation in R spanned by:

$$(x,0) \sim (x,1) \quad \text{and} \quad (0,y) \sim (1,y),$$

for all $x, y \in [0,1]$. It is well known that the quotient space $R/\sim$ is homeomorphic to the torus $\mathbb{T} = S^1 \times S^1$. Let $q : R \to \mathbb{T} \cong R/\sim$ denote the quotient map. We can thus define a cellular decomposition $\mathfrak{E}$ for the torus $\mathbb{T}$ having one open 2-cell e^2, two open 1-cells e^1_1, e^1_2 and one open 0-cell e^0 as follows:

$$e^2 = q\big(\,]0,1[^2\,\big),$$
$$e^1_1 = q\big(\,]0,1[\,\times\{0\}\big), \quad e^1_2 = q\big(\{0\}\times\,]0,1[\,\big),$$
$$e^0 = \{q(0,0)\};$$

namely, the interior $\mathrm{inter}(R) = \,]0,1[^2$ of the square R is a saturated open set for the map q. By Lemma 1.14.4 and Exercise 1.84, q maps $]0,1[^2$ homeomorphically onto e^2, so that e^2 is indeed an open 2-cell; a characteristic map for e^2 is q itself (see Remark 1.15.5 below). The restriction of q to a closed side of the square R is a quotient map by item (4) of Exercise 1.91; the interior of a side is a saturated open set of that side, so that by Lemma 1.14.4 and Exercise 1.84, the map q carries $]0,1[\,\times\{0\}$ homeomorphically onto e^1_1 and $\{0\}\times\,]0,1[$ homeomorphically onto e^1_2. Thus e^1_1 and e^1_2 are indeed open 1-cells and characteristic maps for them are obtained by taking restrictions of q to $[0,1]\times\{0\}$ and to $\{0\}\times[0,1]$ respectively. The remaining properties of a CW-complex listed in Definition 1.15.2 are trivially verified.

1.15.5. REMARK. If a topological space B is homeomorphic to $\overline{\mathrm{B}}^p$ (i.e., if B is a p-cell) and if $f : B \to X$ is a continuous maps taking $\mathrm{inter}(B)$ homeomorphically onto some open p-cell $e \in \mathfrak{E}$ then we will in general (with some abuse) call f a *characteristic map* for e. Obviously a real characteristic map for e can be obtained by considering the composition $f \circ h$, where $h : \overline{\mathrm{B}}^p \to B$ is an arbitrary homeomorphism.

1.15.6. EXAMPLE. Let R be a regular n-agon in the plane $\mathbb{R}^2$ and let $R/\sim$ be a quotient space of R obtained by identifying some of the closed sides of R with each other, generalizing the situation of Example 1.15.4. It is known for instance that every compact surface (possibly with boundary) can be obtained by this construction (see [**96**]). The space $R/\sim$ is always Hausdorff by the result of Exercise 1.90. Moreover, a cellular decomposition for $R/\sim$ can be described in the following way: the image by the quotient map $q : R \to R/\sim$ of the interior of R is an open 2-cell; the images by q of the interiors of the sides of R are open 1-cells and the images by q of the vertices of R are open 0-cells. The characteristic maps for such open cells are all obtained by taking suitable restrictions of q. Detailed arguments that justify that we indeed have obtained a cellular decomposition for $R/\sim$ can be given in analogy with the ones given in Example 1.15.4.

1.15.7. EXAMPLE (CW-complex structure on the real projective space). The *n-dimensional real projective space* $\mathbb{R}P^n$ is the space obtained by identifying antipodal points in S^n, i.e., $\mathbb{R}P^n = S^n/\sim$ where $\sim$ is the equivalence relation spanned by $-x \sim x$, $x \in S^n$. We will prove by induction on n that $\mathbb{R}P^n$ admits a CW-complex structure having exactly one open cell of dimension i for $i = 0, 1, \ldots, n$. The case $n = 0$ is trivial, since $\mathbb{R}P^0$ consists of just one point. To prove the induction step, we think of S^n as the equator of S^{n+1}, i.e., we identify $\mathbb{R}^n$ with the subspace of $\mathbb{R}^{n+1}$ spanned by the first n vectors of the canonical basis. The quotient map $q : S^{n+1} \to \mathbb{R}P^{n+1}$ restricts to a quotient map $q|_{S^n} : S^n \to q(S^n)$ (by item (4) of Exercise 1.91) so that we can identify $q(S^n) \subset \mathbb{R}P^{n+1}$ with $\mathbb{R}P^n$. Obviously, $\mathbb{R}P^{n+1}$ is the union of $\mathbb{R}P^n$ and the homeomorphic image by q of any open hemisphere of S^{n+1}, which is an open $(n+1)$-cell. A characteristic map for such open $(n+1)$-cell is obtained by taking the restriction of q to a closed hemisphere of S^{n+1}.

1.15.8. EXAMPLE (CW-complex structure on the complex projective space). We think of S^{2n+1} as the unit sphere of the complex space $\mathbb{C}^{n+1} \cong \mathbb{R}^{2n+2}$ and we consider the action of the group $S^1 \subset \mathbb{C}$ in S^{2n+1} given by

$$\lambda \cdot (z_1, \ldots, z_{n+1}) = (\lambda z_1, \ldots, \lambda z_{n+1}), \quad \lambda \in S^1,\ (z_1, \ldots, z_{n+1}) \in \mathbb{C}^{n+1}.$$

The corresponding orbit space S^{2n+1}/S^1 is called the *n-dimensional*[15] *complex projective space* and is denoted by $\mathbb{C}P^n$. We now show by induction on n that $\mathbb{C}P^n$ admits a cellular decomposition having exactly one open cell of dimension $2i$ for $i = 0, 1, \ldots, n$. The space $\mathbb{C}P^0$ consists of one single point. To prove the induction step, we identify S^{2n+1} with the subset of S^{2n+3} consisting of those $(n+2)$-tuples in $\mathbb{C}^{2n+2}$ whose last coordinate is zero. The quotient map $q : S^{2n+3} \to \mathbb{C}P^{n+1}$ therefore restricts to a quotient map from S^{2n+1} to $q(S^{2n+1}) \subset \mathbb{C}P^{n+1}$ and so we can identify $q(S^{2n+1})$ with $\mathbb{C}P^n$. The complement of $\mathbb{C}P^n$ in $\mathbb{C}P^{n+1}$ is an open $(2n+2)$-cell; namely, the restriction of q to the set

$$\big\{(z_1, \ldots, z_{n+2}) \in S^{2n+3} : z_{2n+2} \in \,]0, +\infty[\,\big\} \subset S^{2n+3}$$

is a homeomorphism onto $\mathbb{C}P^{n+1} \setminus \mathbb{C}P^n$ and this set can be identified with an open hemisphere of $S^{2n+2} \subset \mathbb{R}^{2n+3} \cong \mathbb{C}^{n+1} \times \mathbb{R}$. A characteristic map for such open cell is obtained by taking the restriction of q to the set:

$$\big\{(z_1, \ldots, z_{n+2}) \in S^{2n+3} : z_{2n+2} \in [0, +\infty[\,\big\} \subset S^{2n+3}.$$

1.15.9. DEFINITION. A *CW-subcomplex* (or simply a *subcomplex*) of a CW-complex X is a closed subset $Y \subset X$ that is the union of some open cells of X. It is easy to see that the cellular decomposition of X induces a cellular decomposition for Y making it a CW-complex (see Exercise 1.122).

1.15.10. EXAMPLE. For $p \geq 0$, the p-th skeleton of a CW-complex X, denoted by X^p, is the subcomplex of X that is the union of all open cells of X of dimension less than or equal to p:

$$X^p = \bigcup_{\substack{e \in \mathfrak{E} \\ \dim(e) \leq p}} e.$$

[15] Actually n is the dimension of $\mathbb{C}P^n$ as a *complex* manifold.

For $p < 0$ we set $X^p = \emptyset$.

1.15.11. DEFINITION. If X, Y are CW-complexes then we say that $f : X \to Y$ is a *cellular map* if f is continuous and maps each skeleton X^p of X into the corresponding skeleton Y^p of Y for every p.

1.15.12. PROPOSITION. *Let X be a CW-complex of dimension p ($p \geq 1$). For each open p-cell $e \in \mathfrak{E}_p$ choose a characteristic map $f_e : \overline{\mathrm{B}}^p \to \overline{e}$ for e. Then the map $q : \big(\sum_{e\in\mathfrak{E}_p} \overline{\mathrm{B}}^p\big) + X^{p-1} \to X$ induced by the f_e's and by the inclusion of X^{p-1} on X is a surjective quotient map. In particular, q induces a homeomorphism from the attachment space $\big(\sum_{e\in\mathfrak{E}_p} \overline{\mathrm{B}}^p\big) \cup_f X^{p-1}$ to X, where $f : \sum_{e\in\mathfrak{E}_p} S^{p-1} \to X^{p-1}$ is the sum of the restrictions of the f_e's to the spheres S^{p-1}.*

PROOF. Observe first that X is the coherent union of the skeleton X^{p-1} and of the closures $\overline{e}$ of the open p-cells $e \in \mathfrak{E}_p$. Moreover, each characteristic map $f_e : \overline{\mathrm{B}}^p \to \overline{e}$ is a quotient map, for $\overline{\mathrm{B}}^p$ is compact and $\overline{e}$ is Hausdorff. The conclusion follows from Exercise 1.105. $\square$

1.15.13. COROLLARY. *Every CW-complex is a T_4 topological space.*

PROOF. Let X be a CW-complex. We first show by induction that every skeleton X^p is T_4. The 0-skeleton is discrete and hence obviously T_4. If X^p is T_4 then by Proposition 1.15.12, the skeleton X^{p+1} is homeomorphic to the attachment of X^p with the topological sum of a family of closed balls $\overline{\mathrm{B}}^p$ along their boundaries. It follows from Lemma 1.14.15 that X^{p+1} is T_4. Now, since all skeletons are T_4 and closed in X and since the union $X = \bigcup_{p\geq 0} X_p$ is coherent, it follows from Lemma 1.14.16 that X is T_4. $\square$

1.15.14. PROPOSITION. *Let X be a CW-complex and let $\sum_{i\in I} \overline{\mathrm{B}}^{p_i}$ be an arbitrary topological sum of closed balls, where the p_i's are arbitrary integers. Let $f : \sum_{i\in I} S^{p_i-1} \to X$ be a continuous map such that $f\big(S^{p_i-1}\big) \subset X^{p_i-1}$ for every $i \in I$. Then the attachment space $X' = \sum_{i\in I} \overline{\mathrm{B}}^{p_i} \cup_f X$ is a CW-complex whose open cells are identified with the open cells of X and with the open balls B^{p_i}.*

PROOF. It follows from Corollary 1.15.13 and Lemma 1.14.15 that X' is T_4 and therefore Hausdorff[16]. The canonical projection $q : \sum_{i\in I} \overline{\mathrm{B}}^{p_i} + X \to X'$ maps X homeomorphically onto a closed subset of X' and $\sum_{i\in I} \mathrm{B}^{p_i}$ homeomorphically onto an open subset of X' (see Exercise 1.108). It follows easily that X' is the disjoint union of the (image by q of) the open cells of X and the image by q of the open balls B^{p_i} (that are the new open cells). The characteristic maps for the cells of X' are obtained using the old characteristic maps for the cells of X and appropriate restrictions of q for the characteristic maps of the new cells. For closure-finiteness we need the closure-finiteness property of X and Exercise 1.123 to conclude that $q(S^{p_i-1})$ is contained in a finite union of open $(p_i - 1)$-cells of

[16]This was the hard part of the proof. Namely, this was the motivation for the development of theory of T_4 spaces in Section 1.14.

X'. Finally, it follows from Exercise 1.106 that X' is the coherent union of the sets $q\big(\overline{\mathrm{B}}^{p_i}\big)$, $i \in I$, and $q(X)$; the weak-topology property of X' follows then from the weak-topology property of X. $\square$

1.15.15. PROPOSITION. *Let X be a topological space and $(X_n)_{n\ge 1}$ an increasing sequence of subspaces of X such that the union $X = \bigcup_{n\ge 1} X_n$ is coherent. Assume that each X_n is endowed with the structure of a CW-complex in such a way that X_n is a subcomplex of X_{n+1} for all n. Then X is a CW-complex whose open cells are precisely the open cells of the X_n's.*

PROOF. Since the union $X = \bigcup_{n\ge 1} X_n$ is coherent, the fact that X_n is closed in X_m for $n \le m$ imply that each X_n is closed in X. It follows that X is T_4 (see Corollary 1.15.13 and Lemma 1.14.16) and, in particular, it is Hausdorff. The other properties of a CW-complex are of straightforward verification. $\square$

1.15.16. LEMMA. *If X is a CW-complex and $e \in \mathfrak{E}$ is an open p-cell then for every $q \in e$ the set $\dot{e}$ is a strong deformation retract of the punctured cell $\overline{e}_\times = \overline{e} \setminus \{q\}$.*

PROOF. Let $f : \overline{\mathrm{B}}^p \to \overline{e}$ be a characteristic map for e; using Lemma 1.10.16 it is easy to see that f can be chosen so that $f(0) = q$. Now the sphere S^{p-1} is a strong deformation retract of the punctured closed ball $\overline{\mathrm{B}}^p_\times$ in the obvious way; since f is a quotient map from $\overline{\mathrm{B}}^p_\times$ to $\overline{e}_\times$ the conclusion follows easily from Corollary 1.14.6. $\square$

1.15.17. COROLLARY. *If one chooses a point q_e in each open p-cell $e \in \mathfrak{E}_p$ then the skeleton X^{p-1} is a strong deformation retract of the set*

$$(X^p)_\times = X^p \setminus \big\{q_e : e \in \mathfrak{E}_p\big\}.$$

PROOF. It is an easy consequence of Exercise 1.102 and the fact that $(X^p)_\times$ is the coherent union of the family consisting of the skeleton X^{p-1} and the punctured p-cells $\overline{e}_\times$, $e \in \mathfrak{E}_p$ (see Exercise 1.100). $\square$

1.15.18. LEMMA. *Let M be a differentiable manifold and let $\phi : M \to \mathbb{R}^n$ be a continuous map. Given a continuous function $\varepsilon : M \to \,]0, +\infty[$, a closed subset $F \subset M$ and an open subset $U \subset M$ with $\overline{U} \cap F = \emptyset$ then there exists a continuous map $\psi : M \to \mathbb{R}^n$ such that $\psi|_F = \phi|_F$, $\psi|_U$ is smooth and $\big\|\psi(x) - \phi(x)\big\| < \varepsilon(x)$ for all $x \in M$.*

PROOF. For every $x \in M$, let $U_x \subset M$ be an open neighborhood of x such that $\big\|\phi(y) - \phi(x)\big\| < \varepsilon(y)$ for all $y \in U_x$. We can subordinate a smooth partition of unity $\sum_{x\in M} \xi_x \equiv 1$ to the open covering $M = \bigcup_{x\in M} U_x$, i.e., each $\xi_x : M \to [0,1]$ is a smooth map whose support $\mathrm{supp}\xi_x$ is contained in U_x and the family $(\mathrm{supp}\xi_x)_{x\in M}$ is locally finite in M. Define a map $\tilde{\phi} : M \to \mathbb{R}$ by $\tilde{\phi}(y) = \sum_{x\in M} \phi(x)\xi_x(y)$; since each ξ_x is smooth and $(\mathrm{supp}\xi_x)_{x\in M}$ is locally finite, it follows that $\tilde{\phi}$ is smooth. Moreover, for every $y \in M$:

$$\begin{aligned}\big\|\tilde{\phi}(y) - \phi(y)\big\| &= \Big\|\sum_{x\in M}\phi(x)\xi_x(y) - \sum_{x\in M}\phi(y)\xi_x(y)\Big\| \\ &\le \sum_{x\in M}\big\|\phi(x) - \phi(y)\big\|\,\xi_x(y) < \varepsilon(y);\end{aligned}$$

the last inequality is obtained by observing that when $\xi_x(y) \neq 0$ then $y \in U_x$. In order to conclude the proof, let $\alpha : M \to [0,1]$ be a smooth map with $\alpha|_F \equiv 0$ and $\alpha|_{\overline{U}} \equiv 1$, set $\psi = \alpha\tilde{\phi} + (1-\alpha)\phi$ and observe that $\big\|\psi(x) - \phi(x)\big\| \le \big\|\tilde{\phi}(x) - \phi(x)\big\|$ for all $x \in M$. □

1.15.19. COROLLARY. *Under the hypothesis of Lemma 1.15.18, if $V \subset \mathbb{R}^n$ is an open subset containing the image of ϕ then the map ψ in the thesis of the lemma can be chosen in such a way that its image is contained in V.*

PROOF. Apply Lemma 1.15.18 replacing $\varepsilon(x)$ with the minimum between $\varepsilon(x)$ and the distance between $\phi(x)$ and the complement of V in $\mathbb{R}^n$. □

1.15.20. PROPOSITION. *Let M be a p-dimensional differentiable manifold, X a CW-complex and $f : M \to X$ a continuous map whose image is contained in some skeleton of X (this happens, for instance, if M is compact). Then, given a subset $S \subset M$ with $f(S) \subset X^p$, there exists a continuous map $g : M \to X$ that is homotopic to f relatively to S and such that the image of g is contained in X^p.*

PROOF. It suffices to show that if $f(M) \subset X^n$ for some $n > p$ then f is homotopic relatively to S to a continuous map whose image is contained in X^{n-1}. Moreover, by Corollary 1.15.17, it suffices to find a continuous map $g : M \to X$ homotopic to f relatively to S such that $g(X) \subset X^n$ and such that $g(M)$ does not contain at least one point in each open n-cell of X, i.e., such that $e \not\subset g(M)$ for every $e \in \mathfrak{E}_n$. We identify every open n-cell $e \in \mathfrak{E}_n$ with the unit open ball in $\mathbb{R}^n$ via an arbitrary homeomorphism; once this identification is made, we denote by e_r, $r \in\,]0,1[$, the open subset in e that corresponds by such homeomorphism to the open ball of radius r. Observe that since e is an open cell of maximal dimension in X^n, then e is indeed an open subset of X^n (see Exercise 1.120) and thus $f^{-1}(e)$ (and each $f^{-1}(e_r)$) is an open subset of M. We now apply Corollary 1.15.19 to the map $f|_{f^{-1}(e)}$ on the differentiable manifold $f^{-1}(e)$, where the open subset $U \subset f^{-1}(e)$ is $f^{-1}(e_{\frac{1}{3}})$, the closed subset $F \subset f^{-1}(e)$ is $f^{-1}(e) \setminus f^{-1}(e_{\frac{1}{2}})$ and $\varepsilon \equiv \frac{1}{6}$; we thus obtain a continuous map $\psi_e : f^{-1}(e) \to e$ that is smooth on $f^{-1}(e_{\frac{1}{3}})$, equals f outside $f^{-1}(e_{\frac{1}{2}})$ and such that $\big\|\psi_e(x) - f(x)\big\| < \frac{1}{6}$ for all $x \in f^{-1}(e)$. Once ψ_e is defined for every $e \in \mathfrak{E}_n$, we define $g : M \to X$ by:

$$g(x) = \begin{cases} f(x), & x \notin \bigcup_{e\in\mathfrak{E}_n} f^{-1}(e), \\ \psi_e(x), & x \in f^{-1}(e). \end{cases}$$

Observe that g actually equals f on $M \setminus \bigcup_{e\in\mathfrak{E}_n} f^{-1}\big(\overline{e_{\frac{1}{2}}}\big)$; this set is open in M, for $\bigcup_{e\in\mathfrak{E}_n} \overline{e_{\frac{1}{2}}}$ is closed in X by the weak-topology axiom. It follows that $g : M \to X$

is continuous. Since g is smooth on $f^{-1}(e_{\frac{1}{3}})$ and $\dim(M) < \dim(e)$, it follows that g maps $f^{-1}(e_{\frac{1}{3}})$ onto a subset of null measure in e. For $x \notin f^{-1}(e_{\frac{1}{3}})$ it cannot be $g(x) \in e_{\frac{1}{6}}$, because $g(x) \notin f^{-1}(e)$ for $x \notin f^{-1}(e)$ and $\big\|g(x) - f(x)\big\| < \frac{1}{6}$ for $x \in f^{-1}(e)$. Since $e_{\frac{1}{6}}$ cannot be contained in a set of null measure, it follows that $e_{\frac{1}{6}}$ (and hence e) is not contained in the image of g. Finally, one can construct a homotopy between f and g that is constant on $M \setminus \bigcup_{e \in \mathfrak{E}_n} f^{-1}\big(\overline{e_{\frac{1}{2}}}\big)$ and "linear" on each $f^{-1}(e)$, $e \in \mathfrak{E}_n$ (see Exercise 1.53). Observe that such homotopy is relative to S because S is disjoint from every $f^{-1}(e)$, $e \in \mathfrak{E}_n$. □

1.16. Homology of CW-complexes

To every CW-complex X we will associate a chain complex called the cellular chain complex of X. We then show that the homology of the cellular chain complex is naturally isomorphic to the singular homology of X.

In what follows, X will always denote a CW-complex and $\mathfrak{E}$ its set of open cells. Recall that X^p denotes the p-th skeleton of X.

1.16.1. DEFINITION. For every $p \in \mathbb{Z}$, we set $\mathcal{D}_p(X) = H_p(X^p, X^{p-1})$ and we consider the homomorphism $\partial_p : \mathcal{D}_p(X) \to \mathcal{D}_{p-1}(X)$ obtained by the composition:

$$H_p(X^p, X^{p-1}) \xrightarrow{\partial_*} H_{p-1}(X^{p-1}) \xrightarrow{i_*} H_{p-1}(X^{p-1}, X^{p-2})$$

where ∂_* is the connecting homomorphism of the long exact homology sequence of the pair (X^p, X^{p-1}) and $i : (X^{p-1}, \emptyset) \to (X^{p-1}, X^{p-2})$ denotes the inclusion map. We call $\big(\mathcal{D}(X), \partial\big)$ the *cellular chain complex* associated to X (see Lemma 1.16.2 below)

Since (by convention) $X^p = \emptyset$ for all $p < 0$, we have $\mathcal{D}_0(X) = H_0(X_0)$ and $\mathcal{D}_p(X) = 0$ for all $p < 0$.

We start by showing that $\big(\mathcal{D}(X), \partial\big)$ is indeed a chain complex.

1.16.2. LEMMA. *$\big(\mathcal{D}(X), \partial\big)$ is a chain complex, i.e., $\partial_{p-1} \circ \partial_p = 0$ for all $p \in \mathbb{Z}$.*

PROOF. The map $\partial_{p-1} \circ \partial_p$ is given by the composition of the following four homomorphisms:

$$H_p(X^p, X^{p-1}) \xrightarrow{\partial_*} H_{p-1}(X^{p-1}) \xrightarrow{i_*} H_{p-1}(X^{p-1}, X^{p-2}) \xrightarrow{\partial_*}$$
$$\xrightarrow{\partial_*} H_{p-2}(X^{p-2}) \xrightarrow{i_*} H_{p-2}(X^{p-2}, X^{p-3})$$

The vanishing of $\partial_{p-1} \circ \partial_p$ follows by observing that the middle part of the sequence above is part of the long exact homology sequence of the pair (X^{p-1}, X^{p-2}). □

The results below will provide a better understanding of how the cellular chain complex $\mathcal{D}(X)$ is related to the cellular structure of X.

1.16.3. LEMMA. *Let $f:\overline{\mathrm{B}}^p\to X$ be a characteristic map for an open p-cell $e\in\mathfrak{E}$. Then, for every $i\in\mathbb{Z}$, the map f induces an isomorphism:*

$$f_*:H_i\big(\overline{\mathrm{B}}^p,S^{p-1}\big)\longrightarrow H_i(\overline{e},\dot{e}); \tag{1.16.1}$$

in particular $H_i(\overline{e},\dot{e})$ is zero for $i\ne p$ and is infinite cyclic for $i=p$.

PROOF. Set $q=f(0)\in e$ and consider the commutative diagram:

$$\begin{tikzcd} H_i(\overline{e},\dot{e}) \arrow[r, "\mathfrak{i}^1_*", "\cong"'] & H_i(\overline{e},\overline{e}_\times) & H_i(e,e_\times) \arrow[l, "\mathfrak{j}^1_*"', "\cong"] \\ H_i\big(\overline{\mathrm{B}}^p,S^{p-1}\big) \arrow[u, "f_*"] \arrow[r, "\cong", "\mathfrak{i}^2_*"'] & H_i\big(\overline{\mathrm{B}}^p,\overline{\mathrm{B}}^p_\times\big) \arrow[u, "f_*"'] & H_i\big(\mathrm{B}^p,\mathrm{B}^p_\times\big) \arrow[l, "\cong"', "\mathfrak{j}^2_*"] \arrow[u, "\cong", "f_*"'] \end{tikzcd} \tag{1.16.2}$$

where $\mathfrak{i}^1,\mathfrak{i}^2,\mathfrak{j}^1,\mathfrak{j}^2$ denote inclusions and $\overline{e}_\times=\overline{e}\setminus\{q\}$, $e_\times=e\setminus\{q\}$, $\overline{\mathrm{B}}^p_\times=\overline{\mathrm{B}}^p\setminus\{0\}$, $\mathrm{B}^p_\times=\mathrm{B}^p\setminus\{0\}$. The fact that $\mathfrak{i}^1_*$ and $\mathfrak{i}^2_*$ are isomorphisms follows from the fact that $\dot{e}$ is a deformation retract of $\overline{e}_\times$ (see Lemma 1.15.16) and S^{p-1} is a deformation retract of $\overline{\mathrm{B}}^p_\times$. The fact that $\mathfrak{j}^1_*$ and $\mathfrak{j}^2_*$ are isomorphisms follows by excision. Finally, the fact that the map f_* on the rightmost column of the diagram is an isomorphism follows by observing that $f:(\mathrm{B}^p,\mathrm{B}^p_\times)\to(e,e_\times)$ is a homeomorphism of pairs. The conclusion now follows by observing that the commutativity of the diagram implies that the other two maps f_* on the vertical arrows are isomorphisms, as well. $\square$

1.16.4. LEMMA. *Let $e\in\mathfrak{E}$ be an open p-cell of X and let β be a generator of the infinite cyclic group $H_p(\overline{e},\dot{e})$. For every $q\in e$, if we set $e_\times=e\setminus\{q\}$ and $\overline{e}_\times=\overline{e}\setminus\{q\}$, then the top row of diagram* (1.16.2) *(with $i=p$) defines an isomorphism from $H_p(\overline{e},\dot{e})$ to $H_p\big(e,e\setminus\{q\}\big)$ that carries β to a generator $\tau(q)$ of the local homology group $H_p\big(e,e\setminus\{q\}\big)$. The map:*

$$e\ni q\longmapsto\tau(q)\in\mathcal{O}(e)$$

thus obtained is a continuous section of the orientation bundle $\mathcal{O}(e)$ and is therefore an orientation for the p-dimensional topological manifold e. Moreover, the correspondence $\beta\mapsto\tau$ just described is a bijection between the (two element) set of generators of $H_p(\overline{e},\dot{e})$ and the set of orientations of the topological manifold e.

PROOF. The case $p=0$ is trivial, so assume $p\ge1$. Let $f:\overline{\mathrm{B}}^p\to X$ be a characteristic map for e and denote by α the generator of the infinite cyclic group $H_p\big(\overline{\mathrm{B}}^p,S^{p-1}\big)$ that is mapped to the canonical orientation $\alpha^{[p-1]}\in\widetilde{H}_{p-1}(S^{p-1})$ of S^{p-1} via the connecting homomorphism ∂_* of the long exact homology sequence of the pair $\big(\overline{\mathrm{B}}^p,S^{p-1}\big)$. The isomorphism (1.16.1) (with $i=p$) takes α to $\pm\beta$; for definiteness, let's assume $f_*(\alpha)=\beta$. For every $v\in\mathrm{B}^p$ we claim that the isomorphism:

$$f_*:H_p\big(\mathrm{B}^p,\mathrm{B}^p\setminus\{v\}\big)\longrightarrow H_p\big(e,e\setminus\{f(v)\}\big),$$

takes the canonical orientation $\tau^{[p]}(v)$ of $\mathbb{R}^p$ to $\tau\big(f(v)\big)$. Once we prove the claim, the continuity of τ will follow (using Proposition 1.10.12). To prove the claim, set

$q = f(v) \in e$, $e_\times = e \setminus \{q\}$, $\overline{e}_\times = \overline{e} \setminus \{q\}$ and consider the commutative diagram (1.16.2) with $\overline{\mathrm{B}}^p_\times$ and $\mathrm{B}^p_\times$ replaced by $\overline{\mathrm{B}}^p \setminus \{v\}$ and by $\mathrm{B}^p \setminus \{v\}$ respectively; more explicitly:

$$\begin{tikzcd} H_i(\overline{e}, \dot{e}) \arrow[r, "\cong"'] & H_i(\overline{e}, \overline{e}_\times) & H_i(e, e_\times) \arrow[l, "\cong"'] \\ H_i\big(\overline{\mathrm{B}}^p, S^{p-1}\big) \arrow[u, "f_*"] \arrow[r, "\cong"] & H_i\big(\overline{\mathrm{B}}^p, \overline{\mathrm{B}}^p \setminus \{v\}\big) \arrow[u, "f_*"'] & H_i\big(\mathrm{B}^p, \mathrm{B}^p \setminus \{v\}\big) \arrow[l, "\cong"'] \arrow[u, "\cong", "f_*"'] \end{tikzcd} \tag{1.16.3}$$

Recalling Convention 1.10.24 (see diagram (1.10.7)), it follows from the result of Exercise 1.71 that the bottom arrow of (1.16.3) carries α to $\tau^{[p]}(v)$. The claim (and the continuity of τ) follows then easily from the commutativity of (1.16.3), since the top row of (1.16.3) takes β to $\tau(q)$.

Finally, the last assertion on the statement of the lemma follows trivially from Proposition 1.10.11. □

1.16.5. DEFINITION. If $e \in \mathfrak{E}$ is an open p-cell of X then a generator of the group $H_p(\overline{e}, \dot{e}) \cong \mathbb{Z}$ will be called an *orientation* for e.

1.16.6. REMARK. According to Lemma 1.16.4, the orientations of e in the sense of Definition 1.16.5 above can be identified with the orientations of the topological manifold e.

1.16.7. REMARK. A nice way of fixing an orientation for an open p-cell $e \in \mathfrak{E}$ consists in choosing a characteristic map $f : \overline{\mathrm{B}}^p \to X$ for e; namely, the homeomorphism $f|_{\mathrm{B}^p} : \mathrm{B}^p \to e$ carries the canonical orientation $\tau^{[p]}$ of B^p to a orientation τ for the manifold e (so that $f|_{\mathrm{B}^p} : \big(\mathrm{B}^p, \tau^{[p]}\big) \to (e, \tau)$ becomes a positively oriented homeomorphism).

During the proof of Lemma 1.16.4, we have actually shown the following fact: for $p \geq 1$, if α denotes the generator of $H_p\big(\overline{\mathrm{B}}^p, S^{p-1}\big)$ that is mapped to $\alpha^{[p-1]}$ via the connecting homomorphism ∂_* of the long exact homology sequence of the pair $\big(\overline{\mathrm{B}}^p, S^{p-1}\big)$ then the generator β of $H_p(\overline{e}, \dot{e})$ corresponding to the orientation τ of e is precisely the image of α by the isomorphism (1.16.1) (with $i = p$). This same statement (obviously) also holds for $p = 0$ if one takes α to be the canonical generator of $H_0\big(\overline{\mathrm{B}}^0, S^{-1}\big) = H_0\big(\{0\}\big)$, i.e., the homology class of the singular 0-simplex determined by the point 0.

We can now finally describe the group $\mathcal{D}_p(X)$. Recall that $\mathfrak{E}_p$ denotes the set of open p-cells of X.

1.16.8. LEMMA. *For any $p \geq 0$, the homomorphism*

$$\bigoplus_{e \in \mathfrak{E}_p} H_i(\overline{e}, \dot{e}) \longrightarrow H_i(X^p, X^{p-1})$$

induced by inclusion is an isomorphism. In particular, $H_i(X^p, X^{p-1}) = 0$ for $i \neq p$ and $H_p(X^p, X^{p-1})$ is free and its rank equals the number of open p-cells of X.

PROOF. Choose a point $q_e \in e$ for every open p-cell $e \in \mathfrak{E}_p$ and define $(X^p)_\times$ and $\overline{e}_\times$ as in Corollary 1.15.17. Consider the commutative diagram:

$$
\begin{tikzcd}
\bigoplus_{e\in\mathfrak{E}_p} H_p(\overline{e}, \dot{e}) \arrow[r] \arrow[d, "\text{by Lemma 1.15.16}"', "\cong"] & H_p(X^p, X^{p-1}) \arrow[d, "\cong"', "\text{by Corollary 1.15.17}"] \\
\bigoplus_{e\in\mathfrak{E}_p} H_p\big(\overline{e}, \overline{e}_\times\big) \arrow[r] & H_p\big(X^p, (X^p)_\times\big) \\
\bigoplus_{e\in\mathfrak{E}_p} H_p\big(e, e_\times\big) \arrow[r, "\cong"'] \arrow[u, "\text{by excision}", "\cong"'] & H_p\Big(\bigcup_{e\in\mathfrak{E}_p} e, \bigcup_{e\in\mathfrak{E}_p} e_\times\Big) \arrow[u, "\cong", "\text{by excision}"']
\end{tikzcd}
\tag{1.16.4}
$$

where all arrows are induced by inclusion. The fact that the bottom arrow of the diagram is an isomorphism follows from the result of Exercise 1.49 (observe that each $e \in \mathfrak{E}_p$ is open in X^p by the result of Exercise 1.120). The commutativity of the diagram now implies that the top arrow is also an isomorphism and that is precisely our thesis. □

Lemma 1.16.8 tells us in particular that the homomorphisms

$$H_p(\overline{e}, \dot{e}) \longrightarrow H_p(X^p, X^{p-1})$$

induced by inclusion are injective. We shall therefore identify $H_p(\overline{e}, \dot{e})$ with a subgroup of $H_p(X^p, X^{p-1})$ for every $e \in \mathfrak{E}_p$. Keeping in mind also the identification between the orientations of the p-dimensional topological manifold e and the generators of the group $H_p(\overline{e}, \dot{e}) \cong \mathbb{Z}$ (see Remark 1.16.6) we obtain the following:

1.16.9. COROLLARY. *For each $p \in \mathbb{Z}$ the group $\mathcal{D}_p(X)$ is free. One obtains a basis for $\mathcal{D}_p(X)$ by choosing orientations for all open p-cells of X.* □

1.16.10. REMARK. If $f : X \to Y$ is a cellular map between CW-complexes X and Y then for every $p \in \mathbb{Z}$, f restricts to a map of pairs:

$$f : (X^p, X^{p-1}) \longrightarrow (Y^p, Y^{p-1});$$

such map of pairs induces a homomorphism in the p-th homology group, i.e., a homomorphism from $\mathcal{D}_p(X)$ to $\mathcal{D}_p(Y)$. We shall denote such homomorphism by:

$$(f_\#)_p : \mathcal{D}_p(X) \longrightarrow \mathcal{D}_p(Y),$$

and we call it the *chain map induced by the cellular map f*. The fact that $f_\#$ indeed defines a chain map from $\mathcal{D}(X)$ to $\mathcal{D}(Y)$ follows easily from the naturality of the long exact homology sequence of a pair. If A is a subcomplex of X then the inclusion $i : A \to X$ is a cellular map; keeping in mind Lemma 1.16.8 and denoting by $\mathfrak{E}'_p \subset \mathfrak{E}_p$ the set of open p-cells of A, we get a commutative diagram:

$$\begin{tikzcd}
\bigoplus_{p\in\mathfrak{E}'_p} H_p(\overline{e},\dot{e}) \arrow[r, "\cong"] \arrow[d] & H_p(A^p, A^{p-1}) \arrow[d, "(i_\#)_p"] \\
\bigoplus_{p\in\mathfrak{E}_p} H_p(\overline{e},\dot{e}) \arrow[r, "\cong"'] & H_p(X^p, X^{p-1})
\end{tikzcd}$$

in which all arrows are induced by inclusion. It follows that the chain map $i_\#$ induced by the inclusion $i : A \to X$ is actually a chain isomorphism from $\mathcal{D}(A)$ onto the subcomplex of $\mathcal{D}(X)$ spanned by the orientations of the open cells of A. Hence, *we can identify the cellular complex of a CW-subcomplex $A \subset X$ with a chain subcomplex of the cellular complex of X by means of the chain map induced by inclusion*.

We are now going to prove that the homology of the chain complex $\mathcal{D}(X)$ is isomorphic to the singular homology of X. Our strategy is to construct two subcomplexes $\mathcal{Z}(X)$ and $\mathcal{B}(X)$ of the singular chain complex $\mathfrak{S}(X)$ in such a way that $\mathcal{D}(X) = \mathcal{Z}(X)/\mathcal{B}(X)$ and that both the inclusion $\mathcal{Z}(X) \to \mathfrak{S}(X)$ and the quotient map $\mathcal{Z}(X) \to \mathcal{D}(X)$ induce an isomorphism in homology.

We start by defining $\mathcal{Z}(X)$ and $\mathcal{B}(X)$ by setting, for every $p \in \mathbb{Z}$:

$$\mathcal{Z}_p(X) = Z_p(X^p, X^{p-1}) = \big\{c \in \mathfrak{S}_p(X^p) : \partial_p\, c \in \mathfrak{S}_{p-1}(X_{p-1})\big\},$$
$$\mathcal{B}_p(X) = B_p(X^p, X^{p-1}) = B_p(X^p) + \mathfrak{S}_p(X^{p-1}).$$

It is easy to see that $\mathcal{Z}(X)$ and $\mathcal{B}(X)$ are subcomplexes of $\mathfrak{S}(X)$; moreover, $\mathcal{D}_p(X) = \mathcal{Z}_p(X)/\mathcal{B}_p(X)$ for every $p \in \mathbb{Z}$. By looking explicitly at the definition of the connecting homomorphism ∂_* of the long exact homology sequence of the pair (X^p, X^{p-1}) (see Corollary 1.3.5) it is easy to see that the boundary homomorphism of the cellular chain complex $\big(\mathcal{D}(X), \partial\big)$ is induced by the boundary homomorphism of $\mathcal{Z}(X)$, i.e., $\big(\mathcal{D}(X), \partial\big)$ is equal to the quotient chain complex $\mathcal{Z}(X)/\mathcal{B}(X)$.

Before establishing the relation between the homologies of $\mathcal{Z}(X)$, $\mathcal{D}(X)$ and $\mathfrak{S}(X)$, we need a few technical lemmas regarding the homologies of the skeletons of X.

1.16.11. LEMMA. *For any $p \in \mathbb{Z}$, the inclusion $X^{p+1} \to X$ induces an isomorphism $H_p(X^{p+1}) \to H_p(X)$.*

PROOF. The long exact homology sequence of the pair (X^{i+1}, X^i) shows that

$$H_{p+1}(X^{i+1}, X^i) \xrightarrow{\partial_*} H_p(X^i) \longrightarrow H_p(X^{i+1}) \longrightarrow H_p(X^{i+1}, X^i)$$

is exact, where the unlabelled arrows are induced by inclusion. For every integer $i \geq p+1$, we conclude from the exactness of the sequence above and from Lemma 1.16.8 that the inclusion $X^i \to X^{i+1}$ induces an isomorphism from $H_p(X^i)$ to $H_p(X^{i+1})$; hence (by composition), the inclusion $X^i \to X^j$ induces an isomorphism from $H_p(X^i)$ to $H_p(X^j)$, for every $j \geq i \geq p+1$. The conclusion now follows from the result of Exercise 1.34; namely, we have a (trivial) filtration:

$$X_{p+1} \subset X_{p+1} \subset X_{p+1} \subset \cdots \subset X_{p+1}$$

of the topological space X_{p+1} and a filtration:

$$X_{p+1} \subset X_{p+2} \subset X_{p+3} \subset \cdots \subset X$$

for the topological space X. The inclusion $X_{p+1} \to X$ is a filtration preserving map. The fact that the hypotheses of the result stated in Exercise 1.34 are indeed satisfied is a consequence of the first part of the proof and of the result of Exercise 1.123. □

1.16.12. LEMMA. *For any $p, i \in \mathbb{Z}$ with $i < p$ we have $H_p(X^i) = 0$.*

PROOF. Given integers $i, j \in \mathbb{Z}$ with $j \leq i$, the long exact homology sequence of the triple (X^i, X^j, X^{j-1}) (see Exercise 1.52) shows that the sequence:

$$H_p(X^j, X^{j-1}) \longrightarrow H_p(X^i, X^{j-1}) \longrightarrow H_p(X^i, X^j) \xrightarrow{\partial_*} H_{p-1}(X^j, X^{j-1})$$

is exact, where the unlabelled arrows are induced by inclusion. For $j \leq p-2$ we conclude from the exactness of the sequence above and from Lemma 1.16.8 that the inclusion of (X^i, X^{j-1}) in (X^i, X^j) induces an isomorphism from $H_p(X^i, X^{j-1})$ to $H_p(X^i, X^j)$. Therefore (by composition), since $X_k = \emptyset$ for $k < 0$, the inclusion of X^i in (X^i, X^j) induces an isomorphism from $H_p(X^i)$ to $H_p(X^i, X^j)$, where $j = \min\{i, p-2\}$.

If $i \leq p-2$ we have $j = i$, so that $H_p(X^i) \cong H_p(X^i, X^j) = 0$ and the proof is complete. Otherwise, $i = p-1$ and $H_p(X^i) \cong H_p(X^i, X^j) = H_p(X^{p-1}, X^{p-2})$; from Lemma 1.16.8, we have $H_p(X^{p-1}, X^{p-2}) = 0$ and so the proof is complete as well. □

We can now prove our two main theorems.

1.16.13. THEOREM. *The inclusion $\mathcal{Z}(X) \to \mathcal{S}(X)$ induces an isomorphism in homology. Such isomorphism is natural, i.e., if $f\colon X \to Y$ is a cellular map then the diagram*

$$\begin{tikzcd} H_p(X) \arrow[r, "f_*"] & H_p(Y) \\ H_p\big(\mathcal{Z}(X)\big) \arrow[u] \arrow[r, "f_*"'] & H_p\big(\mathcal{Z}(Y)\big) \arrow[u] \end{tikzcd} \tag{1.16.5}$$

commutes for every p; the vertical arrows in the diagram above are induced by inclusion and the bottom arrow is induced by the chain map $f_\# : \mathcal{Z}(X) \to \mathcal{Z}(Y)$ obtained by restricting $f_\# : \mathfrak{S}(X) \to \mathfrak{S}(Y)$.

PROOF. For any p, the p-th cycle group of $\mathcal{Z}(X)$ is $Z_p(X^p)$ and the p-th boundary group of $\mathcal{Z}(X)$ is $B_p(X^{p+1}) \cap Z_p(X^p)$. We have to prove that the homomorphism:

$$\frac{Z_p(X^p)}{B_p(X^{p+1}) \cap Z_p(X^p)} \longrightarrow \frac{Z_p(X)}{B_p(X)},$$

induced by inclusion is an isomorphism.

By Lemma 1.16.11, the homomorphism:

$$\frac{Z_p(X^{p+1})}{B_p(X^{p+1})} \longrightarrow \frac{Z_p(X)}{B_p(X)}, \tag{1.16.6}$$

induced by inclusion is an isomorphism. The long exact sequence of the pair (X^{p+1}, X^p) shows that the sequence

$$H_p(X^p) \longrightarrow H_p(X^{p+1}) \longrightarrow H_p(X^{p+1}, X^p) \overset{\text{Lemma 1.16.8}}{=} 0$$

is exact, where all arrows are induced by inclusion. It follows that the homomorphism:

$$\frac{Z_p(X^p)}{B_p(X^p)} \longrightarrow \frac{Z_p(X^{p+1})}{B_p(X^{p+1})},$$

induced by inclusion is surjective; this implies that:

$$Z_p(X^{p+1}) = Z_p(X^p) + B_p(X^{p+1}).$$

By the result of Exercise 1.14, the homomorphism:

$$\frac{Z_p(X^p)}{B_p(X^{p+1}) \cap Z_p(X^p)} \longrightarrow \frac{Z_p(X^p) + B_p(X^{p+1})}{B_p(X^{p+1})} = \frac{Z_p(X^{p+1})}{B_p(X^{p+1})}, \tag{1.16.7}$$

induced by inclusion is an isomorphism.

Since both (1.16.6) and (1.16.7) are isomorphisms, the proof of the first part of the statement is complete. Finally, the commutativity of (1.16.5) follows by observing that such diagram already commutes at the chain level. □

1.16.14. COROLLARY. *The chain map $\mathcal{Z}(X) \otimes G \to \mathfrak{S}(X; G)$ induced by the inclusion of $\mathcal{Z}(X)$ in $\mathfrak{S}(X)$ induces an isomorphism in homology for every abelian group G.*

PROOF. Follows directly from Corollary 1.13.12, observing that $\mathfrak{S}(X)$ is free and hence the subcomplex $\mathcal{Z}(X)$ of $\mathfrak{S}(X)$ is also free. □

1.16.15. THEOREM. *The quotient map $\mathcal{Z}(X) \to \mathcal{D}(X)$ induces an isomorphism in homology. Such isomorphism is natural, i.e., if $f\colon X \to Y$ is a cellular map then the diagram*

$$\begin{tikzcd} H_p\big(\mathcal{Z}(X)\big) \arrow[r, "f_*"] \arrow[d] & H_p\big(\mathcal{Z}(Y)\big) \arrow[d] \\ H_p\big(\mathcal{D}(X)\big) \arrow[r, "f_*"'] & H_p\big(\mathcal{D}(Y)\big) \end{tikzcd} \tag{1.16.8}$$

commutes for every p; the vertical arrows in the diagram above are induced by the quotient map, the top arrow is induced by the chain map $f_\# : \mathcal{Z}(X) \to \mathcal{Z}(Y)$ obtained by restricting $f_\# : \mathfrak{S}(X) \to \mathfrak{S}(Y)$ and the bottom arrow is induced by the chain map $f_\# : \mathcal{D}(X) \to \mathcal{D}(Y)$ induced by f on the cellular complexes.

PROOF. We have a short exact sequence of chain complexes:

$$0 \longrightarrow \mathcal{B}(X) \longrightarrow \mathcal{Z}(X) \longrightarrow \mathcal{D}(X) \longrightarrow 0.$$

The corresponding long exact homology sequence, shows that the quotient map $\mathcal{Z}(X) \to \mathcal{D}(X)$ induces an isomorphism in homology if and only if the homology of $\mathcal{B}(X)$ vanishes. Let's then try to compute such homology.

For any p, the p-th cycle group of $\mathcal{B}(X)$ is $B_p(X^p) + Z_p(X^{p-1})$ and the p-th boundary group of $\mathcal{B}(X)$ is $B_p(X^p)$. We want to show that $Z_p(X^{p-1}) \subset B_p(X^p)$. By Lemma 1.16.12, we have $H_p(X^{p-1}) = 0$ so that:

$$Z_p(X^{p-1}) = B_p(X^{p-1}) \subset B_p(X^p).$$

This concludes the proof of the first part of the statement. The commutativity of diagram (1.16.8) follows by observing that such diagram already commutes at the chain level. □

1.16.16. COROLLARY. *The chain map $\mathcal{Z}(X) \otimes G \to \mathcal{D}(X) \otimes G$ induced by the quotient map $\mathcal{Z}(X) \to \mathcal{D}(X)$ induces an isomorphism in homology for every abelian group G.*

PROOF. Follows directly from Corollary 1.13.12, observing that $\mathcal{Z}(X)$ is free (since it is a subcomplex of $\mathfrak{S}(X)$) and $\mathcal{D}(X)$ is free by Corollary 1.16.9. □

We have proven the following:

1.16.17. THEOREM. *If X is a CW-complex then there exists an isomorphism between the homology of the cellular chain complex $\mathcal{D}(X)$ and the singular homology of X. The same statement holds for reduced homology and for homology with arbitrary coefficients. All the isomorphisms are natural with respect to cellular maps.* □

1.16.18. EXAMPLE. We have seen in Example 1.15.8 that the complex projective space $\mathbb{C}P^n$ admits a cellular decomposition having exactly one cell of dimension $2i$ for $i = 0, \ldots, n$. It follows readily from Theorem 1.16.17 that the homology of $\mathbb{C}P^n$ is given by:

$$H_{2i}(\mathbb{C}P^n) \cong \mathbb{Z}, \quad \text{for } i = 0, 1, \ldots, n,$$

and $H_p(\mathbb{C}P^n) = 0$ otherwise.

We will now prove some results relating the Betti numbers of a CW-complex X and the number of cells of X in each dimension.

1.16.19. PROPOSITION. *Let X be a CW-complex and, for each integer $p \geq 0$, denote by κ_p the number of open p-cells of X. Then, for every coefficient field $\mathbb{K}$ we have:*

$$\beta_p(X; \mathbb{K}) \leq \kappa_p, \tag{1.16.9}$$

for every p.

PROOF. Follows by observing that the $\mathbb{K}$-vector space $H_p(X;\mathbb{K})$ is isomorphic to a quotient of a subspace of $\mathcal{D}_p(X)\otimes\mathbb{K}$ and $\dim_{\mathbb{K}}\big[\mathcal{D}_p(X)\otimes\mathbb{K}\big]=\kappa_p$ for every p. □

1.16.20. PROPOSITION. *Let X be a finite (or, equivalently, compact) CW-complex. Denote by κ_p the number of open p-cells of X. Then the Euler characteristic of X is given by:*

$$\chi(X)=\sum_{p\in\mathbb{Z}}(-1)^p\kappa_p. \tag{1.16.10}$$

PROOF. Apply the result of Exercise 1.78 with f the identity of $\mathcal{D}(X)\otimes\mathbb{K}$ and Lemma 1.13.15 for an arbitrary coefficient field $\mathbb{K}$. □

1.16.21. PROPOSITION. *Let X be a CW-complex and denote by κ_p the number of open p-cells of X. Assume that for some $k\ge 0$ we have $\kappa_p<+\infty$ for all $p\le k$. Then, for any coefficient field $\mathbb{K}$:*

(1.16.11)
$$\beta_k(X;\mathbb{K})-\beta_{k-1}(X;\mathbb{K})+\cdots+(-1)^k\beta_0(X;\mathbb{K})\le\kappa_k-\kappa_{k-1}+\cdots+(-1)^k\kappa_0.$$

PROOF. Define a chain complex of $\mathbb{K}$-vector spaces $\mathfrak{C}$ by setting

$$\mathfrak{C}_p=\mathcal{D}_p(X)\otimes\mathbb{K},$$

for $p\le k$ and $\mathfrak{C}_p=0$ for $p>k$; the boundary operator in $\mathfrak{C}$ is defined so that $\mathfrak{C}$ is a subcomplex of $\mathcal{D}(X)\otimes\mathbb{K}$. If β'_p denotes the dimension over $\mathbb{K}$ of the homology group $H_p(\mathfrak{C})$ then, applying Exercise 1.78 with f the identity of $\mathfrak{C}$, we obtain:

$$\beta'_k-\beta'_{k-1}+\cdots+(-1)^k\beta'_0=\kappa_k-\kappa_{k-1}+\cdots+(-1)^k\kappa_0.$$

The conclusion follows by observing that $\beta'_p=\beta_p(X;\mathbb{K})$ for $p<k$ and $\beta'_p\ge\beta_p(X;\mathbb{K})$. □

1.16.22. PROPOSITION. *Let X be a CW-complex and denote by $\kappa_p\in\mathbb{N}\cup\{+\infty\}$ the number of open p-cells of X. Then, for any coefficient field $\mathbb{K}$, there exists a sequence $(q_p)_{p\ge0}$ in $\mathbb{N}\cup\{+\infty\}$ such that:*

$$\kappa_0=\beta_0(X;\mathbb{K})+q_0,\qquad \kappa_p=\beta_p(X;\mathbb{K})+q_p+q_{p-1},\quad p\ge 1. \tag{1.16.12}$$

PROOF. Denote by $\partial_p:\mathcal{D}_p(X)\otimes\mathbb{K}\to\mathcal{D}_{p-1}(X)\otimes\mathbb{K}$ the p-th boundary operator of the complex $\mathcal{D}(X)\otimes\mathbb{K}$. Set:

$$q_p=\dim\big(\mathcal{D}_{p+1}(X)\otimes\mathbb{K}/\mathrm{Ker}(\partial_{p+1})\big)=\dim\big(\mathrm{Im}(\partial_{p+1})\big),$$

for all $p\ge0$. The conclusion follows by applying the result of Exercise 1.80 to the spaces $\mathrm{Im}(\partial_{p+1})\subset\mathrm{Ker}(\partial_p)\subset\mathcal{D}_p(X)\otimes\mathbb{K}$. □

In spite of the awkward statement of Proposition 1.16.22, it is not hard to show that such Proposition actually implies Propositions 1.16.19, 1.16.20 and 1.16.21 (see Exercise 1.126).

The thesis of Proposition 1.16.22 can be nicely summarized in the following form. Consider the formal "power series" with coefficients in $\mathbb{N}\cup\{+\infty\}$ given by $Q(\lambda) = \sum_{p=0}^{+\infty} q_p \lambda^p$. Then equalities (1.16.12) are equivalent to the following equality of formal "power series" with coefficients in $\mathbb{N} \cup \{+\infty\}$:

$$\sum_{p=0}^{+\infty} \kappa_p \lambda^p = \sum_{p=0}^{+\infty} \beta_p(X;\mathbb{K})\lambda^p + (1+\lambda)Q(\lambda). \tag{1.16.13}$$

The formal power series:

$$\mathfrak{P}_\lambda(X;\mathbb{K}) = \sum_{p=0}^{+\infty} \beta_p(X;\mathbb{K})\lambda^p,$$

appearing in equation (1.16.13) is known as the *Poincaré polynomial* of the topological space X with respect to the field coefficient $\mathbb{K}$.

1.16.23. REMARK. If a singular homology group $H_p(X)$ (with integer coefficients) of a CW-complex X is finitely generated then the Betti number $\beta(X;\mathbb{K})$ is always greater or equal to the Betti number $\beta(X)$ (with integer coefficients) of X (recall Exercise 1.125). Namely, the universal coefficients theorem implies that $H_p(X) \otimes \mathbb{K}$ is a $\mathbb{K}$-vector subspace of $H_p(X;\mathbb{K})$ and therefore:

$$\beta(X;\mathbb{K}) = \dim_{\mathbb{K}}\big(H_p(X;\mathbb{K})\big) \geq \dim_{\mathbb{K}}\big(H_p(X)\otimes\mathbb{K}\big) \geq \beta_p(X).$$

It follows that the lower estimate (1.16.9) for κ_p is always better than (or equivalent to) the estimate $\beta_p(X) \leq \kappa_p$ *if* $H_p(X)$ *is finitely generated*. On the other hand, if $H_p(X)$ is not finitely generated then $\beta_p(X) = +\infty$ by convention and it is indeed true that $\beta_p(X) \leq \kappa_p$, i.e., that κ_p is also equal to $+\infty$. Namely, if κ_p were finite then $\mathcal{D}_p(X)$ would be free of finite rank and hence also $H_p(X)$ (being a quotient of a subgroup of $\mathcal{D}_p(X)$) would be finitely generated. Observe that if $H_p(X)$ is not finitely generated then it may happen that *no coefficient field* $\mathbb{K}$ *will give us the equality* $\kappa_p = +\infty$ *from* (1.16.9) (see Exercise 1.79).

1.17. Explicit Computation of the Cellular Complex

Let X be a CW-complex. We have seen in Section 1.16 that the singular homology of X is isomorphic to the homology of the cellular chain complex $\big(\mathcal{D}(X),\partial\big)$ corresponding to X. The boundary homomorphisms of $\mathcal{D}(X)$ were defined abstractly in terms of the long exact homology sequence of a pair of consecutive dimensional skeletons of X. The goal of this section is to give an explicit geometric method for computing such boundary homomorphisms. Recall from Corollary 1.16.9 that for each $p \geq 0$, the group $\mathcal{D}_p(X)$ is free abelian and a basis for $\mathcal{D}_p(X)$ is obtained by choosing an orientation for each open p-cell of X. More explicitly (recall Lemma 1.16.8), we have an isomorphism:

$$\bigoplus_{e\in\mathfrak{E}_p} H_p(\overline{e},\dot{e}) \longrightarrow \mathcal{D}_p(X) \tag{1.17.1}$$

induced by inclusion (recall that $\mathfrak{E}_p$ denotes the set of open p-cells of X). For every $e \in \mathfrak{E}_p$, the generators of the infinite cyclic group $H_p(\overline{e},\dot{e}) \cong \mathbb{Z}$ are (by definition) called the *orientations* of the open p-cell e; moreover, there is a natural correspondence between the set of generators of $H_p(\overline{e}, \dot{e})$ and the set of actual orientations for the p-dimensional topological manifold e (recall Remark 1.16.6).

In this section we will always identify the group $\mathcal{D}_p(X)$ with the direct sum $\bigoplus_{e\in\mathfrak{E}_p} H_p(\overline{e}, \dot{e})$ via the isomorphism (1.17.1). Moreover, once an orientation for an open p-cell e is fixed, we will simply denote by e the corresponding generator of $H_p(\overline{e}, \dot{e})$. Hence, we write the elements of $\mathcal{D}_p(X)$ simply as (finite) linear combinations of open p-cells of X with integer coefficients; the sign of the coefficient appearing next to some open p-cell e is determined once an orientation for e is fixed.

Let $e^{p+1} \in \mathfrak{E}$ be a fixed open $(p+1)$-cell of X. We choose an orientation for e^{p+1}. A good way of doing that (recall Remark 1.16.7) is choosing a characteristic map $f : \overline{\mathrm{B}}^{p+1} \to X$ for e^{p+1}. The boundary $\partial_{p+1} e^{p+1}$ of e^{p+1} in the chain complex $\mathcal{D}(X)$ equals a finite linear integral combination of open p-cells of X. Let then $e^p \in \mathfrak{E}$ be a fixed open p-cell of X; we want to determine the coefficient next to e^p appearing in $\partial_{p+1} e^{p+1}$. Such coefficient is only determined up to sign; by choosing an orientation for e^p, this coefficient becomes a well-defined integer number. The theorem below tells us how such number can be explicitly computed.

1.17.1. THEOREM. *Let X be a CW-complex and let $e^p, e^{p+1} \in \mathfrak{E}$ be respectively an open p-cell and an open $(p+1)$-cell of X ($p \geq 0$). Assume that $f : \overline{\mathrm{B}}^{p+1} \to X$ is a characteristic map for e^{p+1}, that e^{p+1} has the orientation induced by f and that e^p has a fixed arbitrary orientation. Then the set $f^{-1}(e^p)$ is open in S^p and the map:*

$$f|_{f^{-1}(e^p)} : f^{-1}(e^p) \subset S^p \longrightarrow e^p, \tag{1.17.2}$$

is proper. Moreover, the coefficient appearing next to e^p in the boundary of e^{p+1} in the complex $\mathcal{D}(X)$ equals the degree of the map (1.17.2).

PROOF. The fact that $f^{-1}(e^p)$ is contained in S^p follows by observing that e^p is disjoint from $e^{p+1} = f(\mathrm{B}^{p+1})$; the fact that $f^{-1}(e^p)$ is open in S^p follows from the continuity of $f|_{S^p} : S^p \to X^p$ and from the result of Exercise 1.120. Moreover, the properness of (1.17.2) follows from the result of Exercise 1.73, observing that $f|_{S^p} : S^p \to X^p$ is (obviously) proper.

We have shown so far that it makes sense to talk about the degree of (1.17.2); we now proceed with the proof that such degree equals the coefficient next to e^p in $\partial_{p+1} e^{p+1}$. Let α denote the generator of $H_{p+1}\big(\overline{\mathrm{B}}^{p+1}, S^p\big)$ that is mapped to $\alpha^{[p]} \in \widetilde{H}_p(S^p)$ by the connecting homomorphism ∂_* of the long exact homology sequence of the pair $\big(\overline{\mathrm{B}}^{p+1}, S^p\big)$; by Remark 1.16.7, the basis element of $\mathcal{D}_{p+1}(X)$ that is identified with the oriented $(p+1)$-cell e^{p+1} equals to the image of α by the homomorphism:

$$f_* : H_{p+1}\big(\overline{\mathrm{B}}^{p+1}, S^p\big) \longrightarrow H_{p+1}(X^{p+1}, X^p),$$

induced by f. Such homomorphism f_* is pictured in the leftmost column of the commutative diagram below:

$$\begin{tikzcd}
H_{p+1}(X^{p+1}, X^p) \arrow[r, "\partial_*"] & H_p(X^p) \arrow[r] & H_p(X^p, X^{p-1}) \\
H_{p+1}\big(\overline{\mathrm{B}}^{p+1}, S^p\big) \arrow[u, "f_*"] \arrow[r, "\partial_*"'] & \widetilde{H}_p(S^p) \arrow[u, "f_*"'] \arrow[ur, "f_*"']
\end{tikzcd}$$

the top row of the diagram is the $(p+1)$-th boundary homomorphism of the chain complex $\mathcal{D}(X)$. Since $\partial_*(\alpha) = \alpha^{[p]}$, the boundary of (the basis element of $\mathcal{D}_{p+1}(X)$ that is identified with the) open cell e^{p+1} in the chain complex $\mathcal{D}(X)$ equals the image of $\alpha^{[p]}$ by the homomorphism f_* represented by the slanted arrow in the diagram above; such homomorphism is represented again as the top row of the commutative diagram given below. We choose a point $q_e \in e$ for each open p-cell $e \in \mathfrak{E}_p$ and we set $e_\times = e \setminus \{q_e\}$ and $(X^p)_\times = X^p \setminus \bigcup_{e \in \mathfrak{E}_p} \{q_e\}$; here comes the diagram:

(1.17.3)
$$\begin{tikzcd}
\widetilde{H}_p(S^p) \arrow[r, "f_*"] \arrow[d] \arrow[ddddr, dashed, bend right] & H_p(X^p, X^{p-1}) \arrow[d, "\cong"] \\
H_p\Big(S^p, S^p \setminus \bigcup\limits_{e \in \mathfrak{E}_p} f^{-1}(q_e)\Big) \arrow[r, "f_*"] & H_p\big(X^p, (X^p)_\times\big) \\
H_p\Big(\bigcup\limits_{e \in \mathfrak{E}_p} f^{-1}(e), \bigcup\limits_{e \in \mathfrak{E}_p} f^{-1}(e_\times)\Big) \arrow[u, "\cong"] \arrow[r, "f_*"] \arrow[dr, "f_*"'] & H_p\Big(\bigcup\limits_{e \in \mathfrak{E}_p} e, \bigcup\limits_{e \in \mathfrak{E}_p} e_\times\Big) \arrow[u, "\cong"'] \arrow[d] \\
& H_p\Big(\bigcup\limits_{e \in \mathfrak{E}_p} e, \bigcup\limits_{e \in \mathfrak{E}_p} e \setminus \{q_{e^p}\}\Big) \\
& H_p\big(e^p, e^p_\times\big) \arrow[u, dotted, "\cong"']
\end{tikzcd}$$

as usual, the unlabelled arrows are induced by inclusion. Observe that the top part of the right column of the diagram above is precisely the right column of diagram (1.16.4).

Let $\beta \in H_p(\overline{e^p}, \dot{e^p})$ denote the chosen orientation on e^p; the generator β of $H_p(\overline{e^p}, \dot{e^p})$ corresponds to an orientation $\tau : e^p \to \mathcal{O}(e^p)$ for the topological manifold e^p (see Remark 1.16.6). The left column of diagram (1.16.4) restricts to a homomorphism $H_p(\overline{e^p}, \dot{e^p}) \to H_p(e^p, e^p_\times)$ that carries β to $\tau(q_{e^p})$. Let $d \in \mathbb{Z}$ denote the coefficient appearing next to e^p in the boundary of e^{p+1}, i.e., d is the integer we want to compute. Denote by $f_*\big(\alpha^{[p]}\big)$ the image of $\alpha^{[p]}$ by the top arrow of diagram (1.17.3). If we push $f_*\big(\alpha^{[p]}\big)$ down the right column of diagram (1.16.4) and then pull it back using the bottom arrow of (1.16.4), we will obtain an element of the direct sum $\bigoplus_{e\in\mathfrak{E}_p} H_p(e,e_\times)$ whose component in $H_p(e^p,e^p_\times)$ is $d\cdot\tau(q_{e^p})$. By the result of Exercise 1.50, it then follows that pushing $f_*\big(\alpha^{[p]}\big)$ all the way down the right column of diagram (1.17.3) will give us $d\cdot\tau(q_{e^p})$.

Now let $d'\in\mathbb{Z}$ denote the degree of the map (1.17.2). The proof of the theorem will be concluded if we can show that the dashed path in diagram (1.17.3) takes $\alpha^{[p]}$ to $d'\cdot\tau(q_{e^p})$. Let's observe the following things.

- *The union $\bigcup_{e\in\mathfrak{E}_p} e$ is a p-dimensional topological manifold*; namely each $e\in\mathfrak{E}_p$ is open in X^p (and hence in $\bigcup_{e\in\mathfrak{E}_p} e$) by the result of Exercise 1.120.
- *The set $U=\bigcup_{e\in\mathfrak{E}_p} f^{-1}(e)$ is open in S^p*; as in the item above, we know that each $e\in\mathfrak{E}_p$ is open in X^p. The conclusion follows from the continuity of $f|_{S^p}:S^p\to X^p$.
- *The set $K=\bigcup_{e\in\mathfrak{E}_p} f^{-1}(q_e)$ is compact*; obviously, for each $e\in\mathfrak{E}_p$, the set $f^{-1}(q_e)$ is closed in S^p and therefore compact. Observe now that, by Closure-finiteness, $f^{-1}(q_e)$ is non empty for at most a finite number of e's.
- $U\setminus K=\bigcup_{e\in\mathfrak{E}_p} f^{-1}(e_\times)$; this is obvious.
- *The map:*

$$f\big|_{\left[\bigcup_{e\in\mathfrak{E}_p} f^{-1}(e)\right]}:\bigcup_{e\in\mathfrak{E}_p} f^{-1}(e)\longrightarrow\bigcup_{e\in\mathfrak{E}_p} e \tag{1.17.4}$$

  is proper; this follows from the result of Exercise 1.73, observing that $f|_{S^p}:S^p\to X^p$ is proper.

We (as usual) use the isomorphism given by the dotted arrow of diagram (1.17.3) to identify orientations of the manifold e^p at the point q_{e^p} with orientations of the manifold $\bigcup_{e\in\mathfrak{E}_p} e$ at the point q_{e^p}. Keeping in mind such identification, the items above and Remark 1.11.2, it follows that the dashed path of diagram (1.17.3) takes $\alpha^{[p]}$ to $d''\cdot\tau(q_{e^p})$, where d'' equals the degree of the map (1.17.4) at the point q_{e^p} with respect to the orientation $\tau(q_{e^p})$. We now have only to observe that $d''=d'$; this follows from items (1) and (2) of Proposition 1.11.3. $\square$

We are now going to present a few examples in which all the machinery we have developed will be used to actually compute the singular homology of some spaces. Before that, we make a few remarks that will simplify the practical computations.

1.17.2. REMARK. Sometimes (recall Remark 1.15.5), rather than using a characteristic map $f:\overline{\mathrm{B}}^{p+1}\to X$ for the open $(p+1)$-cell e, we prefer to work with a continuous map $f:B\to X$ that take $\mathrm{inter}(B)$ homeomorphically onto e, where B is an arbitrary topological space homeomorphic to $\overline{\mathrm{B}}^{p+1}$ (i.e., B is a $(p+1)$-cell). Obviously, one can always choose a homeomorphism $h:\overline{\mathrm{B}}^{p+1}\to B$ and then work with the characteristic map $f \circ h$, but it would be nicer to work directly with f. So, how do we adapt Theorem 1.17.1? First, one has to choose an orientation τ for $\mathrm{inter}(B)$ ($\mathrm{inter}(B)$ has no canonical orientation like B^{p+1} does); then the homeomorphism $f|_{\mathrm{inter}(B)} : \mathrm{inter}(B) \to e^{p+1}$ will induce an orientation on the open $(p+1)$-cell e^{p+1} (so that $f|_{\mathrm{inter}(B)}$ becomes positively oriented). The map (1.17.2) will now be replaced by a map defined on an open subset of $\mathrm{Bd}(B)$. In Remark 1.11.10 we have mentioned that there is no problem in using degree theory for maps defined on open subsets of topological spaces that are homeomorphic to the sphere, as long as one *fixes an orientation* for such space. What orientation do we use on $\mathrm{Bd}(B)$? The answer is given in Corollary 1.10.33: we use the orientation τ^{b} that is induced from τ on the boundary of B.

1.17.3. REMARK. If $e^1 \in \mathfrak{E}$ is an open 1-cell then it is particularly simple to determine the boundary of e^1 in the cellular complex $\mathcal{D}(X)$. Namely, let $f : \overline{\mathrm{B}}^1 \to X$ be a characteristic map for e^1; we take on e^1 the orientation induced by f and we fix an open 0-cell $e^0 \in \mathfrak{E}$. Observe that e^0 has a canonical orientation (in the terminology of Example 1.10.21, this is the "+1" orientation). Using Theorem 1.17.1 and Example 1.11.8, we conclude that the coefficient appearing next to e^0 in the boundary of e^1 in $\mathcal{D}(X)$ is equal to:

- zero, if either $f^{-1}(e^0)$ is empty or if $f^{-1}(e^0)$ contains the two points of S^0;
- one, if $f^{-1}(e^0)$ contains only the "north pole" $1 \in S^0$;
- minus one, if $f^{-1}(e^0)$ contains only the "south pole" $-1 \in S^0$.

Regarding Remark 1.17.2, we will in some situations prefer to replace $\overline{\mathrm{B}}^1$ by an arbitrary oriented 1-cell (B, τ). Then the "north pole" (respectively, the "south pole") of S^0 mentioned in the itemization above should be replaced by the point of $\mathrm{Bd}(B)$ in which the orientation τ^{b} induced from τ on the boundary of B is equal to $+1$ (respectively, equal to -1). By Remark 1.10.40, if B is viewed as an oriented 1-dimensional differentiable manifold with boundary, then the point of $\mathrm{Bd}(B)$ where τ^{b} equals $+1$ (respectively, equals -1) is the point where the outward pointing vector defines the positive orientation on the tangent space of B (respectively, the point where the outward pointing vector defines the negative orientation on the tangent space of B).

In the examples below we will use freely the contents of Remarks 1.17.2 and 1.17.3, as well as the basic tools for computing degrees given in Propositions 1.11.3 and 1.11.7.

1.17.4. EXAMPLE. We compute the cellular chain complex of the sphere S^n ($n \geq 1$) endowed with the cellular decomposition explained in Example 1.15.3. We have just an open n-cell and an open 0-cell, so that $\mathcal{D}_p(S^n) \cong \mathbb{Z}$ for $p = 0, n$ and $\mathcal{D}_p(S^n) = 0$ otherwise. The boundary homomorphisms of $\mathcal{D}(S^n)$ are all trivially zero, except when $n = 1$: but in this case the boundary homomorphism $\partial_1 : \mathcal{D}_1(S^1) \to \mathcal{D}_0(S^1)$ is again equal to zero, because a characteristic map for the open 1-cell would collapse both points of the boundary of $\overline{\mathrm{B}}^1$ to the same 0-cell.

Thus, in any case, all boundary homomorphisms of $\mathcal{D}(S^n)$ are zero. We conclude (as we have known already for a long time now) that $H_p(S^n) \cong \mathbb{Z}$ for $p = 0, n$ and $H_p(S^n) = 0$ otherwise.

1.17.5. EXAMPLE. We compute the cellular chain complex of the torus $\mathbb{T} = S^1 \times S^1$ endowed with the cellular decomposition explained in Example 1.15.4. We obviously have:

$$\mathcal{D}_2(\mathbb{T}) \cong \mathbb{Z}, \quad \mathcal{D}_1(\mathbb{T}) \cong \mathbb{Z} \oplus \mathbb{Z}, \quad \mathcal{D}_0(\mathbb{T}) \cong \mathbb{Z};$$

the only non trivial boundary homomorphisms are ∂_2 and ∂_1. Since the characteristic maps for the open 1-cells e_1^1 and e_2^1 collapse both points of the boundary of $\overline{\mathrm{B}}^1$ to the same 0-cell, it follows that $\partial_1 = 0$. Let's now compute $\partial_2(e^2)$. A characteristic map for e^2 is given by the quotient map $q : R \to \mathbb{T}$ itself. We have to choose an orientation τ for $\mathrm{inter}(R) = \,]0,1[^2$; we pick the one induced from the canonical orientation $\tau^{[2]}$ of the plane $\mathbb{R}^2$. The orientation τ of $\mathrm{inter}(R)$ induces an orientation τ^{b} on $\mathrm{Bd}(R)$. Such orientation is described in Example 1.10.41; roughly speaking, this is just the "counter-clockwise" orientation. More explicitly, since the open sides of the rectangle R are (one-dimensional) differentiable manifolds, the orientation τ^{b} can be described as follows:

- the restriction of τ^{b} to the bottom side $]0,1[\, \times \{0\}$ of R is the one that makes the first vector of the canonical basis of $\mathbb{R}^2$ positive;
- the restriction of τ^{b} to the top side $]0,1[\, \times \{1\}$ of R is the one that makes the first vector of the canonical basis of $\mathbb{R}^2$ negative;
- the restriction of τ^{b} to the right side $\{1\} \times]0,1[$ of R is the one that makes the second vector of the canonical basis of $\mathbb{R}^2$ positive;
- the restriction of τ^{b} to the left side $\{0\} \times]0,1[$ is the one that makes the second vector of the canonical basis of $\mathbb{R}^2$ negative.

We now have to choose orientations for the open 1-cells e_1^1 and e_2^1. We choose the ones that makes the homeomorphisms $q|_{]0,1[\times\{0\}} : \,]0,1[\, \times \{0\} \to e_1^1$ and $q|_{\{0\}\times]0,1[} : \{0\} \times]0,1[\, \to e_2^1$ positively oriented (where the open sides of R are oriented by restrictions of τ^{b}). Let's compute the coefficient appearing next to e_1^1 in the boundary of e^2. We have to compute the degree of the map:

$$q|_{q^{-1}(e_1^1)} : q^{-1}(e_1^1) \longrightarrow e_1^1; \tag{1.17.5}$$

such degree is equal to the degree of the map obtained by composing (1.17.5) on the left with the inverse of the positively oriented homeomorphism:

$$q|_{]0,1[\times\{0\}} : \,]0,1[\, \times \{0\} \longrightarrow e_1^1.$$

The map obtained by such composition is described in the figure below:

$$q^{-1}(e_1^1) = \begin{cases}]0,1[\, \times \{0\} \ni (t,0) \longmapsto (t,0) \\ \qquad\quad \cup \\ ]0,1[\, \times \{1\} \ni (t,1) \longmapsto (t,0) \end{cases}$$

But considering the orientations induced by τ^{b} on the open sides of R, we conclude that $(t,0) \mapsto (t,0)$ is a positive diffeomorphism, while $(t,1) \mapsto (t,0)$ is a *negative* diffeomorphism. Hence the degree of (1.17.5) is equal to *zero*. A similar reasoning shows that the degree of the map:

$$q|_{q^{-1}(e_2^1)} : q^{-1}(e_2^1) \longrightarrow e_2^1,$$

is also equal to zero.

1.17.6. EXAMPLE. Let's compute the cellular chain complex of the real projective space $\mathbb{R}P^n$ endowed with the cellular decomposition described in Example 1.15.7. For $p \le n+1$, we identify $\mathbb{R}^p$ with the subspace of $\mathbb{R}^{n+1}$ spanned by the first p vectors of the canonical basis, so that we get a sequence of inclusions $S^0 \subset S^1 \subset S^2 \subset \cdots \subset S^n$ for the unit spheres. If $q : S^n \to \mathbb{R}P^n$ denotes the quotient map that identifies antipodal points then $q(S^p) \subset \mathbb{R}P^n$ is identified with $\mathbb{R}P^p$ for $p = 0, \ldots, n$, so that we also get a sequence of inclusions $\mathbb{R}P^0 \subset \mathbb{R}P^1 \subset \cdots \subset \mathbb{R}P^n$. For $p = 0, \ldots, n-1$, the difference $e^{p+1} = \mathbb{R}P^{p+1} \setminus \mathbb{R}P^p$ is exactly the unique open $(p+1)$-cell of $\mathbb{R}P^n$; a characteristic map for e^{p+1} is obtained by restricting q to the closed northern hemisphere:

$$S_{\mathfrak{n}}^{p+1} = \big\{x \in S^{p+1} \subset \mathbb{R}^{p+2} : x_{p+2} \ge 0\big\}.$$

We orient the northern hemisphere $S_{\mathfrak{n}}^{p+1}$ with the restriction of the canonical orientation $\alpha^{[p+1]}$ (i.e., the outward pointing orientation) of the sphere S^{p+1}. We can now give e^{p+1} the orientation induced from $S_{\mathfrak{n}}^{p+1}$ by the characteristic map $q|_{S_{\mathfrak{n}}^{p+1}}$. Let's compute the boundary of e^{p+1} in the chain complex $\mathcal{D}(\mathbb{R}P^n)$; such boundary is just an integer multiple of e^p. One can check straightforwardly that the orientation τ^{b} that $S_{\mathfrak{n}}^{p+1}$ induces on $\mathrm{Bd}(S_{\mathfrak{n}}^{p+1}) = S^p$ equals the canonical (outward pointing) one if and only if p is odd, i.e., $\tau^{\mathrm{b}} = (-1)^{p+1}\alpha^{[p]}$. Let's now compute the degree of the map:

$$q|_{q^{-1}(e^p)} : q^{-1}(e^p) \longrightarrow e^p, \tag{1.17.6}$$

where $q^{-1}(e^p) = S^p \setminus S^{p-1}$ is endowed with the restriction of the orientation $\tau^{\mathrm{b}} = (-1)^{p+1}\alpha^{[p]}$. The degree of (1.17.6) equals the degree of the map obtained by composing (1.17.6) on the left with the positively oriented homeomorphism $q|_{\mathrm{inter}(S_{\mathfrak{n}}^p)} : \big(\mathrm{inter}(S_{\mathfrak{n}}^p), \alpha^{[p]}\big) \to e^p$. The resulting map is pictured below:

$$q^{-1}(e_p) = \begin{cases} \overbrace{\big(\mathrm{inter}(S_{\mathfrak{n}}^p), (-1)^{p+1}\alpha^{[p]}\big)}^{\text{open northern hemisphere}} \ni x \longmapsto x \in \big(\mathrm{inter}(S_{\mathfrak{n}}^p), \alpha^{[p]}\big) \\ \qquad\qquad \cup \\ \underbrace{\big(S^p \setminus S_{\mathfrak{n}}^p, (-1)^{p+1}\alpha^{[p]}\big)}_{\text{open southern hemisphere}} \ni x \longmapsto -x \in \big(\mathrm{inter}(S_{\mathfrak{n}}^p), \alpha^{[p]}\big) \end{cases}$$

It follows now that the degree of (1.17.6) is equal to $(-1)^{p+1}\big[1 + (-1)^{p+1}\big]$, i.e., it is equal to 0 for even p and it is equal to 2 for odd p. The cellular chain complex of $\mathbb{R}P^n$ is thus given by:

$$\cdots \longrightarrow \underset{n+1}{0} \longrightarrow \underset{n}{\mathbb{Z}} \xrightarrow{\ 2\ } \underset{n-1}{\mathbb{Z}} \xrightarrow{\ 0\ } \underset{n-2}{\mathbb{Z}} \xrightarrow{\ 2\ } \cdots \xrightarrow{\ 0\ } \underset{0}{\mathbb{Z}} \longrightarrow \underset{-1}{0} \longrightarrow \cdots$$

for even n and by:

$$\cdots \longrightarrow \underset{n+1}{0} \longrightarrow \underset{n}{\mathbb{Z}} \xrightarrow{0} \underset{n-1}{\mathbb{Z}} \xrightarrow{2} \underset{n-2}{\mathbb{Z}} \xrightarrow{0} \cdots \xrightarrow{0} \underset{0}{\mathbb{Z}} \longrightarrow \underset{-1}{0} \longrightarrow \cdots$$

for odd n. Finally, the singular homology groups of $\mathbb{R}P^n$ are given by:

$$H_i(\mathbb{R}P^n) \cong \begin{cases} \mathbb{Z}_2 = \mathbb{Z}/2\mathbb{Z}, & \text{if } 1 \le i \le n-1 \text{ and } i \text{ is odd},\\ \mathbb{Z}, & \text{if } i = 0,\\ \mathbb{Z}, & \text{if } i = n \text{ and } n \text{ is odd},\\ 0, & \text{otherwise}. \end{cases}$$

Exercises for Chapter 1

Short review of abelian groups.

EXERCISE 1.1. Let R be a (associative) ring with unit[17]. An *R-module* is an abelian group M (whose operation will be denoted additively) together with a map $R \times M \ni (r, m) \mapsto rm \in M$ satisfying the conditions:

(1) $r(m_1 + m_2) = rm_1 + rm_2$, for all $r \in R$, $m_1, m_2 \in M$;
(2) $(r_1 + r_2)m = r_1 m + r_2 m$, for all $r_1, r_2 \in R$, $m \in M$;
(3) $(r_1 r_2)m = r_1(r_2 m)$, for all $r_1, r_2 \in R$, $m \in M$;
(4) $1m = m$, for all $m \in M$.

The conditions above can also be more economically expressed in the following way: for every $r \in R$, we define $\phi_r : M \to M$ by $\phi_r(m) = rm$ and we ask that $r \mapsto \phi_r$ be a (unit preserving) ring homomorphism from R to the ring $\mathrm{Hom}(M)$ of all group homomorphisms of M (multiplication in $\mathrm{Hom}(M)$ is composition of homomorphisms). Observe also that the definition of R-module is precisely the same as the definition of R-vector space in linear algebra, except for the fact that R need not be a field. Don't be fooled by the superficial similarity though: after the basic definitions, module theory is *much* harder than linear algebra.

Show that if M is an abelian group then there exists a *unique* map $\mathbb{Z} \times M \to M$ that makes M into a $\mathbb{Z}$-module.

EXERCISE 1.2. Let R be a ring with unity. Show that the multiplication of R defines an R-module structure for the subjacent additive abelian group of R. This is called the *canonical* R-module structure of the ring R.

EXERCISE 1.3. If M, N are R-modules then a map $f : M \to N$ is said to be *linear* (over R) if f is a group homomorphism and $f(rm) = rf(m)$ for all $r \in R$, $m \in M$. If G, H are abelian groups, show that every homomorphism $f : G \to H$ is automatically $\mathbb{Z}$-linear when G, H are endowed with their unique $\mathbb{Z}$-module structure.

EXERCISE 1.4. If M is a module over a ring R then a *submodule* of M is a subgroup $N \subset M$ such that $rm \in N$ for all $r \in R$, $m \in N$.

[17] One can also consider modules over rings without unit; in this case, one obviously has to drop condition (4). Even when R has a unit element, one can "forget" about such unit, i.e., drop condition (4) anyway, obtaining then a "weaker" notion of R-module.

- If G is an abelian group (endowed with its unique $\mathbb{Z}$-module structure) show that the submodules of $\mathbb{Z}$ are precisely its subgroups.
- If the ring R is endowed with its canonical R-module structure, show that the submodules of R are precisely its left ideals.
- If $f : M \to N$ is a linear map between R-modules M, N then the *kernel* of f is defined by $\mathrm{Ker}(f) = f^{-1}(0)$ (this is the same as the kernel of f as a group homomorphism). Show that the kernel of f is a submodule of M and that the image of f (denoted by $\mathrm{Im}(f)$) is a submodule of N.

EXERCISE 1.5. Let G be an abelian group. If $S \subset G$ is an arbitrary subset then the *subgroup of G spanned by S*, denoted $\mathbb{Z} \cdot S$, is the smallest subgroup of G containing S (i.e., the intersection of all subgroups of G containing S, which is obviously a subgroup of G). If $(g_i)_{i\in I}$ is a family of elements of G, show that the subgroup spanned by the set $\{g_i : i \in I\}$ equals the set of all linear combinations of the family $(g_i)_{i\in I}$. Conclude that $(g_i)_{i\in I}$ is generating for G if and only if no proper subgroup of G contains the set $\{g_i : i \in I\}$.

EXERCISE 1.6. If G is an abelian group G, show that $(g_i)_{i\in I}$ is a basis for G if and only if every element of G can be written uniquely as a linear combination of the family $(g_i)_{i\in I}$, i.e., if every $g \in G$ is a linear combination of the family $(g_i)_{i\in I}$ and $\sum_{i\in I} n_i g_i = \sum_{i\in I} n_i' g_i$ implies $n_i = n_i'$ for all $i \in I$.

EXERCISE 1.7. Generalize Exercise 1.5 to modules; namely, let M be a module over a ring R.

- Show that the intersection of a (non empty) family of submodules of M is a submodule of M. Define the *submodule spanned by a set $S \subset M$* to be the intersection of all submodules of M containing S (this is the smallest submodule of M containing S). The submodule spanned by S will be denoted by $R \cdot S$.
- If $(m_i)_{i\in I}$ is a family of elements of M then a *linear combination* of $(m_i)_{i\in I}$ is a sum of the form $\sum_{i\in I} r_i m_i \in M$, where $(r_i)_{i\in I}$ is an essentially zero family of elements of R (i.e., $r_i = 0$ except for a finite number of indices $i \in I$). Show that the submodule spanned by the set $\{m_i : i \in I\}$ equals the set of all linear combinations of the family $(m_i)_{i\in I}$ (the fact that R *has a unit element* is crucial here!).

EXERCISE 1.8. If M is a module over a ring R the a family $(m_i)_{i\in I}$ of elements of M is called *linearly independent* if for every essentially zero family $(r_i)_{i\in I}$ in R, $\sum_{i\in I} r_i m_i = 0$ implies $r_i = 0$ for all $i \in I$. A family that is not linearly independent is called *linearly dependent*. If the set $\{m_i : i \in I\}$ spans M as a module we sat that the family $(m_i)_{i\in I}$ is *generating* for M. A family $(m_i)_{i\in I}$ that is both linearly independent and generating is called a *basis* for M (we also say that M is *free* over $(m_i)_{i\in I}$).

- Generalize Exercise 1.6 to modules.
- Generalize Proposition 1.1.3 to modules.
- If A is an arbitrary set, define the *free R-module spanned by A* to be the set $\mathrm{Free}_R[A]$ of all essentially zero maps $f : A \to R$ endowed with

the R-module structure given by pointwise addition and multiplication by elements of R. Identify each $a \in A$ with the map that takes a to the unit of R and $A \setminus \{a\}$ to zero. Show that the R-module $\mathrm{Free}_R[A]$ is free over A.

EXERCISE 1.9. Show that if G is a finite abelian group then every non empty family in G is linearly dependent; conclude that a non zero finite abelian group cannot be free.

EXERCISE 1.10. Let G be an abelian group. Let $\mathrm{T}(G)$ be the set of elements of G having finite order, i.e.:

$$\mathrm{T}(G) = \big\{ g \in G : ng = 0 \text{ for some positive integer } n \big\}.$$

Show that $\mathrm{T}(G)$ is a subgroup of G; we call it the *torsion subgroup* of G. If $\mathrm{T}(G) = 0$ we say that G is *torsion free*. Show that the quotient $G/\mathrm{T}(G)$ is torsion free for every abelian group G.

EXERCISE 1.11. Let G be an abelian group and $(g_i)_{i \in I}$ a family of elements of G. Assume that the property given in the statement of Proposition 1.1.3 holds, i.e., that for *every* abelian group H and every family $(h_i)_{i \in I}$ of elements of G there exists a unique homomorphism $f : G \to H$ with $f(g_i) = h_i$ for all $i \in I$. Show that $(g_i)_{i \in I}$ is a basis for G (*hint*: let H be the free abelian group spanned by $\{g_i\}_{i \in I}$).

EXERCISE 1.12. Let A be an arbitrary set, G an abelian group and $i : A \to G$ an injective map such that $i(A)$ is a basis for G. Show that there exists a unique homomorphism $\phi : \mathrm{Free}[A] \to G$ such that the diagram:

$$\begin{tikzcd} \mathrm{Free}[A] \arrow[rr, "\phi"] & & G \\ & A \arrow[ul] \arrow[ur, "i"'] & \end{tikzcd}$$

commutes, where the unlabelled arrow denotes the canonical inclusion of A in $\mathrm{Free}[A]$. Show that ϕ is an isomorphism.

EXERCISE 1.13. Let G be a free abelian group with basis $(g_i)_{i \in I}$. Given subsets $I_1, I_2 \subset I$ show that the intersection of the subgroups spanned by $\{g_i : i \in I_1\}$ and $\{g_i : i \in I_2\}$ is the subgroup spanned by $\{g_i : i \in I_1 \cap I_2\}$. Generalize this result for an arbitrary (non empty) family of subsets of I.

EXERCISE 1.14. Let G be an abelian group and let $H, K \subset G$ be subgroups with $G = H + K$. Show that the inclusion of K in G induces an isomorphism

$$\frac{K}{H \cap K} \xrightarrow{\cong} \frac{G}{H}.$$

EXERCISE 1.15. Let $f : G \to H$ be an epimorphism (i.e., a surjective homomorphism) between abelian groups. Show that if $S_1 \subset S_2 \subset H$ are subgroups of H then f induces an isomorphism between the quotients $f^{-1}(S_2)/f^{-1}(S_1)$ and

S_2/S_1. Show that this isomorphism is *natural* in the following sense: given epimorphisms $f : G \to H$, $f' : G' \to H'$ and homomorphisms $\phi : G \to G'$, $\psi : H \to H'$ such that the diagram

$$\begin{array}{ccc} G & \xrightarrow{f} & H \\ {\scriptstyle\phi}\downarrow & & \downarrow{\scriptstyle\psi} \\ G' & \xrightarrow[f']{} & H' \end{array}$$

commutes then the diagram

$$\begin{array}{ccc} f^{-1}(S_2)/f^{-1}(S_1) & \xrightarrow{\cong} & S_2/S_1 \\ \downarrow & & \downarrow \\ f'^{-1}(S_2')/f'^{-1}(S_1') & \xrightarrow[\cong]{} & S_2'/S_1' \end{array}$$

also commutes, where the horizontal arrows are induced by f and f', the vertical arrows are induced by ϕ and ψ, S_1, S_2 are subgroups of H and $S_1' = \psi(S_1)$, $S_2' = \psi(S_2)$.

EXERCISE 1.16. Let G, H be abelian groups and $f : G \to H$ a homomorphism. Show that:

(a) f has a left inverse that is a homomorphism if and only if f is injective and $\mathrm{Im}(f)$ is a direct summand of H;
(b) f has a right inverse that is a homomorphism if and only if f is surjective and $\mathrm{Ker}(f)$ is a direct summand of H.

EXERCISE 1.17. Given abelian groups G, H with H free, show that every surjective homomorphism $f : G \to H$ has a right inverse.

EXERCISE 1.18. Given homomorphisms $h_i : G_i \to G_i'$ and $k_i : G_i' \to G_i''$, $i = 1, 2$, show that $(k_1 \otimes k_2) \circ (h_1 \otimes h_2) = (k_1 \circ h_1) \otimes (k_2 \circ h_2)$.

EXERCISE 1.19. Denote by $e_1 = (1, 0)$, $e_2 = (0, 1)$ the canonical basis of the free abelian group $G = \mathbb{Z} \otimes \mathbb{Z}$. Show that $e_1 \otimes e_2 + e_2 \otimes e_1 \in G \otimes G$ is not of the form $g_1 \otimes g_2$ with $g_1, g_2 \in G$.

EXERCISE 1.20. Show that if $(g_i)_{i \in I}$ is a basis for an abelian group G and if $(h_j)_{j \in J}$ is a basis for an abelian group H then $(g_i \otimes h_j)_{(i,j) \in I \times J}$ is a basis for the tensor product $G \otimes H$.

EXERCISE 1.21. Let R be a commutative ring with unity and let M, N be R-modules. A map $B : M \times N \to P$ taking values in an R-module P is called *bilinear* (over R) if for every $m \in M$, $n \in N$, the maps $B(m, \cdot) : N \to P$ and $B(\cdot, n) : M \to P$ are linear over R. A *tensor product* of M and N (over R) is a pair (T, b) where T is an R-module, $b : M \times N \to T$ is bilinear over R and the following property holds: given an arbitrary R-module P and an arbitrary bilinear map $B : M \times N \to P$ over R there exists a unique linear map $\overline{B} : T \to P$ over R such that the diagram:

$$\begin{tikzcd} M \times N \arrow[r, "B"] \arrow[d, "b"'] & P \\ T \arrow[ur, dotted, "\overline{B}"'] & \end{tikzcd}$$

commutes. Generalize Lemmas 1.1.11 and 1.1.12 to the context of R-modules by proving the uniqueness (up to isomorphisms) and the existence of tensor products over R. The (essentially unique) tensor product between M and N over R is denoted by $M \otimes_R N$.

EXERCISE 1.22. Let R be a (non necessarily commutative) ring with unity. A *right R-module* is an abelian group M together with a map $R \times M \to M$ satisfying the properties (1)–(4) given in the statement of Exercise 1.1, except for the fact that property (3) is replaced by:

$$(r_2 r_1)m = r_1(r_2 m),$$

for all $r_1, r_2 \in R$, $m \in M$; actually, one usually writes mr rather than rm, so that the property above becomes a natural "associativity law" $m(r_2 r_1) = (mr_2)r_1$. If M is an R-module (according to the definition given in the statement of Exercise 1.1) then one sometimes call M a *left R-module*, if it is necessary to make the distinction from right R-modules clearer.

If M is a right R-module and N is a left R-module then a map $B : M \times N \to G$ taking values in an abelian group G is called *balanced* if B is bilinear as a map of abelian groups and if:

$$B(mr, n) = B(m, rn),$$

for all $m \in M$, $n \in N$, $r \in R$. A *tensor product* of M and N (over R) is a pair (T, b) where T is an abelian group, $b : M \times N \to T$ is balanced and the following property holds: given an arbitrary abelian group H and an arbitrary balanced map $B : M \times N \to H$ there exists a unique group homomorphism $\overline{B} : T \to H$ such that the diagram:

$$\begin{tikzcd} M \times N \arrow[r, "B"] \arrow[d, "b"'] & H \\ T \arrow[ur, dotted, "\overline{B}"'] & \end{tikzcd}$$

commutes. Generalize the result of Exercise 1.21 by proving the uniqueness (up to isomorphisms) and the existence of tensor products in this context.

Singular homology.

EXERCISE 1.23. Identify explicitly the singular chain groups $\mathfrak{S}_p(X)$ and the boundary homomorphisms ∂_p when X is empty and when X consists of just one point. Obtain the singular homology groups $H_p(X)$. What happens more generally when X is an arbitrary discrete space?

EXERCISE 1.24. Show that (1.2.4) defines a continuous map in Δ_{p+1} (*hint:* $[0,1]\times\Delta_p \ni (t,u) \mapsto (1-t)u + te_{p+1} \in \Delta_{p+1}$ is a quotient map).

EXERCISE 1.25. Given a family $(\mathfrak{C}^i, \partial^i)_{i\in I}$ of chain complexes then their *external direct sum* is the chain complex $(\mathfrak{C}, \partial)$ defined by:

$$\mathfrak{C}_p = \bigoplus_{i\in I} \mathfrak{C}^i_p, \quad \partial_p = \bigoplus_{i\in I} \partial^i_p,$$

for all $p \in \mathbb{Z}$. Show that for every $p \in \mathbb{Z}$ we obtain an isomorphism:

$$\bigoplus_{i\in I} H_p(\mathfrak{C}^i) \longrightarrow H_p(\mathfrak{C})$$

whose restriction to $H_p(\mathfrak{C}^i)$ is induced by the inclusion of $\mathfrak{C}^i$ in $\mathfrak{C}$.

EXERCISE 1.26. Let X be a topological space and assume that we have a disjoint union $X = \bigcup_{\lambda\in\Lambda} X_\lambda$ where each X_λ is a union of arc-connected components of X (this happens for instance when all X_λ's are open). For every $\lambda \in \Lambda$, denote by $i_\lambda : X_\lambda \to X$ the inclusion; for some fixed $p \in \mathbb{Z}$, consider the homomorphism

$$\phi : \bigoplus_{\lambda\in\Lambda} H_p(X_\lambda) \longrightarrow H_p(X)$$

whose restriction to $H_p(X_\lambda)$ equals $(i_\lambda)_*$. Show that ϕ is an isomorphism (*hint*: use Exercise 1.25).

EXERCISE 1.27. Given a topological space X, show that $H_0(X)$ is free and that a basis for $H_0(X)$ is obtained by choosing a point of each arc-connected component of X and taking their homology classes (*hint*: use Exercise 1.26).

EXERCISE 1.28. Let X be a non empty topological space.

- show that the augmentation map $\epsilon : \mathfrak{S}_0(X) \to \mathbb{Z}$ is surjective;
- show that there exists a homomorphism $\bar\epsilon : H_0(X) \to \mathbb{Z}$ such that:

  $$\bar\epsilon\big(c + B_0(X)\big) = \epsilon(c),$$

  for all $c \in \mathfrak{S}_0(X)$;
- show that we have a short exact sequence

  $$0 \longrightarrow \widetilde{H}_0(X) \xrightarrow{\text{inclusion}} H_0(X) \xrightarrow{\bar\epsilon} \mathbb{Z} \longrightarrow 0$$

  and that such sequence splits;
- conclude from the item above that $H_0(X) \cong \widetilde{H}_0(X) \oplus \mathbb{Z}$;
- denote by $(X_i)_{i\in I}$ the family of all arc-connected components of X and choose $x_i \in X_i$ for each $i \in I$. Show that for any fixed $i_0 \in I$, the set $\big\{x_i - x_{i_0} : i \in I,\ i \ne i_0\big\}$ is mapped by the quotient map:

  $$Z_0\big(\widetilde{\mathfrak{S}}(X)\big) = \mathrm{Ker}(\epsilon) \longrightarrow \widetilde{H}_0(X)$$

  onto a basis of $\widetilde{H}_0(X)$ (*hint*: use Exercise 1.27);
- conclude from the item above that $\widetilde{H}_0(X)$ is free and that its rank equals the number of arc-connected components of X minus 1. Conclude also that X is arc-connected if and only if $\widetilde{H}_0(X) = 0$.

EXERCISE 1.29. A chain complex $\mathfrak{C}$ is called *nonnegative* if $\mathfrak{C}_p = 0$ for $p < 0$. Given a nonnegative chain complex $\mathfrak{C}$ then an *augmentation* for $\mathfrak{C}$ is a homomorphism $\epsilon : \mathfrak{C}_0 \to \mathbb{Z}$ such that $\epsilon \circ \partial_1 = 0$. The complex $(\mathfrak{C}, \epsilon)$ obtained from $\mathfrak{C}$ by replacing ∂_0 with ϵ and $\mathfrak{C}_{-1} = \{0\}$ with $\mathbb{Z}$ is called the *augmented chain complex* corresponding to $\mathfrak{C}$ and ϵ. Assuming that the augmentation map ϵ is surjective, generalize the result of Exercise 1.28 by obtaining a splitting short exact sequence:

$$0 \longrightarrow H_0(\mathfrak{C}, \epsilon) \xrightarrow{\text{inclusion}} H_0(\mathfrak{C}) \xrightarrow{\bar{\epsilon}} \mathbb{Z} \longrightarrow 0$$

where $\bar{\epsilon}$ is induced by ϵ.

EXERCISE 1.30. Given non negative chain complexes $\mathfrak{C}$, $\mathfrak{C}'$ with augmentations ϵ and ϵ' respectively (see Exercise 1.29) then a chain map $f : \mathfrak{C} \to \mathfrak{C}'$ is called *augmentation preserving* if $\epsilon' \circ f_0 = \epsilon$. Show that a chain map $f : \mathfrak{C} \to \mathfrak{C}'$ is augmentation preserving iff we obtain a chain map $\tilde{f} : (\mathfrak{C}, \epsilon) \to (\mathfrak{C}', \epsilon')$ by setting $\tilde{f}_{-1} = \mathrm{Id} : \mathbb{Z} \to \mathbb{Z}$ and $\tilde{f}_p = f_p$ for $p \neq -1$. Assuming that both augmentations ϵ and ϵ' are surjective, show that the map:

$$f_* : H_0(\mathfrak{C}) \longrightarrow H_0(\mathfrak{C}')$$

induced by f is an isomorphism iff the map:

$$\tilde{f}_* : H_0(\mathfrak{C}, \epsilon) \longrightarrow H_0(\mathfrak{C}', \epsilon')$$

induced by $\tilde{f}$ is an isomorphism.

EXERCISE 1.31. Let $f : X \to Y$ be a continuous map between topological spaces. Show that the homomorphisms $(f_\#)_p : \mathfrak{S}_p(X) \to \mathfrak{S}_p(Y)$ define a chain map $f_\#$ from $\mathfrak{S}_p(X)$ to $\mathfrak{S}_p(Y)$. Show that $f_\#$ is augmentation preserving (see Exercise 1.30) when one considers the standard augmentations for the singular chain complexes (see (1.2.2)).

EXERCISE 1.32. Show that *homology classes are compactly supported*, i.e., given a topological space X and a homology class $\alpha \in H_p(X)$ then there exists a compact subspace $K \subset X$ such that α belongs to the image of the homomorphism $H_p(K) \to H_p(X)$ induced by inclusion.

EXERCISE 1.33. Show that *homology relations are compactly supported*, i.e., given a topological space X, a subspace $Y \subset X$ and a homology class $\alpha \in H_p(Y)$ such that the homomorphism induced by inclusion $H_p(Y) \to H_p(X)$ maps α to zero, then there exists a compact subspace $K \subset X$ such that the homomorphism induced by inclusion $H_p(Y) \to H_p(Y \cup K)$ also maps α to zero.

EXERCISE 1.34. A *filtration* for a topological space X is a sequence $(X_n)_{n \geq 0}$ of subspaces of X with $X_n \subset X_{n+1}$ for all n; a topological space together with a given filtration is called a *filtered space*. Given filtered spaces X, Y with filtrations $(X_n)_{n\geq 0}$, $(Y_n)_{n\geq 0}$ for X and Y respectively then a continuous map $f : X \to Y$ is called *filtration preserving* if $f(X_n) \subset Y_n$ for all n. Suppose that we are given filtered spaces X, Y and a filtration preserving map $f : X \to Y$. Assume that every compact subset of X is contained in some X_n and that every compact subset of Y is contained in some Y_n. Show that if (for some fixed p) $f|_{X_n} : X_n \to Y_n$

induces homology isomorphisms $H_p(X_n) \to H_p(Y_n)$ for all n then f induces a homology isomorphism $f_* : H_p(X) \to H_p(Y)$.

EXERCISE 1.35. Let $f : \mathfrak{C} \to \mathfrak{D}$ be a chain map. Show that if $f_p : \mathfrak{C}_p \to \mathfrak{D}_p$ is an isomorphism for every $p \in \mathbb{Z}$ then $(f^{-1})_p = (f_p)^{-1}$ defines a chain map $f^{-1} : \mathfrak{D} \to \mathfrak{C}$. A chain map which is an isomorphism in every dimension is called a *chain isomorphism*.

EXERCISE 1.36. Let $\mathfrak{C}$ be a chain complex. A chain complex $\mathfrak{C}'$ is called a *subcomplex* of $\mathfrak{C}$ if for each $p \in \mathbb{Z}$, $\mathfrak{C}'_p$ is a subgroup of $\mathfrak{C}_p$ and the boundary homomorphism of $\mathfrak{C}'$ is the restriction of the boundary homomorphism of $\mathfrak{C}$.

(a) Show that if $\mathfrak{C}'$ is a subcomplex of $\mathfrak{C}$ then we obtain another chain complex whose p-th dimensional group is the quotient $\mathfrak{C}_p/\mathfrak{C}'_p$ and the boundary homomorphism is induced by the boundary homomorphism of $\mathfrak{C}$.
(b) Show that if $f : \mathfrak{C} \to \mathfrak{D}$ is a chain map then we obtain a subcomplex of $\mathfrak{C}$ whose p-th dimensional group is $\mathrm{Ker}(f_p)$ and a subcomplex of $\mathfrak{D}$ whose p-th dimensional group is $\mathrm{Im}(f_p)$. These complexes are called respectively the *kernel* and the *image* of the chain map f.
(c) Given a chain map $f : \mathfrak{C} \to \mathfrak{D}$, show that f induces a chain isomorphism between $\mathfrak{C}/\mathrm{Ker}(f)$ and $\mathrm{Im}(f)$ (see Exercise 1.35).

EXERCISE 1.37. Let $\mathfrak{C}$ be a chain complex and let $\mathfrak{D}, \mathfrak{D}'$ be subcomplexes of $\mathfrak{C}$. For every $p \in \mathbb{Z}$ set $(\mathfrak{D} \cap \mathfrak{D}')_p = \mathfrak{D}_p \cap \mathfrak{D}'_p$ and $(\mathfrak{D} + \mathfrak{D}')_p = \mathfrak{D}_p + \mathfrak{D}'_p$. Show that $\mathfrak{D} \cap \mathfrak{D}'$ and $\mathfrak{D} + \mathfrak{D}'$ are subcomplexes of $\mathfrak{C}$. More generally, for an arbitrary[18] family $(\mathfrak{D}^i)_{i\in I}$ of subcomplexes of $\mathfrak{C}$ we set $\left(\bigcap_{i\in I} \mathfrak{D}^i\right)_p = \bigcap_{i\in I} \mathfrak{D}^i_p$ and $\left(\sum_{i\in I} \mathfrak{D}^i\right)_p = \sum_{i\in I} \mathfrak{D}^i_p$ for every $p \in \mathbb{Z}$. Show that $\bigcap_{i\in I} \mathfrak{D}^i$ and $\sum_{i\in I} \mathfrak{D}^i$ are subcomplexes of $\mathfrak{C}$.

EXERCISE 1.38. Let $f : \mathfrak{C} \to \mathfrak{D}$ be a chain map and let $\mathfrak{C}'$, $\mathfrak{D}'$ be subcomplexes of $\mathfrak{C}$ and $\mathfrak{D}$ respectively with $f(\mathfrak{C}') \subset \mathfrak{D}'$. Consider the chain maps

$$f|_{\mathfrak{C}'} : \mathfrak{C}' \longrightarrow \mathfrak{D}', \quad f : \mathfrak{C} \longrightarrow \mathfrak{D}, \quad \bar{f} : \mathfrak{C}/\mathfrak{C}' \longrightarrow \mathfrak{D}/\mathfrak{D}'$$

induced by f. Show that if any two of the chain maps above induce isomorphisms in homology (in all dimensions) then also the third one does.

EXERCISE 1.39. Let $\mathfrak{C}$ be a free chain complex. Show that the cycle group $Z_p(\mathfrak{C})$ is a direct summand in $\mathfrak{C}_p$ for every $p \in \mathbb{Z}$ (*hint*: observe that the image of ∂_p is free and use the result of Exercises 1.17 and 1.16).

EXERCISE 1.40. Let X be a topological space and $\mathfrak{A}$ a non empty collection of subspaces of X. Show that

$$\bigcap_{A\in\mathfrak{A}} \mathfrak{S}(A) = \mathfrak{S}\Big(\bigcap\nolimits_{A\in\mathfrak{A}} A\Big).$$

[18] For the finicky reader we observe that if the index set I is empty then the intersection $\bigcap_{i\in I} \mathfrak{D}^i$ is not defined whereas the sum $\sum_{i\in I} \mathfrak{D}^i$ is defined and equals the zero subcomplex of $\mathfrak{C}$.

EXERCISE 1.41. Let $f, g : \mathfrak{C} \to \mathfrak{D}$, $h : \mathfrak{D} \to \mathfrak{D}'$, $l : \mathfrak{C}' \to \mathfrak{C}$ be chain maps and assume that f and g are chain homotopic. Show that $f \circ l$ is chain homotopic to $g \circ l$ and that $h \circ f$ is chain homotopic to $h \circ g$.

EXERCISE 1.42. Let $C \subset \mathbb{R}^n$ be a convex subset.

(1) Show that $\overline{C}$ is convex.
(2) If $x \in \mathrm{int}(C)$, $y \in C$, show that the line segment $[x, y[$ is contained in the interior of C.
(3) Show that $\mathrm{int}(C)$ is convex.
(4) If $x \in \mathrm{int}(C)$, $y \in \overline{C}$, show that $[x, y[\subset \mathrm{int}(C)$ (*hint*: show first that $[x, y[\subset C$).
(5) Let $x_0 \in \mathrm{int}(C)$ and $v \in \mathbb{R}^n \setminus \{0\}$ be fixed; consider the ray $r(t) = x_0 + tv$, $t > 0$. Show that $r^{-1}(C)$ is an interval whose left endpoint is zero; denote the right endpoint of $r^{-1}(C)$ by $b \in]0, +\infty]$. Show that if $b = +\infty$ then $r(t) \in \mathrm{int}(C)$ for all $t > 0$. Moreover, show that if $b < +\infty$, then $r(t) \in \mathrm{int}(C)$ for $t \in]0, b[$, $r(b) \in \partial C$ and $r(t) \notin \overline{C}$ for $t > b$.
(6) If C is bounded and $x_0 \in \mathrm{int}(C)$, show that the map
$$\partial C \ni x \longmapsto \frac{x - x_0}{\|x - x_0\|} \in S^{n-1}$$
is a homeomorphism. Show that such homeomorphism extends to a homeomorphism from $\overline{C}$ to the closed unit ball $\overline{\mathrm{B}}^n$.
(7) If $\mathrm{int}(C) = \emptyset$ but $C \neq \emptyset$, show that there exists an affine subspace $\mathbb{A} \subset \mathbb{R}^n$ with $C \subset \mathbb{A}$ and such that the interior of C relatively to $\mathbb{A}$ is non empty.
(8) Show that $\mathrm{int}\big(\overline{C}\big) = \mathrm{int}(C)$, $\partial\overline{C} = \partial C$. Moreover, assuming $\mathrm{int}(C) \neq \emptyset$, show that $\overline{\mathrm{int}(C)} = \overline{C}$ and $\partial\big(\mathrm{int}(C)\big) = \partial C$.

EXERCISE 1.43. Let V be a vector space over a field $\mathbb{K}$. A family $(v_i)_{i \in I}$ of vectors in V is called *geometrically independent* if for every essentially zero family $(c_i)_{i \in I}$ of scalars in $\mathbb{K}$, the equalities:
$$\sum_{i \in I} c_i = 0 \quad \text{and} \quad \sum_{i \in I} c_i v_i = 0,$$
imply $c_i = 0$ for all $i \in I$.

- Show that a non empty family $(v_i)_{i \in I}$ in V is geometrically independent if and only if there exists $i_0 \in I$ such that the family $(v_i - v_{i_0})_{i \in I \setminus \{i_0\}}$ is linearly independent.
- Show that a family $(v_i)_{i \in I}$ in V is geometrically independent if and only if for every $i_0 \in I$ the family $(v_i - v_{i_0})_{i \in I \setminus \{i_0\}}$ is linearly independent.
- Show that every family having less than two elements is geometrically independent. Show that a pair of points is geometrically independent if and only if the two points are distinct.
- An *affine subspace* $P \subset V$ is a translation $S + v$ ($v \in V$) of a vector subspace $S \subset V$; the *dimension* $\dim(P)$ of P is by definition equal to

the dimension of S. Show that a finite family $(v_0, \ldots, v_p)$ in V having $p+1$ elements is geometrically independent if and only if there is no affine subspace $P \subset V$ containing $\{v_0, \ldots, v_p\}$ with $\dim(P) < p$.

EXERCISE 1.44. If V is a real vector space then a subset $\sigma \subset V$ is called an *affine simplex* if it is the convex hull of a non empty finite geometrically independent family $(v_0, \ldots, v_p)$. We say also that σ is a *p-dimensional simplex* or, more simply, a p-simplex; the vectors v_i, $i = 0, \ldots, p$, are called the *vertices* of σ. Show that the dimension and the set of vertices of a simplex are well-defined, i.e., that if $(v_0, \ldots, v_p)$ and $(w_0, \ldots, w_q)$ are geometrically independent families with the same convex hull σ then $p = q$ and $\{v_0, \ldots, v_p\} = \{w_0, \ldots, w_q\}$ (*hint*: show that $x \in \sigma$ is a vertex if and only if there is no open line segment that contains x and is contained in σ).

EXERCISE 1.45. Let V be a real vector space. If σ is a p-simplex in V having vertices $\{v_0, \ldots, v_p\}$ then a *face* s of σ is an affine simplex whose set of vertices is contained in $\{v_0, \ldots, v_p\}$ (i.e., s is the convex hull of a non empty subset of $\{v_0, \ldots, v_p\}$). If s is a face of σ but $s \neq \sigma$ then we call s a *proper face* of σ. The *boundary* of σ, denoted $\mathrm{Bd}(\sigma)$, is defined as the union of all the proper faces of σ; the *interior* of σ, denoted $\mathrm{inter}(\sigma)$, is defined by $\mathrm{inter}(\sigma) = \sigma \setminus \mathrm{Bd}(\sigma)$ (observe that if σ is a zero-dimensional simplex, i.e., a single point, then $\mathrm{inter}(\sigma) = \sigma$ and $\mathrm{Bd}(\sigma) = \emptyset$). Show that every affine simplex σ equals the disjoint union of the interiors of its faces.

EXERCISE 1.46. A *simplicial complex* on V is a set K of affine simplexes on V for which the following two properties hold:

(a) if $\sigma \in K$ then all the faces of σ are in K;
(b) if $\sigma, \tau \in K$ then either $\sigma \cap \tau = \emptyset$ or $\sigma \cap \tau$ is a face of both σ and τ.

Show that a set K of affine simplexes on V is a simplicial complex if and only if (a) holds and for every pair of distinct simplexes $\sigma, \tau \in K$ the interiors $\mathrm{inter}(\sigma)$ and $\mathrm{inter}(\tau)$ are disjoint. Show also that if K is a set of affine simplexes on V for which (b) holds then the set:

$$K' = \big\{ s : s \text{ is a face of some } \sigma \in K \big\}$$

is a simplicial complex (we call it the *simplicial complex spanned by* K).

Relative homology.

EXERCISE 1.47. Generalize the result of Exercise 1.32 to the case of relative homology as follows. Given a pair of topological spaces (X, A) and a homology class $\alpha \in H_p(X, A)$ show that there exists a pair (K_1, K_2) of *compact* topological spaces with $K_1 \subset X$, $K_2 \subset A$ and such that α belongs to the image of the homomorphism $H_p(K_1, K_2) \to H_p(X, A)$ induced by inclusion.

EXERCISE 1.48. Generalize the result of Exercise 1.33 to the case of relative homology as follows. Let (X, A), (Y, B) be pairs of topological spaces with $Y \subset X$ and $B \subset A$. Let $\alpha \in H_p(Y, B)$ be a homology class and assume that α is mapped to zero by the homomorphism $H_p(Y, B) \to H_p(X, A)$

induced by inclusion. Show that there exist compact subsets $K_1 \subset X$, $K_2 \subset A$ with $K_2 \subset K_1$ and such that α is also mapped to zero by the homomorphism $H_p(Y, B) \to H_p(Y \cup K_1, B \cup K_2)$ induced by inclusion.

EXERCISE 1.49. Generalize the result of Exercise 1.26 to the case of relative homology, as follows. Suppose that we have a disjoint union $X = \bigcup_{\lambda \in \Lambda} X_\lambda$, where each X_λ is a union of arc-connected components of X (this happens for instance when all X_λ's are open). For each $\lambda \in \Lambda$, let A_λ be an arbitrary subset of X_λ and set $A = \bigcup_{\lambda \in \Lambda} A_\lambda$. Show that (for fixed $p \in \mathbb{Z}$) the homomorphism

$$\bigoplus_{\lambda \in \Lambda} H_p(X_\lambda, A_\lambda) \longrightarrow H_p(X, A) \tag{1.17.7}$$

induced by inclusion is an isomorphism.

EXERCISE 1.50. Let X, $(X_\lambda)_{\lambda \in \Lambda}$ and $(A_\lambda)_{\lambda \in \Lambda}$ be as in Exercise 1.49. For fixed $\lambda_0 \in \Lambda$, show that the inclusion

$$i : (X_{\lambda_0}, A_{\lambda_0}) \longrightarrow \Big(X, A_{\lambda_0} \cup \bigcup_{\lambda \neq \lambda_0} X_\lambda\Big)$$

induces an isomorphism in homology. Moreover, for fixed $p \in \mathbb{Z}$, prove that the composite map:

$$H_p(X, A) \xrightarrow{\text{induced by inclusion}} H_p\Big(X, A_{\lambda_0} \cup \bigcup_{\lambda \neq \lambda_0} X_\lambda\Big) \xrightarrow{(i_*)^{-1}} H_p(X_{\lambda_0}, A_{\lambda_0})$$

equals projection onto the direct summand $H_p(X_{\lambda_0}, A_{\lambda_0})$, when one identifies $H_p(X, A)$ with $\bigoplus_{\lambda \in \Lambda} H_p(X_\lambda, A_\lambda)$ via (1.17.7).

hint: consider the commutative diagram:

$$\begin{tikzcd}
H_p(X_{\lambda_0}, A_{\lambda_0}) \oplus \bigoplus_{\lambda \neq \lambda_0} H_p(X_\lambda, X_\lambda) \arrow[rr, "\text{by Exercise 1.49}", "\cong"'] & & H_p\Big(X, A_{\lambda_0} \cup \bigcup_{\lambda \neq \lambda_0} X_\lambda\Big) \\
& H_p(X_{\lambda_0}, A_{\lambda_0}) \arrow[ul, "\cong"] \arrow[ur, "i_*"'] &
\end{tikzcd}$$

in which all arrows are induced by inclusion.

EXERCISE 1.51 (the serpent lemma). Consider the commutative diagram of abelian groups and homomorphisms given below:

$$\begin{tikzcd}
0 \arrow[r] & A \arrow[r, "\sigma"] \arrow[d, "f_1"'] & B \arrow[r, "\tau"] \arrow[d, dashed, "f_2"] & C \arrow[r] \arrow[d, "f_3"] & 0 \\
0 \arrow[r] & A' \arrow[r, "\sigma'"'] & B' \arrow[r, "\tau'"'] & C' \arrow[r] & 0
\end{tikzcd} \tag{1.17.8}$$

Assume that the rows of (1.17.8) are exact. Show that there exists a exact sequence (called the *serpent sequence*):

$$0 \longrightarrow \mathrm{Ker}(f_1) \xrightarrow{\sigma|_{\mathrm{Ker}(f_1)}} \mathrm{Ker}(f_2) \xrightarrow{\tau|_{\mathrm{Ker}(f_2)}} \mathrm{Ker}(f_3) \longrightarrow$$
$$\longrightarrow \mathrm{Coker}(f_1) \xrightarrow{\bar{\sigma}} \mathrm{Coker}(f_2) \xrightarrow{\bar{\tau}} \mathrm{Coker}(f_3) \longrightarrow 0$$

where $\bar{\sigma}$ and $\bar{\tau}$ are induced by σ and τ respectively[19]. State and prove the naturality of such sequence (*hint*: think of the columns of (1.17.8) as being chain complexes having at most two non zero groups. Then (1.17.8) may be thought as a short exact sequence of chain complexes and chain maps. Apply the zig-zag lemma (Lemma 1.3.4) to such sequence). The homomorphism $\mathrm{Ker}(f_3) \to \mathrm{Coker}(f_1)$ appearing in the serpent sequence is called the *connecting homomorphism*.

EXERCISE 1.52. We call (X, A, B) a *triple* of topological spaces if X is a topological space and $B \subset A \subset X$; we call $f : (X, A, B) \to (X', A', B')$ a *map of triples* if $f : X \to X'$ is a continuous map and $f(A) \subset A'$, $f(B) \subset B'$. The goal of this exercise is to produce a long exact sequence in homology associated to every triple (X, A, B) of topological spaces.

- Consider (for fixed $p \in \mathbb{Z}$) the commutative diagram:

$$\begin{tikzcd}
0 \arrow[r] & \mathfrak{S}_p(B) \arrow[r, "\mathrm{Id}"] \arrow[d] & \mathfrak{S}_p(B) \arrow[r] \arrow[d] & 0 \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \mathfrak{S}_p(A) \arrow[r] & \mathfrak{S}_p(X) \arrow[r] & \mathfrak{S}_p(X, A) \arrow[r] & 0
\end{tikzcd}$$

  where the unlabelled arrows are induced by inclusion. Apply the serpent Lemma (Exercise 1.51) to it and conclude that there exists a short exact sequence of chain complexes:

$$0 \longrightarrow \mathfrak{S}(A, B) \longrightarrow \mathfrak{S}(X, B) \longrightarrow \mathfrak{S}(X, A) \longrightarrow 0$$

  with arrows induced by inclusion.
- Apply the zig-zag lemma (Lemma 1.3.4) to the sequence obtained above to produce a long exact sequence:

$$\cdots \xrightarrow{\partial_*} H_p(A, B) \longrightarrow H_p(X, B) \longrightarrow H_p(X, A) \xrightarrow{\partial_*} H_{p-1}(A, B) \longrightarrow \cdots$$

  where unlabelled arrows are induced by inclusion. The sequence above is known as the *long exact homology sequence* of the triple (X, A, B); the homomorphism ∂_* is called the *connecting homomorphism*. State and prove the naturality of the long exact homology sequence of a triple with respect to maps of triples $f : (X, A, B) \to (X', A', B')$.
- Show that the long exact homology sequence of a triple $(X, A, \emptyset)$ is equal to the long exact homology sequence of the pair (X, A).

[19]The *co-kernel* of a homomorphism $f : G \to H$, denoted $\mathrm{Coker}(f)$, is defined to be the quotient $H/\mathrm{Im}(f)$.

- Considering the inclusion $(X, A, \emptyset) \to (X, A, B)$ and the naturality of the long exact homology sequence of a triple, show that the connecting homomorphism:
$$H_p(X, A) \xrightarrow{\partial_*} H_{p-1}(A, B)$$
appearing in the long exact homology sequence of the triple (X, A, B) equals the composition:
$$H_p(X, A) \xrightarrow{\partial_*} H_{p-1}(A) \longrightarrow H_{p-1}(A, B)$$
of the connecting homomorphism ∂_* appearing in the long exact homology sequence of the pair (X, A) with the homomorphism induced by the inclusion of A in (A, B).

Rudiments of homotopy theory.

EXERCISE 1.53. Let X, Y be topological spaces such that Y is homeomorphic to a convex subset of some normed vector space. Show that any two continuous maps $f, g : X \to Y$ are homotopic relatively to the set $\big\{x \in X : f(x) = g(x)\big\}$).

EXERCISE 1.54. Let X be a topological space, Y a normed real vector space and $f, g : X \to Y \setminus \{0\}$ continuous maps such that:
$$\big\|f(x) - g(x)\big\| < \big\|f(x)\big\|,$$
for all $x \in X$. Show that:
$$H(s, x) = (1 - s)f(x) + sg(x), \quad s \in [0, 1],\ x \in X,$$
defines a homotopy from $f : X \to Y \setminus \{0\}$ to $g : X \to Y \setminus \{0\}$.

EXERCISE 1.55. If $T, S : \mathbb{R}^n \to \mathbb{R}^n$ are linear isomorphisms whose determinants have the same sign, show that the maps $T : \big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big) \to \big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big)$ and $S : \big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big) \to \big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big)$ are homotopic (*hint*: T and S can be connected by a continuous curve in the general linear group of $\mathbb{R}^n$).

EXERCISE 1.56. Let X be a convex subset of a normed vector space. Show that a subset $S \subset X$ is a retract of X if and only if S is a strong deformation retract of X.

EXERCISE 1.57. Let $\sigma : [0, 1] \to [0, 1]$ be a continuous map. Given topological spaces X, Y and a homotopy $H : X \times [0, 1] \to Y$ we define the *reparameterization* of H by σ by:
$$H_\sigma(x, t) = H\big(x, \sigma(t)\big),$$
for all $x \in X$, $t \in [0, 1]$.

- Show that if $\sigma(0) = 0$ and $\sigma(1) = 1$ then the map H_σ is homotopic[20] to H relatively to $X \times \{0, 1\}$ (*hint*: look at $H\big(x, (1 - s)t + s\sigma(t)\big)$).
- Show that if $\sigma(0) = \sigma(1) = t_0 \in [0, 1]$ then H_σ is homotopic relatively to $X \times \{0, 1\}$ to the constant homotopy given by $(x, t) \mapsto H(x, t_0)$.

[20] Yes, we mean a homotopy of homotopies!

- If $H, H', H'' : X \times [0,1] \to Y$ are homotopies with $H_1 = H'_0$, $H'_1 = H''_0$, conclude from the items above that $(H \cdot H') \cdot H''$ is homotopic relatively to $X \times \{0,1\}$ to $H \cdot (H' \cdot H'')$.
- Conclude from the items above that $H \cdot H^{-1}$ and $H^{-1} \cdot H$ are both homotopic relatively to $X \times \{0,1\}$ to constant homotopies.
- If $H^0, H^1 : X \times [0,1] \to Y$ are the constant homotopies defined by $H^0(x,t) = H(x,0)$, $H^1(x,t) = H(x,1)$ then $H^0 \cdot H$ and $H \cdot H^1$ are both homotopic relatively to $X \times \{0,1\}$ to the homotopy H.
- If $H, K, H' : X \times [0,1] \to Y$ are homotopies such that $H_1 = H'_0$ and such that H and K are homotopic relatively to $X \times \{0,1\}$ then ($K_1 = H'_0$ and) $H \cdot H'$ is homotopic to $K \cdot H'$ relatively to $X \times \{0,1\}$. Similarly, if $H_1 = K_0$ and K is homotopic to K' relatively to $X \times \{0,1\}$ then $H \cdot K$ and $H \cdot K'$ are homotopic relatively to $X \times \{0,1\}$.

EXERCISE 1.58. A non empty topological space X is called *contractible* if the identity of X is homotopic to some constant map $\mathfrak{c} : X \to X$. Show that the following are equivalent:

- X is contractible;
- X has the same homotopy type of a one point space.

Show that a contractible space is arc-connected; conclude that in a contractible space X the identity is homotopic to *every* constant map $\mathfrak{c} : X \to X$.

A computation of the singular homology of spheres and the torus.

EXERCISE 1.59. Let $R : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ denote the reflection map (1.8.5). Show that the homomorphisms:

$$R_* : H_{n+1}\big(\overline{\mathrm{B}}^{n+1}, S^n\big) \longrightarrow H_{n+1}\big(\overline{\mathrm{B}}^{n+1}, S^n\big),$$
$$R_* : H_{n+1}\big(\mathbb{R}^{n+1}, \mathbb{R}^{n+1} \setminus \{0\}\big) \longrightarrow H_{n+1}\big(\mathbb{R}^{n+1}, \mathbb{R}^{n+1} \setminus \{0\}\big),$$

are both equal to minus the identity (*hint*: use Example 1.8.2).

EXERCISE 1.60. The goal of this exercise is to compute the homology of the Klein bottle by an argument similar to the one used in Example 1.8.3. Denote by K the *Klein bottle* that is obtained as the quotient of the square $[0,1]^2$ by the equivalence relation $\sim$ spanned by $(x,0) \sim (x,1)$ and $(0,y) \sim (1,1-y)$ for all $x, y \in [0,1]$; we denote by $q : [0,1]^2 \to K$ the quotient map. Consider the following objects:

- $A = q\Big(\big(\,[0,\tfrac{1}{2}[\,\cup\,]\tfrac{1}{2},1]\,\big) \times [0,1]\Big)$, $C = q\big([\tfrac{1}{4},\tfrac{3}{4}] \times [0,1]\big)$;
- $D_i = q\big(\{x_i\} \times [0,1]\big)$, $i = 0,1,2,3$, where $x_0 = \frac{1}{2}$, $x_1 = 0$, $x_2 = \frac{1}{4}$ and $x_3 = \frac{3}{4}$;
- $D = D_2 \cup D_3$;
- the reflection $r : [0,1]^2 \to [0,1]^2$ defined by $r(x,y) = (1-x,y)$.

Prove the following:

(b) D_i is homeomorphic to a circle for all i.

(a) The circles D_2 and D_3 are strong deformation retracts of both A and C; D_0 is a strong deformation retract of C, D_1 is a strong deformation retract of A and D is a strong deformation retract of $C \setminus D_0$.
(b) r defines by passage to the quotient a homeomorphism $R : K \to K$; R leaves C, A, D_1 and D invariant and it fixes D_0.
(c) $R_* : H_1(D_1) \to H_1(D_1)$ equals minus the identity;
(d) Repeat the steps used in Example 1.8.3 after observing that one can use the commutative diagram (1.8.8) with $\mathbb{T}$ replaced by K; using item (c) conclude that the homomorphism $H_2(K, A) \cong \mathbb{Z} \to \mathbb{Z} \cong H_1(A)$ is multiplication by 2 (instead of zero).
(e) Prove that $H_0(K) \cong \mathbb{Z}$, $H_1(K) \cong \mathbb{Z} \oplus \mathbb{Z}_2$ and $H_p(K) = 0$ for $p \geq 2$.

Local homology.

EXERCISE 1.61 (invariance of dimension). Show that if a non empty open subset of $\mathbb{R}^m$ is homeomorphic to an open subset of $\mathbb{R}^n$ then $m = n$.

EXERCISE 1.62 (invariance of the boundary). Show that if $h : U \to V$ is a homeomorphism between open subsets of the closed half-space H^n then:

$$h(U \cap \mathrm{Bd}(\mathsf{H}^n)) = V \cap \mathrm{Bd}(\mathsf{H}^n).$$

EXERCISE 1.63. An *n-dimensional topological manifold with boundary* is a non empty Hausdorff second countable topological space M such that each point in M has an open neighborhood in M that is homeomorphic to an open subset of the closed half-space H^n. The *interior* of M, denoted $\mathrm{inter}(M)$, is the set of points of M that admit an open neighborhood that is homeomorphic to an open subset of $\mathbb{R}^n$; the other points of M constitute the *boundary* of M, which we denote[21] by $\mathrm{Bd}(M)$. If the boundary of M is empty we say that M is a *topological manifold without boundary* (or simply a *topological manifold*). When $n = 0$, an n-dimensional topological manifold with boundary is simply a discrete topological space M and actually *there can be no boundary*.

Let M be an n-dimensional topological manifold with boundary.

- If $U \subset M$ is a non empty open subset show that U is also an n-dimensional topological manifold with boundary and that:

$$\mathrm{inter}(U) = \mathrm{inter}(M) \cap U, \quad \mathrm{Bd}(U) = \mathrm{Bd}(M) \cap U.$$

[21] One should be careful to distinguish between the notions of interior and boundary just defined for manifolds and the usual notions of interior and boundary of a subset of a topological space. If M is a subset of a larger topological space X then it is indeed quite possible that the interior of M (respectively, the boundary of M) as a manifold do not coincide with the interior of M (respectively, the boundary of M) as a subset of the topological space X; for instance, if $X = M$ then the interior of M (respectively, the boundary of M) as a subset of X would always be the entire manifold M (respectively, the empty set). Unfortunately, this messy terminology is quite useful in the literature. We somewhat attempt to reduce the confusion by using the following notation: the "manifold" interior and boundary of M will be denoted respectively by $\mathrm{inter}(M)$ and $\mathrm{Bd}(M)$, while the "topological" interior and boundary of M will be denoted respectively by $\mathrm{int}(M)$ and ∂M.

- If N is a topological space and $h : M \to N$ is a homeomorphism show that N is also an n-dimensional topological manifold with boundary and that h takes boundary to boundary and interior to interior, i.e.:
$$h\big(\mathrm{inter}(M)\big) = \mathrm{inter}(N), \quad h\big(\mathrm{Bd}(M)\big) = \mathrm{Bd}(N).$$
- Show that the local homology groups of M at an interior point $x \in \mathrm{inter}(M)$ are given by:
$$H_p\big(M, M \setminus \{x\}\big) \cong \begin{cases} 0, & p \neq n, \\ \mathbb{Z}, & p = n; \end{cases}$$
if $x \in \mathrm{Bd}(M)$ is a boundary point, show that $H_p\big(M, M \setminus \{x\}\big) = 0$ for all $p \in \mathbb{Z}$.
- Show that the *dimension* $\dim(M)$ of M is well-defined, i.e., that if M is also an m-dimensional topological manifold with boundary then $m = n$.
- Show that H^n is an n-dimensional topological manifold with boundary whose boundary is:
$$\mathrm{Bd}(\mathsf{H}^n) = \big\{x = (x_1, \ldots, x_n) \in \mathsf{H}^n : x_n = 0\big\},$$
i.e., the notation $\mathrm{Bd}(\mathsf{H}^n)$ used so far is indeed compatible with the notation for a boundary of a manifold just introduced.
- Conclude from the items above that if $\varphi : U \to \widetilde{U}$ is an arbitrary homeomorphism between an open subset $U \subset M$ and an open subset $\widetilde{U} \subset \mathsf{H}^n$ then:
$$\varphi\big(U \cap \mathrm{inter}(M)\big) = \big\{x = (x_1, \ldots, x_n) \in \widetilde{U} : x_n > 0\big\},$$
$$\varphi\big(U \cap \mathrm{Bd}(M)\big) = \big\{x = (x_1, \ldots, x_n) \in \widetilde{U} : x_n = 0\big\}.$$
- Show that the interior of M is an open dense subset of M and that the boundary of M is either empty or an $(n-1)$-dimensional topological manifold (without boundary).

Orientation on manifolds.

EXERCISE 1.64. Let X be a topological space. A *pre-sheaf of abelian groups* on X is a rule $\mathcal{PS}$ that associates to each open set $U \subset X$ an abelian group $\mathcal{PS}(U)$ and to each pair of open sets $U, V \subset X$ with $V \subset U$ a homomorphism $\rho_{UV} : \mathcal{PS}(U) \to \mathcal{PS}(V)$ such that given open sets $W \subset V \subset U \subset X$ we have a commutative diagram:

$$\begin{tikzcd} & \mathcal{PS}(U) \arrow[dl, "\rho_{UV}"'] \arrow[dr, "\rho_{UW}"] & \\ \mathcal{PS}(V) \arrow[rr, "\rho_{VW}"'] & & \mathcal{PS}(W) \end{tikzcd}$$

The homomorphisms ρ_{UV} are usually called the *restriction maps* of the pre-sheaf $\mathcal{PS}$. Show that the following are examples of pre-sheaves:

- X an arbitrary topological space,
$$\mathcal{PS}(U) = \text{the set of continuous real-valued maps on } U,$$
and $\rho_{UV}(f) = f|_V$ for all $f \in \mathcal{PS}(U)$.
- X a differentiable manifold,
$$\mathcal{PS}(U) = \text{the set of smooth real-valued maps on } U,$$
and $\rho_{UV}(f) = f|_V$ for all $f \in \mathcal{PS}(U)$.
- X a differentiable manifold,
$$\mathcal{PS}(U) = \text{the set of smooth differential } r\text{-forms on } U,$$
and $\rho_{UV}(\omega) = \omega|_V$ for all $\omega \in \mathcal{PS}(U)$.
- $X = \mathbb{C}^n$,
$$\mathcal{PS}(U) = \text{the set of holomorphic complex-valued maps on } U,$$
and $\rho_{UV}(f) = f|_V$ for all $f \in \mathcal{PS}(U)$.
- X an arbitrary topological space, G an arbitrary fixed abelian group, $\mathcal{PS}(U) = G$ for all U and ρ_{UV} equals the identity of G for all U, V.

EXERCISE 1.65. Let X be a topological space. Recall that a *basis of open sets* for X is a collection $\mathfrak{B}$ of open subsets of X such that every open subset of X can be written as a union of elements of $\mathfrak{B}$. Given a set X and a collection $\mathfrak{B}$ of subsets of X show that the following conditions are equivalent:

- there exists a (automatically unique) topology for X having $\mathfrak{B}$ as a basis of open sets;
- $X = \bigcup_{B \in \mathfrak{B}} B$ and for every $B_1, B_2 \in \mathfrak{B}$ and every $x \in B_1 \cap B_2$ there exists $C \in \mathfrak{B}$ with $x \in C \subset B_1 \cap B_2$.

EXERCISE 1.66. Let X be a topological space and let $\mathcal{PS}$ be a pre-sheaf of abelian groups on X. The *group of germs* $\mathcal{S}_x$ of the pre-sheaf $\mathcal{PS}$ at the point $x \in X$ is defined as follows. Consider the quotient:
$$\mathcal{S}_x = \bigcup_{\substack{U \text{ an open} \\ \text{neighborhood of } x}} \mathcal{PS}(U) \Big/ \sim,$$
where the equivalence relation $\sim$ is defined by declaring that $f \in \mathcal{PS}(U)$ is equivalent to $g \in \mathcal{PS}(V)$ if and only if $\rho_{UW}(f) = \rho_{VW}(g)$ for some open neighborhood W of x contained in $U \cap V$. We define an abelian group structure in $\mathcal{S}_x$ by taking the sum of the class of $f \in \mathcal{PS}(U)$ with the class of $g \in \mathcal{PS}(V)$ to be the class of $\rho_{UW}(f) + \rho_{VW}(g) \in \mathcal{PS}(W)$, for some open neighborhood W of x contained in $U \cap V$. If $f \in \mathcal{PS}(U)$ and $x \in U$ then we denote by $[f]_x$ the class of f in $\mathcal{S}_x$; we call $[f]_x$ the *germ* of f at x.

- Check that all the construction above actually makes sense.
- Let $\mathcal{S}$ be the disjoint union of the abelian groups $\mathcal{S}_x$, $x \in X$. Consider the collection $\mathfrak{B}$ consisting of the sets $\big\{[f]_x : x \in U\big\}$, where U runs over all the open subsets of X and f runs through $\mathcal{PS}(U)$. Show that there exists a (unique) topology for $\mathcal{S}$ having $\mathfrak{B}$ as a basis of open sets.

The topological space $\mathcal{S}$ is called the *sheaf of germs* corresponding to the pre-sheaf $\mathcal{PS}$.

EXERCISE 1.67. Let M be a topological manifold. Show that the map:

$$\mathcal{O}(M) \ni \alpha \longmapsto -\alpha \in \mathcal{O}(M),$$

is a homeomorphism. Conclude that if $\tau : M \to \mathcal{O}(M)$ is an orientation for M then also $(-\tau)(x) = -\tau(x)$, $x \in M$, defines an orientation for M.

EXERCISE 1.68. Let M, N be differentiable manifolds of the same dimension and let $f : M \to N$ be a diffeomorphism. Let $\tau : M \to \mathcal{O}(M)$, $\tau' : N \to \mathcal{O}(N)$ be homological orientations for M and N respectively and let $\bar{\tau}$, $\bar{\tau}'$ be the differentiable orientations that correspond respectively to τ and τ'. Show that f is a positively oriented (respectively, negatively oriented) homeomorphism in the sense of Definition 1.10.13 if and only if f is *positively oriented (*respectively, *negatively oriented) in the differentiable sense*, i.e., if and only if the isomorphism $\mathrm{d}f_x : T_xM \to T_{f(x)}N$ takes $\bar{\tau}_x$ to $\bar{\tau}'_{f(x)}$ (respectively, takes $\bar{\tau}_x$ to $-\bar{\tau}'_{f(x)}$) for every $x \in M$ (*hint*: use local charts and Corollary 1.10.19).

EXERCISE 1.69 (perturbation of identity). Recall that a *contraction* is a map between metric spaces that is Lipschitz continuous with a Lipschitz constant less than 1. Show that if $g : \mathbb{R}^n \to \mathbb{R}^n$ is a contraction then the map $h : \mathbb{R}^n \to \mathbb{R}^n$ defined by $h(x) = x + g(x)$ is a (global) homeomorphism of $\mathbb{R}^n$ (*hint*: for fixed $y \in \mathbb{R}^n$, solving the equation $h(x) = y$ is equivalent to find a fixed point for the contraction $x \mapsto y - g(x)$).

EXERCISE 1.70. Let $\varphi : S^n \setminus \{\mathfrak{n}\} \to \mathbb{R}^n$ denote the stereographic projection from the north pole $\mathfrak{n} = e_{n+1}$ onto the hyper-plane $\mathbb{R}^n \times \{0\}$ containing the equator of S^n. By identifying the tangent space $T_{\mathfrak{s}}S^n$ at the south pole with $\mathbb{R}^n$ via the projection onto the first n coordinates, show that the differential $\mathrm{d}\varphi_{\mathfrak{s}} : T_{\mathfrak{s}}S^n \to \mathbb{R}^n$ equals multiplication by $\frac{1}{2}$. Conclude that, if S^n is endowed with the outward pointing orientation, then φ is a positive chart if and only if n is odd.

EXERCISE 1.71. Use the commutative diagram (in which all arrows are induced by inclusion):

$$\begin{tikzcd}
H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \mathrm{B}^n\big) \arrow[dr] & \\
H_n\big(\overline{\mathrm{B}}^n, S^{n-1}\big) \arrow[u] \arrow[r] & H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{v\}\big)
\end{tikzcd}$$

and Remark 1.10.20 to prove the following fact: if $\alpha \in H_n\big(\overline{\mathrm{B}}^n, S^{n-1}\big)$ is mapped to $\tau^{[n]}(0)$ by the homomorphism:

$$H_n\big(\overline{\mathrm{B}}^n, S^{n-1}\big) \longrightarrow H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{0\}\big)$$

induced by inclusion then for every $v \in \mathrm{B}^n$ the homomorphism:

$$H_n\big(\overline{\mathrm{B}}^n, S^{n-1}\big) \longrightarrow H_n\big(\mathbb{R}^n, \mathbb{R}^n \setminus \{v\}\big)$$

induced by inclusion maps α to $\tau^{[n]}(v)$.

EXERCISE 1.72. Let M be an n-dimensional topological manifold with non empty boundary and W an open subset of M with $\mathrm{Bd}(W) \neq \emptyset$. Assume that $U \subset W$ is open and that $\alpha \in H_n\big(M, M \setminus \mathrm{inter}(U)\big)$ is the image of $\beta \in H_n\big(W, W \setminus \mathrm{inter}(U)\big)$ under the homomorphism induced by inclusion. Recalling the usual identifications of $\mathcal{O}\big(\mathrm{inter}(W)\big)$ (respectively, of $\mathcal{O}\big(\mathrm{Bd}(W)\big)$) with a subset of $\mathcal{O}\big(\mathrm{inter}(M)\big)$ (respectively, with a subset of $\mathcal{O}\big(\mathrm{Bd}(M)\big)$), prove that:

$$\mathcal{O}_{\mathrm{i}}(\alpha; U, M) = \mathcal{O}_{\mathrm{i}}(\beta; U, W), \quad \mathcal{O}_{\mathrm{b}}(\alpha; U, M) = \mathcal{O}_{\mathrm{b}}(\beta; U, W).$$

Degree theory.

EXERCISE 1.73. Let $f : X \to Y$ be a continuous map. Show that:

- if f is proper and $f(X) \subset Y' \subset Y$ then $f : X \to Y'$ is proper;
- if $f(X) \subset Y'$, Y' is closed in Y and $f : X \to Y'$ is proper then $f : X \to Y$ is proper;
- if $X' \subset X$ is closed and $f : X \to Y$ is proper then $f|_{X'} : X' \to Y$ is proper;
- for any $A \subset Y$, if $f : X \to Y$ is proper then $f|_{f^{-1}(A)} : f^{-1}(A) \to A$ is proper;
- if $f : X \to Y$ and $g : Y \to Z$ are proper then so is $g \circ f : X \to Z$.

EXERCISE 1.74. If $f : X \to Y$ and $f' : X' \to Y'$ are proper then so is $f \times f' : X \times X' \to Y \times Y'$.

EXERCISE 1.75. Let $H : X \times [0,1] \to X$ be a *homotopy with compact support*, i.e., there exists a compact subset $K \subset X$ such that $H(t, x) = x$ for all $t \in [0,1]$ and all $x \in X \setminus K$. Show that if X is Hausdorff then H is a proper map.

EXERCISE 1.76. Assume that X, Y are Hausdorff and that Y satisfies either one of the following:

- Y is first countable, i.e., every point has a countable fundamental system of neighborhoods;
- Y is locally compact.

Then every proper map $f : X \to Y$ is closed.

EXERCISE 1.77. Let $n \geq 1$ be fixed and denote by $(S^n)_{\mathfrak{n}}$ and $(S^n)_{\mathfrak{s}}$ respectively the northern and the southern hemispheres of S^n, i.e.:

$$(S^n)_{\mathfrak{n}} = \big\{x \in S^n : x_{n+1} \geq 0\big\}. \quad (S^n)_{\mathfrak{s}} = \big\{x \in S^n : x_{n+1} \leq 0\big\}.$$

Let $f : S^n \to S^n$ be a continuous map that preserves $(S^n)_{\mathfrak{n}}$ and $(S^n)_{\mathfrak{s}}$, i.e., $f\big((S^n)_{\mathfrak{n}}\big) \subset (S^n)_{\mathfrak{n}}$ and $f\big((S^n)_{\mathfrak{s}}\big) \subset (S^n)_{\mathfrak{n}}$; in particular, f must preserve the equator $S^{n-1} = \big\{x \in S^n : x_{n+1} = 0\big\}$ of S^n. If all spheres are endowed with their canonical orientation, show that the degree of f equals the degree of its restriction $f|_{S^{n-1}} : S^{n-1} \to S^{n-1}$ to the equator.

hint: use Example 1.11.6 and the commutative diagram:

$$\begin{tikzcd}
H_n(S^n) \arrow[r, "\cong"] \arrow[d, "f_*"'] & H_n\big(S^n,(S^n)_{\mathfrak s}\big) \arrow[d, "f_*"'] & H_n\big((S^n)_{\mathfrak n},S^{n-1}\big) \arrow[l, "\cong", "\text{by excision and retraction}"'] \arrow[r, "\cong", "\partial_*"'] \arrow[d, "f_*"] & H_{n-1}(S^{n-1}) \arrow[d, "f_*"] \\
H_n(S^n) \arrow[r, "\cong"'] & H_n\big(S^n,(S^n)_{\mathfrak s}\big) & H_n\big((S^n)_{\mathfrak n},S^{n-1}\big) \arrow[l, "\cong"', "\text{by excision and retraction}"] \arrow[r, "\partial_*", "\cong"'] & H_{n-1}(S^{n-1})
\end{tikzcd}$$

Homology with general coefficients.

EXERCISE 1.78. Let $\mathbb{K}$ be an arbitrary field and $\mathfrak{C}$ a chain complex of $\mathbb{K}$-vector spaces. Assume that every $\mathfrak{C}_p$ is finite-dimensional and that $\mathfrak{C}_p \neq 0$ for at most a finite number of indices $p \in \mathbb{Z}$. Given a $\mathbb{K}$-linear chain map $f : \mathfrak{C} \to \mathfrak{C}$ show that:

$$\sum_{p\in\mathbb{Z}} (-1)^p \operatorname{tr}(f_p) = \sum_{p\in\mathbb{Z}} (-1)^p \operatorname{tr}(f_*)_p,$$

where $(f_*)_p : H_p(\mathfrak{C}) \to H_p(\mathfrak{C})$ is the $\mathbb{K}$-linear map induced in homology by f and $\operatorname{tr}(T)$ denotes the *trace* of an endomorphism T of a finite dimensional vector space (*hint*: use the equalities:

$$\operatorname{tr}(f_*)_p = \operatorname{tr}\big(f_p|_{Z_p(\mathfrak{C})}\big) - \operatorname{tr}\big(f_p|_{B_p(\mathfrak{C})}\big),$$

$$\operatorname{tr}\big(f_{p-1}|_{B_{p-1}(\mathfrak{C})}\big) = \operatorname{tr}(f_p) - \operatorname{tr}\big(f_p|_{Z_p(\mathfrak{C})}\big). \tag{1.17.9}$$

The equality (1.17.9) follows from the commutative diagram:

$$\begin{tikzcd}
\mathfrak{C}_p/Z_p(\mathfrak{C}) \arrow[r] \arrow[d, "\cong"'] & \mathfrak{C}_p/Z_p(\mathfrak{C}) \arrow[d, "\cong"] \\
B_{p-1}(\mathfrak{C}) \arrow[r] & B_{p-1}(\mathfrak{C})
\end{tikzcd}$$

where the horizontal arrows are induced by f and the vertical arrows are induced by ∂).

EXERCISE 1.79. Given a field $\mathbb{K}$, consider the tensor product of abelian groups $\mathbb{Q}\otimes\mathbb{K}$ endowed with the standard $\mathbb{K}$-vector space structure. Show that $\dim_{\mathbb{K}}(\mathbb{Q}\otimes\mathbb{K}) = 1$ if $\mathbb{K}$ has characteristic zero and that $\mathbb{Q}\otimes\mathbb{K} = 0$ otherwise.

EXERCISE 1.80. Let V be a (possibly infinite dimensional) vector space over a field $\mathbb{K}$ and let $V_2 \subset V_1 \subset V$ be any two subspaces of V. Show that:

$$\dim(V) = \dim(V/V_1) + \dim(V_2) + \dim(V_1/V_2).$$

Quotient topology and attachment spaces.

EXERCISE 1.81. Given a topological space X, a set Y and a map $q : X \to Y$, show that there exists a unique topology τ_q in Y such that $q : X \to (Y, \tau_q)$ is a quotient map; we call τ_q the *quotient topology* in Y corresponding to q. Show that if τ is a topology for Y that makes $q : X \to (Y, \tau)$ continuous then τ_q is *finer* (or *stronger*) than τ, i.e., $\tau_q \supset \tau$.

EXERCISE 1.82. Show that $q : X \to Y$ is a quotient map if and only if the following condition hold: $F \subset Y$ is closed if and only if $q^{-1}(F) \subset X$ is closed.

EXERCISE 1.83. Show that $q : X \to Y$ is a quotient map iff the following three conditions hold:

- $q : X \to q(X)$ is a quotient map;
- $q(X)$ is both open and closed in Y;
- $Y \setminus q(X)$ is discrete.

EXERCISE 1.84. If $q : X \to Y$ is bijective then q is a quotient map iff q is a homeomorphism.

EXERCISE 1.85. Show that the composition of quotient maps is a quotient map.

EXERCISE 1.86. Show the converse of Proposition 1.14.2: if $q : X \to Y$ is a continuous map and if the continuity of f implies the continuity of $\overline{f}$ in the commutative diagram (1.14.1) (for every topological space Z and every map $f : X \to Z$) then q is a quotient map.

EXERCISE 1.87. In the commutative diagram (1.14.1), show that if q and $\overline{f}$ are continuous and f is a quotient map then $\overline{f}$ is a quotient map; conclude that, if q is a quotient map, then f is a quotient map iff $\overline{f}$ is a quotient map.

EXERCISE 1.88. Given a map $q : X \to Y$ and subsets $S_1, S_2 \subset X$, show that $q(S_1 \cap S_2) = q(S_1) \cap q(S_2)$ if either S_1 or S_2 is saturated.

EXERCISE 1.89. Let X be a T_1 topological space. Show that the following are equivalent:

- given two disjoint closed subsets $F, G \subset X$, there exists a continuous map $\phi : X \to \mathbb{R}$ with $\phi|_F \equiv 0$ and $\phi|_G \equiv 0$;
- every continuous map $\phi : F \to \mathbb{R}$ defined in a closed subset $F \subset X$ admits a continuous ($\mathbb{R}$-valued) extension;
- X is T_4.

EXERCISE 1.90. Show that if $q : X \to Y$ is a continuous, closed surjective map and if X is T_4 then also Y is T_4 (*hint*: given disjoint closed subsets $F, G \subset Y$, choose disjoint open neighborhoods U and V of $q^{-1}(F)$ and $q^{-1}(G)$ respectively on X; look at $Y \setminus q(X \setminus U)$, $Y \setminus q(X \setminus V)$).

EXERCISE 1.91. Let $q : X \to Y$ be a surjective continuous map. Prove the following.

(1) q is a quotient map iff q maps saturated open (resp., closed) subsets of X to open (resp., closed) subsets of Y.
(2) If q is a quotient map then q is an open (resp., closed) map iff the saturation of open (resp., closed) subsets of X is open (resp., closed) in X.
(3) If q is either an open map or a closed map then q is a quotient map.
(4) If X is compact and Y is Hausdorff then q is a closed map and therefore a quotient map.

EXERCISE 1.92. If a group G acts on a topological space X by homeomorphisms then the quotient map $q : X \to X/G$ is open.

EXERCISE 1.93. Given a quotient map $q : X \to Y$ and a subset $S \subset X$, show that $q|_S : S \to q(S)$ is a quotient map iff the following property holds: given a subset $A \subset S$ that is open (resp., closed) in S and saturated with respect to $q|_S$ then there exists an open (resp., a closed) subset A' in X, saturated with respect to q, such that $A = A' \cap S$.

EXERCISE 1.94. Given a family $(X_i)_{i\in I}$ of topological spaces, denote by X their external topological sum. Prove the following:

- for every $i \in I$ the map $x \mapsto (i, x)$ is a homeomorphism from X_i onto the subspace $\{i\} \times X_i$ of X;
- X is the internal topological sum of the subspaces $\{i\} \times X_i$.

Moreover, if a topological space X is the internal topological sum of a family $(X_i)_{i\in I}$ of subspaces of X and if X' denotes the external topological sum of the family $(X_i)_{i\in I}$ then we have a homeomorphism from X' to X that carries (i, x) to x for every $i \in I$, $x \in X_i$.

EXERCISE 1.95. Assume that X is the disjoint union of a family of subspaces $(X_i)_{i\in I}$. If for every topological space Y and every map $f : X \to Y$ we have

$$f \text{ is continuous} \iff f|_{X_i} \text{ is continuous for every } i \in I$$

then X is the topological sum of the spaces X_i.

EXERCISE 1.96. Show that a union $X = \bigcup_{i\in I} X_i$ is coherent iff the following property holds: given a subset $F \subset X$, if $F \cap X_i$ is closed in X_i for every $i \in I$ then F is closed in X.

EXERCISE 1.97. Show that the union $X = \bigcup_{i\in I} X_i$ is coherent in the following situations:

(a) the interiors of the sets X_i cover X;
(b) I is finite and each X_i is closed;
(c) $(X_i)_{i\in I}$ is a locally finite family of closed subsets (recall that $(X_i)_{i\in I}$ is *locally finite* if every point of X has a neighborhood that intercepts X_i for at most a finite number of indices i).

EXERCISE 1.98. If $X = \bigcup_{i\in I} X_i$, consider the canonical map

$$q : \sum_{i\in I} X_i \longrightarrow X$$

such that $q|_{X_i}$ is the inclusion of X_i in X for every $i \in I$. Show that $X = \bigcup_{i\in I} X_i$ is a coherent union iff q is a quotient map.

EXERCISE 1.99. Let X be a set and let $X = \bigcup_{i\in I} X_i$ be a covering of X. Assume that, for every $i \in I$, we are given a topology τ_i for the set X_i such that the following conditions are satisfied:

- for every $i, j \in I$, $X_i \cap X_j$ inherits the same topology from (X_i, τ_i) and from (X_j, τ_j);

- for every $i, j \in I$, $X_i \cap X_j$ is closed in both (X_i, τ_i) and (X_j, τ_j).

Let τ be the topology for X that makes the canonical projection

$$q : \sum_{i \in I} (X_i, \tau_i) \longrightarrow X$$

a quotient map, i.e., $F \subset X$ is closed in (X, τ) iff $F \cap X_i$ is closed in (X_i, τ_i) for every $i \in I$. Show that:

- τ_i is precisely the topology that X_i inherits from (X, τ) and every X_i is closed in (X, τ).
- $X = \bigcup_{i \in I} X_i$ is a coherent union.

Repeat the exercise stated above with the word "closed" replaced by the word "open" throughout.

EXERCISE 1.100. Let $X = \bigcup_{i \in I} A_i$ be a coherent union. For each $i \in I$ let B_i be an open set in A_i and assume that for every $i, j \in I$ we have $B_i \cap A_j \subset B_j$. Show that the union of the B_i's is coherent (*hint*:prove that the topological sum $\sum_{i \in I} B_i$ is a saturated open set in $\sum_{i \in I} A_i$ with respect to the canonical map $q : \sum_{i \in I} A_i \to X$; use Exercise 1.98).

EXERCISE 1.101. Show that if $X = \bigcup_{i \in I} X_i$ is a coherent union and if Z is a locally compact Hausdorff space then also the union $X \times Z = \bigcup_{i \in I} (X_i \times Z)$ is coherent.

EXERCISE 1.102. Let $X = \bigcup_{i \in I} X_i$ be a coherent union and let for each $i \in I$, $H_i : X_i \times [0,1] \to Y$ be a continuous map. Assume that for every $i, j \in I$ the maps H_i and H_j agree on $(X_i \cap X_j) \times [0,1]$. Show that there exists a (unique) continuous map $H : X \times [0,1] \to Y$ such that $H|_{X_i \times [0,1]} = H_i$ for all $i \in I$.

EXERCISE 1.103. Show that a disjoint union $X = \bigcup_{i \in I} X_i$ is coherent iff X is the topological sum of the spaces X_i.

EXERCISE 1.104. Show the converse of Proposition 1.14.9: if $X = \bigcup_{i \in I} X_i$ and if for every topological space Y and every map $f : X \to Y$ the condition

$$f \text{ is continuous} \quad \Longleftrightarrow \quad f|_{X_i} \text{ is continuous for every } i \in I$$

holds then $X = \bigcup_{i \in I} X_i$ is a coherent union.

EXERCISE 1.105. If $X = \bigcup_{i \in I} X_i$, $Y = \bigcup_{i \in I} Y_i$ are coherent unions and $q : X \to Y$ is a map with $q(X_i) \subset Y_i$ and such that $q|_{X_i} : X_i \to Y_i$ is a quotient map for every $i \in I$ then q is a quotient map.

EXERCISE 1.106. If $q : X \to Y$ is a quotient surjective map and $X = \bigcup_{i \in I} X_i$ is a coherent union then $Y = \bigcup_{i \in I} q(X_i)$ is a coherent union.

EXERCISE 1.107. Let V be a real vector space and let K be a simplicial complex in V. Consider each affine simplex $\sigma \in K$ endowed with the topology it inherits from any finite dimensional subspace of V containing σ (finite dimensional real vector spaces have a canonical topology, which is induced by any norm). Show that the set:

$$|K| = \bigcup_{\sigma \in K} \sigma$$

admits a unique topology that makes the union above coherent. The resulting topological space is called the *polyhedron* of the simplicial complex K. If X is a topological space then a homeomorphism $h : X \to |K|$ onto the polyhedron of some simplicial complex K is called a *triangulation* for X; if X admits a triangulation then we say that X is triangulable.

EXERCISE 1.108. Let X, Y be topological spaces and $f : A \subset X \to Y$ a continuous map. Show that every open (respectively, closed) subset of X disjoint from A is mapped homeomorphically onto an open (respectively, closed) subset of $X \cup_f Y$ via the canonical projection $X + Y \to X \cup_f Y$.

EXERCISE 1.109. The goal of this exercise is to prove that if $\alpha : S^{n-1} \to S^{n-1}$ is a homeomorphism then the attachment space $\overline{\mathrm{B}}^n \cup_\alpha \overline{\mathrm{B}}^n$ is homeomorphic to the sphere S^n.

(a) Show that every homeomorphism $\beta : S^{n-1} \to S^{n-1}$ extends to a homeomorphism from $\overline{\mathrm{B}}^n$ to $\overline{\mathrm{B}}^n$ (*hint*: the ball is the cone of the sphere!).
(b) Let $h : \overline{\mathrm{B}}^n \to \overline{\mathrm{B}}^n$ be a homeomorphism that extends α^{-1} and consider the homeomorphism from $\overline{\mathrm{B}}^n + \overline{\mathrm{B}}^n$ to itself and gives the identity on the first ball and h on the second. Show that the latter homeomorphism induces a homeomorphism from $\overline{\mathrm{B}}^n \cup_\alpha \overline{\mathrm{B}}^n$ to $\overline{\mathrm{B}}^n \cup_i \overline{\mathrm{B}}^n$, where $i : S^{n-1} \to \overline{\mathrm{B}}^n$ is the inclusion.
(c) Conclude the proof by showing that $\overline{\mathrm{B}}^n \cup_i \overline{\mathrm{B}}^n$ is homeomorphic to S^n.

EXERCISE 1.110. Show that the attachment space $X \cup_f Y$ is the coherent union of the images of X and Y by the quotient map $X + Y \to X \cup_f Y$.

EXERCISE 1.111. Let X, Y be topological spaces and $A, X' \subset X$ subspaces of X with $A \subset X'$. Given a continuous map $f : A \to Y$, denote by $q : X + Y \to X \cup_f Y$ the canonical quotient map.

- If X' is closed in X, show that q restricts to a quotient map $q : X' + Y \to X' \cup_f Y$, when $X' \cup_f Y$ is considered with the topology induced by $X \cup_f Y$ (*hint*: $X' + Y$ is a saturated closed subset of $X + Y$).
- Generalize the item above to the case that X' is not necessarily closed in X (*hint*: if F_1 is closed in X', F_2 is closed in Y, $F_1 + F_2$ is saturated and F is a closed subset of X with $F_1 = F \cap X'$ then $F + F_2$ is saturated; use Exercise 1.93).
- If X' is a strong deformation retract of X, show that $X' \cup_f Y$ is a strong deformation retract of $X \cup_f Y$ (the items above showed that $X' \cup_f Y$ is indeed a subspace of $X \cup_f Y$!).

EXERCISE 1.112. Let X, Y be topological spaces, $f : A \to Y$ a continuous map defined on a subspace of $A \subset X$ and A' an open subset of X contained in A. Show that the quotient map $q : X + Y \to X \cup_f Y$ restricted to $(X \setminus A') + Y$ is still a quotient map onto $X \cup_f Y$. Obtain a homeomorphism from $(X \setminus A') \cup_{f|_{(A\setminus A')}} Y$ to $X \cup_f Y$.

EXERCISE 1.113. Show that the quotient map $q : X + Y \to X \cup_f Y$ maps Y homeomorphically onto $q(Y)$ (even if A is not closed in X).

EXERCISE 1.114. Show that if $f : A \to f(A) \subset Y$ is a quotient map then $q : X + Y \to X \cup_f Y$ restricts to a quotient map $q|_X : X \to q(X)$. Conclude that if $f : A \to f(A)$ is a homeomorphism then also $q|_X : X \to q(X)$ is a homeomorphism.

EXERCISE 1.115. Consider a topological sum $X = \sum_{i\in I} X_i$ of topological spaces X_i. Show that the map:

$$X \times [0,1] \ni \big((x_i)_{i\in I}, t\big) \longmapsto \big((x_i,t)\big)_{i\in I} \in \sum_{i\in I} \big(X_i \times [0,1]\big)$$

induces a homeomorphism from the cone C_X to the topological sum $\sum_{i\in I} C_{X_i}$.

EXERCISE 1.116. Let X, Y be topological spaces and $f : X \to Y$ a continuous map. Consider the commutative diagram:

$$\begin{tikzcd}
 & \big(X \times [0,1]\big) + Y \arrow[dl] \arrow[dr] & \\
C_X + Y \arrow[d] & & M_f \arrow[d] \\
C_X \cup_f Y \arrow[rr, dotted, "h"'] & & C_f
\end{tikzcd}$$

where the unlabelled arrows are the canonical quotient maps. Show that the dotted arrow defines a homeomorphism from the attachment space $C_X \cup_f Y$ to the mapping cone C_f that fixes Y.

EXERCISE 1.117. Generalize Lemma 1.14.21 as follows. Consider a family $(\nu_i)_{i\in I}$ of positive integers and let $f : \sum_{i\in I} S^{\nu_i - 1} \to Y$ be a continuous map defined in the topological sum of the spheres S^{ν_i-1}, taking values in a topological space Y. Show that the mapping cone C_f is homeomorphic to the attachment space $\sum_{i\in I} \overline{\mathrm{B}}^{\nu_i} \cup_f Y$ by a homeomorphism that fixes Y (*hint*: use Exercise 1.115).

CW-complexes.

EXERCISE 1.118. If X is a CW-complex and $f : \overline{\mathrm{B}}^p \to X$ is a characteristic map for an open p-cell $e \in \mathfrak{E}$ then f maps the unit sphere S^{p-1} onto $\dot{e}$ (*hint*: if f maps some $x \in S^{p-1}$ to a point of e then $f(x) = f(y)$ for some $y \in \mathrm{B}^p$. Choose a sequence $(x_n)_{n\ge1}$ in B^p with $x_n \to x$; conclude that $f(x_n) \to f(y)$ and obtain a contradiction from the continuity of $\big(f|_{\mathrm{B}^p}\big)^{-1} : e \to \mathrm{B}^p$).

EXERCISE 1.119. For $p \ge 1$, construct explicitly a continuous surjective map $q : \overline{\mathrm{B}}^p \to S^p$ that is constant on the sphere $S^{p-1} \subset \overline{\mathrm{B}}^p$ and that maps the open ball

B^p homeomorphically onto $S^p \setminus q(S^{p-1})$ (*hint*: take $q|_{\mathrm{B}^p}$ to be the composition of the map $\mathrm{B}^p \ni x \mapsto \frac{x}{1-\|x\|} \in \mathbb{R}^p$ with the inverse of the stereographic projection $S^p \setminus \{\mathfrak{n}\} \to \mathbb{R}^p$).

EXERCISE 1.120. Show that if X is a CW-complex having dimension $n < +\infty$ then every open n-cell $e \in \mathfrak{E}$ is an open subset of X (*hint*: check that $e \cap \overline{e'} = \emptyset$ for all $e' \in \mathfrak{E}$ with $e' \neq e$).

EXERCISE 1.121. Let X be a CW-complex and let $F \subset X$ be a subset such that $F \cap e$ has at most one point for every $e \in \mathfrak{E}$. Show that F is closed and discrete (*hint*: use the Weak-topology property of X to conclude that every subset of F is closed in X).

EXERCISE 1.122. Let X be a CW-complex and let $Y \subset X$ be a subset. Show that the following conditions are equivalent:

- Y is a subcomplex of X;
- for every open cell e of X, if $e \cap Y \neq \emptyset$ then $\overline{e} \subset Y$.

EXERCISE 1.123. Let X be a CW-complex and let $K \subset X$ be a compact subset. Show that K intercepts at most a finite number of open cells of X. Conclude that K is contained in a finite subcomplex of X (*hint*: let $F \subset K$ contain precisely one point from each open cell of X that intercepts K. Use the result of Exercise 1.121 to conclude that F is compact, discrete and hence finite).

EXERCISE 1.124. Show that the polyhedron of a simplicial complex K can be given the structure of a CW-complex whose open cells are the interiors of the affine simplexes belonging to K. Conclude that also every triangulable space can be made into a CW-complex.

EXERCISE 1.125. Given a topological space X (respectively, a pair (X, A) of topological spaces), we define the *p-th Betti number* of X (respectively, of (X, A)) to be the Betti number of the abelian group $H_p(X)$ (respectively, the Betti number of the abelian group $H_p(X, A)$); we denote by $\beta_p(X)$ the p-th Betti number of X (respectively, by $\beta_p(X, A)$ the p-th Betti number of (X, A)). If $H_p(X)$ is finitely generated for every p and $H_p(X) = 0$ for p sufficiently large, we define the *Euler characteristic* $\chi(X)$ of X to be the integer number:

$$\chi(X) = \sum_{p \in \mathbb{Z}} (-1)^p \beta_p(X);$$

similarly, one can define the Euler characteristic $\chi(X, A)$ of a pair (X, A) provided that $H_p(X, A)$ is finitely generated for every p and zero for p sufficiently large.

Compute the Euler characteristic of the sphere S^n, of the torus $\mathbb{T} = S^1 \times S^1$ and of the Klein bottle (see Exercise 1.60).

Homology of CW-complexes.

EXERCISE 1.126. Show that equalities (1.16.12) (or, equivalently, (1.16.13)) imply (1.16.11), (1.16.10) and (1.16.9). More precisely, let $(\beta_p)_{p \geq 0}$ and $(\kappa_p)_{p \geq 0}$ be sequences in $\mathbb{N} \cup \{+\infty\}$.

- Assume the existence of a sequence $(q_p)_{p\geq 0}$ in $\mathbb{N}\cup\{+\infty\}$ such that:

$$\kappa_0 = \beta_0 + q_0, \qquad \kappa_p = \beta_p + q_p + q_{p-1}, \quad p \geq 1. \tag{1.17.10}$$

  Show that, if $\kappa_p < +\infty$ for $p = 0, \ldots, k$ then:

$$\beta_k - \beta_{k-1} + \cdots + (-1)^k\beta_0 \leq \kappa_k - \kappa_{k-1} + \cdots + (-1)^k\kappa_0. \tag{1.17.11}$$

- Assume that for some $r \geq 0$ we have that $\kappa_p < +\infty$ for $p = 0, \ldots, r$ and that (1.17.11) is satisfied for $k = 0, \ldots, r$. Show that there exists (finite) natural numbers q_0, $\ldots$, q_r such that (1.17.10) is satisfied for $p \leq r$. Conclude that $\beta_p \leq \kappa_p$ for $p = 0, \ldots, r$.
- Assume that $\kappa_p < +\infty$ for all p and that $\kappa_p = 0$ for p sufficiently large. Assuming that (1.17.11) is satisfied for all k show that:

$$\sum_{p=0}^{+\infty}(-1)^p\beta_p = \sum_{p=0}^{+\infty}(-1)^p\kappa_p.$$

CHAPTER 2

Morse Theory on Compact Manifolds

2.1. Critical Points and Morse Functions

If $f : U \subset \mathbb{R}^n \to \mathbb{R}$ is a smooth map on an open subset $U \subset \mathbb{R}^n$ then the *Hessian* of f at a point x is the symmetric bilinear map $\mathrm{Hess} f_x : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ that is canonically identified with the second order differential $\mathrm{d}(\mathrm{d}f)_x : \mathbb{R}^n \to \mathbb{R}^{n*}$. If we replace U by an arbitrary differentiable manifold M, then one cannot give a canonical definition for the Hessian of f at an arbitrary point $x \in M$; namely, the Hessian of a function in an open subset of $\mathbb{R}^n$ does not transform correctly with respect to a change of coordinates (see Exercise 2.23). However, it is indeed possible to have a well defined notion of Hessian of f *at the critical points*; recall that for a real valued function $f : M \to \mathbb{R}$, a critical point $x \in M$ is simply a point with $\mathrm{d}f(x) = 0$. We set:

$$\mathrm{Crit}_f = \big\{ x \in M : \mathrm{d}f(x) = 0 \big\},$$
$$\mathrm{Crit}_f(a) = \mathrm{Crit}_f \cap f^{-1}(a), \quad a \in \mathbb{R};$$

obviously, Crit_f and $\mathrm{Crit}_f(a)$ are closed subsets of M and the set of regular values of f is equal to $\mathbb{R} \setminus f(\mathrm{Crit}_f)$. As we have already observed, the set of regular values is open if f is proper (this happens, for instance, if M is compact).

There are several equivalent ways of defining the Hessian of a function $f : M \to \mathbb{R}$ at a critical point $x \in M$. We give the following:

2.1.1. DEFINITION. If $f : M \to \mathbb{R}$ is a smooth function and $x \in M$ is a critical point then the *Hessian* of f at the point x is the symmetric bilinear form $\mathrm{Hess} f_x : T_xM \times T_xM \to \mathbb{R}$ defined by:

$$\mathrm{Hess} f_x(v, w) = v\big(W(f)\big),$$

where W is an arbitrary smooth vector field around $x \in M$ with $W(x) = w$.

The fact that $v\big(W(f)\big)$ is symmetric and independent of the extension W of w follows directly from the fact that:

$$v\big(W(f)\big) - w\big(V(f)\big) = [V, W]_x(f) = \mathrm{d}f_x\big([V, W]\big) = 0,$$

for every smooth vector fields V, W around $x \in M$ with $V(x) = v$, $W(x) = w$. For other equivalent definitions of the Hessian of a function at a critical point see Exercise 2.24. In particular, we observe that the above definition of Hessian when written in local coordinates gives the usual Hessian of functions in open subsets of $\mathbb{R}^n$.

Obviously, the local maxima and the local minima of $f : M \to \mathbb{R}$ are critical points. Using the Taylor polynomial of order 2 of f in local coordinates around a critical point $x \in M$, it is easy to see that f increases along the directions $v \in T_xM$ with $\mathrm{Hess} f_x(v, v) > 0$ and that f decreases in the directions v with $\mathrm{Hess} f_x(v, v) < 0$. Moreover, if $\mathrm{Hess} f_x$ is positive definite then x is a local minimum of f and if $\mathrm{Hess} f_x$ is negative definite then x is a local maximum of f. If there exists directions $v \in T_xM$ with $\mathrm{Hess} f_x(v, v) > 0$ and directions $v \in T_xM$ with $\mathrm{Hess} f_x(v, v) < 0$ then x is called a *saddle point* of f; obviously a saddle point is neither a local minimum nor a local maximum.

Before proceeding with the development of Morse theory, we need to recall a few things from linear algebra.

2.1.2. DEFINITION. Let V be a real (possibly infinite-dimensional) vector space and $B : V \times V \to \mathbb{R}$ a symmetric bilinear form. The *index* of B, denoted $n_-(B)$, is the (possibly infinite) natural number defined by:

$$n_-(B) = \sup\big\{\dim(W) : W \text{ subspace of } V,\ B|_{W\times W} \text{ negative definite}\big\}.$$

The *co-index* of B, denoted $n_+(B)$, is defined by:

$$n_+(B) = n_-(-B).$$

The *degeneracy* of B, denoted $\mathrm{dgn}(B)$, is defined as the (possibly infinite) dimension of the kernel of the map $V \ni v \mapsto B(v, \cdot) \in V^*$. If $\mathrm{dgn}(B)$ is equal to zero we say that B is *nondegenerate*.

The following is a very simple result of linear algebra.

2.1.3. THEOREM (Sylvester's theorem of inertia). *Let V be a finite-dimensional real vector space and $B : V \times V \to \mathbb{R}$ a symmetric bilinear form. Then there exists a basis of V on which B is represented by a diagonal matrix of the form:*

$$B \sim \begin{pmatrix} \mathrm{I}_p & 0 & 0 \\ 0 & -\mathrm{I}_q & 0 \\ 0 & 0 & 0_r \end{pmatrix}$$

where I_α denotes the $\alpha \times \alpha$ identity matrix, 0_α denotes the $\alpha \times \alpha$ zero matrix. Moreover, if B is represented by a matrix in the form above in some basis of V then $p = n_+(B)$, $q = n_-(B)$ and $r = \mathrm{dgn}(B)$. □

We are now ready to give the following:

2.1.4. DEFINITION. A critical point $x \in M$ of a smooth map $f : M \to \mathbb{R}$ is called *nondegenerate* if the symmetric bilinear form $\mathrm{Hess} f_x : T_xM \times T_xM \to \mathbb{R}$ is nondegenerate. The *Morse index* of a critical point $x \in M$ is defined as the index of the symmetric bilinear form $\mathrm{Hess} f_x$. By a *Morse function* $f : M \to \mathbb{R}$ we mean a smooth map all of whose critical points are nondegenerate.

It follows easily from the Taylor polynomial expansion of f that nondegenerate critical points of Morse index zero (resp., of Morse index equal to $\dim(M)$) are strict local minima (resp., strict local maxima) of f. A critical point that is neither a minimal nor a local maximum is called a *saddle point*. Observe that a nondegenerate critical point is a saddle point if and only if its Morse index is positive and less than $\dim(M)$.

Around a nondegenerate critical point, a function can be locally identified with a quadratic form in a suitable coordinate chart. This the the content of the following:

2.1.5. THEOREM (Morse lemma). *Let $f : M \to \mathbb{R}$ be a smooth map on an arbitrary manifold M and let $p \in M$ be a nondegenerate critical point of f. There exists a diffeomorphism $\varphi : U \to \widetilde{U}$ from an open neighborhood U of p in M to an open neighborhood of the origin in T_pM such that $\varphi(p) = 0$ and $f \circ \varphi^{-1} - f(p)$ equals (the restriction to $\widetilde{U}$ of) the quadratic form $v \mapsto \frac{1}{2}\mathrm{Hess} f_p(v, v)$.*

PROOF. Let $\psi : V \to \widetilde{V}$ be an arbitrary diffeomorphism from an open neighborhood V of p in M to an open neighborhood $\widetilde{V}$ of the origin in T_pM; we may choose ψ with $\psi(p) = 0$. Set $\tilde{f} = f \circ \psi^{-1} : \widetilde{V} \to \mathbb{R}$, so that $0 \in T_pM$ is a critical point of $\tilde{f}$ and $\mathrm{Hess}\tilde{f}_0 = \mathrm{Hess} f_x$. We will determine a diffeomorphism α between open neighborhoods of the origin in T_pM with $\alpha(0) = 0$ and $\tilde{f} \circ \alpha = \mathrm{Hess}\tilde{f}_0$ around the origin.

Since $\mathrm{d}\tilde{f}(0) = 0$, the first order Taylor expansion of $\tilde{f}$ around 0 with remainder in integral form gives:

$$\tilde{f}(v) = \tilde{f}(0) + \int_0^1 (1 - t)\mathrm{Hess}\tilde{f}_{tv}(v, v)\,\mathrm{d}t,$$

for $v \in T_pM$ in a neighborhood of 0. We may represent the symmetric bilinear form $\int_0^1 (1 - t)\mathrm{Hess}\tilde{f}_{tv}\,\mathrm{d}t$ with respect to some arbitrarily fixed inner product $\langle \cdot, \cdot \rangle$ in T_pM, obtaining a symmetric linear endomorphism $A_v \in \mathrm{Lin}(T_pM)$ such that:

$$\tilde{f}(v) = \tilde{f}(0) + \big\langle A_v(v), v \big\rangle, \tag{2.1.1}$$

for $v \in T_pM$ in a neighborhood of 0; obviously $v \mapsto A_v$ is a smooth $\mathrm{Lin}(T_pM)$-valued map. The nondegeneracy of $\mathrm{Hess}\tilde{f}_0$ means that the linear map $A_0 : T_pM \to T_pM$ is an isomorphism; since $v \mapsto A_0^{-1}A_v$ takes values in a neighborhood of the identity of T_pM for v near zero, we may define a smooth map

$$v \longmapsto B_v \in \mathrm{Lin}(T_pM)$$

with $B_0 = \mathrm{Id}$ and $B_v^2 = A_0^{-1}A_v$ for v near zero (see Exercise 2.1). Thus:

$$A_v = A_0 B_v^2. \tag{2.1.2}$$

Since A_0 and A_v are symmetric, we may take the transpose with respect to $\langle \cdot, \cdot \rangle$ in both sides of the equality (2.1.2) obtaining $A_v = (B_v^*)^2 A_0$ and thus $B_v^2 = (A_0^{-1}B_v^*A_0)^2$. By taking v in a sufficiently small neighborhood of zero, we have both B_v and $A_0^{-1}B_v^*A_0$ in a neighborhood of the identity in $\mathrm{Lin}(T_pM)$ where the square function is injective; then:

$$A_0 B_v = B_v^* A_0, \tag{2.1.3}$$

for v sufficiently close to zero. From (2.1.1), (2.1.2) and (2.1.3) we obtain:

$$\tilde{f}(v) = \tilde{f}(0) + \big\langle B_v^* A_0 B_v(v), v \big\rangle = f(p) + \mathrm{Hess} f_p\big(B_v(v), B_v(v)\big),$$

for v sufficiently close to zero. Once we show that the map $v \mapsto B_v(v)$ is a diffeomorphism in a neighborhood of the origin, the conclusion will follow from the above equality by taking α to be the inverse of such diffeomorphism. The fact that $v \mapsto B_v(v)$ is a diffeomorphism in a neighborhood of the origin is easily stablished by the inverse function theorem, observing that the differential of such map at 0 equals $B_0 = \mathrm{Id}$. □

Observe that the origin is the unique critical point of a nondegenerate quadratic form in a vector space. We thus obtain the following immediate corollary.

2.1.6. COROLLARY. *The nondegenerate critical points of a smooth map $f : M \to \mathbb{R}$ are isolated in Crit_f. In particular, if f is a Morse function then Crit_f is discrete.* □

As a matter of fact, the fact that nondegenerate critical points are isolated is a rather elementary fact that follows from the inverse function theorem (see Exercise 2.25).

2.1.7. REMARK. It can be proven that every differentiable manifold M admits a Morse function. Actually, one can show that Morse functions are dense in the space of continuous maps $: M \to \mathbb{R}$ with respect to the topology of uniform convergence, i.e., every continuous map is the uniform limit of Morse functions (see Exercise 2.26).

We will apply the Morse Lemma in order to study the change of the topology of the sublevels of a Morse function when passing a critical value. The precise statement (and most of all the proof) of such result is quite involved and will be given in Section 2.5. For now we will just give an example of how the Morse Lemma can be used to study the topology of the levels f^a when a is slightly bigger then the minimum of f.

2.1.8. PROPOSITION. *Let M be a compact differentiable n-dimensional manifold and $f : M \to \mathbb{R}$ a smooth function whose minimum points are non degenerate critical points. Then there exists $\varepsilon > 0$ such that for $a \in \left]\min f, \min f + \varepsilon\right[$ the sublevel f^a is homeomorphic to a topological sum of r closed n-balls, where r is the number of minimum points of f.*

PROOF. let $x_1, \dots, x_r \in M$ be the minimum points of f and let $m \in \mathbb{R}$ be the minimum value of f. By the Morse lemma, for every $i = 1, \dots, r$, we can find an open neighborhood U_i of x_i in M and a diffeomorphism $\varphi_i : U_i \to \widetilde{U}_i$ onto an open neighborhood $\widetilde{U}_i$ of the origin in $T_{x_i}M$ such that $f \circ \varphi_i^{-1}(v) = m + \frac{1}{2}\mathrm{Hess} f_{x_i}(v, v)$ for all $v \in \widetilde{U}_i$. We can assume that the open sets U_i are disjoint. Since each x_i is a nondegenerate minimum point of f, the symmetric bilinear form $\mathrm{Hess} f_{x_i}$ in T_{x_i} is a positive definite inner product and hence there exists $\varepsilon_i > 0$ such that $\frac{1}{2}\mathrm{Hess} f_{x_i}(v, v) < \varepsilon_i$ implies $v \in \widetilde{U}_i$. Choose $\varepsilon > 0$ less than the minimum of the

ε_i's and less than the minimum of the positive function $f - m$ in the compact set $M \setminus \bigcup_{i=1}^r U_i$. We have then:

$$f^{m+\varepsilon} = \bigcup_{i=1}^{r} \big(f^{m+\varepsilon} \cap U_i\big) = \bigcup_{i=1}^{r} \varphi_i^{-1}(B_i),$$

where $B_i \subset \widetilde{U}_i$ denotes the closed ball:

$$B_i = \big\{v \in T_{x_i}M : \tfrac{1}{2}\mathrm{Hess} f_{x_i}(v,v) \le \varepsilon\big\}.$$

This concludes the proof. $\square$

2.2. An Instructive Example: the Height Function on the Torus

Given a Morse function $f : M \to \mathbb{R}$ on a compact manifold M, then using the critical points of f one is able to determine information on the homotopy type of M. For every $c \in \mathbb{R}$, we define the *closed c-sublevel* of f by:

$$f^c = \big\{x \in M : f(x) \le c\big\} = f^{-1}\big(-\,]-\infty, c]\big);$$

when c is a regular value for f then f^c is a smooth submanifold with boundary in M whose boundary is the level surface $f^{-1}(c)$. When c is a critical level, the level surface $f^{-1}(c)$ may become singular. Usually, it is better to picture the situation in the following way: we identify M with the graph of f in $M \times \mathbb{R}$ and then f is identified with the "height function" $M \times \mathbb{R} \ni (m,t) \mapsto t \in \mathbb{R}$. With such identification, the critical points of f become the valleys, passes and mountain summits of the graph of f. The basic idea is that the topological type of the sublevel f^c does not change when c runs through a non critical interval $[a,b]$, i.e., an interval that does not contain critical values. This can be shown by considering the flow of minus the gradient field ∇f of f (with respect to some arbitrary Riemannian metric). This flow gives the direction of "steepest descent" in the graph of f and can be used to deform the sublevel f^b onto the sublevel f^a. Clearly, the presence of a critical value on the interval $[a,b]$ is an obstruction to such argument, because some lines of flow of $-\nabla f$ do not go all the way from the level b to the level a. We will show indeed that when c passes through a critical value, the topological type (and also the homotopy type) of f^c changes, according to the number of critical points in $f^{-1}(c)$ and their Morse indexes.

Before we get into the details of the theory, it will be useful to describe a very simple example, which served as a motivation in many classical textbooks on the subject (see for instance [**98, 119**]).

Let us consider a torus $M = \mathbb{T}$ in $\mathbb{R}^3$ tangent to a horizontal plane as in Figure 1; in the language of [**119**], this is described as a "tire standing in a ready to roll position". Define $f : M \to \mathbb{R}$ to be the function that assigns to each point of M its height above the "floor". By an elementary analysis of the picture, one sees that the function f has exactly four critical points that are all nondegenerate: P_1 is a global minimum point, P_2 and P_3 are saddle points (having Morse index equal to one), P_4 is a global maximum. Set $c_i = f(P_i)$, $i = 1,2,3,4$; in Figures 2—6 we give a picture of the closed sublevels f^{a_1}, f^{c_1}, f^{a_2}, f^{c_2}, f^{a_3}, f^{c_3}, f^{a_4}, with $c_i < a_i < c_{i+1}$, $i = 1,2,3$.

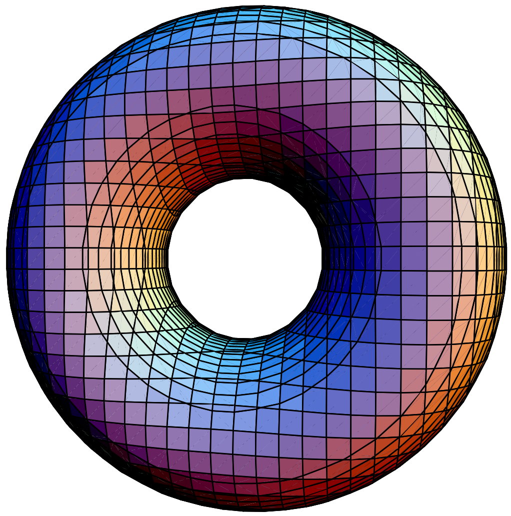

FIGURE 1. A “tire in a ready to roll position”

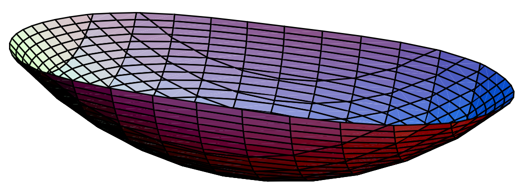

FIGURE 2. The sublevel f^{a_1}

- The closed sublevel f^{a_1} is homeomorphic to a closed disc, i.e., f^{a_1} is a closed 2-cell; observe that the Morse index of the critical point P_1 is precisely 2.
- The closed sublevel f^{c_2} is no longer homeomorphic to f^{a_1}, but it is a strong deformation retract of f^{a_2} which is homeomorphic to f^{a_1} with a handle $[-1,1]\times[-1,1]$ attached along $[-1,1]\times\{-1,1\}$. Observe that P_2 is a critical point of Morse index 1.
- The closed sublevel f^{c_3} is no longer homeomorphic to f^{a_2}, but it is a strong deformation retract of f^{a_3} which is homeomorphic to f^{a_2} with a

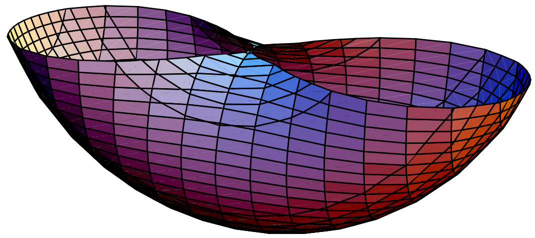

FIGURE 3. The sublevel f^{c_1}

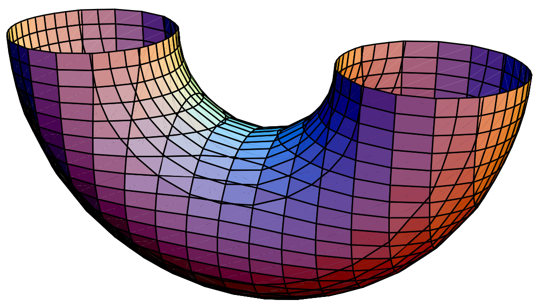

FIGURE 4. The sublevel f^{a_2}

handle $[-1,1] \times [-1,1]$ attached along $[-1,1] \times \{-1,1\}$. Observe that P_3 is a critical point of Morse index 1.

- The closed sublevel $f^{c_4} = \mathbb{T}$ is no longer homeomorphic to f^{a_3}, but it is homeomorphic to f^{a_3} with a closed 2-cell attached along its boundary. Observe that P_4 is a critical point of Morse index 2.

In this chapter we will show that the sublevels of a general Morse function on a compact manifold satisfy relations that are similar to the ones described for the height function on the torus.

2.3. Dynamics of the Gradient Flow

In this section (M,g) denotes a compact Riemannian manifold, $f : M \to \mathbb{R}$ a smooth map and $F : \mathbb{R} \times M \to M$ the flow of $-\nabla f$, i.e., $F(0,x) = x$ and:

$$\frac{\mathrm{d}}{\mathrm{d}t} F(t,x) = -\nabla f\big(F(t,x)\big),$$

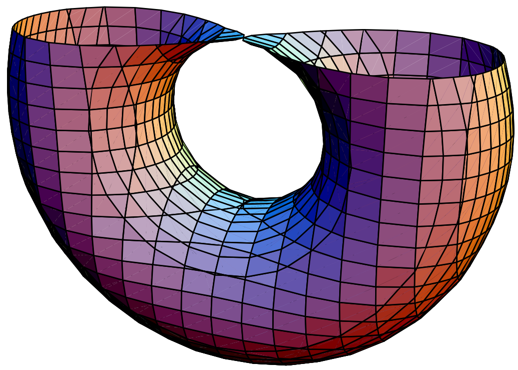

FIGURE 5. The sublevel f^{c_3}

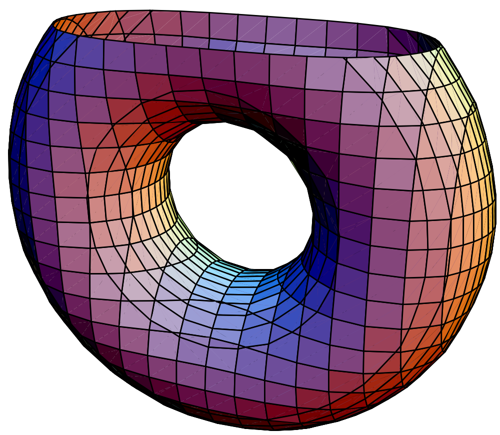

FIGURE 6. The sublevel f^{a_3}

for all $t \in \mathbb{R}$, $x \in M$. We also use the short notation:

$$t \cdot x = F(t, x), \tag{2.3.1}$$

for all $t \in \mathbb{R}$, $x \in M$; then (2.3.1) defines an action of the additive group $\mathbb{R}$ in M.

Obviously, if $x \in M$ is a critical point of f then $F(t, x) = x$ for all $t \in \mathbb{R}$; if x is not a critical point of f then:

$$\frac{\mathrm{d}}{\mathrm{d}t} f(t \cdot x) = -\mathrm{d}f_{t \cdot x}\big(\nabla f(t \cdot x)\big) = -g\big(\nabla f(t \cdot x), \nabla f(t \cdot x)\big) < 0, \quad t \in \mathbb{R},$$

so that $t \mapsto f(t \cdot x)$ is a strictly decreasing function on $\mathbb{R}$.

2.3.1. LEMMA. *Given an isolated critical point $x \in M$ of f and a neighborhood $U \subset M$ of x then there exists an open neighborhood $V \subset M$ of x contained in U and $\varepsilon > 0$ such that $f(t \cdot x) \geq c - \varepsilon$ implies $t \cdot x \in U$ for all $t \geq 0$.*

The *ω-limit* (resp., the *α-limit*) of a flow line $x \mapsto t \cdot x$ is the set of points $y \in M$ for which there exists a sequence $(t_n)_{n \geq 1}$ of real numbers with $\lim_{n \to +\infty} t_n = +\infty$ (resp., $\lim_{n \to +\infty} t_n = -\infty$) and $\lim_{n \to +\infty} t_n \cdot x = y$.

In what follows we prove a series of lemmas concerning the asymptotic behavior of the flow lines of $-\nabla f$. For simplicity, we only state the results concerning limits as $t \to +\infty$; by replacing f with $-f$ one can obviously obtain analogous statements for the limits as $t \to -\infty$.

2.3.2. LEMMA. *The ω-limit of any flow line $x \mapsto t \cdot x$ consists only of critical points of f.*

PROOF. Assume by contradiction that there exists a noncritical point $y_0 \in M$ belonging to the ω-limit of $x \mapsto t \cdot x$. Set $c = f(y_0)$. Then f is a submersion near y_0 and thus we can find an open neighborhood $V \subset M$ of y_0 such that $S = V \cap f^{-1}(c)$ is a submanifold of M (orthogonal to ∇f). By the inverse function theorem, we can find an open subset S_0 in S containing y_0 and $\varepsilon > 0$ such that the map:

$$S_0 \times \,]-\varepsilon, \varepsilon[\, \ni (y, s) \longmapsto s \cdot y \in M$$

is a diffeomorphism onto an open neighborhood $U \subset M$ of y_0. Since y_0 is in the ω-limit of $t \mapsto t \cdot x$, we can find $t_1, t_2 > 0$, with $t_1 \cdot x \in U$, $t_2 \cdot x \in U$ and $t_2 \geq t_1 + 2\varepsilon$. We can now find $y_1, y_2 \in S_0$, $s_1, s_2 \in \,]-\varepsilon, \varepsilon[$ with $t_1 \cdot x = s_1 \cdot y_1$ and $t_2 \cdot x = s_2 \cdot y_2$. This implies:

$$(t_1 - s_1) \cdot x = y_1 \in S_0 \subset f^{-1}(c), \quad (t_2 - s_2) \cdot x = y_2 \in S_0 \subset f^{-1}(c);$$

since $t \mapsto f(t \cdot x)$ is strictly increasing, we have $t_1 - s_1 = t_2 - s_2$. Hence:

$$|t_1 - t_2| = |s_1 - s_2| < 2\varepsilon,$$

which contradicts $t_2 \geq t_1 + 2\varepsilon$. $\square$

2.3.3. LEMMA. *Let y_0 belong to the ω-limit of a flow line $t \mapsto t \cdot x$ (so that y_0 is a critical point of f, by Lemma 2.3.2). If y_0 is an isolated critical point of f then $\lim_{t \to +\infty} t \cdot x = y_0$.*

PROOF. Let $U \subset M$ be a neighborhood of y_0; let us show that $t \cdot x \in U$ for t sufficiently large. Choose a sequence $(t_n)_{n \geq 1}$ with $\lim_{n \to +\infty} t_n = +\infty$ and $\lim_{n \to +\infty} t_n \cdot x = y_0$. Set $c = f(y_0)$. Then $\lim_{n \to +\infty} f(t_n \cdot x) = c$ and, since f is decreasing along the flow line $t \mapsto t \cdot x$, it follows that $f(t \cdot x) \geq c$ for all t. Choose V and $\varepsilon > 0$ as in the statement of Lemma 2.3.1 and $n \geq 1$ with $t_n \cdot x \in V$. Then $t \cdot x \in U$ for all $t \geq t_n$. This concludes the proof. $\square$

2.3.4. COROLLARY. *If all the critical points of f are isolated (in particular, if f is a Morse function) then each flow line of $-\nabla f$ converges to a critical point of f, i.e., given an arbitrary point $x \in M$ then the limit $\lim_{t \to +\infty} t \cdot x$ exists and it is a critical point of f.*

PROOF. The compactness of M obviously implies that the ω-limit of any flow line is nonempty. The conclusion follows. $\square$

Assuming that all critical points of f are isolated, Corollary 2.3.4 allows us to extend the flow $F : \mathbb{R} \times M \to M$ of $-\nabla f$ to $[-\infty, +\infty] \times M$ by setting:

$$F(-\infty, x) = -\infty \cdot x = \lim_{t \to -\infty} F(t, x), \quad F(+\infty, x) = +\infty \cdot x = \lim_{t \to +\infty} F(t, x),$$

for all $x \in M$. Since $\mathbb{R} \times M$ is open in $[-\infty, +\infty] \times M$, the extension of F defined above is continuous at the points of $\mathbb{R} \times M$. However, one should be very careful about the continuity of F at the points of $\{-\infty, +\infty\} \times M$ (in fact, F is *not* continuous in general at those points: see Exercise 2.30). The following weaker continuity condition holds: $F(t_n, x_n)$ tends to $F(+\infty, x)$ when (t_n, x_n) tends to $(+\infty, x)$ *provided that* $f(t_n \cdot x_n) \geq f(+\infty \cdot x)$ for all n. This is proven in the following:

2.3.5. LEMMA. *Choose $x \in M$; set $y = +\infty \cdot x$ and $c = f(y)$. The restriction of F to the set:*

$$(f \circ F)^{-1}\big([c, +\infty[\big) = \big\{ (t, z) \in [-\infty, +\infty] \times M : f(t \cdot z) \geq c \big\}$$

is continuous at the point $(+\infty, x)$.

PROOF. Let U be a neighborhood of y. We have to show that if t is sufficiently large and z is sufficiently close to x then $t \cdot z \in U$, provided that $f(t \cdot z) \geq c$. By Lemma 2.3.1, we can find an open neighborhood V of y contained in U such that the flow lines starting in V remain in U, as long as they don't go below the level c. Choose $t_0 > 0$ such that $t_0 \cdot x \in V$. By the continuity of F on $\mathbb{R} \times M$, we have $t_0 \cdot z \in V$ for z in some neighborhood of x. But then $t \cdot z \in U$ for all $t \geq t_0$ with $f(t \cdot z) \geq c$. $\square$

Given $a \in \mathbb{R}$ then each nonconstant flow line of $-\nabla f$ meets the level a at most once; it will be useful to look at the "arrival time function" defined as follows. Set:

$$D_a = \big\{ x \in M \setminus \mathrm{Crit}_f : f(t \cdot x) = a, \text{ for some } t \in [-\infty, +\infty] \big\}$$

and define $\lambda_a : D_a \to [-\infty, +\infty]$ by the equality:

$$f\big(\lambda_a(x) \cdot x\big) = a,$$

for all $x \in D_a$. We also set:

$$D = \big\{ (a, x) \in \mathbb{R} \times M : x \in D_a \big\}$$

and we define $\lambda : D \to [-\infty, +\infty]$ by:

$$\lambda(a, x) = \lambda_a(x).$$

We will now study the regularity of the map λ. We start with the points where λ is finite.

2.3.6. LEMMA. *The set $\lambda^{-1}(\mathbb{R}) \subset D$ is open in $\mathbb{R} \times M$ and the map:*

$$\lambda|_{\lambda^{-1}(\mathbb{R})} : \lambda^{-1}(\mathbb{R}) \longrightarrow \mathbb{R}$$

is smooth.

PROOF. Observe that $\lambda|_{\lambda^{-1}(\mathbb{R})}$ is the map obtained by solving the equation:

$$f(t\cdot x)-a=0,\quad x\in M\setminus \mathrm{Crit}_f,\ a,t\in\mathbb{R},$$

for t. The derivative with respect to t of the lefthand side of the equation above is $-\big\|\nabla f(t\cdot x)\big\|^2$, which is nonzero when x is noncritical. The conclusion follows from the implicit function theorem. □

Now we look at the points where λ is infinite. We will show that the map λ is continuous. In fact, we show a little more. We define an extension:

$$\overline{\lambda}:\big\{(a,x)\in\mathbb{R}\times M : x\not\in \mathrm{Crit}_f(a)\big\}\longrightarrow[-\infty,+\infty],$$

of λ by setting:

$$\overline{\lambda}(a,x)=\begin{cases}\lambda(a,x), & \text{if } (a,x)\in D,\\ +\infty, & \text{if } (a,x)\not\in D \text{ and } f(x)>a,\\ -\infty, & \text{if } (a,x)\not\in D \text{ and } f(x)<a.\end{cases}$$

Obviously the domain of $\overline{\lambda}$ is open in $\mathbb{R}\times M$. We now prove the following:

2.3.7. LEMMA. *The map $\overline{\lambda}$ (and in particular the map λ) is continuous.*

PROOF. By Lemma 2.3.6, it suffices to prove that $\overline{\lambda}$ is continuous at those points where $\overline{\lambda}$ is infinite. Let thus $(a,x)\in\mathbb{R}\times M$ be fixed with $x\not\in\mathrm{Crit}_f(a)$ and $\overline{\lambda}(a,x)=\pm\infty$. For definiteness, we assume $\overline{\lambda}(a,x)=+\infty$; the case $\overline{\lambda}(a,x)=-\infty$ is handled in a similar way. If either $(a,x)\in D$ and $\lambda(a,x)=+\infty$ or $(a,x)\not\in D$ and $f(x)>a$, we have $f(t\cdot x)>a$ for all $t\in\mathbb{R}$. Then given $t_0>0$ we have $f(t_0\cdot x)>a$ and by continuity we have $f(t_0\cdot y)>a+\varepsilon$ for some $\varepsilon>0$ and for all y in a neighborhood V of x. Thus, for $y\in V$ and $|a-b|<\varepsilon$ we have $\overline{\lambda}(b,y)>t_0$. This concludes the proof. □

In Exercises 2.31 and 2.32 it will become clear that one cannot hope to find a continuous extension of $\overline{\lambda}$ to the pairs (a,x) with $x\in\mathrm{Crit}_f(a)$.

2.3.8. LEMMA. *Let $x\in M$ be a point that is not critical and choose $a\in\mathbb{R}$. If there are no critical values of f in the open interval with endpoints $f(x)$ and a then $x\in D_a$. Moreover, if a is not a critical value of f then $\lambda_a(x)$ is finite.*

PROOF. If $f(x)=a$ there is nothing to prove. We may assume that $f(x)>a$ (the case $f(x)<a$ can be obtained simply by replacing f with $-f$). If x were not in D_a, then it would be $f(t\cdot x)>a$ for all $t\in[0,+\infty]$; then $+\infty\cdot x$ would be a critical point of f with $a<f(+\infty\cdot x)<f(x)$, contradicting our hypothesis. Thus $x\in D_a$. If $\lambda_a(x)=+\infty$ then $+\infty\cdot x$ is a critical point of f at the level a and thus $\lambda_a(x)$ is finite if a is noncritical. □

2.3.9. PROPOSITION. *Choose real numbers $a<b$ such that f has no critical values in the open interval $]a,b[$. Then $f^{-1}(a)$ is a strong deformation retract of $f^{-1}\big([a,b]\big)\setminus\mathrm{Crit}_f(b)$.*

PROOF. It follows from Lemma 2.3.8 that every point $x \notin \mathrm{Crit}_f(a)$ in the strip

$$S = f^{-1}\big([a,b]\big) \setminus \mathrm{Crit}_f(b)$$

belongs to D_a. We define a map $G : [0,+\infty] \times S \to S$ by:

$$G(t,x) = \begin{cases} F\big(\min\{t, \lambda_a(x)\}, x\big), & \text{if } x \notin \mathrm{Crit}_f(a), \\ x, & \text{if } x \in \mathrm{Crit}_f(a). \end{cases}$$

Since the restriction of F to

$$\big\{(t,x) : x \in S \setminus \mathrm{Crit}_f(a),\ t \in \big[0, \lambda_a(x)\big]\big\},$$

is continuous (Lemma 2.3.5) and so is λ_a (Lemma 2.3.7) it follows that G is continuous in $[0,+\infty] \times \big(S \setminus \mathrm{Crit}_f(a)\big)$. The continuity of G in $[0,+\infty] \times \mathrm{Crit}_f(a)$ follows easily from Lemma 2.3.1 (see also Exercise 2.33). The desired deformation retraction $H : [0,1] \times S \to S$ is now obtained from G by setting $H(t,x) = G\big(\alpha(t), x\big)$, where $\alpha : [0,1] \to [0,+\infty]$ is an increasing homeomorphism. $\square$

2.3.10. COROLLARY. *Under the assumptions of Proposition 2.3.9, the sublevel f^a is a strong deformation retract of $f^b \setminus \mathrm{Crit}_f(b)$.*

PROOF. Extend the map H given in the proof of Proposition 2.3.9 by setting $H(t,x) = x$ for all $x \in f^a$ and all t. $\square$

2.3.11. PROPOSITION (non-critical neck principle). *Choose real numbers $a < b$ such that f has no critical values in the closed interval $[a,b]$. Then for every $t_0 \in [a,b]$, there exists a homeomorphism $H : f^{-1}\big([a,b]\big) \to [a,b] \times f^{-1}(t_0)$ whose first coordinate is f, i.e., such that the diagram:*

$$\begin{tikzcd} f^{-1}\big([a,b]\big) \arrow[rr, "H"] \arrow[dr, "f"'] & & {[a,b] \times f^{-1}(t_0)} \arrow[dl, "\mathrm{pr}_1"] \\ & {[a,b]} & \end{tikzcd}$$

commutes, where pr_1 denotes the projection onto the first coordinate. Moreover, H can be chosen in such a way that $H(x) = (t_0, x)$ for all $x \in f^{-1}(t_0)$.

PROOF. Since f has no critical values on $[a,b]$, Lemma 2.3.8 implies that the set $[a,b] \times f^{-1}\big([a,b]\big)$ is contained in D and that λ takes finite values in such set. The map H can be explicitly defined by:

$$H(x) = \big(f(x), \lambda_{t_0}(x) \cdot x\big), \quad x \in f^{-1}\big([a,b]\big);$$

we exhibit a continuous inverse for H:

$$H^{-1}(c,y) = \lambda(c,y) \cdot y, \quad c \in [a,b],\ y \in f^{-1}(t_0). \qquad \square$$

2.3.12. COROLLARY. *If $[a,b] \subset \mathbb{R}$ does not contain critical points of f then the sublevel f^b is homeomorphic to f^a; moreover, for every $a_1 < a$ we can find a homeomorphism from f^b to f^a that is the identity on f^{a_1}.*

PROOF. Choose $\varepsilon > 0$ small enough so that $a - \varepsilon > a_1$ and such that the interval $[a-\varepsilon, b]$ does not contain critical values of f. Consider the unique affine increasing bijection:

$$\sigma : [a-\varepsilon, b] \longrightarrow [a-\varepsilon, a]$$

and the corresponding homeomorphism $\tilde{\sigma} = \sigma \times \mathrm{Id}$ from $[a-\varepsilon, b] \times f^{-1}(a)$ to $[a-\varepsilon, a] \times f^{-1}(a)$. By the non-critical neck principle we can find homeomorphisms

$$H_1 : f^{-1}\big([a-\varepsilon, b]\big) \longrightarrow [a-\varepsilon, b] \times f^{-1}(a-\varepsilon),$$
$$H_2 : f^{-1}\big([a-\varepsilon, a]\big) \longrightarrow [a-\varepsilon, a] \times f^{-1}(a-\varepsilon),$$

both with first coordinate equal to f and such that $H_1(x) = H_2(x) = (a-\varepsilon, x)$ for all $x \in f^{-1}(a-\varepsilon)$. The composition $H_2^{-1} \circ \tilde{\sigma} \circ H_1$ gives a homeomorphism from $f^{-1}\big([a-\varepsilon, b]\big)$ to $f^{-1}\big([a-\varepsilon, a]\big)$ that is the identity on $f^{-1}(a-\varepsilon)$. The conclusion is obtained by extending $H_2^{-1} \circ \tilde{\sigma} \circ H_1$ to be the identity on $f^{a+\varepsilon}$. □

Using Corollary 2.3.12 we can now prove one of the most classical results of Morse theory.

2.3.13. THEOREM (Reeb). *Let M be a compact differentiable manifold. If M admits a Morse function having precisely two critical points then M is homeomorphic to a sphere.*

PROOF. Let $f : M \to \mathbb{R}$ be a Morse function having precisely two critical points. Since M is compact, one of then is the global minimum and the other is the global maximum. Write $c_0 = \min f$, $c_1 = \max f$ and choose any a in the open interval $]c_0, c_1[$. From Proposition 2.1.8 and Corollary 2.3.12 we conclude that the sublevel f^a is homeomorphic to the closed ball $\overline{\mathrm{B}}^n$, where $n = \dim(M)$. Since f^a is a manifold with boundary whose boundary is $f^{-1}(a)$ (see Exercise 2.11), a homeomorphism $h : f^a \to \overline{\mathrm{B}}^n$ takes $f^{-1}(a)$ to S^{n-1}, which is the boundary of $\overline{\mathrm{B}}^n$ (see Exercise 1.63). By a similar argument we get a homeomorphism $\tilde{h}$ from $(-f)^{(-a)} = f^{-1}\big([a, c_1]\big)$ to $\overline{\mathrm{B}}^n$; such homeomorphism also maps $f^{-1}(a)$ to S^{n-1}. Now consider the homeomorphism $\alpha : S^{n-1} \to S^{n-1}$ given by the "transition map" $\tilde{h} \circ \big(h|_{f^{-1}(a)}\big)^{-1}$. We now obtain that M is homeomorphic to the attachment space $\overline{\mathrm{B}}^n \cup_\alpha \overline{\mathrm{B}}^n$ (see Lemma 1.14.12) and such attachment space is homeomorphic to the sphere S^n (see Exercise 1.109). □

2.4. The Morse Relations

2.4.1. DEFINITION. If $x \in M$ is an isolated critical point of $f : M \to \mathbb{R}$ then the *critical numbers* of f at x with respect to a field $\mathbb{K}$ are defined by:

$$\mu_k(x, f; \mathbb{K}) = \beta_k\big(f^c, f^c \setminus \{x\}; \mathbb{K}\big) = \dim_{\mathbb{K}}\big(H_k\big(f^c, f^c \setminus \{x\}\big)\big),$$

where $c = f(x)$.

Recall that $H_k\big(f^c, f^c \setminus \{x\}\big)$ is the local homology group of the space f^c at the point x; thus, for any neighborhood V of x in M we have an isomorphism:

$$H_k\big(V \cap f^c, (V \cap f^c) \setminus \{x\}\big) \cong H_k\big(f^c, f^c \setminus \{x\}\big)$$

induced by inclusion.

2.4.2. LEMMA. *Given reals numbers $a < b$ such that there exists at most one critical value of f in the interval $]a, b]$ then, for any field $\mathbb{K}$, we have*[1]*:*

$$\beta_k(f^b, f^a; \mathbb{K}) = \dim_{\mathbb{K}}\big(H_k(f^b, f^a; \mathbb{K})\big) = \sum_{\substack{x \in \mathrm{Crit}_f \\ a < f(x) \le b}} \mu_k(x, f; \mathbb{K}), \tag{2.4.1}$$

for all $k \ge 0$.

PROOF. By Corollary 2.3.10, if there are no critical values of f in $]a, b]$ then f^a is a strong deformation retract of f^b and thus $H_k(f^b, f^a; \mathbb{K}) \cong H_k(f^a, f^a; \mathbb{K}) = 0$ and hence both sides of (2.4.1) vanish. Assume now that $c \in\,]a, b]$ is the unique critical value of f in $]a, b]$. By Corollary 2.3.10, f^c is a strong deformation retract of f^b and f^a is a strong deformation retract of $f^c \setminus \mathrm{Crit}_f(c)$; thus:

$$H_k(f^b, f^a; \mathbb{K}) \cong H_k(f^c, f^a; \mathbb{K}) \cong H_k\big(f^c, f^c \setminus \mathrm{Crit}_f(c); \mathbb{K}\big).$$

Write $\mathrm{Crit}_f(c) = \{x_1, \ldots, x_r\}$ and choose disjoint open sets $(U_i)_{i=1}^r$ in M such that $x_i \in U_i$, $i = 1, \ldots, r$; set $U = \bigcup_{i=1}^r U_i$. Since $\mathrm{Crit}_f(c)$ is a closed set contained in the open subset $U \cap f^c$ relatively to f^c, by excision, we have:

$$H_k\big(f^c, f^c \setminus \mathrm{Crit}_f(c); \mathbb{K}\big) \cong H_k\big(U \cap f^c, (U \cap f^c) \setminus \mathrm{Crit}_f(c); \mathbb{K}\big).$$

Moreover:

$$\begin{aligned} H_k\big(U \cap f^c, (U \cap f^c) \setminus \mathrm{Crit}_f(c); \mathbb{K}\big) &\cong \bigoplus_{i=1}^r H_k\big(U_i \cap f^c, (U_i \cap f^c) \setminus \{x_i\}; \mathbb{K}\big) \\ &\cong \bigoplus_{i=1}^r H_k\big(f^c, f^c \setminus \{x_i\}; \mathbb{K}\big). \end{aligned}$$

The conclusion follows. □

2.4.3. THEOREM. *Let $f : M \to \mathbb{R}$ be a smooth function on a compact manifold M having only a finite number of critical points. Then, for any field $\mathbb{K}$, the sequences given by:*

$$\mu_k = \sum_{x \in \mathrm{Crit}_f} \mu_k(x, f; \mathbb{K}), \quad \beta_k = \beta_k(M; \mathbb{K}),$$

satisfy the Morse relations.

PROOF. Let $a_1 < a_2 < \cdots < a_r$ be the critical values of f and choose arbitrarily $a_0 < a_1$. Observe that, since M is compact, f has a global minimum and a global maximum and therefore a_1 must be the minimum value of f and a_r must be the maximum value of f. We define a filtration $(X_n)_{n \ge 0}$ in M by setting $X_n = f^{a_n}$ for $n = 0, \ldots, r$ and $X_n = M$ for $n > r$; observe that $X_0 = \emptyset$ and $X_n = M$ for all $n \ge r$. Obviously the filtration $(X_n)_{n \ge 0}$ satisfies the hypothesis

[1] The sum in (2.4.1) is understood to be zero if f has no critical values in $]a, b]$.

of Proposition **??**. To conclude the proof we simply apply Lemma 2.4.2 to compute as follows:

$$\sum_{n=0}^{+\infty} \beta_k(X_{n+1}, X_n; \mathbb{K}) = \sum_{n=0}^{r-1} \beta_k(f^{a_{n+1}}, f^{a_n}; \mathbb{K}) = \sum_{x \in \mathrm{Crit}_f} \mu_k(x, f; \mathbb{K}). \qquad \square$$

2.4.4. LEMMA. *If $x \in M$ is a nondegenerate critical point of $f : M \to \mathbb{R}$ then, for any field $\mathbb{K}$, the critical numbers of f at x are given by:*

$$\mu_k(x, f; \mathbb{K}) = \begin{cases} 1, & k = \mu(x), \\ 0, & k \neq \mu(x). \end{cases}$$

PROOF. Let r denote the Morse index of x, n the dimension of M and set $c = f(x)$. By the Morse Lemma, some neighborhood of x in f^c is homeomorphic to a neighborhood of the origin in the cone:

$$C = \big\{(y_1, y_2) \in \mathbb{R}^r \times \mathbb{R}^{n-r} : \|y_2\|^2 - \|y_1\|^2 \leq 0\big\} \subset \mathbb{R}^n,$$

by a homeomorphism that sends x to the origin. Thus:

$$H_k\big(f^c, f^c \setminus \{x\}; \mathbb{K}\big) \cong H_k\big(C, C \setminus \{0\}; \mathbb{K}\big),$$

for all k. It is easy to see that $\mathbb{R}^r \times \{0\}$ and $\big(\mathbb{R}^r \times \{0\}\big) \setminus \{0\}$ are strong deformation retracts respectively of C and $C \setminus \{0\}$; therefore:

$$H_k\big(C, C \setminus \{0\}; \mathbb{K}\big) \cong H_k\big(\mathbb{R}^r, \mathbb{R}^r \setminus \{0\}; \mathbb{K}\big).$$

This concludes the proof. $\square$

2.4.5. COROLLARY. *If $f : M \to \mathbb{R}$ is a Morse function on a compact manifold M then, for any field $\mathbb{K}$, the sequences given by:*

$\mu_k =$ *number of critical points of f having Morse index equal to k,*

$\beta_k = \beta_k(M; \mathbb{K})$,

satisfy the Morse relations.

PROOF. It follows from Theorem 2.4.3 and Lemma 2.4.4. $\square$

2.5. The CW-Complex Associated to a Morse Function

In Section **??** we have seen that the the sublevels f^a and f^b of a smooth map $f : M \to \mathbb{R}$ are homeomorphic if $[a, b]$ is a non critical interval for f. In this section we will study the relation between the topology of f^b and f^a when $[a, b]$ contains critical values of f. More precisely, we will show the following:

2.5.1. PROPOSITION. *Let $f : M \to \mathbb{R}$ be a smooth map where M is a compact n dimensional manifold. Assume that $c \in \,]a, b[$ is the unique critical value of f in $[a, b]$ and that all the critical points of f at the level c are nondegenerate. Hence, there is only a finite number (say, r) of such critical points; denote by $\nu_1, \ldots, \nu_r$ their Morse indexes. Then, there exists a continuous map $\alpha : \sum_{i=1}^{r} S^{\nu_i - 1} \times \overline{\mathrm{B}}^{n-\nu_i} \to f^a$ and a homeomorphism from f^b to the attachment* *space* $\Big(\sum_{i=1}^{r}\overline{\mathrm{B}}^{\nu_i}\times\overline{\mathrm{B}}^{n-\nu_i}\Big)\cup_\alpha f^a$*; moreover, given* $a_1<a$*, such homeomorphism can be chosen to be the identity on* f^{a_1}*.*

The proof of Proposition 2.5.1 will take the rest of this section. By adding a constant to f, we can assume without loss of generality that $c=0$. Moreover, for $\varepsilon>0$ sufficiently small, we may assume that $a=-\varepsilon$ and $b=\varepsilon$; namely, from Corollary 2.3.12, we can find homeomorphisms $f^b\to f^\varepsilon$ and $f^{-\varepsilon}\to f^a$ that are the identity on f^{a_1}. Furthermore, in order to simplify the proof we will assume that there exists a unique critical point $p\in M$ at the level c; we denote by ν the Morse index of such critical point. The proof in this case illustrates the technique that can be applied with straightforward adaptations to the general case. We left the details to the reader.

The idea of the proof of the proposition is to determine a smooth function $g:M\to\mathbb{R}$ satisfying the following conditions:

- $g\le f$;
- $g^\varepsilon=f^\varepsilon$;
- $[-\varepsilon,\varepsilon]$ is a non critical interval for g;
- there exists a homeomorphism $\chi:\mathrm{Dom}(\chi)\subset M\to\overline{\mathrm{B}}^\nu\times\overline{\mathrm{B}}^{n-\nu}$ such that $\chi^{-1}\big(S^{\nu-1}\times\overline{\mathrm{B}}^{n-\nu}\big)\subset f^{-\varepsilon}$ and $\mathrm{Dom}(\chi)$ is a closed subset of M with $g^{-\varepsilon}=f^{-\varepsilon}\cup\mathrm{Dom}(\chi)$.

Once we show the existence of such g, the proof of Proposition 2.5.1 will follow easily by applying Corollary 2.3.12 to g. Namely, since $[-\varepsilon,\varepsilon]$ is a non critical interval for g, there exists a homeomorphism from $g^\varepsilon=f^\varepsilon$ onto $g^{-\varepsilon}$ that fixes $g^{a_1}\supset f^{a_1}$. Moreover, by Lemma 1.14.12, $g^{-\varepsilon}=f^{-\varepsilon}\cup\mathrm{Dom}(\chi)$ is homeomorphic (by a homeomorphism that is the identity on $f^{-\varepsilon}$) to the attachment space $\big(\overline{\mathrm{B}}^\nu\times\overline{\mathrm{B}}^{n-\nu}\big)\cup_\alpha f^{-\varepsilon}$, where $\alpha=\chi^{-1}|_{S^{\nu-1}\times\overline{\mathrm{B}}^{n-\nu}}$.
In order to define g, we consider a diffeomorphism $\varphi:U\to\widetilde{U}$ as in Theorem 2.1.5; we will define g to be a perturbation of f inside U. Let's now go for the technical details. Consider a $\mathrm{Hess}f_p$-orthogonal direct sum decomposition

$$T_pM=\mathcal{H}_+\oplus\mathcal{H}_- \tag{2.5.1}$$

with $\mathrm{Hess}f_p$ positive definite on $\mathcal{H}_+$ and negative definite on $\mathcal{H}_-$. We will write the points of T_pM as pairs (x,y) with $x\in\mathcal{H}_+$ and $y\in\mathcal{H}_-$. Define an inner product $\langle\cdot,\cdot\rangle$ on T_pM by setting:

$$\langle(x_1,y_1),(x_2,y_2)\rangle=\mathrm{Hess}f_p(x_1,x_2)-\mathrm{Hess}f_p(y_1,y_2). \tag{2.5.2}$$

Denoting by $\|\cdot\|$ the norm corresponding to $\langle\cdot,\cdot\rangle$ we have:

$$\big(f\circ\varphi^{-1}\big)(x,y)=\|x\|^2-\|y\|^2,$$

for all $(x,y)\in\widetilde{U}$. The number $\varepsilon>0$ must be chosen so that:

$$\overline{\mathrm{B}}\big(0;\sqrt{6\varepsilon}\big)\subset\widetilde{U}.$$

Let $\lambda:\mathbb{R}\to\mathbb{R}$ be a smooth function such that $\lambda(t)=1$ for $t\le\frac12$, $\lambda(t)=0$ for $t\ge1$ and $-3\le\lambda'(t)\le0$ for all $t\in\mathbb{R}$. We define g to be equal to f outside

U; in U we define g by:

$$\begin{aligned}\big(g\circ\varphi^{-1}\big)(x,y) &= \big(f\circ\varphi^{-1}\big)(x,y)-\frac{3\varepsilon}{2}\lambda\left(\frac{\|x\|^2}{\varepsilon}\right)\lambda\left(\frac{\|x\|^2+\|y\|^2}{6\varepsilon}\right)\\ &= \|x\|^2-\|y\|^2-\frac{3\varepsilon}{2}\lambda\left(\frac{\|x\|^2}{\varepsilon}\right)\lambda\left(\frac{\|x\|^2+\|y\|^2}{6\varepsilon}\right),\end{aligned}$$

for all $(x,y)\in\widetilde{U}$. Obviously g equals f outside the closed set $\varphi^{-1}\big(\overline{\mathrm{B}}(0;\sqrt{6\varepsilon})\big)$, so that g is indeed smooth; moreover, since λ is non negative, we have $g\le f$ everywhere. We now observe that f equals g outside f^{ε} since for $(x,y)\in U\setminus f^{\varepsilon}$ we have $\|x\|^2>\varepsilon$. It follows that $f^{\varepsilon}=g^{\varepsilon}$.

Let us now prove that $[-\varepsilon,\varepsilon]$ is a non critical interval for g. Observe first that f equals g on a neighborhood of $M\setminus U$ and therefore f and g have the same critical points outside U; since $p\in U$ is the unique critical point of f in $f^{-1}\big([-\varepsilon,\varepsilon]\big)$, we conclude that the critical points of g in $g^{-1}\big([-\varepsilon,\varepsilon]\big)$ must be inside U. The differential of g in U is easily computed[2] as:

$$\mathrm{d}\big(g\circ\varphi^{-1}\big)(x,y)=\big(\delta_1(x,y)\langle x,\cdot\rangle,\delta_2(x,y)\langle y,\cdot\rangle\big), \tag{2.5.3}$$

for all $(x,y)\in\widetilde{U}$, where δ_1 and δ_2 are given by:

$$\delta_1(x,y)=2-3\lambda'\left(\frac{\|x\|^2}{\varepsilon}\right)\lambda\left(\frac{\|x\|^2+\|y\|^2}{6\varepsilon}\right)-\frac{1}{2}\lambda\left(\frac{\|x\|^2}{\varepsilon}\right)\lambda'\left(\frac{\|x\|^2+\|y\|^2}{6\varepsilon}\right),$$

$$\delta_2(x,y)=-2-\frac{1}{2}\lambda\left(\frac{\|x\|^2}{\varepsilon}\right)\lambda'\left(\frac{\|x\|^2+\|y\|^2}{6\varepsilon}\right).$$

Since $\lambda\ge 0$ and $-3\le\lambda'\le 0$, it is easily seen that $\delta_1\ge 2$ and $\delta_2\le-\frac{1}{2}$; this implies that the only critical point of g in U is p. However, $g(p)=-\frac{3\varepsilon}{2}$ and $[-\varepsilon,\varepsilon]$ is a non critical interval for g.

To prove the last item of our scheme, we start by observing that

$$g^{-\varepsilon}=f^{-\varepsilon}\cup\varphi^{-1}(\widehat{Q}), \tag{2.5.4}$$

where $\widehat{Q}\subset\overline{\mathrm{B}}\big(0,\sqrt{3\varepsilon}\big)\subset\widetilde{U}\subset T_pM$ is defined by:

$$\begin{aligned}\widehat{Q}=&\Big\{(x,y)\in\mathcal{H}_+\times\mathcal{H}_-:\|x\|^2\le\frac{\varepsilon}{2},\ \|y\|^2\le\|x\|^2+\varepsilon\Big\}\\ &\cup\Big\{(x,y)\in\mathcal{H}_+\times\mathcal{H}_-:\frac{\varepsilon}{2}\le\|x\|^2\le\varepsilon,\ \tau\big(\|x\|\big)\le\|y\|^2\le\|x\|^2+\varepsilon\Big\},\end{aligned}$$

and $\tau:\Big[\sqrt{\frac{\varepsilon}{2}},\sqrt{\varepsilon}\Big]\to\mathbb{R}$ is defined by:

$$\tau(t)=t^2+\varepsilon\Big[1-\frac{3}{2}\lambda\Big(\frac{t^2}{\varepsilon}\Big)\Big].$$

The verification of the equality (2.5.4) requires a number of elementary arguments among which we single out the following:

- since $\widehat{Q}\subset\overline{\mathrm{B}}\big(0,\sqrt{3\varepsilon}\big)$, the quantity $\lambda\big(\frac{\|x\|^2+\|y\|^2}{6\varepsilon}\big)$ appearing in the definition of g is equal to 1 when $(x,y)\in\widehat{Q}$;

[2] We identify T_pM^* with $\mathcal{H}_+^*\oplus\mathcal{H}_-^*$.

- the set $g^{-\varepsilon} \setminus f^{-\varepsilon}$ is contained in U;
- for $(x, y) \in \widetilde{U}$, if $\varphi^{-1}(x, y)$ is in $g^{-\varepsilon} \setminus f^{-\varepsilon}$ then $\|x\|^2 < \varepsilon$ and $\|y\|^2 < 2\varepsilon$, so that again $\lambda\big(\frac{\|x\|^2+\|y\|^2}{6\varepsilon}\big)$ is equal to 1.

To complete the proof of Proposition 2.5.1 we now must exhibit a homeomorphism $\hat{h} : \widehat{Q} \to \overline{\mathrm{B}}^{\nu} \times \overline{\mathrm{B}}^{n-\nu}$ such that:

$$\hat{h}^{-1}\big(S^{\nu-1} \times \overline{\mathrm{B}}^{n-\nu}\big) \subset \big\{(x, y) \in \mathcal{H}_+ \times \mathcal{H}_- : \|x\|^2 - \|y\|^2 \le -\varepsilon\big\}.$$

The homeomorphism $\hat{h}$ is defined with the help of the following:

2.5.2. LEMMA. *Given subsets $Q_1, Q_2 \subset [0, +\infty[^2$ and normed real vector spaces $\mathcal{H}_+$, $\mathcal{H}_-$ set*

$$\widehat{Q}_i = \big\{(x, y) \in \mathcal{H}_+ \times \mathcal{H}_- : \big(\|x\|, \|y\|\big) \in Q_i\big\},$$

for $i = 1, 2$. Assume that $h : Q_1 \to Q_2$ is a homeomorphism satisfying the following conditions:

(1) *for $i = 1, 2$, the map*

$$\big\{(u_1, u_2) \in Q_1 : u_i \neq 0\big\} \ni u \longmapsto \frac{h_i(u)}{u_i} \in \mathbb{R}$$

admits a continuous extension to a map $\overline{h}_i : Q_1 \to \mathbb{R}$, where $h = (h_1, h_2)$;

(2) *for $i = 1, 2$, the map*

$$\big\{(v_1, v_2) \in Q_2 : v_i \neq 0\big\} \ni v \longmapsto \frac{k_i(v)}{v_i} \in \mathbb{R}$$

admits a continuous extension to a map $\overline{k}_i : Q_2 \to \mathbb{R}$, where $k = (k_1, k_2)$ and $k = h^{-1} : Q_2 \to Q_1$.

Then the map:

$$\hat{h} : \widehat{Q}_1 \ni (x, y) \longmapsto \big(\overline{h}_1(\|x\|, \|y\|)x, \overline{h}_2(\|x\|, \|y\|)y\big) \in \widehat{Q}_2$$

is a homeomorphism.

PROOF. Observe that the map:

$$\hat{k} : \widehat{Q}_2 \ni (z, w) \longmapsto \big(\overline{k}_1(\|z\|, \|w\|)z, \overline{k}_2(\|z\|, \|w\|)w\big) \in \widehat{Q}_1$$

is the continuous inverse of $\hat{h}$. $\square$

Finally, we define the homeomorphism h:

2.5.3. LEMMA. *Consider the region $Q_1 \subset [0, +\infty[^2$ given by:*

$$Q_1 = \Big\{(u_1, u_2) \in [0, +\infty[^2 : u_1^2 \le \frac{\varepsilon}{2},\ u_2^2 \le u_1^2 + \varepsilon\Big\}$$
$$\cup \Big\{(u_1, u_2) \in [0, +\infty[^2 : \frac{\varepsilon}{2} \le u_1^2 \le \varepsilon,\ \tau(u_1) \le u_2^2 \le u_1^2 + \varepsilon\Big\},$$

and the unit square $Q_2 = [0, 1]^2$. There exists a homeomorphism $h : Q_1 \to Q_2$ satisfying conditions (1) *and* (2) *in the statement of Lemma 2.5.2 and mapping the graph of $\big[0, \sqrt{\varepsilon}\big] \ni u_1 \mapsto \sqrt{u_1^2 + \varepsilon}$ to the upper side $[0, 1] \times \{1\}$ of Q_2.*

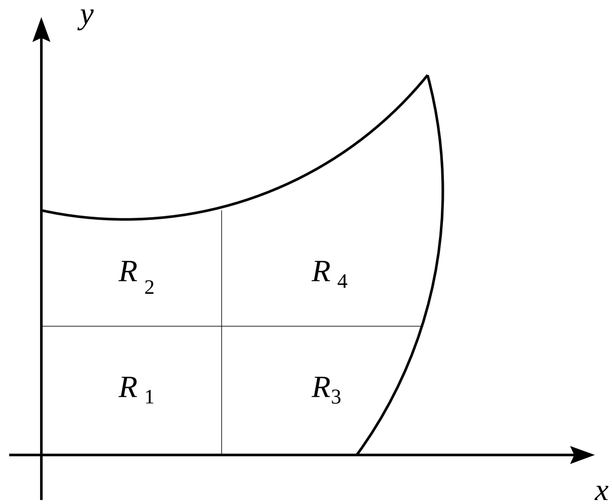

FIGURE 7. The regions R_1, R_2, R_3, R_4

PROOF. Consider the regions (see figure 7):

$$
\begin{aligned}
R_1 &= \Big[0, \sqrt{\tfrac{\varepsilon}{8}}\Big] \times \big[0, \tfrac{\sqrt{\varepsilon}}{2}\big],\\
R_2 &= \Big\{(u_1,u_2) \in [0,+\infty[^2 : u_1^2 \le \frac{\varepsilon}{8},\ \frac{\varepsilon}{4} \le u_2^2 \le u_1^2 + \varepsilon\Big\},\\
R_3 &= \Big\{(u_1,u_2) \in [0,+\infty[^2 : u_2^2 \le \frac{\varepsilon}{4},\ \frac{\varepsilon}{8} \le u_1^2 \le \sigma(u_2)^2\Big\},\\
R_4 &= \Big\{(u_1,u_2) \in [0,+\infty[^2 : \frac{\varepsilon}{8} \le u_1^2 \le \sigma\big(\tfrac{\sqrt{\varepsilon}}{2}\big)^2,\ \frac{\varepsilon}{4} \le u_2^2 \le u_1^2 + \varepsilon\Big\}\\
&\quad \cup \Big\{(u_1,u_2) \in [0,+\infty[^2 : \sigma\big(\tfrac{\sqrt{\varepsilon}}{2}\big)^2 \le u_1^2 \le \varepsilon,\ \tau(u_1) \le u_2^2 \le u_1^2 + \varepsilon\Big\},
\end{aligned}
$$

where $\sigma : \big[0, \sqrt{2\varepsilon}\,\big] \to \big[\sqrt{\frac{\varepsilon}{2}}, \sqrt{\varepsilon}\big]$ is the inverse of $\sqrt{\tau} : \big[\sqrt{\frac{\varepsilon}{2}}, \sqrt{\varepsilon}\big] \to \big[0, \sqrt{2\varepsilon}\,\big]$; observe that $Q_1 = R_1 \cup R_2 \cup R_3 \cup R_4$.

We will construct a homeomorphism h from the region Q_1 to the rectangle $Q_2' = \big[0, \sigma\big(\frac{\sqrt{\varepsilon}}{2}\big)\big] \times \big[0, \frac{3\sqrt{\varepsilon}}{\sqrt{8}}\big]$ satisfying conditions (1) and (2) in the statement of Lemma 2.5.2 and mapping the graph of $\big[0, \sqrt{\varepsilon}\big] \ni u_1 \mapsto \sqrt{u_1^2 + \varepsilon}$ to the upper side of Q_2'. The desired homeomorphism from Q_1 to Q_2 is obtained by composing h with the map $(u_1,u_2) \mapsto \Big(\sigma\big(\frac{\sqrt{\varepsilon}}{2}\big)^{-1} u_1, \frac{\sqrt{8}}{3\sqrt{\varepsilon}} u_2\Big)$.

The homeomorphism h will be defined by describing its restriction to each region R_i.

- $h|_{R_1}$ is the identity.
- $h|_{R_2} : R_2 \to \big[0, \sqrt{\frac{\varepsilon}{8}}\big] \times \big[\frac{\sqrt{\varepsilon}}{2}, \frac{3\sqrt{\varepsilon}}{\sqrt{8}}\big]$ is the homeomorphism

$$(u_1,u_2) \longmapsto \big(u_1, h_2(u_1,u_2)\big),$$

where $h_2(u_1,\cdot)$ is an increasing affine map for all u_1.

- $h|_{R_3} : R_3 \to \big[\sqrt{\frac{\varepsilon}{8}}, \sigma\big(\frac{\sqrt{\varepsilon}}{2}\big)\big] \times \big[0, \frac{\sqrt{\varepsilon}}{2}\big]$ is the homeomorphism
$$(u_1, u_2) \longmapsto \big(h_1(u_1, u_2), u_2\big),$$
where $h_1(\cdot, u_2)$ is an increasing affine map for all u_2.
- $h|_{R_4} : R_4 \to \big[\sqrt{\frac{\varepsilon}{8}}, \sigma\big(\frac{\sqrt{\varepsilon}}{2}\big)\big] \times \big[\frac{\sqrt{\varepsilon}}{2}, \sqrt{\varepsilon}\big]$ is a homeomorphism that equals the identity on the left and bottom sides of the rectangle $\big[\sqrt{\frac{\varepsilon}{8}}, \sigma\big(\frac{\sqrt{\varepsilon}}{2}\big)\big] \times \big[\frac{\sqrt{\varepsilon}}{2}, \sqrt{\varepsilon}\big]$ and that maps the graph of $\big[\sqrt{\frac{\varepsilon}{8}}, \sqrt{\varepsilon}\big] \ni u_1 \mapsto \sqrt{u_1^2 + \varepsilon}$ to the upper side of $\big[\sqrt{\frac{\varepsilon}{8}}, \sigma\big(\frac{\sqrt{\varepsilon}}{2}\big)\big] \times \big[\frac{\sqrt{\varepsilon}}{2}, \sqrt{\varepsilon}\big]$. For a explicit construction of such homeomorphism see Exercise 2.28.

It is easy to see that h is well-defined and that it is a homeomorphism from Q_1 to Q_2' that maps the graph of $\big[0, \sqrt{\varepsilon}\big] \ni u_1 \mapsto \sqrt{u_1^2 + \varepsilon}$ to the upper side of Q_2'. For conditions (1) and (2), observe that $\frac{h_i(u_1,u_2)}{u_i}$ equals 1 near the axis $u_i = 0$ and that also $\frac{k_i(v_1,v_2)}{v_i}$ equals 1 near the axis $v_i = 0$ for $i = 1, 2$. $\square$

2.5.4. LEMMA. *Let $\alpha : \sum_{i=1}^r S^{\nu_i - 1} \times \overline{\mathrm{B}}^{\mu_i} \to Y$ be a continuous map, where Y is a topological space and ν_i, μ_i, $i = 1, \ldots, r$, are non negative integers. Let $\tilde{\alpha}$ be the restriction of α to $\sum_{i=1}^r S^{\nu_i - 1}$, where we identify $\overline{\mathrm{B}}^{\nu_i}$ with the subspace $\overline{\mathrm{B}}^{\nu_i} \times \{0\}$ of $\overline{\mathrm{B}}^{\nu_i} \times \overline{\mathrm{B}}^{\mu_i}$. Then $\Big(\sum_{i=1}^r \overline{\mathrm{B}}^{\nu_i}\Big) \cup_{\tilde{\alpha}} Y$ is a strong deformation retract of $\Big(\sum_{i=1}^r \overline{\mathrm{B}}^{\nu_i} \times \overline{\mathrm{B}}^{\mu_i}\Big) \cup_\alpha Y$.*

PROOF. Set:
$$X = \sum_{i=1}^r \overline{\mathrm{B}}^{\nu_i} \times \overline{\mathrm{B}}^{\mu_i}, \quad A = \sum_{i=1}^r S^{\nu_i - 1} \times \overline{\mathrm{B}}^{\mu_i},$$
$$X' = \sum_{i=1}^r \big(\overline{\mathrm{B}}^{\nu_i} \times \{0\}\big) \cup \big(S^{\nu_i - 1} \times \overline{\mathrm{B}}^{\mu_i}\big),$$
$$A' = \sum_{i=1}^r S^{\nu_i - 1} \times \big(\overline{\mathrm{B}}^{\mu_i} \setminus \{0\}\big),$$
so that $X' \setminus A' = \sum_{i=1}^r \overline{\mathrm{B}}^{\nu_i}$, $A \setminus A' = \sum_{i=1}^r S^{\nu_i - 1}$ and $\tilde{\alpha} = \alpha|_{(A \setminus A')}$. It is easy to see that X' is a strong deformation retract of X (see Exercise 2.29). It follows from Exercise 1.111 that $X' \cup_\alpha Y$ is a strong deformation retract of $X \cup_\alpha Y$; finally, Exercise 1.112 implies that $X' \cup_\alpha Y = (X' \setminus A') \cup_{\alpha|_{(A \setminus A')}} Y$. $\square$

2.5.5. THEOREM. *Let M be a compact differentiable manifold and $f : M \to \mathbb{R}$ a smooth Morse function. Then M has the same homotopy type of a (finite) CW-complex Y such that for every $\nu = 0, 1, \ldots, \dim(M)$, the number of open ν-cells of Y equals the number of critical points of f having Morse index ν.*

PROOF. Since f is a Morse function and M is compact, the number of critical points (and hence of critical values) of f is finite (see Corollary 2.1.6). Denote by $c_1 < c_2 < \cdots < c_p$ the critical values of f; choose $b_0 < c_1$, $b_p > c_p$ and for every

$i = 1, \ldots, p-1$ choose $a_i, b_i \in \mathbb{R}$ with $c_i < a_i < b_i < c_{i+1}$. Clearly, $f^{b_0} = \emptyset$ and $f^{b_p} = M$.

We will construct inductively a sequence of homotopy equivalences $h_i : f^{b_i} \to Y_i$, $i = 0, 1, \ldots, p$, where Y_i is a CW-complex and for each $i = 0, 1, \ldots, p-1$ we have:

- Y_i is a subcomplex of Y_{i+1};
- for every integer $\nu \geq 0$, the number of ν-cells of Y_{i+1} not in Y_i equals the number of critical points of f at the level c_{i+1} having Morse index ν;
- h_{i+1} coincides with h_i on f^{a_i}.

After such construction we will have a homotopy equivalence h_p from $f^{b_p} = M$ to the CW-complex $Y = Y_p$ which has the desired number of cells on each dimension.

For $i = 0$ we have $f^{b_0} = \emptyset$ and there is nothing to do. Now assume that for some $i = 0, \ldots, p-1$ the homotopy equivalence $h_i : f^{b_i} \to Y_i$ has been constructed. Assume that f has r critical points at the level c_{i+1} whose Morse indexes are denoted by $(\nu_i)_{i=1}^r$; set $\mu_i = \dim(M) - \nu_i$. Since c_{i+1} is the unique critical value of f on $[b_i, b_{i+1}]$, Proposition 2.5.1 gives us a homeomorphism:

$$f^{b_{i+1}} \longrightarrow \left(\sum_{j=1}^r \overline{\mathrm{B}}^{\nu_j} \times \overline{\mathrm{B}}^{\mu_j} \right) \cup_\alpha f^{b_i}, \tag{2.5.5}$$

that fixes the points of f^{a_i}, where $\alpha : \sum_{i=1}^r S^{\nu_i - 1} \times \overline{\mathrm{B}}^{\mu_i} \to f^{b_i}$ is a continuous map. By Lemma 2.5.4, we have a strong deformation retraction:

$$\left(\sum_{j=1}^r \overline{\mathrm{B}}^{\nu_j} \times \overline{\mathrm{B}}^{\mu_j} \right) \cup_\alpha f^{b_i} \longrightarrow \left(\sum_{j=1}^r \overline{\mathrm{B}}^{\nu_j} \right) \cup_{\tilde{\alpha}} f^{b_i}, \tag{2.5.6}$$

where $\overline{\mathrm{B}}^{\nu_j}$ is identified with $\overline{\mathrm{B}}^{\nu_j} \times \{0\} \subset \overline{\mathrm{B}}^{\nu_j} \times \overline{\mathrm{B}}^{\mu_j}$ and $\tilde{\alpha}$ is the restriction of α to $\sum_{j=1}^r S^{\nu_j - 1}$. By Exercise 1.117, we have a homeomorphism:

$$\left(\sum_{j=1}^r \overline{\mathrm{B}}^{\nu_j} \right) \cup_{\tilde{\alpha}} f^{b_i} \longrightarrow C_{\tilde{\alpha}}, \tag{2.5.7}$$

that fixes the points of f^{b_i}, where $C_{\tilde{\alpha}}$ denotes the cone of the map

$$\tilde{\alpha} : \sum_{j=1}^r S^{\nu_j - 1} \longrightarrow f^{b_i}.$$

Using Corollary 1.14.25, we obtain a homotopy equivalence:

$$C_{\tilde{\alpha}} \longrightarrow C_{h_i \circ \tilde{\alpha}}, \tag{2.5.8}$$

that extends $h_i : f^{b_i} \to Y_i$. Applying Proposition 1.15.20 to the restriction of $h_i \circ \tilde{\alpha}$ to each sphere $S^{\nu_j - 1}$, we obtain a continuous map $k : \sum_{j=1}^r S^{\nu_j - 1} \to Y_i$ that is homotopic to $h_i \circ \tilde{\alpha}$ and such that $k\big(S^{\nu_j - 1}\big)$ is contained in the $(\nu_j - 1)$-skeleton $Y_i^{\nu_j - 1}$ of the CW-complex Y_i. Now Corollary 1.14.23 gives us a homotopy equivalence:

$$C_{h_i \circ \tilde{\alpha}} \longrightarrow C_k, \tag{2.5.9}$$

that fixes the points of Y_i. Again by Exercise 1.117 we have a homeomorphism:

$$C_k \longrightarrow \left(\sum_{j=1}^{r} \overline{\mathrm{B}}^{\nu_j}\right) \cup_k Y_i, \tag{2.5.10}$$

that fixes the points of Y_i. By Proposition 1.15.14 the topological space:

$$Y_{i+1} = \left(\sum_{j=1}^{r} \overline{\mathrm{B}}^{\nu_j}\right) \cup_k Y_i,$$

can be endowed with the structure of a CW-complex having Y_i as a subcomplex and the open balls B^{ν_j}, $j = 1, \ldots, r$, as open cells. To complete the induction step and the proof of the theorem, now take h_{i+1} to be the composition of the homotopy equivalences (2.5.5)—(2.5.10). □

NEW PROOF OF COROLLARY 2.4.5. By Theorem 2.5.5, M has the same homotopy type (and hence the same homology) of a CW-complex Y having μ_k open cells of dimension k for every $k \geq 0$. But the singular homology of Y with coefficients in $\mathbb{K}$ is isomorphic to the homology of the cellular complex $\mathcal{D}(Y;\mathbb{K})$ of Y, which is a nonnegative chain complex of $\mathbb{K}$-vector spaces whose k-th chain space has the dimension equal to the number of k-th dimensional open cells of Y. The conclusion follows from Lemma **??**. □

2.6. The Morse–Witten Complex

2.6.1. DEFINITION. Given a critical point $p \in M$ of f then the *stable* and the *unstable manifold* of p are defined respectively by:

$$W_{\mathrm{s}}(p,f) = \big\{x \in M : \lim_{t \to +\infty} t \cdot x = p\big\},$$
$$W_{\mathrm{u}}(p,f) = \big\{x \in M : \lim_{t \to -\infty} t \cdot x = p\big\}.$$

When f is fixed by the context, we will write simply $W_{\mathrm{s}}(p)$ and $W_{\mathrm{u}}(p)$.

The concepts of stable and unstable manifolds are standard in the theory of dynamical systems (see [**?**]). More generally, one can define the stable and unstable manifolds for *hyperbolic singularities* of an arbitrary vector field. In Appendix B we present a summary of the basic concepts of such theory, as well as the proof of the following:

2.6.2. THEOREM. *Let $f : M \to \mathbb{R}$ be a smooth function on a Riemannian manifold (M, g) and let $p \in M$ be a nondegenerate critical point of f. Then the stable and the unstable manifold of p are connected embedded submanifolds of M, whose dimensions are respectively equal to the coindex and the index of $\mathrm{Hess} f_p$. The tangent spaces $T_pW_{\mathrm{s}}(p)$ and $T_pW_{\mathrm{u}}(p)$ are given respectively by the positive and the negative eigenspaces of $\mathrm{Hess} f_p$.* □

Obviously if x belongs to the stable (resp., the unstable) manifold of a critical point p then $t \cdot x$ also belongs to the stable (resp., the unstable) manifold of p. Thus

$W_{\rm s}(p)$ and $W_{\rm u}(p)$ are unions of flow lines of $-\nabla f$. In particular, for $x \in W_{\rm s}(p)$ we have:

$$-\nabla f(x) = \frac{\mathrm{d}}{\mathrm{d}t}\, t \cdot x\Big|_{t=0} \in T_x W_{\rm s}(p), \tag{2.6.1}$$

for all $x \in W_{\rm s}(p)$ and similarly $-\nabla f(x) \in T_x W_{\rm u}(p)$, for all $x \in W_{\rm u}(p)$. Obviously, the unique critical point of f in $W_{\rm s}(p)$ or in $W_{\rm u}(p)$ is p itself. Since f is strictly decreasing along the nonconstant flow lines of $-\nabla f$, it follows that p is a strict global maximum of $f|_{W_{\rm u}(p)}$ and a strict global minimum of $f|_{W_{\rm s}(p)}$. In particular, $W_{\rm s}(p) \cap W_{\rm u}(p) = \{p\}$.

Given two critical points $p, q \in M$, we will be interested in the set of flow lines going from p to q, i.e., the flow lines contained in $W_{\rm u}(p) \cap W_{\rm s}(q)$. We have the following:

2.6.3. LEMMA. *Let $p, q \in M$ be critical points of f and assume that $W_{\rm u}(p)$ and $W_{\rm s}(q)$ are transversal and nondisjoint. Then the intersection $W_{\rm u}(p) \cap W_{\rm s}(q)$ is an embedded submanifold of M having dimension $\mu(p) - \mu(q)$. Moreover, for any $x \in W_{\rm u}(p) \cap W_{\rm s}(q)$ we have:*

$$T_x\big(W_{\rm u}(p) \cap W_{\rm s}(q)\big) = T_x W_{\rm u}(p) \cap T_x W_{\rm s}(q); \tag{2.6.2}$$

in particular (see (2.6.1)*), $\nabla f(x) \in T_x\big(W_{\rm u}(p) \cap W_{\rm s}(q)\big)$.*

PROOF. The intersection of embedded transversal submanifolds is an embedded submanifold; moreover, the tangent space of the intersection is equal to the intersection of the tangent spaces, which proves (2.6.2). As for the dimension of $W_{\rm u}(p) \cap W_{\rm s}(q)$ we compute:

$$\begin{aligned} \dim\big(W_{\rm u}(p) \cap W_{\rm s}(q)\big) &= \dim\big(W_{\rm u}(p)\big) + \dim\big(W_{\rm s}(q)\big) - \dim(M) \\ &= \mu(p) + \dim(M) - \mu(q) - \dim(M) = \mu(p) - \mu(q). \quad \square \end{aligned}$$

2.6.4. COROLLARY. *Let $p, q \in M$ be distinct critical points of f and assume that $W_{\rm u}(p)$ and $W_{\rm s}(q)$ are transversal and nondisjoint. Then $\mu(p) > \mu(q)$.*

PROOF. Since $p \neq q$, there must exist a regular point x of f in $W_{\rm u}(p) \cap W_{\rm s}(q)$, so that $0 \neq \nabla f(x) \in T_x\big(W_{\rm u}(p) \cap W_{\rm s}(q)\big)$. Then:

$$\mu(p) - \mu(q) = \dim\big(W_{\rm u}(p) \cap W_{\rm s}(q)\big) \geq 1. \qquad \square$$

2.6.5. COROLLARY. *Let $p, q \in M$ be critical points of f such that $W_{\rm u}(p)$ and $W_{\rm s}(q)$ are transversal and let $a \in \mathbb{R}$ be a regular value of f such that the intersection $W_{\rm u}(p) \cap W_{\rm s}(q) \cap f^{-1}(a)$ is nonempty. Then $W_{\rm u}(p) \cap W_{\rm s}(q) \cap f^{-1}(a)$ is an embedded submanifold of M having dimension $\mu(p) - \mu(q) - 1$. Its tangent space is given by:*

$$T_x\big(W_{\rm u}(p) \cap W_{\rm s}(q) \cap f^{-1}(a)\big) = T_x W_{\rm u}(p) \cap T_x W_{\rm s}(q) \cap \nabla f(x)^{\perp},$$

for all $x \in W_{\rm u}(p) \cap W_{\rm s}(q) \cap f^{-1}(a)$.

PROOF. Since $T_x f^{-1}(a) = \nabla f(x)^\perp$ and $\nabla f(x)$ is tangent to $W_\mathrm{u}(p) \cap W_\mathrm{s}(q)$, we have that $W_\mathrm{u}(p) \cap W_\mathrm{s}(q)$ and $f^{-1}(a)$ are transversal. The conclusion follows (as in the proof of Lemma 2.6.3) from general facts about the intersection of transversal submanifolds. $\square$

2.6.6. DEFINITION. Given $k \in \mathbb{Z}$, we say that $f : (M, g) \to \mathbb{R}$ *satisfies the Morse–Smale condition of order* k if for every pair of critical points $p, q \in M$ with $\mu(p) - \mu(q) \le k$, the unstable manifold of p and the stable manifold of q are transversal. If $f : (M, g) \to \mathbb{R}$ satisfies the Morse–Smale condition for all $k \in \mathbb{Z}$ (i.e., if $W_\mathrm{u}(p)$ and $W_\mathrm{s}(q)$ are transversal for every $p, q \in \mathrm{Crit}_f$) then we say simply that f *satisfies the Morse–Smale condition.*

The following lemma is just a restatement of Corollary 2.6.4.

2.6.7. LEMMA. *Assume that $f : (M, g) \to \mathbb{R}$ satisfies the Morse–Smale condition of order zero. Then the Morse index decreases strictly when one goes through a flow line of $-\nabla f$, i.e., if $p, q \in M$ are critical points of f such that there exists a flow line of $-\nabla f$ going from p to q then $\mu(p) > \mu(q)$.* $\square$

We now consider fixed a Morse function $f : M \to \mathbb{R}$ on a compact Riemannian manifold (M, g) satisfying the Morse–Smale condition of order 1. Our goal is to associate a chain complex $\mathfrak{C}$ to f (or, more precisely, to ∇f) which can be roughly described as follows. For every $k \ge 0$ we define $\mathfrak{C}_k$ to be the free abelian group spanned by the set of critical points of f having Morse index equal to k; for $k < 0$ we set $\mathfrak{C}_k = 0$. Now, if $p, q \in M$ are critical points with $\mu(p) = k$ and $\mu(q) = k - 1$, we have to define the coefficient for q in the expression for the boundary of p in $\mathfrak{C}$. Since $\mu(p) - \mu(q) = 1$, by Lemma 2.6.3, the manifold $W_\mathrm{u}(p) \cap W_\mathrm{s}(q)$ of flow lines going from p to q is one-dimensional, i.e., the flow lines going from p to q are isolated (see Exercise 2.34). We will prove that the number of flow lines going from p to q is indeed finite. The coefficient for q in the expression for the boundary of p in $\mathfrak{C}$ will then be given by an algebraic count of the number of flow lines going from p to q.

Before giving the details of the construction, we need some technical lemmas.

2.6.8. LEMMA. *Let $p \in \mathrm{Crit}_f$. If x is in the closure of $W_\mathrm{s}(p)$ then $t \cdot x$ is also in the closure of $W_\mathrm{s}(p)$ for all $t \in [-\infty, +\infty]$. In particular, by the continuity of f, we have $f(t \cdot x) \ge f(p)$, for all $t \in [-\infty, +\infty]$.*

PROOF. Given $t \in \mathbb{R}$, we have $F_t\big(W_\mathrm{s}(p)\big) \subset W_\mathrm{s}(p)$; this implies, by the continuity of F_t, that $F_t\big(\overline{W_\mathrm{s}(p)}\big) \subset \overline{W_\mathrm{s}(p)}$. Thus $x \in \overline{W_\mathrm{s}(p)}$ implies $t \cdot x \in \overline{W_\mathrm{s}(p)}$ for all $t \in \mathbb{R}$ and hence also $t \cdot x \in \overline{W_\mathrm{s}(p)}$ for $t = \pm\infty$. $\square$

2.6.9. LEMMA. *Let $p \in \mathrm{Crit}_f$ and set $f(p) = c$. Then the intersection of the closure of $W_\mathrm{s}(p)$ with the level $f^{-1}(c)$ contains only p, i.e., $\overline{W_\mathrm{s}(p)} \cap f^{-1}(c) = \{p\}$.*

PROOF. Choose $x \in \overline{W_\mathrm{s}(p)}$ with $f(x) = c$ and let us show that $x = p$. First, by Lemma 2.6.8, we have $f(t \cdot x) \ge c$ for all $t \in [-\infty, +\infty]$. On the other hand, $f(t \cdot x) \le f(x) = c$ for $t \ge 0$, so $f(t \cdot x) = c$ for $t \ge 0$ and x must be a critical point of f. Thus $+\infty \cdot x = x$ and by Lemma 2.3.5, the restriction of F to $(f\circ F)^{-1}\big([c,+\infty[\big)$ is continuous at the point $(+\infty,x)$. But F is constant and equal to p in $\{+\infty\}\times W_{\mathrm{s}}(p)$ and $\{+\infty\}\times W_{\mathrm{s}}(p)$ is contained in $(f\circ F)^{-1}\big([c,+\infty[\big)$, so it must be $F(+\infty,x)=p$, i.e., $x=p$. $\square$

2.6.10. LEMMA. *Given distinct critical points $p,q\in\mathrm{Crit}_f$ then q is in the closure of $W_{\mathrm{s}}(p)$ if and only if there exists $x\in W_{\mathrm{u}}(q)$, $x\neq q$, which is in the closure of $W_{\mathrm{s}}(p)$.*

PROOF. If there exists $x\in W_{\mathrm{u}}(q)$ with $x\in\overline{W_{\mathrm{s}}(p)}$ then $q=-\infty\cdot x$ is in $\overline{W_{\mathrm{s}}(p)}$, by Lemma 2.6.8. Conversely, assume that $q\in\overline{W_{\mathrm{s}}(p)}$. If there were no point $x\in W_{\mathrm{u}}(q)\setminus\{q\}$ in the closure of $W_{\mathrm{s}}(p)$ then $Z=M\setminus\overline{W_{\mathrm{s}}(p)}$ would be an open set in M containing $W_{\mathrm{u}}(q)\setminus\{q\}$. By Lemma B.22, there exists a neighborhood V of q with the following property; for $y\in V$, either $y\in W_{\mathrm{s}}(q)$ or $t\cdot y\in Z$ for some $t>0$. Since $q\in\overline{W_{\mathrm{s}}(p)}$, we can find $y\in V\cap W_{\mathrm{s}}(p)$ and since $p\neq q$, it cannot be $y\in W_{\mathrm{s}}(q)$. Thus, there must exist $t>0$ with $t\cdot y\in Z$. But this contradicts the fact that $t\cdot y\in W_{\mathrm{s}}(p)$. $\square$

2.6.11. DEFINITION. A *broken flow line* is a sequence $\gamma=(\gamma_1,\ldots,\gamma_k)$ of flow lines $\gamma_i:\mathbb{R}\to M$ of $-\nabla f$ such that $\lim_{t\to+\infty}\gamma_i(t)=\lim_{t\to-\infty}\gamma_{i+1}(t)$, for $i=1,\ldots,k-1$. We say that k is the *number of steps* of γ or that γ is a *k-step broken flow line*. If $p=\lim_{t\to-\infty}\gamma_1(t)$ and $q=\lim_{t\to+\infty}\gamma_k(t)$ then we say that γ is a (k-step) broken flow line *from p to q*. If $p=q$ we also say that there exists a 0-step broken flow line from p to q.

Given distinct critical points $p,q\in\mathrm{Crit}_f$ then obviously $W_{\mathrm{u}}(p)\cap W_{\mathrm{s}}(q)\neq\emptyset$ if and only if there exists a 1-step broken flow line from p to q.

2.6.12. LEMMA. *Let $p\in\mathrm{Crit}_f$. If $x\in\overline{W_{\mathrm{s}}(p)}$ then there exists a broken flow line from $+\infty\cdot x$ to p.*

PROOF. Set $q_1=+\infty\cdot x\in\mathrm{Crit}_f$. By Lemma 2.6.8, $q_1\in\overline{W_{\mathrm{s}}(p)}$. If $q_1=p$, we are done. Otherwise, by Lemma 2.6.10, we can find $x_1\in W_{\mathrm{u}}(q_1)$, $x_1\neq q_1$, with $x_1\in\overline{W_{\mathrm{s}}(p)}$. Now set $q_2=+\infty\cdot x_1\in\mathrm{Crit}_f$. Observe that there exists a flow line of $-\nabla f$ from q_1 to q_2 and $f(q_2)<f(q_1)$. Moreover, $q_2\in\overline{W_{\mathrm{s}}(p)}$, by Lemma 2.6.8. If $q_2=p$, we are done. Otherwise, we can continue this process inductively until some $q_n=p$; otherwise, we would obtain a sequence $(q_n)_{n\geq1}$ of critical points with $f(q_1)>f(q_2)>\cdots$, which contradicts the fact that f has only a finite number of critical points. $\square$

2.6.13. LEMMA. *Let $p\in\mathrm{Crit}_f$ and set $f(p)=c$. If $a<c$ is such that there are no critical values of f on $[a,c[$ then every nonconstant flow line contained in $W_{\mathrm{u}}(p)$ intersects the level $f^{-1}(a)$, i.e., for every $x\in W_{\mathrm{u}}(p)\setminus\{p\}$ there exists $t\in\mathbb{R}$ with $f(t\cdot x)=a$.*

PROOF. Choose $x\in W_{\mathrm{u}}(p)$, $x\neq p$. Then $f(x)<c$. If $f(x)\leq a$, then, since $f(t\cdot x)\to c>a$ as $t\to-\infty$, there exists $t\leq0$ with $f(t\cdot x)=a$. Now assume that $f(x)>a$. It suffices to show that $f(t\cdot x)\leq a$ for some $t\geq0$. If we had $f(t\cdot x)>a$ for all $t\geq0$ then $y=+\infty\cdot x$ would be a critical point of f with $a\leq f(y)<c$, which is a contradiction. $\square$

2.6.14. LEMMA. *Assume that* $f : (M, g) \to \mathbb{R}$ *satisfies the Morse–Smale condition of order* 1*. Then, given* $p, q \in \mathrm{Crit}_f$ *with* $\mu(p) - \mu(q) = 1$*, there exists only a finite number of flow lines of* $-\nabla f$ *from* p *to* q*.*

PROOF. Choose $a < f(p)$ such that there are no critical values of f on the interval $[a, f(p)[$. Then, by Lemma 2.6.13, every nonconstant flow line of $-\nabla f$ contained in $W_{\mathrm{u}}(p)$ intersects $f^{-1}(a)$ (precisely once), so there exists bijection between the set of nonconstant flow lines of $-\nabla f$ contained in $W_{\mathrm{u}}(p)$ and $W_{\mathrm{u}}(p) \cap f^{-1}(a)$. Thus, there exists a bijection between the set of flow lines of $-\nabla f$ from p to q and $W_{\mathrm{u}}(p) \cap W_{\mathrm{s}}(q) \cap f^{-1}(a)$. We have to prove that $W_{\mathrm{u}}(p) \cap W_{\mathrm{s}}(q) \cap f^{-1}(a)$ is finite. By Corollary 2.6.5, $W_{\mathrm{u}}(p) \cap W_{\mathrm{s}}(q) \cap f^{-1}(a)$ is a zero-dimensional embedded submanifold of M, i.e., it is a discrete subset of M. $\square$

2.6.15. DEFINITION. A smooth map $f : M \to \mathbb{R}$ is said to be *self-indexing* if for every critical point $p \in M$ of f, we have $f(p) = \mu(p)$.

2.6.16. PROPOSITION. *If* $f : M \to \mathbb{R}$ *is a self-indexing Morse function on a compact Riemannian manifold* (M, g) *then the sublevels* $(f^k)_{k \geq 0}$ *of* f *form a cellular filtration of* M *whose corresponding cellular complex is isomorphic to the Morse–Witten complex of* f*.*

2.6.17. PROPOSITION. *Let* $f : M \to \mathbb{R}$ *be a smooth Morse function on a compact Riemannian manifold* (M, g) *satisfying the Morse–Smale condition of order zero. Then there exists a self-indexing Morse function* $\widetilde{f} : M \to \mathbb{R}$ *and a Riemannian metric* $\widetilde{g}$ *on* M *such that the gradient of* f *with respect to* g *is equal to the gradient of* $\widetilde{f}$ *with respect to* $\widetilde{g}$*.*

2.6.18. LEMMA. *Let* $f : M \to \mathbb{R}$ *be a smooth function on a compact Riemannian manifold* (M, g)*. Let* $a < b$ *be noncritical levels of* f *such that* $f^{-1}\big([a, b]\big)$ *contains precisely two critical points* $p, q \in M$ *of* f*. Assume that* p *and* q *are nondegenerate and that there are no flow lines of* $-\nabla f$ *connecting* p *and* q*. Then, given* $c_1, c_2 \in \mathbb{R}$*, there exists a smooth function* $\widetilde{f} : M \to \mathbb{R}$ *and a Riemannian metric* $\widetilde{g}$ *on* M *such that:*

- f *and* g *are respectively equal to* $\widetilde{f}$ *and* $\widetilde{g}$ *outside* $f^{-1}\big([a + \varepsilon, b - \varepsilon]\big)$*, for some* $\varepsilon > 0$*;*
- $f^{-1}\big([a, b]\big) = \widetilde{f}^{-1}\big([a, b]\big)$*;*
- *the gradient of* f *with respect to* g *is equal to the gradient of* $\widetilde{f}$ *with respect to* $\widetilde{g}$*;*
- $\widetilde{f}(p) = c_1$, $\widetilde{f}(q) = c_2$.

Exercises for Chapter 2

Calculus on manifolds: basic terminology.

EXERCISE 2.1. Let V be a finite dimensional real vector space and let $q : \mathrm{Lin}(V) \to \mathrm{Lin}(V)$ be defined by $q(T) = T^2$. Show that the differential of q is given by:

$$\mathrm{d}q(T) \cdot H = T \circ H + H \circ T;$$

conclude that q restricts to a diffeomorphism between open neighborhoods of the identity.

EXERCISE 2.2. Show that every differentiable manifold admits a Riemannian metric. More generally, given a vector bundle E over a differentiable manifold M, an open subset $A \subset E$ whose intersection with every fiber of E is convex and non empty, show that E admits a global smooth section whose image is contained in A (*hint*: use partitions of unity). Obtain the existence of a Riemannian metric on M as a consequence of this more general result (*hint*: let E be the subbundle of $TM^* \otimes TM^*$ consisting of symmetric bilinear forms and let A be the subset of E consisting of positive definite forms). Where does the argument fail in the case of Lorentzian metrics?

EXERCISE 2.3. Let $\mathcal{E}$ be a fiber bundle over a differentiable manifold M with projection $\pi : \mathcal{E} \to M$. Assume that $f : N \to M$ is a smooth map defined on another differentiable manifold N. The *pull-back* of the fiber bundle $\mathcal{E}$ by f is defined by:

$$f^*\mathcal{E} = \bigcup_{x \in N} \{x\} \times \mathcal{E}_{f(x)};$$

we have a canonical map $\hat{\pi} : f^*\mathcal{E} \to N$ that sends $\{x\} \times \mathcal{E}_{f(x)}$ to $x \in N$. If $\alpha : \mathcal{E}|_U \to U \times \mathcal{E}_0$ is a trivialization of E then we define a trivialization:

$$\hat{\alpha} : \hat{\pi}^{-1}\big(f^{-1}(U)\big) \longrightarrow f^{-1}(U) \times \mathcal{E}_0$$

of $f^*\mathcal{E}$ by setting:

$$\hat{\alpha}(x, e) = \big(x, \alpha_{f(x)}(e)\big),$$

for all $(x, e) \in \hat{\pi}^{-1}\big(f^{-1}(U)\big)$ (so that $x \in f^{-1}(U)$ and $e \in \mathcal{E}_{f(x)}$). Show that:

- $f^*\mathcal{E}$ is a fiber bundle over N;
- the map $F : f^*\mathcal{E} \to \mathcal{E}$ defined by $F(x, e) = e$ is smooth and the diagram:

$$\begin{tikzcd} f^*\mathcal{E} \arrow[r, "F"] \arrow[d, "\hat{\pi}"'] & \mathcal{E} \arrow[d, "\pi"] \\ N \arrow[r, "f"'] & M \end{tikzcd}$$

  commutes;
- given a smooth map $s : N \to \mathcal{E}$ with $\pi \circ s = f$ show that there exists a unique smooth section $\hat{s} : N \to f^*\mathcal{E}$ of $f^*\mathcal{E}$ for which the diagram:

$$\begin{tikzcd} f^*\mathcal{E} \arrow[r, "F"] & \mathcal{E} \arrow[d, "\pi"] \\ N \arrow[u, "\hat{s}"] \arrow[ur, "s"'] \arrow[r, "f"'] & M \end{tikzcd}$$

  commutes;
- if $N \subset M$ is a submanifold and $f : N \to M$ is the inclusion then $f^*\mathcal{E}$ can be naturally identified with the restriction $\mathcal{E}|_N$;

- if $\mathcal{E}$ has the structure of a vector bundle then so does $f^*\mathcal{E}$ (more precisely, if α is fiber-linear then also $\hat{\alpha}$ is).

EXERCISE 2.4. Let E_1, E_2 be vector bundles over a differentiable manifold M. A map $T : E_1 \to E_2$ is called a *vector bundle morphism* if for every $x \in M$, T maps $(E_1)_x$ linearly into $(E_2)_x$, i.e.:

$$T\big[(E_1)_x\big] \subset (E_2)_x, \quad \text{for all } x \in M,$$

and

$$T|_{(E_1)_x} : (E_1)_x \longrightarrow (E_2)_x \text{ is linear, for all } x \in M.$$

Show that if $T : E_1 \to E_2$ is a smooth bijective vector bundle morphism then $T^{-1} : E_2 \to E_1$ is also a smooth vector bundle morphism; we then say that T is a *vector bundle isomorphism* (*hint*: to prove that T^{-1} is smooth, use local trivializations and the inverse function theorem).

EXERCISE 2.5. Let E_1, E_2 be vector bundles over a differentiable manifold M and let $T : E_1 \to E_2$ be a smooth vector bundle morphism such that the rank of $T_x : (E_1)_x \to (E_2)_x$ is independent of $x \in M$. Show that $\mathrm{Ker}(T) = \bigcup_{x\in M} \mathrm{Ker}(T_x)$ is a vector subbundle of E_1 and $\mathrm{Im}(T) = \bigcup_{x\in M} \mathrm{Im}(T_x)$ is a vector subbundle of E_2.

EXERCISE 2.6. If E is a vector bundle over a differentiable manifold M and E' is a vector subbundle of E, show that E' is closed in E.

EXERCISE 2.7. Let E be a vector bundle over a differentiable manifold M with projection $\pi : E \to M$. For every $e \in E$, the vertical space $\mathrm{Ver}_e\, E$ may be identified with the fiber E_x containing x (as usual, one can identify the tangent space to a vector space with the vector space it self). Use the identification $\mathrm{Ver}_e\, E \cong E_x$ to construct an isomorphism of vector bundles from $\mathrm{Ver}\, E$ to the pull-back π^*E.

EXERCISE 2.8. Let E be a vector bundle over a differentiable manifold M with projection π. Given horizontal spaces $\mathrm{Hor}^i_e\, E$, $i = 1, 2, 3$ at a point $e \in E$, show that:

$$\begin{aligned}
\mathrm{Comp}\big(\mathrm{Hor}^1_e\, E, \mathrm{Hor}^1_e\, E\big) &= 0,\\
\mathrm{Comp}\big(\mathrm{Hor}^1_e\, E, \mathrm{Hor}^2_e\, E\big) &= -\mathrm{Comp}\big(\mathrm{Hor}^2_e\, E, \mathrm{Hor}^1_e\, E\big),\\
\mathrm{Comp}\big(\mathrm{Hor}^1_e\, E, \mathrm{Hor}^3_e\, E\big) &= \mathrm{Comp}\big(\mathrm{Hor}^1_e\, E, \mathrm{Hor}^2_e\, E\big) +\\
&\quad + \mathrm{Comp}\big(\mathrm{Hor}^2_e\, E, \mathrm{Hor}^3_e\, E\big);
\end{aligned}$$

conclude that affine compatibility is an equivalence relation on the set of all horizontal bundles of E.

EXERCISE 2.9. Let E be a vector bundle over a differentiable manifold M and let $f : N \to M$ be a smooth map defined in another differentiable manifold N. Assume that $\mathrm{Hor}\, E$ is a connection on E and consider the map $F : f^*E \to E$ defined in Exercise 2.3. For every $(x, e) \in f^*E$, set:

$$\mathrm{Hor}_{(x,e)}(f^*E) = \mathrm{d}F^{-1}_{(x,e)}(\mathrm{Hor}_e\, E); \tag{2.6.3}$$

show that (2.6.3) defines a connection on f^*E. This is called the *pull-back* of the connection $\mathrm{Hor}\,E$ by the map f. Denoting by $f^*\nabla$ the covariant derivative operator of (2.6.3), show that for every $\hat{s} \in \Gamma(f^*E)$ and every $v \in TM$ we have:

$$(f^*\nabla)_v \hat{s} = \nabla^f_v s,$$

where $s = F \circ \hat{s}$.

EXERCISE 2.10. Given vector bundles E_1, E_2 over a differentiable manifold M, define a natural vector bundle structure on the set:

$$E_1 \otimes E_2 = \bigcup_{x \in M} (E_1)_x \otimes (E_2)_x.$$

Given connection ∇^1 and ∇^2 in E_1 and E_2 respectively, show that:

$$\nabla_V(s_1 \otimes s_2) = \big(\nabla_V s_1\big) \otimes s_2 + s_1 \otimes \big(\nabla_V s_2\big), \qquad s_1 \in \Gamma(E_1),\ s_2 \in \Gamma(E_2),\ V \in \Gamma(TM),$$

defines a connection on $E_1 \otimes E_2$. If E is a vector bundle over M define also a natural vector structure on the set:

$$E^* = \bigcup_{x \in M} (E_x)^*;$$

if ∇ is a connection on E, show that the formula:

$$\big(\nabla^*_V s\big)(s') = V\big(s \cdot s'\big) - s\big(\nabla_V s'\big), \quad s \in \Gamma(E^*),\ s' \in \Gamma(E),\ V \in \Gamma(TM),$$

defines a connection on E^*.

EXERCISE 2.11. Let M be an n-dimensional differentiable manifold and $D \subset M$ a subset. We call D a *domain with smooth boundary* (or a *submanifold with boundary of codimension zero*) if for every $x \in D \cap \partial D$ there exists a chart $\varphi : U \to \widetilde{U}$ of M with $x \in U$ and $\varphi(U \cap D) = \widetilde{U} \cap \mathsf{H}^n$ (by ∂D we mean the boundary of D as a subset of the topological space M). Show that:

- if D is a domain with smooth boundary in M then D is a topological manifold with boundary (in the sense of Exercise 1.63), whose interior points coincide with the interior points of D as a subset of the topological space M.
- If $f : M \to \mathbb{R}$ is a smooth map and $a \in \mathbb{R}$ is a regular value for f, show that $f^a = f^{-1}\big(\,]-\infty, a]\,\big)$ is a domain with smooth boundary in M whose boundary is $f^{-1}(a)$.
- If $f : M \to \mathbb{R}$ is a smooth map and $a, b \in \mathbb{R}$ are regular values of f with $a < b$, show that $f^{-1}\big([a, b]\big)$ is a domain with smooth boundary in M whose boundary is $f^{-1}(a) \cup f^{-1}(b)$.

EXERCISE 2.12. A smooth map $f : M \to N$ between differentiable manifolds M, N is said to be *transversal* to a submanifold $P \subset N$ if for every $x \in f^{-1}(P)$ the (not necessarily direct) sum $\mathrm{Im}(\mathrm{d}f_x) + T_{f(x)}P$ equals the whole tangent space $T_{f(x)}N$. Show that if $f : M \to N$ is transversal to $P \subset N$ then $f^{-1}(P)$ is a submanifold of M whose codimension in M equals the codimension of P in N; show that $T_x f^{-1}(P) = \mathrm{d}f_x^{-1}(T_{f(x)}P)$ for every $x \in f^{-1}(P)$.

EXERCISE 2.13 (transversality theorem). Let $f : U \subset M \times N \to P$ be a smooth map, where $U \subset M \times N$ is open and M, N, P are differentiable manifolds. For every $y \in N$, consider the map $f_y : U_y \subset M \to P$ defined by $f_y = f(x, y)$, where $U_y = \big\{x \in M : (x, y) \in U\big\}$. Show that if f is transversal to P then f_y is transversal to P for almost every $y \in N$ (*hint*: apply Sard's theorem to the restriction of the projection $M \times N \to N$ to the submanifold $f^{-1}(P) \subset U$).

EXERCISE 2.14. Let X, Y be topological spaces. A map $f : X \to Y$ is called a *covering map* if every $y \in Y$ admits an open neighborhood $V \subset Y$ such that $f^{-1}(V)$ can be written as a disjoint union $f^{-1}(V) = \bigcup_{i \in I} U_i$ where each U_i is open in X and f maps U_i homeomorphically *onto* V. Show that:

(a) every covering map is a local homeomorphism.
(b) If Y is connected and $f : X \to Y$ is a covering map then the cardinality of $f^{-1}(y)$ is independent of $y \in Y$. In particular, if Y is connected and $X \neq \emptyset$ then every covering map $f : X \to Y$ is surjective.
(c) Assume that X and Y are Hausdorff and that Y satisfies either one of the following:
   - Y is first countable, i.e., every point has a countable fundamental system of neighborhoods;
   - Y is locally compact;

   then every proper map $f : X \to Y$ which is a local homeomorphism is a covering map (*hint*: f is closed by Exercise 1.76).

EXERCISE 2.15. Let S_1, S_2 be finite dimensional real vector spaces and let ω_1, ω_2 be volume forms for S_1 and S_2 respectively. Set $S = S_1 \oplus S_2$ and denote by $\pi_1 : S \to S_1$, $\pi_2 : S \to S_2$ the projections. Show that $\omega = (\pi_1^*\omega_1) \wedge (\pi_2^*\omega_2)$ is a volume form on S such that if $(b_i)_{i=1}^k$ is a basis for S_1 and $(b_i')_{i=1}^l$ is a basis for S_2 then:

$$\omega(b_1, \ldots, b_k, b_1', \ldots, b_l') = \omega_1(b_1, \ldots, b_k)\omega_2(b_1', \ldots, b_l').$$

We call ω the *direct sum* of the volume forms ω_1 and ω_2 and we write $\omega = \omega_1 \oplus \omega_2$. Prove a version of the result above for volume densities in the following sense: if $\mathcal{O}_i$ is an orientation for S_i, $i = 1, 2$, then there exists a unique orientation $\mathcal{O}_1 \times \mathcal{O}_2$ in S for which the concatenation of an $\mathcal{O}_1$-positive basis of S_1 with an $\mathcal{O}_2$-positive basis of S_2 is $(\mathcal{O}_1 \times \mathcal{O}_2)$-positive. Show that if $\delta_i = [\mathcal{O}_i, \omega_i]$ is a volume density in S_i, $i = 1, 2$, then $\delta = [\omega_1 \oplus \omega_2, \mathcal{O}_1 \times \mathcal{O}_2]$ is a well-defined volume density in S. We call δ the *direct sum* of the volume densities δ_1 and δ_2 and we write $\delta = \delta_1 \oplus \delta_2$.

EXERCISE 2.16. Let $T : V \to W$ be a linear operator, where V, W are finite dimensional real vector spaces; set $k = \dim(\mathrm{Ker}T)$, $l = \dim(\mathrm{Im}T)$ and $n = k + l = \dim(V)$. Suppose we are given volume form ω_1 on $\mathrm{Ker}(T)$ and a volume form ω_2 on $\mathrm{Im}(T)$. For every subspace $W \subset V$ complementary to $\mathrm{Ker}T$ define a volume form ω on $V = \mathrm{Ker}(T) \oplus W$ and the direct sum of ω_1 and $(T|_W)^*\omega_2$. Show that:

- $\omega = \pi_{\mathrm{Ker}}^*\omega_1 \wedge T^*\omega_2$, where $\pi_{\mathrm{Ker}} : V \to \mathrm{Ker}(T)$ denotes the projection with respect to the decomposition $V = \mathrm{Ker}(T) \oplus W$;
- ω is the pull-back of $\omega_1 \oplus \omega_2$ by the isomorphism $\phi_W : V \to \mathrm{Ker}T \oplus \mathrm{Im}(T)$ defined by $\phi_W = (\pi_{\mathrm{Ker}}, T)$;
- ω does not depend on the choice of W (*hint*: given another complementary subspace W', the determinant of $\phi_{W'} \circ \phi_W^{-1}$ is equal to 1);

Prove a version of the result above for volume densities.

EXERCISE 2.17. Let M be a (semi-)Riemannian manifold, δ the canonical volume density of M and X a smooth vector field on M. The *divergence* of X is the scalar function $\mathrm{div}X : M \to \mathbb{R}$ defined by $\mathrm{div}X(x) = \mathrm{tr}\nabla X(x)$. Show that the Lie derivative $\mathbb{L}_X\delta$ equals $(\mathrm{div}X)\delta$ (*hint*: if $(X_i)_{i=1}^n$ is a local orthonormal frame for M and ω is the n-form that corresponds to δ in the orientation determined by $(X_i)_{i=1}^n$, show that $\mathbb{L}_X\omega(X_1,\dots,X_n) = \mathrm{div}X$).

EXERCISE 2.18. Let M be a Riemannian manifold with boundary and X a smooth vector field on M with compact support. Show that:

$$\int_M \mathrm{div}X \, \mathrm{d}\mu_\delta = \int_{\partial M} \langle X, N\rangle \, \mathrm{d}\mu_{\delta'},$$

where δ and δ' denote the canonical volume densities of M and ∂M respectively and $N : \partial M \to TM$ is the unit outward pointing normal vector field along ∂M (*hint*: apply Stoke's theorem to the density $i_X\delta$).

EXERCISE 2.19. Show that the volume of the unit ball $\overline{\mathrm{B}}^N$ is given by:

$$\mathrm{vol}\big(\overline{\mathrm{B}}^N\big) = \begin{cases} \dfrac{\pi^{N/2}}{\big(N/2\big)!}, & \text{if } N \text{ is even},\\[2ex] \dfrac{2^N \pi^{\frac{N-1}{2}}\left(\dfrac{N-1}{2}\right)!}{N!}, & \text{if } N \text{ is odd}.\end{cases}$$

Apply Divergence's theorem to the identity vector field of $\overline{\mathrm{B}}^N$ to conclude that:

$$\mathrm{vol}\big(S^{N-1}\big) = N \cdot \mathrm{vol}\big(\overline{\mathrm{B}}^N\big).$$

EXERCISE 2.20. Let $B : \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{k \text{ times}} \to \mathbb{R}$ be a k-linear map. The *trace* of B is the $(k-2)$-linear map $\mathrm{tr}B$ defined by:

$$\mathrm{tr}B(x_1,\dots,x_{k-2}) = \sum_{i=1}^m B(x_1,\dots,x_{k_2},e_i,e_i),$$

where $(e_i)_{i=1}^n$ is an orthonormal basis of $\mathbb{R}^n$.

(a) show that if k is odd then:

$$\int_{S^{n-1}} B\big((x)^{(k)}\big) \, \mathrm{d}\mu_\sigma(x) = 0,$$

where $x^{(k)} = (\underbrace{x, \ldots, x}_{k \text{ times}})$ and σ denotes the canonical volume density of S^{n-1}.

(b) Assume that B is symmetric. Define a vector field X on $\mathbb{R}^n$ such that:

$$\langle X(x), v\rangle = B(\underbrace{x, \ldots, x}_{k-1 \text{ times}}, v),$$

for all $x, v \in \mathbb{R}^n$, where $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product. Apply divergence's Theorem for X on the ball $\overline{\mathrm{B}}^n$ to conclude that:

$$\int_{S^{n-1}} B\big(x^{(k)}\big)\,\mathrm{d}\mu_\sigma(x) = (k-1)\int_{\overline{\mathrm{B}}^n} (\mathrm{tr}B)\big(x^{(k-2)}\big)\,\mathrm{d}x.$$

(c) given an integrable map $\phi : \overline{\mathrm{B}}^n \to \mathbb{R}$, show that:

$$\int_{\overline{\mathrm{B}}^n} \phi = \int_0^1 \left(\int_{S^{n-1}} \phi(xt)t^{n-1}\,\mathrm{d}\mu_\sigma(x)\right)\,\mathrm{d}t.$$

(d) Assuming B symmetric and $k \geq 2$, show that:

$$\int_{S^{n-1}} B\big(x^{(k)}\big)\,\mathrm{d}\mu_\sigma(x) = \frac{k-1}{k+n-2}\int_{S^{n-1}} (\mathrm{tr}B)\big(x^{(k-2)}\big)\,\mathrm{d}\mu_\sigma(x).$$

(e) Conclude that if $k \geq 2$ is even and B is symmetric:

$$\begin{aligned}\int_{S^{n-1}} & B\big(x^{(k)}\big)\,\mathrm{d}\mu_\sigma(x) \\ &= \left(\frac{k-1}{k+n-2}\right)\left(\frac{k-3}{k+n-4}\right)\cdots\left(\frac{3}{n+2}\right)\frac{1}{n}\big(\mathrm{tr}^{k/2}B\big)\mathrm{vol}\big(S^{n-1}\big).\end{aligned}$$

EXERCISE 2.21. Let $\pi : E \to M$ be a Riemannian vector bundle over a differentiable manifold M. Show that

$$E^1 = \big\{\xi \in E : \|\xi\| = 1\big\}$$

is a submanifold of E and that $\pi|_{E^1} : E^1 \to M$ is a fiber bundle over M. If ∇ is a connection on E for which the Riemannian structure is parallel, show that for all $\xi \in E^1$ the tangent space $T_\xi E^1$ is given by:

$$T_\xi E^1 = \mathrm{Hor}_\xi\, E \oplus \xi^\perp \subset \mathrm{Hor}_\xi\, E \oplus E_{\pi(\xi)} = T_\xi E.$$

EXERCISE 2.22. Let E be a vector bundle over a differentiable manifold M and let ∇ be a connection on E. If $(\xi_i)_{i=1}^k$ is a local referential of E defined in an open subset U of M then we define $\mathrm{gl}(k, \mathbb{R})$-valued differential forms ω and Ω on U by setting:

$$\omega_{ij}(v) = \theta_i\big(\nabla_v \xi_j\big),$$
$$\Omega_{ij}(v, w) = \theta_i\big(R(v, w)\xi_j\big),$$

for $i, j = 1, \ldots, k$, $v, w \in T_xM$, $x \in M$, where $(\theta_i)_{i=1}^k$ denotes the dual referential of $(\xi_i)_{i=1}^k$ and R denotes the curvature tensor of ∇. The forms ω and Ω are called respectively the *connection form* and the *curvature form* of ∇ with respect to the referential $(\xi_i)_{i=1}^k$. We write $\theta = (\theta_i)_{i=1}^k$ and if ξ is a section of E we will denote by $\theta(\xi)$ the map $U \ni x \mapsto \theta_x(\xi_x) \in \mathbb{R}^k$ which is simply the coordinate representation of ξ in the referential $(\xi_i)_{i=1}^k$. Show that:

(a) if $\xi \in \Gamma(E|_U)$ then the covariant derivative of ξ in a direction $v \in T_xM$, $x \in U$, is given in coordinates by the formula:

$$\theta_x\big(\nabla_v \xi\big) = v\big[\theta(\xi)\big] + \omega(v) \cdot \theta(\xi),$$

where v acts on the $\mathbb{R}^k$-valued map $\theta(\xi)$ as a directional derivative operator and $\omega(v)$ is thought of as a linear endomorphism of $\mathbb{R}^k$.

(b) The following identity holds:

$$\Omega_{ij} = \mathrm{d}\omega_{ij} + \sum_{r=1}^{k} \omega_{ir} \wedge \omega_{rj}, \tag{2.6.4}$$

for $i, j = 1, \ldots, k$.

(c) If a vector bundle morphism $\iota : TM \to E$ is given, we define the *torsion* of ∇ with respect to ι as the tensor $T \in \Gamma(TM^* \otimes TM^* \otimes E)$:

$$T(X, Y) = \nabla_X \iota(Y) - \nabla_Y \iota(X) - \iota([X, Y]), \quad X, Y \in \Gamma(TM).$$

The *torsion form* of ∇ (and ι) with respect to the local referential $(\xi_i)_{i=1}^k$ is the $\mathbb{R}^k$-valued 2-form Θ on U defined by:

$$\Theta_i(v, w) = \theta_i\big(T(v, w)\big),$$

for $i = 1, \ldots, k$. The following identity holds:

$$\Theta_i = \mathrm{d}(\theta_i \circ \iota) + \sum_{r=1}^{k} \omega_{ir} \wedge (\theta_r \circ \iota), \quad i = 1, \ldots, k, \tag{2.6.5}$$

where $\theta_i \circ \iota$ is regarded as a 1-form on U.

(d) If E is endowed with a Riemannian structure which is parallel with respect to ∇, show that ω and Ω take values in $\mathrm{so}(k)$, i.e., $\omega_{ij} = -\omega_{ji}$ and $\Omega_{ij} = -\Omega_{ji}$ for all $i, j = 1, \ldots, k$.

Critical points and Morse functions.

EXERCISE 2.23. Let $f : \widetilde{V} \to \mathbb{R}$, $\alpha : \widetilde{U} \to \mathbb{R}^n$ be smooth maps, where $\widetilde{V}$ is open in $\mathbb{R}^n$, $\widetilde{U}$ is open in $\mathbb{R}^m$ and $\alpha(\widetilde{U}) \subset \widetilde{V}$. Show that for every $x \in \widetilde{U}$ the following holds:

$$\mathrm{Hess}(f \circ \alpha)_x = \mathrm{d}\alpha(x)^*\big(\mathrm{Hess} f_{\alpha(x)}\big) + \mathrm{d}f\big(\alpha(x)\big) \circ \mathrm{Hess}\alpha_x.$$

Conclude that if $f : M \to \mathbb{R}$ is a smooth map on a manifold M, $x \in M$ is a point and $\varphi : U \subset M \to \widetilde{U} \subset \mathbb{R}^n$ is a local chart with $x \in U$ then the symmetric bilinear form $\mathrm{d}\varphi(x)^*\big(\mathrm{Hess}(f \circ \varphi^{-1})_{\varphi(x)}\big)$ in T_xM does not depend on the choice of the chart φ if and only if x is a critical point of f.

EXERCISE 2.24. Let $f : M \to \mathbb{R}$ be a smooth map and $x \in M$ a critical point of f.

- Show that the Hessian of f at x (introduced in Definition 2.1.1) equals the symmetric bilinear form $\mathrm{d}\varphi(x)^*\big(\mathrm{Hess}(f\circ\varphi^{-1})_{\varphi(x)}\big)$, given in Exercise 2.23;
- Show that for any smooth curve γ in M with $\gamma(0)=x$ we have
$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}(f\circ\gamma)(0)=\mathrm{Hess}f_x\big(\gamma'(0),\gamma'(0)\big).$$
- Identify M with the zero section of TM^* and consider the canonical decomposition $T_xTM^*=T_xM\oplus T_xM^*$. Show that the second coordinate of $\mathrm{d}(\mathrm{d}f)_x:T_xM\to T_xM\oplus T_xM^*$ is identified with $\mathrm{Hess}f_x$.

EXERCISE 2.25. Let $f:U\to\mathbb{R}$ be a smooth map defined on an open subset $U\subset\mathbb{R}^n$. Show that if $\mathrm{Hess}f_x$ is nondegenerate for some $x\in U$ then $\mathrm{d}f:U\to\mathbb{R}^{n*}$ is a diffeomorphism in an open neighborhood of x in U. Conclude that nondegenerate critical points are isolated in the set of critical points.

EXERCISE 2.26. Let $f:M\to\mathbb{R}$ be a smooth map on a differentiable manifold M.

- Show that f is a Morse functions if and only if the map $\mathrm{d}f:M\to TM^*$ is transversal to the zero section.
- Let $\phi:M\to\mathbb{R}^n$ be a smooth immersion[3]. Define $F:\mathbb{R}^{n*}\times M\to\mathbb{R}$ by $F(\alpha,x)=f(x)+\alpha\big(\phi(x)\big)$. Show that the map $\frac{\partial F}{\partial x}:\mathbb{R}^{n*}\times M\to TM^*$ is transversal to the zero section of TM^*.
- Conclude from the Transversality Theorem (see Exercise 2.13) that the map $F(\alpha,\cdot):M\to\mathbb{R}$ is a Morse function for almost every $\alpha\in\mathbb{R}^{n*}$.
- By observing that one can choose ϕ to be bounded, show that every smooth function $f:M\to\mathbb{R}$ is the uniform limit of Morse functions.
- Recalling that every continuous map $f:M\to\mathbb{R}$ is the uniform limit of smooth maps, conclude that the set of Morse functions is dense in the space of continuous maps $f:M\to\mathbb{R}$ endowed with the uniform convergence topology.

The passage through a critical level.

EXERCISE 2.27. Let $\rho:[0,1]\to\mathbb{R}$ be a continuous function such that $\rho(x)<1$ for $x\in[0,1[$ and $\rho(1)=1$. Consider the triangle T with vertices $(0,0)$, $(1,1)$ and $(0,1)$ and let $\widetilde{T}$ be the region:
$$\widetilde{T}=\big\{(x,y)\in\mathbb{R}^2:0\le x\le 1,\ \rho(x)\le y\le 1\big\}.$$
Let $h:\widetilde{T}\to T$ be the unique map such that $h(x,\rho(x))=(x,x)$, $h(x,1)=(x,1)$ and $h(x,\cdot)$ is affine for every $x\in[0,1]$. Show that h is a homeomorphism.

EXERCISE 2.28. Let $\sigma_1:[0,1]\to\,]0,+\infty[$ and $\sigma_2:\big[\frac12,1\big]\to\mathbb{R}$ be continuous functions such that $\sigma_1(0)=1$, $\sigma_2\big(\frac12\big)=0$, $\sigma_1(1)=\sigma_2(1)$ and $\sigma_2(x)<\sigma_1(x)$

[3] Whitney's Theorem yields the existence of a smooth immersion $\phi:M\to\mathbb{R}^n$ for $n\ge 2\dim(M)$ and the existence of a smooth embedding $\phi:M\to\mathbb{R}^n$ for $n\ge 2\dim(M)+1$.

for all $x \in \left[\frac{1}{2}, 1\right[$. Consider the region R given by:

$$R = \Big\{(x,y) \in \mathbb{R}^2 : 0 \le x \le \frac{1}{2},\ 0 \le y \le \sigma_1(x)\Big\} \\ \cup \Big\{(x,y) \in \mathbb{R}^2 : \frac{1}{2} \le x \le 1,\ \sigma_2(x) \le y \le \sigma_1(x)\Big\}.$$

Show that there exists a homeomorphism $h : R \to \left[0, \frac{1}{2}\right] \times [0,1]$ that fixes the points of $\left[0, \frac{1}{2}\right] \times \{0\} \cup \{0\} \times [0,1]$.

Hint:

- Consider the map $h_1 : R \to \mathbb{R}^2$ such that $h_1(x,\cdot)$ is affine, $h_1(x,0) = (x,0)$ and $h_1(x,\sigma_1(x)) = (x,1)$ for every $x \in [0,1]$. Then h_1 is a homeomorphism onto the region $R' = \big(\left[0, \frac{1}{2}\right] \times [0,1]\big) \cup \widetilde{T}$, where $\widetilde{T}$ is given by:
$$\widetilde{T} = \Big\{(x,y) \in \mathbb{R}^2 : \frac{1}{2} \le x \le 1,\ \frac{\sigma_2(x)}{\sigma_1(x)} \le y \le 1\Big\}.$$
- Use Exercise 2.27 to obtain a homeomorphism $h_2 : \widetilde{T} \to T$ that fixes the points of $\left\{\frac{1}{2}\right\} \times [0,1]$, where T is the triangle with vertices $\left(\frac{1}{2}, 0\right)$, $\left(\frac{1}{2}, 1\right)$ and $(1,1)$. Extend h_2 to R' by setting $h_2 = \mathrm{Id}$ on $\left[0, \frac{1}{2}\right] \times [0,1]$, obtaining a homeomorphism from R' to $R'' = \big(\left[0, \frac{1}{2}\right] \times [0,1]\big) \cup T$.
- Define $h_3 : R'' \to \left[0, \frac{1}{2}\right] \times [0,1]$ to be the homeomorphism such that $h_3(\cdot, y)$ is affine, $h_3(0,y) = (0,y)$ and $h_3\big(\frac{y+1}{2}, y\big) = \big(\frac{1}{2}, y\big)$ for all $y \in [0,1]$.
- Set $h = h_3 \circ h_2 \circ h_1$.

The CW-complex associated to a Morse function.

EXERCISE 2.29. Given non negative integers ν, μ, show that $\big(\overline{\mathrm{B}}^{\nu} \times \{0\}\big) \cup \big(S^{\nu-1} \times \overline{\mathrm{B}}^{\mu}\big)$ is a strong deformation retract of $\overline{\mathrm{B}}^{\nu} \times \overline{\mathrm{B}}^{\mu}$.

EXERCISE 2.30. Let $f : M \to \mathbb{R}$ be a Morse function on a compact manifold M. Show that the map $M \ni x \mapsto F(+\infty, x) \in M$ is not continuous.

EXERCISE 2.31. If x is a nondegenerate saddle point of f with $f(x) = a$, show that the map λ_a has no continuous extension to x.

EXERCISE 2.32. Consider the map $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x,y) = \frac{1}{2}(x^2 - y^2)$. Compute the arrival time map λ_0, identifying its domain.

EXERCISE 2.33. The goal of this exercise is to give a fancier proof of the continuity of the map G in the proof of Proposition 2.3.9. Let ω be an arbitrary point not in $[-\infty, +\infty]$ and define a topology on the set $\widetilde{\mathbb{R}} = [-\infty, +\infty] \cup \{\omega\}$ as follows; the open subsets of $\widetilde{\mathbb{R}}$ are the open subsets of $[-\infty, +\infty]$ and the sets of the form $U \cup \{\omega\}$ with U an open subset of $[-\infty, +\infty]$ containing $[0, +\infty]$.

- given $a \in \mathbb{R}$, then setting $\lambda_a(x) = \omega$ for $x \in \mathrm{Crit}_f(a)$, show that the map:
$$\lambda_a : \big\{x \in D_a : f(x) \ge a\big\} \cup \mathrm{Crit}_f(a) \longrightarrow \widetilde{\mathbb{R}}$$

is continuous;

- set $F(\omega, x) = x$ for $x \in \mathrm{Crit}_f(a)$ and, under the notations and hypothesis of Proposition 2.3.9, show that the restriction of F to the set:

  $$\big\{(t, x) : x \in S \setminus \mathrm{Crit}_f(a),\ t \in \big[0, \lambda_a(x)\big]\big\} \cup \big(\widetilde{\mathbb{R}} \times \mathrm{Crit}_f(a)\big),$$

  is continuous;
- show that G is continuous.

EXERCISE 2.34. Let $f : M \to \mathbb{R}$ be a Morse function on a compact Riemannian manifold (M, g) and let $p, q \in M$ be critical points with $\mu(p) - \mu(q) = 1$ and $W_{\mathrm{u}}(p)$ transversal to $W_{\mathrm{s}}(q)$. Show that if $\gamma : \mathbb{R} \to M$ is a flow line of $-\nabla f$ going from p to q then there exists an open subset $U \subset M$ with $\mathrm{Im}(\gamma) = U \cap \big(W_{\mathrm{u}}(p) \cap W_{\mathrm{s}}(q)\big)$.

CHAPTER 3

Applications of Morse Theory in the Compact Case

In this chapter we will present some applications of Morse Theory for compact manifolds to the theory of submanifolds of a Euclidean spaces.

The first application is a generalized version of the standard Gauss–Bonnet theorem for compact surfaces. Recall that the Gauss–Bonnet theorem states that the integral of the Gaussian curvature of a compact surface M equals $2\pi\chi(M)$, where $\chi(M)$ is the Euler characteristic of M. The generalized version of this result, called the Gauss–Bonnet–Chern theorem, holds for an arbitrary even-dimensional compact Riemannian manifolds and it states that the Euler characteristic of M can be obtained as the integral of a suitable density on M defined in terms of the curvature tensor of M.

Then we present the Theorem of Chern and Lashof, which gives a characterization the isometric immersions of Riemannian manifolds in a Euclidean space having minimal *absolute total curvature*.

We then give a topological characterization of those compact Riemannian manifolds having positive sectional curvature and that admit an isometric immersion in codimension one and two. Finally we discuss generalizations of the above situations to a class of hypersurfaces, that we call *quasi-convex*, that includes the *conformally flay hypersurfaces* and the hypersurfaces with *nonnegative isotropic curvature*.

3.1. The Fundamental Equations of an Isometric Immersion

Let (M,g), $(\overline{M},\overline{g})$ be Riemannian manifolds and let $f:(M,g)\to(\overline{M},\overline{g})$ be an isometric immersion, i.e., $f:M\to\overline{M}$ is an immersion and g is the pull-back of $\overline{g}$ by f. The inner products g and $\overline{g}$ will be usually denoted simply by $\langle\cdot,\cdot\rangle$. We denote by ∇ the Levi–Civita connection of M and by $\overline{\nabla}$ the Levi–Civita connection of $\overline{M}$. For every $x\in M$, the tangent space $T_{f(x)}\overline{M}$ is the direct sum of the spaces $\mathrm{d}f_x(T_xM)\cong T_xM$ and its orthogonal complement $\mathrm{d}f_x(T_xM)^\perp$ in $T_{f(x)}M$. The space $\mathrm{d}f_x(T_xM)$ will be identified with the tangent space T_xM and the space $\mathrm{d}f_x(T_xM)^\perp$, denoted by ν_xM, is called the *normal space* corresponding to the immersion f at the point x. In the language of vector bundles we can describe this situation as follows. The differential of f induces an injective vector bundle morphism from the tangent bundle TM of M to the pull-back $f^*T\overline{M}$; this vector bundle morphism gives an isomorphism from TM to a vector subbundle of $f^*T\overline{M}$, that will be identified with TM. The spaces $\nu_xM\subset(f^*T\overline{M})_x=T_{f(x)}\overline{M}$ form another vector subbundle νM of $f^*T\overline{M}$ and we have a $\overline{g}$-orthogonal direct sum decomposition of vector bundles:

$$f^*T\overline{M} = TM \oplus \nu M.$$

We call νM the *normal bundle* of the immersion f. Given $x \in M$ and $z \in T_{f(x)}\overline{M}$ we denote by z^T and by $z^\perp$ respectively the components of z in $\mathrm{d}f_x(T_xM) \cong T_xM$ and in ν_xM.

Let X, Y be smooth local sections of TM and ξ a smooth local section of νM. It is easily seen that:

- $\nabla_X Y = (\overline{\nabla}_X Y)^T$.
- $\nabla^\perp_X \xi := (\overline{\nabla}_X \xi$ is a Riemannian connection on νM called the *normal connection*.

Set $\alpha(X,Y) = (\overline{\nabla}_X Y)^\perp$, and $A_\xi X = -(\overline{\nabla}_X \xi)^T$. An easy computation gives:

- $\alpha(X,Y) = \alpha(Y,X)$ and $\alpha(X,Y)$ at $x \in M$ depends only on $X(x)$ and $Y(x)$. In particular it defines, $\forall x \in M$, a symmetric bilinear map:
$$\alpha_x : T_xM \oplus T_xM \to \nu_xM,$$
called the *second fundamental form* of f at x.
- $A_\xi X$ at $x \in M$ depends only on $\xi(x)$ and $X(x)$, hence define a symmetric linear map:
$$A_\xi : T_xM \to T_xM,$$
called the *Weingarten (or shape) operator* in the ξ direction. The eigenvalues of A_ξ are called the *principal curvatures*.
- $< \alpha(X,Y), \xi >=< A_\xi X, Y >$.

Resuming the situation, we have the so called *Formulas of Gauss and Weingarten*:

$$\overline{\nabla}_X Y = \nabla_X Y + \alpha(X,Y),$$
$$\overline{\nabla}_X \xi = -A_\xi X + \nabla^\perp_X \xi.$$

A simple computation yields the following:

3.1.1. PROPOSITION. *If R, $\overline{R}$ and $R^\perp$ denote respectively the curvature tensors of ∇, $\overline{\nabla}$ and $\nabla^\perp$ then the following identities hold:*

$$\begin{aligned}&\big\langle \overline{R}(X,Y)Z,T\big\rangle \\ &\quad= \big\langle R(X,Y)Z,T\big\rangle + \big\langle \mathrm{II}(X,Z),\mathrm{II}(Y,T)\big\rangle - \big\langle \mathrm{II}(X,T),\mathrm{II}(Y,Z)\big\rangle,\end{aligned} \tag{Gauss}$$

$$\big\langle \overline{R}(X,Y)Z,\eta\big\rangle = \big\langle \big(\nabla^\otimes_X \mathrm{II}\big)(Y,Z) - \big(\nabla^\otimes_Y \mathrm{II}\big)(X,Z),\eta\big\rangle, \tag{Codazzi}$$

$$\big\langle \overline{R}(X,Y)\xi,\eta\big\rangle = \big\langle R^\perp(X,Y)\xi,\eta\big\rangle + \big\langle [A_\eta, A_\xi]X,Y\big\rangle, \tag{Ricci}$$

for all $X,Y,Z,T \in T_xM$, $\xi,\eta \in \nu_xM$, $x \in M$, where $[A_\eta, A_\xi] = A_\eta A_\xi - A_\xi A_\eta$ and $\nabla^\otimes$ denotes the connection induced by ∇ and $\nabla^\perp$ in the tensor bundle $TM^* \otimes TM^* \otimes \nu M$,i.e.:

$$\langle\big(\nabla^\otimes_X \mathrm{II}\big)(Y,Z) = \nabla^\perp_X(\mathrm{II}(Y,Z)) - \mathrm{II}(\nabla_X Y, Z) - \mathrm{II}(Y, \nabla_X Z).$$

□

For a generalization of Proposition 3.1.1 to the context of general direct sum decompositions of vector bundles endowed with connections, see Exercise **??**.

The equations of Gauss–Codazzi–Ricci are called the *fundamental equations* of the isometric immersion due to the following:

3.1.2. THEOREM. *Let M be a simply-connected (and connected) Riemannian manifold and let E be a Riemannian vector bundle over M; we denote by $\langle\cdot,\cdot\rangle$ both the inner product on the tangent spaces of M and on the fibers of E. We also denote by $\langle\cdot,\cdot\rangle$ the inner product on the fibers of $TM \oplus E$ that correspond to the orthogonal direct sum of the Riemannian structures of TM and E. Suppose we are given a connection ∇^E on E and a smooth tensor field $\mathrm{II}^E \in \Gamma(TM^* \otimes TM^* \otimes E)$ that is symmetric with respect to the first two variables. We denote by ∇ the Levi–Civita connection of M and also the connection on $TM^* \otimes TM^* \otimes E$ induced by ∇ and ∇^E; by R^E we denote the curvature tensor of ∇^E. For $x \in M$, $\xi \in E_x$, we denote by II^E_ξ the symmetric bilinear form on T_xM given by $\big\langle \mathrm{II}^E(\cdot,\cdot),\xi\big\rangle$ and by A^E_ξ the symmetric linear endomorphism of T_xM that represents such bilinear form. Fix $c \in \mathbb{R}$ and denote by $\mathbb{S}^c$ the complete, simply connected space of constant sectional curvature c and dimension $n+dimE$; for $x \in M$, $v_1, v_2, v_3 \in T_xM \oplus E_x$ set:*

$$\overline{R}(v_1,v_2)v_3 = c\big[\langle v_2,v_3\rangle v_1 - \langle v_1,v_3\rangle v_2\big].$$

Assume that the equations of Gauss, Codazzi and Ricci are satisfied with II, II_ξ, A_ξ and $R^\perp$ replaced by II^E, II^E_ξ, A^E_ξ and R^E respectively. Then, there exists an isometric immersion $f : M \to \mathbb{S}^c$ and a Riemannian vector bundle isometry ϕ from E to the normal bundle νM of the immersion f that carries ∇^E to the normal connection $\nabla^\perp$ and II^E to the second fundamental form of the immersion f. Any other such pair $(\tilde{f},\tilde{\phi})$ differs from (f,ϕ) only by left composition with a global isometry of $\mathbb{S}^c$.

3.1.3. REMARK. For the uniqueness part of Theorem 3.1.2 (up to global isometries of the space form) it suffices to assume that M is connected; simply-connectedness is used only for the existence.

3.1.4. REMARK. The theorem above tell us that, similarly to what happen with curves whose local geometry is completely described by the Frenet formulas, the local geometry of an isometric immersion into a space form is completely determined by the fundamental equations.

3.2. Absolute Total Curvature and Height Functions

In this section M denotes an n-dimensional Riemannian manifold and $f : M \to \mathbb{R}^{n+p}$ denotes an isometric immersion. In this case the equations of Gauss, Codazzi and Ricci can be written in a simplified form using the fact that $\overline{R} = 0$.

3.2.1. DEFINITION. Denote by $\nu^1 M$ the unitary normal bundle of the immersion f, i.e.:

$$\nu^1 M = \big\{(x,\xi) \in \nu M : \|\xi\| = 1\big\}.$$

The unitary normal bundle is a submanifold of νM (see Exercise 2.21) and the map

$$\mathfrak{G} : \nu^1 M \longrightarrow S^{n+p-1} \subset \mathbb{R}^{n+p}$$

defined by $\mathfrak{G}(x, \xi) = \xi$ is smooth. We call $\mathfrak{G}$ the (generalized) *Gauss map* of the immersion f.

Observe that $\mathfrak{G} : \nu^1 M \to S^{n+p-1}$ is a map between manifolds of the same dimension.

3.2.2. REMARK. If $p = 1$, $\nu^1 M$ is the orientation covering of M. So the manifold is orientable if and only if $\nu^1 M$ is disconnected hence diffeomorphic to two copies of M and, in this case, the classical Gauss map is the restriction of $\mathfrak{G}$ to a connected component of $\nu^1 M$.

The normal bundle νM is a Riemannian vector bundle with the inner product in the fibers induced from $\mathbb{R}^{n+p}$; considering such Riemannian vector bundle structure, the Riemannian metric of M and the normal connection $\nabla^\perp$, we can construct a Riemannian metric in the manifold νM as explained in Remark **??**. The unitary normal bundle $\nu^1 M$ will be considered with the Riemannian metric induced from νM. Recall from Exercise 2.21 that for $\xi \in \nu^1 M$ we have:

$$T_{(x,\xi)}\nu^1 M = \mathrm{Hor}_{(x,\xi)}\, \nu M \oplus (\xi^\perp \cap \nu_x M) \cong T_x M \oplus (\xi^\perp \cap \nu_x M). \tag{3.2.1}$$

Observe that by identifying $T_x M$ with $\mathrm{d}f_x(T_x M)$, the tangent space $T_{(x,\xi)}\nu^1 M$ is identified with $T_\xi S^{n+p-1} = \xi^\perp$.

Also the normal bundle and the unit normal bundle can be naturally immersed into $\mathbb{R}^{2(n+p)}$ by the map:

$$F : \nu M \to \mathbb{R}^{2(n+p)}, \quad F(x, \xi) = (f(x), \xi) \in \mathbb{R}^{n+p} \times \mathbb{R}^{n+p} = \mathbb{R}^{2(n+p)},$$

and the induced metric is the one described above. Is then clear that the tangent space $T_{(x,\xi)}\nu^1 M$ is spanned by frames of the type $\{X_1, \ldots, X_n, \xi_1, \ldots, \xi_{p-1}\}$ where the $X_i's$ are tangent to M and the $\xi_i's$ are a basis for $\xi^\perp \cap \nu_x M$. Choosing such a basis orthonormal and extendig it locally to an orthonormal frame field of the same type, since $\overline{\nabla}_{\xi_i}\xi_j = -\delta_{ij}\xi$, we get:

- $< X_i(\mathfrak{G}), X_j >=< \overline{\nabla}_{X_i}\xi, X_j >= - < A_\xi X_i, X_j >$,
- $< \xi_i(\mathfrak{G}), X_j >=< \overline{\nabla}_{\xi_i}\xi, X_j >= 0$,
- $< \xi_i(\mathfrak{G}), \xi_j >=< \overline{\nabla}_{\xi_i}\xi, \xi_j >= - < \overline{\nabla}_{\xi_i}\xi_j, \xi >= \delta_{ij}$.

Hence, the differential of the Gauss map has a matrix representation, in a basis of the form above, of the type:

$$\mathrm{d}\mathfrak{G}_{(x,\xi)} = \begin{pmatrix} -A_{(x,\xi)} & * \\ 0 & \mathrm{Id} \end{pmatrix}.$$

3.2.3. COROLLARY. *If $\bar{\delta}$ denotes the canonical volume density of $\nu^1 M$ and σ denotes the canonical volume density of the unit sphere S^{n+p-1} then*

$$(\mathfrak{G}^*\sigma)_{(x,\xi)} = \left|\det A_{(x,\xi)}\right| \bar{\delta}_{(x,\xi)},$$

for all $(x,\xi) \in \nu^1 M$*. In particular,* $(x,\xi) \in \nu^1 M$ *is a regular point for* $\mathfrak{G}$ *if and only if* $A_{(x,\xi)}$ *is invertible.*

We are now ready to define the absolute total curvature of an isometric immersion $f : M \to \mathbb{R}^{n+p}$, which gives a sort of global measure of how much f "bends" the manifold M inside $\mathbb{R}^{n+p}$.

3.2.4. DEFINITION. The *absolute total curvature* of the isometric immersion f is defined by:

$$\tau(f) = \frac{1}{\mathrm{vol}\big(S^{n+p-1}\big)} \int_{\nu^1 M} \mathfrak{G}^*\sigma. \tag{3.2.2}$$

From Lemma 3.2.3 we obtain immediately the following formula for $\tau(f)$:

$$\tau(f) = \frac{1}{\mathrm{vol}\big(S^{n+p-1}\big)} \int_{\nu^1 M} \big|\det A_{(x,\xi)}\big| \,\mathrm{d}\mu_{\bar{\delta}}(x,\xi) \in [0,+\infty], \tag{3.2.3}$$

where $\bar{\delta}$ denotes the canonical volume density of $\nu^1 M$. Moreover, using Fubini's theorem (Theorem **??**) it follows that:

$$\tau(f) = \frac{1}{\mathrm{vol}\big(S^{n+p-1}\big)} \int_M \left(\int_{\nu^1_x M} \big|\det A_{(x,\xi)}\big| \,\mathrm{d}\mu_{\sigma_x}(\xi) \right) \mathrm{d}\mu_\delta(x),$$

where σ_x denotes the canonical volume density of the unit sphere $\nu^1_x M$ and δ denotes the canonical volume density of M.

3.2.5. REMARK. If M is an oriented surface in $\mathbb{R}^3$ and $f : M \to \mathbb{R}^3$ is the inclusion map, then $\tau(f)$ coincides with the integral over M of the absolute value of the Gaussian curvature of M divided by 2π (see remark ??).

When M is compact, $\tau(f)$ is finite; we make the following:

3.2.6. ASSUMPTION. *In the rest of the section we will assume that* M *is compact.*

We will naw relate the absolute total curvature to the critical point of certain important functions.

3.2.7. DEFINITION. Let $\xi \in S^{n+p-1}$ be a fixed vector. We define the *height function in the* ξ *direction* as:

$$h_\xi : M \to \mathbb{R}, \quad h_\xi(x) =< f(x), \xi > .$$

Geometrically, $h_\xi(x)$ is the projection of $f(x)$ onto the oriented line $\{t\xi, t \in \mathbb{R}\}$ or, equivalently, the (oriented) height of $f(x)$ in relation to the hyperplane $\xi^\perp$. An easy computation gives:

- $dh_\xi(x)X =< df(x)X, \xi >$,
- $d^2(h_\xi)(x)(X,Y) =< \overline{\nabla}_Y X, \xi >=< A_\xi X, Y >$ if $(x,\xi) \in \nu^1_x M$.

Hence we have:

3.2.8. COROLLARY. *A point $x \in M$ is critical for h_ξ if and only if $(x, \xi) \in \nu_x^1 M$. Moreover such a critical point is nondegenerate if and only if A_ξ is nonsingular. Finally, h_ξ is a Morse function if and only if ξ is a regular value of the Gauss map.*

Let $D \subset S^{n+p-1}$ denote the set of regular values of $\mathfrak{G}$. Since M is compact, D is open; moreover, by Sard's theorem, the complement of D in S^{n+p-1} has null measure. For $k \geq 0$, we define integer valued maps $\kappa_k : D \to \mathbb{N}$ by setting:

$$\kappa_k(\xi) = \text{number of critical points of } h_\xi \text{ having index } k;$$

observe thath_ξ is a Morse function, for $\xi \in D$, and hence has only a finite number of critical points in the compact manifold M. We also set $\kappa(\xi) = \sum_{k=0}^{n} \kappa_k(\xi)$, so that:

$$\kappa(\xi) = \text{number of elements of } \mathfrak{G}^{-1}(\xi),$$

for all $\xi \in D$.

3.2.9. LEMMA. *The restriction of the Gauss map to $\mathfrak{G}^{-1}(D)$ is a (smooth) covering map onto D.*

PROOF. Follows easily from the observation that $\mathfrak{G}|_{\mathfrak{G}^{-1}(D)} : \mathfrak{G}^{-1}(D) \to D$ is a proper local diffeomorphism (see Exercise 2.14). □

3.2.10. LEMMA. *The functions κ_k and κ are continuous in D, i.e., constant in every connected component of D.*

PROOF. Let $\xi \in D$ be fixed; we write:

$$\mathfrak{G}^{-1}(\xi) = \big\{(x_i, \xi) : i = 1, \dots, r\big\},$$

where each $x_i \in M$. We are going to show that the maps κ_k are constant in a neighborhood of ξ. To this aim, we can assume that $r \geq 1$, otherwise all κ_k's are zero around ξ. Since $\mathfrak{G}|_{\mathfrak{G}^{-1}(D)} : \mathfrak{G}^{-1}(D) \to D$ is a covering map (Lemma 3.2.9), we can find an open neighborhood $V \subset D$ of ξ and, for each $i = 1, \dots, r$, an open neighborhood $U_i \subset \nu^1 M$ of (x_i, ξ) such that $\mathfrak{G}$ maps each U_i diffeomorphically onto V and $\mathfrak{G}^{-1}(V)$ is the disjoint union of the U_i's. Since $\mathrm{II}_{(x_i,\xi)}$ is nondegenerate, by continuity one can find an open neighborhood Z_i of x_i in M and an open neighborhood W_i of ξ in S^{n+p-1} such that $n_-\big(\mathrm{II}_{(x,\eta)}\big) = n_-\big(\mathrm{II}_{(x_i,\xi)}\big)$ for all $(x, \eta) \in \nu^1 M$ with $x \in Z_i$ and $\eta \in W_i$. We can obviously assume that the W_i's, $i = 1, \dots, r$ are disjoint. Now it is easy to check that the functions κ_k are constant on the open neighborhood W of ξ in S^{n+p-1} defined by:

$$W = \bigcap_{i=1}^{r} \mathfrak{G}\big(\pi^{-1}(Z_i) \cap U_i\big) \cap W_i,$$

where $\pi : \nu^1 M \to M$ denotes the canonical projection. □

For every $k = 0, \dots, n$, we set:

$$\tau_k(f) = \frac{1}{\mathrm{vol}\big(S^{n+p-1}\big)} \int_D \kappa_k(\xi)\, \mathrm{d}\mu_\sigma(\xi) \in [0, +\infty]; \tag{3.2.4}$$

it will follow from Corollary 3.2.12 below that $\tau_k(f)$ is indeed finite for all k. Moreover, observing that $\kappa_k(\xi) = \kappa_{n-k}(-\xi)$ for all $\xi \in D$ we get:

$$\tau_k(f) = \tau_{n-k}(f), \tag{3.2.5}$$

for all $k = 0, \dots, n$.

3.2.11. LEMMA. *Let $\phi : \nu^1 M \to \mathbb{R}$ be a $\mu_{\bar\delta}$-integrable function. Then the function $D \ni \xi \mapsto \sum_{x \in \mathfrak{G}^{-1}(\xi)} \phi(x,\xi) \in \mathbb{R}$ is μ_σ-integrable and the following identity holds:*

$$\int_{\nu^1 M} \big|\det A_{(x,\xi)}\big| \phi(x,\xi)\,\mathrm{d}\mu_{\bar\delta}(x,\xi) = \int_D \left(\sum_{x \in \mathfrak{G}^{-1}(\xi)} \phi(x,\xi) \right) \mathrm{d}\mu_\sigma(\xi). \tag{3.2.6}$$

PROOF. Since $\det A_{(x,\xi)}$ vanishes when (x,ξ) is a critical point of $\mathfrak{G}$ and since the set of regular points of $\mathfrak{G}$ outside $\mathfrak{G}^{-1}(D)$ has null measure (see Proposition **??**) we have:

$$\int_{\nu^1 M} \big|\det A_{(x,\xi)}\big| \phi(x,\xi)\,\mathrm{d}\mu_{\bar\delta}(x,\xi) = \int_{\mathfrak{G}^{-1}(D)} \big|\det A_{(x,\xi)}\big| \phi(x,\xi)\,\mathrm{d}\mu_{\bar\delta}(x,\xi).$$

Keeping in mind Lemmas 3.2.3 and 3.2.9, the conclusion follows by applying Fubini's theorem for covering maps (Corollary **??**) to compute the righthand side of the equality above. □

For every $k = 0, \dots, n$, we set:

$$\mathcal{U}_k = \big\{ (x,\xi) \in \nu^1 M : \mathrm{II}_{(x,\xi)} \text{ is nondegenerate and has index } k \big\};$$

it is easy to see that $\mathcal{U}_k$ is open for all k.

3.2.12. COROLLARY. *The following equalities hold:*

$$\tau_k(f) = \frac{1}{\mathrm{vol}\big(S^{n+p-1}\big)} \int_{\mathcal{U}_k} \big|\det A_{(x,\xi)}\big|\,\mathrm{d}\mu_{\bar\delta}(x,\xi) \in [0,+\infty[, \quad k = 0,\dots,n, \tag{3.2.7}$$

$$\tau(f) = \frac{1}{\mathrm{vol}\big(S^{n+p-1}\big)} \int_D \kappa(\xi)\,\mathrm{d}\mu_\sigma(\xi) = \sum_{k=0}^n \tau_k(f), \tag{3.2.8}$$

$$\begin{aligned}\int_{\nu^1 M} \det A_{(x,\xi)}\,\mathrm{d}\mu_{\bar\delta}(x,\xi) &= \frac{1}{\mathrm{vol}\big(S^{n+p-1}\big)} \int_D \sum_{k=0}^n (-1)^k \kappa_k(\xi)\,\mathrm{d}\mu_\sigma(\xi) \\ &= \sum_{k=0}^n (-1)^k \tau_k(f)\end{aligned} \tag{3.2.9}$$

where σ denotes the canonical volume density of S^{n+p-1} and $\bar\delta$ denotes the canonical volume density of $\nu^1 M$.

PROOF. Equalities (3.2.7), (3.2.8) and (3.2.9) follow respectively by taking ϕ to be the characteristic function of $\mathcal{U}_k$, ϕ constant and equal to 1 and ϕ equal to the sign of $\det A_\xi$ in Lemma 3.2.11. □

3.2.13. COROLLARY. *The Euler characteristic of M is given by:*

$$\chi(M) = \frac{1}{\mathrm{vol}\big(S^{n+p-1}\big)} \int_{\nu^1 M} \det A_{(x,\xi)} \,\mathrm{d}\mu_{\bar{\delta}}(x,\xi).$$

PROOF. Given $\xi \in D$, since h_ξ is a Morse function on M, Proposition **??** implies that $\sum_{k=0}^{n}(-1)^k \kappa_k(\xi) = \chi(M)$. The conclusion follows from (3.2.9), observing that the complement of D in S^{n+p-1} has null measure. □

3.3. The Gauss–Bonnet–Chern Theorem

Recall that the classical Gauss–Bonnet theorem states that the integral of the curvature of a compact two-dimensional Riemannian manifold M equals 2π times the Euler characteristic of M.

In this section we generalize this result to the case of a compact Riemannian manifold M whose dimension n is an arbitrary even number; we set $n = 2s$.

Recalling Exercise 2.22, we can associate to a local tangent frame field $(X_i)_{i=1}^n$ the curvature form Ω of the Levi–Civita connection, which is a $\mathrm{gl}(n,\mathbb{R})$-valued 2-form defined on the domain of the X_i's. If $(X_i)_{i=1}^n$ is an orthonormal frame field, we set:

$$\Omega_{ij}(v,w) = \big\langle R(v,w)X_j, X_i\big\rangle,$$

for $i, j = 1, \ldots, n$. Observe that $\Omega_{ij} = -\Omega_{ji}$, i.e., Ω is a $\mathrm{so}(n,\mathbb{R})$-valued 2-form. Set:

$$\gamma_0 = \frac{1}{s!2^n\pi^s} \sum_{\sigma\in\mathfrak{S}_n} (-1)^\sigma \Omega_{\sigma(1)\sigma(2)} \wedge \Omega_{\sigma(3)\sigma(4)} \wedge \ldots \wedge \Omega_{\sigma(n-1)\sigma(n)}, \tag{3.3.1}$$

where $\mathfrak{S}_n$ denotes the symmetric group on n elements and $(-1)^\sigma$ denotes the sign of the permutation σ. We denote by γ the n-density corresponding to γ_0 and to the orientation defined by $(X_i)_{i=1}^n$. Next we compute what happens with γ when one changes the orthonormal frame $(X_i)_{i=1}^n$.

3.3.1. LEMMA. *The n-density γ does not depend on the choice of the orthonormal frame $(X_i)_{i=1}^n$.*

PROOF. Let $(X_i')_{i=1}^n$ be another local orthonormal frame and write $T_{ij} = \langle X_j', X_i\rangle$, so that T is an orthogonal $n \times n$ matrix and $X_j' = \sum_{i=1}^n T_{ij}X_i$. The curvature form Ω' corresponding to $(X_i')_{i=1}^n$ is related to Ω by:

$$\Omega'_{ij} = \sum_{k_1,k_2=1}^{n} T_{k_1 i}T_{k_2 j}\Omega_{k_1 k_2}.$$

Let γ_0' be the version of γ_0 defined for the orthonormal frame $(X_i')_{i=1}^n$. We have to show that $\gamma_0' = (\det T)\gamma_0$. We compute as follows:

$$\begin{aligned}\gamma_0' &= \frac{1}{s!2^n\pi^s}\sum_{k_1,\ldots,k_n=1}^{n}\Big[\sum_{\sigma\in\mathfrak{S}_n}(-1)^\sigma T_{k_1\sigma(1)}\cdots T_{k_n\sigma(n)}\Big]\Omega_{k_1k_2}\wedge\ldots\wedge\Omega_{k_{n-1}k_n}\\ &= \frac{1}{s!2^n\pi^s}\sum_{k_1,\ldots,k_n=1}^{n}\det T^{(k_1,\ldots,k_n)}\Omega_{k_1k_2}\wedge\ldots\wedge\Omega_{k_{n-1}k_n},\end{aligned}$$

where $T^{(k_1,\ldots,k_n)}$ is the $n\times n$ matrix defined by $T_{ij}^{(k_1,\ldots,k_n)} = T_{k_ij}$. Since the determinant of $T^{(k_1,\ldots,k_n)}$ is zero when the k_i's are not distinct we can replace k_i by $\tau(i)$ with $\tau\in\mathfrak{S}_n$ and then we get:

$$\begin{aligned}\gamma_0' &= \frac{1}{s!2^n\pi^s}\sum_{\tau\in\mathfrak{S}_n}\det T^{(\tau(1),\ldots,\tau(n))}\Omega_{\tau(1)\tau(2)}\wedge\ldots\wedge\Omega_{\tau(n-1)\tau(n)}\\ &= \frac{1}{s!2^n\pi^s}\sum_{\tau\in\mathfrak{S}_n}(-1)^\tau(\det T)\Omega_{\tau(1)\tau(2)}\wedge\ldots\wedge\Omega_{\tau(n-1)\tau(n)} = (\det T)\gamma_0. \quad\square\end{aligned}$$

We have proven that γ is a (global) smooth n-density on M; formula (3.3.1) is sometimes used to define the so called *Euler class* of the tangent bundle TM (see [**77**, §5, Chap. XII]). Since γ is an n-density, there exists a smooth function $K : M \to \mathbb{R}$ such that $\gamma = K\delta$, where δ is the canonical volume density of M. We have the following:

3.3.2. LEMMA. *If $f : M \to \mathbb{R}^{n+p}$ is an isometric immersion then for every $x\in M$:*

$$K(x) = \frac{1}{\operatorname{vol}\big(S^{n+p-1}\big)}\int_{\nu_x^1 M}\det A_{(x,\xi)}\,\mathrm{d}\mu_{\sigma_x}(\xi), \tag{3.3.2}$$

where σ_x denotes the canonical volume density of the sphere $\nu_x^1 M$.

PROOF. Using the Gauss equation in the language of differential forms (see Exercise 3.4), since $\overline{\Omega} = 0$, we obtain:

$$\Omega_{ij} = \sum_{\alpha=1}^{p} A_{\xi_\alpha}(X_i)\wedge A_{\xi_\alpha}(X_j), \quad i,j = 1,\ldots,n,$$

where $(\xi_\alpha)_{\alpha=1}^p$ is a local orthonormal frame of νM around x and the vector $A_{\xi_\alpha}(X_i)$ is identified with the covector $\big\langle A_{\xi_\alpha}(X_i),\cdot\big\rangle$. We can now write γ_0 as:

$$\begin{aligned}\gamma_0 = \frac{1}{s!2^n\pi^s}&\sum_{\sigma\in\mathfrak{S}_n}\sum_{\alpha_1,\ldots,\alpha_s=1}^{p}\\ &(-1)^\sigma A_{\xi_{\alpha_1}}(X_{\sigma(1)})\wedge A_{\xi_{\alpha_1}}(X_{\sigma(2)})\wedge\ldots\wedge A_{\xi_{\alpha_s}}(X_{\sigma(n-1)})\wedge A_{\xi_{\alpha_s}}(X_{\sigma(n)});\end{aligned}$$

hence:

$$K(x) = \gamma_0(X_1,\ldots,X_n) = \frac{1}{s!2^n\pi^s}\sum_{\sigma,\tau\in\mathfrak{S}_n}\sum_{\alpha_1,\ldots,\alpha_s=1}^{p}(-1)^{\sigma\tau} \langle A_{\xi_{\alpha_1}}(X_{\sigma(1)},X_{\tau(1)}\rangle\langle A_{\xi_{\alpha_1}}(X_{\sigma(2)},X_{\tau(2)}\rangle \cdots\langle A_{\xi_{\alpha_s}}(X_{\sigma(n-1)},X_{\tau(n-2)}\rangle\langle A_{\xi_{\alpha_s}}(X_{\sigma(n)},X_{\tau(n)}\rangle. \tag{3.3.3}$$

Consider the n-linear form $B:\nu_xM\times\cdots\times\nu_xM\to\mathbb{R}$ defined by:

$$B(\eta_1,\ldots,\eta_n) = \sum_{\sigma\in\mathfrak{S}_n}(-1)^{\sigma}\prod_{i=1}^{n}\langle A_{\eta_i}(X_i),X_{\sigma(i)}\rangle,$$

and observe that $B(\underbrace{\xi,\ldots,\xi}_{n\text{ times}}) = \det A_\xi$. Consider the symmetrization of B which is the unique symmetric n-linear form $\tilde{B}$ on ν_xM such that $\tilde{B}(\xi,\ldots,\xi)=\det A_\xi$; $\tilde{B}$ is computed explicitly as:

$$\tilde{B}(\eta_1,\ldots,\eta_n) = \frac{1}{n!}\sum_{\sigma,\tau\in\mathfrak{S}_n}(-1)^{\sigma\tau}\prod_{i=1}^{n}\langle A_{\eta_i}(X_{\tau(i)}),X_{\sigma(i)}\rangle.$$

Using Exercise 2.20, we can compute the integral on the righthand side of (3.3.2) as:

$$\begin{aligned}\int_{\nu_x^1M}\det A_{(x,\xi)}\,\mathrm{d}\mu_\sigma(\xi) &= \left(\frac{n-1}{n+p-2}\right)\left(\frac{n-3}{n+p-4}\right)\cdots\left(\frac{3}{p+2}\right)\frac{1}{p}\mathrm{vol}\big(S^{p-1}\big) \\ &\qquad\sum_{\alpha_1,\ldots,\alpha_s=1}^{p}\tilde{B}(\xi_{\alpha_1},\xi_{\alpha_1},\ldots,\xi_{\alpha_s},\xi_{\alpha_s}) \\ = \left(\frac{n-1}{n+p-2}\right)\left(\frac{n-3}{n+p-4}\right)\cdots\frac{3}{p+2}\frac{1}{p}&\mathrm{vol}\big(S^{p-1}\big)\frac{1}{n!}\sum_{\alpha_1,\ldots,\alpha_s=1}^{p}\sum_{\sigma,\tau\in\mathfrak{S}_n}(-1)^{\sigma\tau} \\ &\langle A_{\xi_{\alpha_1}}(X_{\tau(1)},X_{\sigma(1)}\rangle\langle A_{\xi_{\alpha_1}}(X_{\tau(2)},X_{\sigma(2)}\rangle \\ &\cdots\langle A_{\xi_{\alpha_s}}(X_{\tau(n-1)},X_{\sigma(n-1)}\rangle\langle A_{\xi_{\alpha_s}}(X_{\tau(n)},X_{\sigma(n)}\rangle.\end{aligned} \tag{3.3.4}$$

The conclusion follows using formulas (3.3.3), (3.3.4), the formula for the volume of the sphere (see Exercise 2.19) and a lot of patience in handling nasty coefficients. □

3.3.3. COROLLARY (Gauss–Bonnet–Chern theorem). *If M is a compact even dimensional manifold then $\int_M\gamma=\chi(M)$.*

PROOF. By a well-known result of Nash, every Riemannian manifold can be isometrically embedded in some Euclidean space. In particular, we can find an isometric immersion $f : M \to \mathbb{R}^{n+p}$. By Fubini's theorem (Theorem **??**) we have:

$$\int_M \left(\int_{\nu_x^1 M} \det A_{(x,\xi)} \, \mathrm{d}\mu_{\sigma_x}(\xi) \right) \mathrm{d}\mu_\delta(x) = \int_{\nu^1 M} \det A_{(x,\xi)} \, \mathrm{d}\mu_{\bar{\delta}}(x,\xi). \tag{3.3.5}$$

By Corollary 3.2.13, the righthand side of (3.3.5) equals the Euler characteristic of M times $\mathrm{vol}\big(S^{n+p-1}\big)$. Using Lemma 3.3.2 we get that the lefthand side of (3.3.5) is equal to:

$$\mathrm{vol}\big(S^{n+p-1}\big) \int_M K(x) \, \mathrm{d}\mu_\delta(x) = \mathrm{vol}\big(S^{n+p-1}\big) \int_M \gamma.$$

The conclusion follows. □

3.4. The Chern–Lashof Theorem

The following theorem is the main result of the section. It gives a characterization of isometric immersions in Euclidean spaces having minimal total absolute curvature.

3.4.1. THEOREM (Chern-Lashof). *Let $f : M \to \mathbb{R}^{n+p}$ be an isometric immersion of the n-dimensional compact Riemannian manifold M. Then:*

(1) $\tau(f) \geq 2$*;*
(2) *if $\tau(f) < 3$ then M is homeomorphic to the sphere S^n;*
(3) *if $\tau(f) = 2$ then f is an embedding, $f(M)$ is contained in an $(n+1)$-dimensional affine subspace $\mathbb{A}$ of $\mathbb{R}^{n+p}$ and it is the boundary in $\mathbb{A}$ of a bounded convex open subset of $\mathbb{A}$.*

PROOF. Since M is compact, for every $\xi \in D$, the height function h_ξ has at least two critical points, so that $\kappa(\xi) \geq 2$. It follows from (3.2.8) that $\tau(f) \geq 2$, which proves part (1). If $\tau(f) < 3$ then $\kappa(\xi) = 2$ for some $\xi \in D$ and therefore h_ξ is a Morse function with precisely two critical points. The proof of part (2) follows then from Reeb's Theorem (Theorem 2.3.13). If $\tau(f) = 2$ then $\kappa(\xi) = 2$ for all $\xi \in D$, since κ is locally constant (see Lemma 3.2.10). The proof of part (3) will be divided into various steps. We will keep the hypothesis of the theorem and, in order to simplify the language, we give the following:

3.4.2. DEFINITION. A pair $(x,\xi) \in \nu^1 M$ is called *separating* if there are points of $f(M)$ in both sides of the affine hyper-plane $f(x) + \xi^\perp$, i.e., if there exists $x_1, x_2 \in M$ with $h_\xi(x_1) < h_\xi(x) < h_\xi(x_2)$.

We make the following simple observations:

(1) *if $(x,\xi) \in \nu^1 M$ is separating then h_ξ has at least three critical points*; namely, x is a critical point of h_ξ that is neither the minimum nor the maximum.
(2) *The set of separating pairs $(x,\xi) \in \nu^1 M$ is open in $\nu^1 M$*; this follows from an obvious continuity argument.
(3) *The set $\mathfrak{G}^{-1}(D)$ is an open dense subset of the open set of all regular points $(x,\xi) \in \nu^1 M$ of $\mathfrak{G}$*; this follows from Proposition **??**.

(4) *There are no separating pairs* $(x,\xi)\in\nu^1M$ *with* $A_{(x,\xi)}$ *invertible*; recall from Lemma 3.2.3 that $A_{(x,\xi)}$ is invertible iff (x,ξ) is a regular point for $\mathfrak{G}$. If there were separating regular points of $\mathfrak{G}$ then by items (2) and (3) above there would exist a separating point $(x,\xi)\in\mathfrak{G}^{-1}(D)$. Then, by item (1) on page 177, h_ξ would be a Morse function having more than two critical points, a contradiction.

STEP 1. *The image of f is contained in an $(n+1)$-dimensional affine subspace of $\mathbb{R}^{n+p}$.*

It suffices to show that if $p\ge 2$ then the image of f is contained in some affine hyper-plane (i.e., an affine subspace of dimension $n+p-1$). The conclusion will follow then by an obvious induction argument[1]. Suppose that $f(M)$ is not contained in any affine hyper-plane. Let $\xi\in D$ and let $x\in M$ be a critical point of h_ξ, so that $(x,\xi)\in\mathfrak{G}^{-1}(D)\subset\nu^1M$. Since $p\ge2$, there exists $\eta\in\nu^1_xM$ with $\langle\xi,\eta\rangle=0$. For $\theta\in\mathbb{R}$, define:

$$\xi_\theta=\xi\cos\theta+\eta\sin\theta\in\nu^1_xM,$$

and denote by $\mathbb{A}_\theta$ the affine hyper-plane $f(x)+\xi_\theta^\perp$. Since for every $u\in\mathbb{R}^{n+p}$, $\big(u-f(x)\big)^\perp$ intercepts the plane spanned by ξ and η, it follows that:

$$\mathbb{R}^{n+p}=\bigcup_{\theta\in\mathbb{R}}\mathbb{A}_\theta. \tag{3.4.1}$$

Our aim is to show that there exists $\theta\in\mathbb{R}$ with $A_{\xi_\theta}(x)$ invertible and such that (x,ξ_θ) is separating; this will give us a contradiction, by item (4) on page 178. Since $\theta\mapsto\det A_{\xi_\theta}$ is real-analytic and does not vanish at $\theta=0$ then A_{ξ_θ} is singular only for a discrete set of θ's in $\mathbb{R}$. It follows that, towards our goal, it suffices to determine one value of θ for which (x,ξ_θ) is separating; namely, by item (2) on page 177, the set of such θ's is open in $\mathbb{R}$ and hence (if it is non empty) it must contain a point θ with $A_{\xi_\theta}(x)$ non singular.

Choose $x_1\in M$ with $f(x_1)$ outside $\mathbb{A}_0=f(x)+\xi^\perp$. By (3.4.1), there exists $\theta_1\in\mathbb{R}$ with $f(x_1)\in\mathbb{A}_{\theta_1}$. Choose $x_2\in M$ with $f(x_2)$ outside $\mathbb{A}_{\theta_1}$. The proof of Step 1 will be completed if we can find $\theta\in\mathbb{R}$ for which the functions:

$$h_{\xi_\theta}(x_1)-h_{\xi_\theta}(x)=\big\langle f(x_1)-f(x),\xi\big\rangle\cos\theta+\big\langle f(x_1)-f(x),\eta\big\rangle\sin\theta, \tag{3.4.2}$$

$$h_{\xi_\theta}(x_2)-h_{\xi_\theta}(x)=\big\langle f(x_2)-f(x),\xi\big\rangle\cos\theta+\big\langle f(x_2)-f(x),\eta\big\rangle\sin\theta, \tag{3.4.3}$$

have opposite signs. The coefficient of $\cos\theta$ in (3.4.2) is not zero because $f(x_1)\notin\mathbb{A}_0$; moreover, the coefficients of $\cos\theta$ and $\sin\theta$ in (3.4.3) cannot both be zero, because $f(x_2)\notin\mathbb{A}_{\theta_1}$. We can thus rewrite the righthand sides of (3.4.2) and (3.4.3) respectively in the form $k_1\cos(\theta+\varphi_1)$, $k_2\cos(\theta+\varphi_2)$, with $k_1,k_2>0$; the difference $\varphi_1-\varphi_2$ cannot be an integer multiple of π because the functions in (3.4.2) and (3.4.3) do not vanish simultaneously at $\theta=\theta_1$. It is now an easy

[1] If $f(M)$ is contained in some affine hyper-plane $\mathbb{A}$ in $\mathbb{R}^{n+p}$ then obviously the isometric immersion $f:M\to\mathbb{A}\cong\mathbb{R}^{n+p-1}$ has again the property that all height functions that are Morse functions have exactly two critical points.

exercise to show the existence of $\theta \in \mathbb{R}$ with $\cos(\theta+\varphi_1)\cos(\theta+\varphi_2) < 0$. This concludes the proof of Step 1.

Step 1 allows us to assume from now on that $p = 1$. In this case, we will say that a point $x \in M$ is *separating* if (x,ξ) is separating for one (hence both) the ξ's in $\nu_x^1 M$, i.e., if there are points of $f(M)$ on both sides of the affine tangent space $f(x) + \mathrm{Im}(\mathrm{d}f_x)$. Obviously the set of separating points is open in M (recall item (2) on page 177).

STEP 2. *Assume that M has no separating points. Then f is an embedding and $f(M)$ is the boundary of a bounded convex open subset of* $\mathbb{R}^{n+1}$.

Observe that since M is compact, $f(M)$ cannot be contained in an affine hyper-plane $\mathbb{A}$ of $\mathbb{R}^{n+1}$; otherwise, f would be a local diffeomorphism onto a (compact) open subset of $\mathbb{A}$.

For each $x \in M$, denote by $\mathbb{A}_x$ the affine hyper-plane $f(x)+\mathrm{Im}(\mathrm{d}f_x)$ of $\mathbb{R}^{n+1}$ and by H_x the (unique) open half-space determined by $\mathbb{A}_x$ such that $f(M) \subset \overline{\mathsf{H}_x}$. Set $\mathsf{H} = \bigcap_{x\in M} \mathsf{H}_x$. Clearly, H is convex. Let us now prove the following facts.

(a) *For $x \in M$, the open half-space $\mathbb{R}^{n+1} \setminus \overline{\mathsf{H}_x}$ is disjoint from the closure of* H;

for, obviously $\mathsf{H} \subset \mathsf{H}_x$ and hence $\overline{\mathsf{H}} \subset \overline{\mathsf{H}_x}$.

(b) *The union $\bigcup_{x\in M} \mathbb{A}_x$ is closed in $\mathbb{R}^{n+1}$ and disjoint from* H;

the fact that $\bigcup_{x\in M} \mathbb{A}_x$ is disjoint from H is obvious. For each $k \ge 1$, let $x_k \in M$, $v_k \in T_{x_k}M$ be given and assume that $f(x_k)+\mathrm{d}f_{x_k}(v_k)$ converges to $u \in \mathbb{R}^{n+1}$. Since M is compact, we may assume that x_k converges to $x \in M$; hence $\mathrm{d}f_{x_k}(v_k)$ converges to $u - f(x) \in \mathbb{R}^{n+1}$. The set

$$E = \bigcup_{y\in M} \{y\} \times \mathrm{Im}(\mathrm{d}f_y)$$

is a smooth vector subbundle of the trivial bundle $M \times \mathbb{R}^{n+1}$ and therefore is closed in $M \times \mathbb{R}^{n+1}$ (see Exercise 2.6); since $\big(x_k, \mathrm{d}f_{x_k}(v_k)\big)$ is a sequence in E that converges to $\big(x, u - f(x)\big) \in M \times \mathbb{R}^{n+1}$, it follows that $u - f(x) \in \mathrm{Im}(\mathrm{d}f_x)$ and therefore $u \in \mathbb{A}_x$. This concludes the argument.

(c) H *is open in* $\mathbb{R}^{n+1}$;

let $u \in \mathsf{H}$ be given. It follows from item (b) that there exists $\varepsilon > 0$ such that the ball $\mathrm{B}(u;\varepsilon)$ is disjoint from $\bigcup_{x\in M} \mathbb{A}_x$. Then, for $x \in M$, the $\mathrm{B}(u;\varepsilon)$ intercepts H_x and is disjoint from $\mathbb{A}_x$; hence $\mathrm{B}(u,\varepsilon) \subset \mathsf{H}_x$.

(d) H *is bounded in* $\mathbb{R}^{n+1}$;

for $\xi \in S^n$, the function $\mathbb{R}^{n+1} \ni u \mapsto g_\xi(u) = \langle \xi, u\rangle$ is bounded in H. Namely, let $x_0, x_1 \in M$ be respectively a minimum and a maximum of $h_\xi = g_\xi \circ f$. Then x_0 and x_1 are critical points of h_ξ and hence $\mathrm{Im}(\mathrm{d}f_{x_0})$ and $\mathrm{Im}(\mathrm{d}f_{x_1})$ are both orthogonal to ξ. It follows that:

$$\mathsf{H} \subset \mathsf{H}_{x_0} \cap \mathsf{H}_{x_1} = g_\xi^{-1}\big(\,]h_\xi(x_0), h_\xi(x_1)[\,\big).$$

This proves that g_ξ is bounded in H for all $\xi \in S^n$. In particular, the coordinate functions of $\mathbb{R}^{n+1}$ are bounded in H.

(e) *If $x, y \in M$ are such that $f(x) \in \mathbb{A}_y$ then $\mathbb{A}_x = \mathbb{A}_y$*;

let $\xi \in S^n$ be a unit vector that is normal to $\mathrm{Im}(\mathrm{d}f_y)$. Since $f(M)$ is contained in one half-space determined by $\mathbb{A}_y$, it follows that the height function h_ξ has either a minimum or a maximum at $y \in M$. But $f(x) \in \mathbb{A}_y$ implies $h_\xi(x) = h_\xi(y)$ and hence x is also an extremum of h_ξ of the same kind as y. Then x is a critical point of h_ξ, ξ is orthogonal to $\mathrm{Im}(\mathrm{d}f_x)$ and $\mathbb{A}_x = \mathbb{A}_y$ because they are parallel to the same vector space and have the common point $f(x)$.

(f) *If $x, x_0 \in M$ are such that $f(x_0) \notin \mathbb{A}_x$ then the open line segment with endpoints $f(x_0)$ and $f(x)$ is contained in H and the open half-line issuing from $f(x_0)$ in the direction of $f(x)$ intersects $\partial\mathsf{H}$ only at $f(x)$*;

denote by $]f(x_0), f(x)[$ the open line segment with endpoints $f(x_0)$ and $f(x)$. For every $y \in M$, the endpoints $f(x_0)$ and $f(x)$ are both in $\overline{\mathsf{H}_y}$ and therefore $]f(x_0), f(x)[$ is contained in $\overline{\mathsf{H}_y}$. We claim that $]f(x_0), f(x)[$ is indeed contained in H_y; for, otherwise, $f(x_0)$ and $f(x)$ would be both on $\mathbb{A}_y$. By item (e) this implies $\mathbb{A}_x = \mathbb{A}_y$ and therefore $f(x_0) \in \mathbb{A}_x$, contradicting our hypothesis.

For $t > 0$, denote by u_t the point $f(x_0) + t\big(f(x) - f(x_0)\big)$ on the half-line issuing from $f(x_0)$ in the direction of $f(x)$. We have shown that $u_t \in \mathsf{H}$ for $t \in \,]0,1[$; by item (c), H is open and therefore $u_t \notin \partial\mathsf{H}$. For $t > 1$, u_t is in $\mathbb{R}^{n+1} \setminus \overline{\mathsf{H}_x}$ and therefore outside $\overline{\mathsf{H}}$, by item (a). It is now obvious that $u_1 = f(x) \in \partial\mathsf{H}$.

(g) *$f(M) \subset \partial\mathsf{H}$ and $f : M \to \partial\mathsf{H}$ is an open map*;

for every $x \in M$ we can find $x_0 \in M$ with $f(x_0) \notin \mathbb{A}_x$ and therefore $f(x)$ is indeed in $\partial\mathsf{H}$, by item (f). To prove that $f : M \to \partial\mathsf{H}$ is an open map, it suffices to show that if V is a sufficiently small open neighborhood of x in M then $f(V)$ is open in $\partial\mathsf{H}$. Let ξ be a smooth unit normal vector field defined in a neighborhood of x in M and choose an open neighborhood V of x small enough so that $\big\langle \xi(y), f(x_0) - f(y) \big\rangle \neq 0$ for all $y \in V$; then $f(x_0) \notin \mathbb{A}_y$ for $y \in V$. Consider the map:

$$]0, +\infty[\, \times V \ni (t, y) \longmapsto \phi(t, y) = f(x_0) + t\big(f(y) - f(x_0)\big) \in \mathbb{R}^{n+1};$$

it follows from item (f) that $\mathrm{Im}(\phi) \cap \partial\mathsf{H} = f(V)$. Moreover, a simple computation using the inverse function theorem, shows that $\mathrm{Im}(\phi)$ is open in $\mathbb{R}^{n+1}$. This concludes the argument.

(h) *f is an embedding and $f(M) = \partial\mathsf{H}$*;

since H is an open bounded convex subset of $\mathbb{R}^{n+1}$, $\partial\mathsf{H}$ is homeomorphic to the sphere S^n by Exercise 1.42. By item (g), $f(M)$ is both open and closed in H and therefore $\mathsf{H} = f(M)$. Since $f : M \to \partial\mathsf{H}$ is locally injective, continuous and open, then it is a local homeomorphism. Since M is compact,

f is proper and hence a covering map (see Exercise 2.14). The conclusion follows by observing that $\partial \mathsf{H} \cong S^n$ is simply connected.

The following step will conclude the proof of the Theorem.

STEP 3. *There are no separating points $x \in M$.*

For the proof of the last step we need the following technical fact, and we refer to [**?**] for a proof:

3.4.3. PROPOSITION. *For every $x \in M$ set:*

$$\mathcal{D}_x = \mathrm{Ker} A_{(x,\xi)} \subset T_x M,$$
$$d(x) = \textit{dimension of } \mathcal{D}_x \in \mathbb{N}, \mathcal{U}_k = \{x \in M : d(x) = k\}$$

where ξ denotes any one of the two elements of $\nu_x^1 M$. Let $U \subseteq M$ be an open set contained in $\mathcal{U}_k$. Then:

(1) *$\mathcal{D}$ is an integrable distribution in U and it's leaves are totally geodesic in $\mathbb{R}^{n+1}$.*
(2) *If $\gamma : [0, b] \to M$ is a geodesic such that $\gamma([0, b[)$ is contained in a leaf of $\mathcal{D} \subset U$, then $\gamma(b) \in \mathcal{U}_k$ and the (affine) tangent space of M is constant along γ.*

We will prove now Step 3. By item (4) on page 178, there are no separating points $x \in M$ with $d(x) = 0$. So will be enough to prove that the existence of a separating point x with $d(x) \neq 0$, implies the existence of another separating point $y \in M$ with $d(y) < d(x)$. Let $x \in \mathcal{U}_k$ be a separating point. If $x \in \partial \mathcal{U}_k$, since $d : M \to \mathbb{N}$ is an upper semi-continuous function, there are, arbitrarily near x, points in $\mathcal{U}_l$, $l < k$. Suppose now that x belongs to the interior of $\mathcal{U}_k$. Let $\mathcal{S}$ be a maximal leaf of the distribution $\mathcal{D}, x \in \mathcal{S}$. Let $\gamma : \mathbb{R} \to M$ be a geodesic with $\gamma(0) = x$, $\dot\gamma(0) \in T_x \mathcal{S}$. Since $\mathcal{U}_k$ is bounded and γ is s stright line, as long as $\gamma(t) \in \mathcal{U}_k$, there exists a smallest $b \in \mathbb{R}$ such that $\gamma(b) \in \partial \mathcal{U}_k$. Since the (affine) tangent space is constant along $\gamma([0, b]), \gamma(b)$ is again a separating point and $d(\gamma(b)) = k$. Arguing as above we get, arbitrarely near $\gamma(b)$, a separating point y with $d(y), k$. $\square$

We introduce naw an other class of functions which turns out to be very usefull in the study of the geometry and topology of submanifolds of Euclidean spaces.

3.4.4. DEFINITION. Let $f : M \to \mathbb{R}^{n+p}$ be an isometric immersion of an n-dimensional Riemannian manifold. Fix $q \in \mathbb{R}^{n+p}$. The *distance function from q* is the function:

$$L_q : M \to \mathbb{R}, \quad L_q(x) =< q - x, q - x > .$$

We study now the critical points of L_q. Set $\xi(x) = q - f(x)$. Identifing (locally) M with $f(M)$, we have $\overline{\nabla}_X \xi = -X, X\imath T_x M$. Hence:

- $dL_q(x)X = -2 < X, \xi >$.
  In particular, $x \in M$ is a critical point of L_q if and only if $\xi(x) \in \nu_n M$.
- If x is a critical poin of L_q, we have:

$$d^2 L_q(x)(X, Y) = -2Y < X, \xi >= 2 < (Id - A_\xi)X, Y > .$$

We want to characterize the points $q \in \mathbb{R}^{n+p}$ such that L_q is a Morse function. This will be done in terms of the *endpoint map* or normal exponential map:

$$E : \nu M \to \mathbb{R}^{n+p}, \quad E(x, \eta) = f(x) + \eta.$$

We compiute the differential of E. Let $(x, \eta) \in \nu M$ and $\gamma(t) = (x(t), \eta(t))$ be a curve in νM such that $x(0) = x,\ \eta(0) = \eta$. Then:

$$dE(x, \eta)(\dot\gamma(0)) = (x(t) + \eta(t))'(0) = \dot x(0) + (\dot\eta(0))^T + (\dot\eta(0))^\perp,$$

where, as before, for $z \in T_x\mathbb{R}^{n+p}, z^T$ and $z^\perp$ denote the projections of z onto T_xM and ν_xM respectively. In particular, taking $x(t) = x,\ \eta(t) = x + t\eta$, we get that the differential of E along the fibres is the identity (which was geometrically obvious). In particular $dE(x, \eta)$ and $(Id - A_\eta)$ have kernels of the same dimension. In particular:

3.4.5. LEMMA. *L_q has only nondegenerate critical poins if and only if q is a regular value of E.*

A critical value of E is called a *focal point*.

3.4.6. REMARK. If M is non compact and but $f(M)$ is closed, then L_q is a proper function. So, using the Whitney's theorem on the existence of closed embeddings and Sard's theorem, the above lemma gives the existence of (proper) Morse functions on every differentiable manifold.

The following result, due to Nomizu and Rodrigues (see [**?**]), can be seen as the version of the Chern–Lashof theorem for distance functions:

3.4.7. THEOREM. *Let $f : M \to \mathbb{R}^{n+p}$ be an isometric immersion of a compact, connected Riemannian manifold of dimension $n \geq 2$. Suppose that, for every non focal point $q \in \mathbb{R}^{n+p}$, the function L_q has only two critical points. Then f is totally umbilical[2].In particular f embeds M as a round sphere in some $(n + 1)$-dimensional affine subspace.*

PROOF. Let $(x, \eta) \in \nu M$ and $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of A_η. We want to show that $\lambda_1 = \ldots = \lambda_n$. Suppose $\lambda_1 < \lambda_2$, Choose $t \in \mathbb{R}$ such that $1 - t\lambda_1 > 0 > 1 - t\lambda_2$ and $1 - t\lambda_i \neq 0$. Then $(Id - A_\eta)$ is non singular with index different from $0, n$. In particular $(x, t\eta)$ is a regular point for the endpoint map E, hence E maps an open neigborhood of $(x, t\eta)$ diffeomorfically onto an open neigborhood of $q = E(x, t\eta) \in \mathbb{R}^{n+p}$. By Sard's theorem there exist a regular value of E, $q' = E(x', \eta')$ arbitrarely close to q, with (x', η') arbitrarely close

[2]Recall that f is *totally umbilical* if for every $(x, \eta) \in \nu M$, the Weingarten operator A_η is a multiple of the identity. It is a classical fact that for such an immersion, if $n \geq 2$, the connected components of $f(M)$ are open parts of n-dimensional affine spaces or n-dimensional round spheres, in some $(n + 1)$-dimensional affine subspace (see [**?**]).

to (x, η). Then $L_{q'}$ is a Morse function and x' is a critical point of $L_{q'}$ which, by continuity has index $\neq 0, n$. Therefore $L_{q'}$ has at least three critical poins, a contraddiction. □

For a "convex embedding", the Morse height functions have two critical points, but the distance functions have, if the embedding is not a round sphere, more then two critical points. So, in general, "height functions have moore critical points than distance functions". Depending on the problem may be more convinient to work with one or the other class of functions. However, in an interesting case, the two classes coincide:

3.4.8. PROPOSITION. *Let $f : M \to \mathbb{R}^{n+p}$ be an isometric immersion such that $\|f(x)\|^2 = r^2 > 0$. Then $L_q(x) = (1 + \|q\|^2) - 2h_q$.*

PROOF. An easy calculation. □

The results of this section lead naturally to consider two clases of immersions: The ones for which the Morse height functions Have the minimum number of critical points allowed by the (weak) Morse inequalities, an the clas for which the same appens for the distance functions. Immersions of the first type are called *tight*, and the one of the second type are called *tought*. Properties of those classes will be discussed in the Appendix.

3.5. Low Co-dimensional Isometric Immersions of Compact Manifolds with non Negative Curvature

In this section we will study the topology of compact Riemannian manifolds with nonnegative sectional curvature, isometrically immersed in Euclidean spaces in codimension one and two.

The case of codimension one is an easy consequence of the theorem of Chern and Lashof:

3.5.1. THEOREM. *Let M be a compact connected n-dimensional Riemannian manifold ($n \geq 2$) and $f : M \to \mathbb{R}^{n+1}$ an isometric immersion. If the sectional curvature of M is nonnegative then M is homeomorphic to the sphere S^n, f is an embedding and $f(M)$ is the boundary of a bounded convex open subset of $\mathbb{R}^{n+1}$.*

PROOF. Let $\xi \in S^n$ be such that the height function $h_\xi : M \to \mathbb{R}$ is a Morse function. We will show that h_ξ has exactly two critical points and then the conclusion will follow the Chern–Lashof theorem. For every critical point $x \in M$ of h_ξ, the Hessian of h_ξ at x is the second fundamental form II_ξ at the point x. If $(E_i)_{i=1}^n$ is an orthonormal basis of T_xM that diagonalizes the Weingarten operator A_ξ, say $A_\xi E_i = \lambda_i E_i$. Then, by the Gauss equation, the sectional curvature of the plane spanned by E_i and E_j ($i \neq j$) is $\lambda_i\lambda_j$. Since M has nonnegative sectional curvature, it follows that $\lambda_i\lambda_j > 0$, so all λ_i's have the same sign. This means that the Morse index of h_ξ at x is either 0 or n. Using Corollary **??**, it follows that$\tau(f) = 2$. □

The case of codimension two was considered, between others, by J. D. Moore in [**?**] who proved the following:

3.5.2. THEOREM. *Let (M, g) be an n-dimensional compact connected Riemannian manifold with* positive *sectional curvature and $n \geq 3$. If M admits an isometric immersion in $\mathbb{R}^{n+2}$ then M has the homotopy type of the sphere S^n.*

During the proof of Theorem 3.5.2 we will need some results from algebraic topology that will be stated without proof.

3.5.3. THEOREM (Poincaré duality). *Let $\mathbb{K}$ be an arbitrary field. If M is a compact topological oriented n-dimensional manifold then for every i the homology groups $H_i(M; \mathbb{K})$ and $H_{n-i}(M; \mathbb{K})$ are isomorphic. If $\mathbb{K} = \mathbb{Z}_2$, the same result holds without the assumption that M is orientable.*

3.5.4. THEOREM. *Let M be a compact n-dimensional differentiable manifold with $n \geq 1$. If there exists a natural number k for which the Whitney sum:*

$$TM \oplus (M \times \mathbb{R}^k)$$

is a trivial vector bundle then the Euler characteristic of M is even.

3.5.5. THEOREM. *Let M be a compact, connected, simply-connected n-dimensional differentiable manifold. If $H_i(M; \mathbb{K}) = 0$ for $i = 1, \ldots, n-1$ and any field $\mathbb{K}$, then M has the same homotopy type that the sphere S^n.*

PROOF. We will divide the proof in several steps and will assume the hypothesis and notations of the theorem. The starting point is the following observation due to A. Weinstein:

STEP 4. *Given $x \in M$, $\mathbb{I}_x(v, v) \in \nu_x M$ is non zero whenever v is non zero; in particular, the map:*

$$T_x M \setminus \{0\} \ni v \longmapsto \frac{\mathbb{I}_x(v, v)}{\left\|\mathbb{I}_x(v, v)\right\|} \in \nu_x^1 M \tag{3.5.1}$$

is well-defined. Its image $\mathcal{S}_x \subset \nu_x^1 M$ is a closed arc of length less than $\frac{\pi}{2}$.

PROOF. For $v, w \in T_x M$, the Gauss equation gives us:

$$\big\langle R(v, w)v, w \big\rangle = \left\|\mathbb{I}_x(v, w)\right\|^2 - \big\langle \mathbb{I}_x(v, v), \mathbb{I}_x(w, w) \big\rangle;$$

since M has positive sectional curvature, it follows that if $v, w \in T_x M$ are linearly independent then:

$$\left\|\mathbb{I}_x(v, w)\right\|^2 - \big\langle \mathbb{I}_x(v, v), \mathbb{I}_x(w, w) \big\rangle < 0. \tag{3.5.2}$$

Since $n \geq 2$, equation (3.5.2) implies that $\mathbb{I}_x(v, v) \neq 0$ if $v \neq 0$, so that the map (3.5.1) is indeed well-defined. Obviously $\mathcal{S}_x$ equals the image of the restriction of (3.5.1) to the unit sphere of $T_x M$. This implies that $\mathcal{S}_x$ is compact and connected, i.e., a closed arc. Finally, (3.5.2) implies that the angle between $\mathbb{I}_x(v, v)$ and $\mathbb{I}_x(w, w)$ is less than $\frac{\pi}{2}$ whenever $v, w \in T_x M$ are linearly independent. It follows that the length of $\mathcal{S}_x$ is less than $\frac{\pi}{2}$. □

The following step is the basic algebraic fact that will allow us to estimate the absolute total curvature:

STEP 5. *Let A_0, $A_{\frac{\pi}{2}}$ be two $n\times n$ positive definite symmetric matrices. Then:*

$$\big|\det(A_0 - A_{\frac{\pi}{2}})\big| < \big|\det(A_0 + A_{\frac{\pi}{2}})\big|.$$

PROOF. The result is obvious if A_0 and $A_{\frac{\pi}{2}}$ are diagonal matrices. We reduce the general case to this case by the following argument. We identify A_0 and $A_{\frac{\pi}{2}}$ with positive definite symmetric bilinear forms in $\mathbb{R}^n$; observe that both A_0 and $A_{\frac{\pi}{2}}$ are inner products. Denote by T the linear endomorphism of $\mathbb{R}^n$ that represents $A_{\frac{\pi}{2}}$ with respect to the inner product A_0, i.e., $A_{\frac{\pi}{2}}(\cdot,\cdot) = A_0(T\cdot,\cdot)$. Then T is a A_0-symmetric linear operator and therefore there exists a A_0-orthonormal basis in $\mathbb{R}^n$ for which the matrix representation of T is diagonal. Hence, we can find an invertible $n\times n$ matrix P such that P^*A_0P is the identity and $P^*A_{\frac{\pi}{2}}P$ is diagonal (and positive). The conclusion follows from the computation below:

$$\begin{aligned}(\det P)^2\,\big|\det(A_0 - A_{\frac{\pi}{2}})\big| &= \big|\det(P^*A_0P - P^*A_{\frac{\pi}{2}}P)\big| \\ &\le \big|\det(P^*A_0P + P^*A_{\frac{\pi}{2}}P)\big| = (\det P)^2\,\big|\det(A_0 + A_{\frac{\pi}{2}})\big|. \quad \square\end{aligned}$$

As a consequence we get:

STEP 6. *Let A_0, $A_{\frac{\pi}{2}}$ be two $n\times n$ positive definite symmetric matrices and for $\theta\in\mathbb{R}$ set:*

$$A(\theta) = A_0\cos\theta + A_{\frac{\pi}{2}}\sin\theta.$$

Then:

$$\big|\det\big(A(\theta)\big)\big| \ge \big|\det\big(A(\pi-\theta)\big)\big|,$$

for all $\theta\in\big[0,\frac{\pi}{2}\big]$.

PROOF. If $\theta = 0$ or $\theta = \frac{\pi}{2}$ the result is trivial; otherwise, apply Lemma 5 to the positive definite symmetric matrices $A_0\cos\theta$ and $A_{\frac{\pi}{2}}\sin\theta$. $\square$

We are now ready to estimate the absolute total curvature:

STEP 7.

$$\tau_0(f) + \tau_n(f) > \sum_{k=0}^{n-1}\tau_k(f).$$

PROOF. Using formula (3.2.7) and Fubini's Theorem (Theorem **??**) we get:

$$\tau_k(f) = \frac{1}{\operatorname{vol}\big(S^{n+1}\big)}\int_M\left(\int_{\nu_x^1M\cap\mathcal{U}_k}\big|\det A_{(x,\xi)}\big|\,\mathrm{d}\mu_{\sigma_x}(\xi)\right)\,\mathrm{d}\mu_\delta(x),$$

for $k = 0,\ldots,n$. Now let $x\in M$ be fixed. The proof will be completed once we prove that:

$$\begin{aligned}\int_{\nu_x^1M\cap\mathcal{U}_0}\big|\det A_{(x,\xi)}\big|\,\mathrm{d}\mu_{\sigma_x}(\xi) &+ \int_{\nu_x^1M\cap\mathcal{U}_n}\big|\det A_{(x,\xi)}\big|\,\mathrm{d}\mu_{\sigma_x}(\xi) \\ &> \sum_{k=1}^{n-1}\int_{\nu_x^1M\cap\mathcal{U}_k}\big|\det A_{(x,\xi)}\big|\,\mathrm{d}\mu_{\sigma_x}(\xi).\end{aligned} \tag{3.5.3}$$

Since the closed arc $\mathcal{S}_x$ has length less than $\frac{\pi}{2}$ (Lemma 4), we can choose an orthonormal basis ξ_1, ξ_2 of $\nu_x M$ which leaves $\mathcal{S}_x$ in the first quadrant, i.e., such that $\langle \xi, \xi_1 \rangle$ and $\langle \xi, \xi_2 \rangle$ are positive for all $\xi \in \mathcal{S}_x$. Observe that with such choice of ξ_1 and ξ_2, A_{ξ_i} are positive definite. For $\theta \in \mathbb{R}$ we set $\xi_\theta = \xi_1 \cos\theta + \xi_2 \sin\theta$, and $A(\theta) = A_{\xi_\theta}$. Observing that $A(\theta)$ is positive definite for $\theta \in \left[0, \frac{\pi}{2}\right]$ and negative definite for $\theta \in \left[\pi, \frac{3\pi}{2}\right]$ we get:

$$\left\{\xi_\theta : \theta \in \left[0, \tfrac{\pi}{2}\right]\right\} \subset \mathcal{U}_0, \quad \left\{\xi_\theta : \theta \in \left[\pi, \tfrac{3\pi}{2}\right]\right\} \subset \mathcal{U}_n, \tag{3.5.4}$$

and hence:

$$\begin{aligned} \int_{\nu_x^1 M \cap \mathcal{U}_0} \left|\det A_{(x,\xi)}\right| \mathrm{d}\mu_{\sigma_x}(\xi) &+ \int_{\nu_x^1 M \cap \mathcal{U}_n} \left|\det A_{(x,\xi)}\right| \mathrm{d}\mu_{\sigma_x}(\xi) \\ &> \int_0^{\frac{\pi}{2}} \left|\det\big(A(\theta)\big)\right| \mathrm{d}\theta + \int_\pi^{\frac{3\pi}{2}} \left|\det\big(A(\theta)\big)\right| \mathrm{d}\theta; \end{aligned} \tag{3.5.5}$$

the fact that the inequality above is strict follows by observing that the continuous function $\theta \mapsto \left|\det\big(A(\theta)\big)\right|$ is positive on $\left[0, \frac{\pi}{2}\right] \cup \left[\pi, \frac{3\pi}{2}\right]$ and that $\mathcal{U}_0$ (respectively, $\mathcal{U}_n$) contains ξ_θ for θ in an interval which is strictly larger than $\left[0, \frac{\pi}{2}\right]$ (respectively, strictly larger than $\left[\pi, \frac{3\pi}{2}\right]$).

Observing that both integrals in the righthand side of (3.5.5) are equal and using Corollary 6, we get:

$$\begin{aligned} \int_0^{\frac{\pi}{2}} \left|\det\big(A(\theta)\big)\right| \mathrm{d}\theta &+ \int_\pi^{\frac{3\pi}{2}} \left|\det\big(A(\theta)\big)\right| \mathrm{d}\theta \\ &\geq \int_{\frac{\pi}{2}}^{\pi} \left|\det\big(A(\theta)\big)\right| \mathrm{d}\theta + \int_{\frac{3\pi}{2}}^{2\pi} \left|\det\big(A(\theta)\big)\right| \mathrm{d}\theta. \end{aligned}$$

Finally, (3.5.4) implies $\bigcup_{k=1}^{n-1} \left(\nu_x^1 M \cap \mathcal{U}_k\right) \subset \left\{\xi_\theta : \theta \in \left[\frac{\pi}{2}, \pi\right] \cup \left[\frac{3\pi}{2}, 2\pi\right]\right\}$ and hence:

$$\begin{aligned} \int_{\frac{\pi}{2}}^{\pi} \left|\det\big(A(\theta)\big)\right| \mathrm{d}\theta &+ \int_{\frac{3\pi}{2}}^{2\pi} \left|\det\big(A(\theta)\big)\right| \mathrm{d}\theta \\ &\geq \sum_{k=1}^{n-1} \int_{\nu_x^1 M \cap \mathcal{U}_k} \left|\det A_{(x,\xi)}\right| \mathrm{d}\mu_{\sigma_x}(\xi). \end{aligned}$$

This proves (3.5.3) and concludes the proof. $\square$

The latter result and Morse inequalities will allow us to estimate the Betti Numbers:

STEP 8. *Let $\mathbb{K}$ be a field with $\mathbb{K} = \mathbb{Z}_2$ or M orientable then:*

$$\sum_{k=1}^{n-1} \beta_k(M;\mathbb{K}) < \beta_0(M;\mathbb{K}) + \beta_n(M;\mathbb{K}) = 2.$$

.

PROOF. Using the strong Morse inequality (**??**) with $k = 1$ and recalling (3.2.4) we get:

$$\begin{aligned}(3.5.6)\quad \tau_1(f) - \tau_0(f) &= \frac{1}{\mathrm{vol}\big(S^{n+1}\big)}\int_D \kappa_1(\xi) - \kappa_0(\xi)\,\mathrm{d}\mu_\sigma(\xi)\\ &\ge \frac{\mathrm{vol}(D)}{\mathrm{vol}\big(S^{n+1}\big)}\big(\beta_1(M;\mathbb{K}) - \beta_0(M;\mathbb{K})\big) = \beta_1(M;\mathbb{K}) - \beta_0(M;\mathbb{K}),\end{aligned}$$

where the last equality follows from the fact that $S^{n+1}\setminus D$ has null measure. Using (3.2.5) and Poincaré duality (Theorem 3.5.3) we get:

$$(3.5.7)\qquad \tau_{n-1}(f) - \tau_n(f) \ge \beta_{n-1}(M;\mathbb{K}) - \beta_n(M;\mathbb{K}).$$

¿From (3.5.6) and (3.5.7) we get:

$$\begin{aligned}(3.5.8)\quad &\tau_1(f) + \tau_{n-1}(f) - \tau_0(f) - \tau_n(f)\\ &\qquad\ge \beta_1(M;\mathbb{K}) + \beta_{n-1}(M;\mathbb{K}) - \beta_0(M;\mathbb{K}) - \beta_n(M;\mathbb{K}).\end{aligned}$$

Using the weak Morse inequalities (**??**) we get:

$$(3.5.9)\qquad \sum_{k=2}^{n-2}\tau_k(f) \ge \sum_{k=2}^{n-2}\beta_k(M;\mathbb{K}).$$

Adding (3.5.8), (3.5.9) and using Lemma 7 we get:

$$\sum_{k=1}^{n-1}\beta_k(M;\mathbb{K}) - \big(\beta_0(M;\mathbb{K}) + \beta_n(M;\mathbb{K})\big) \le \sum_{k=1}^{n-1}\tau_k(f) - \big(\tau_0(f) + \tau_n(f)\big) < 0.$$

Since M is connected, $\beta_0(M;\mathbb{K}) = 1$; moreover, Poincaré duality implies also $\beta_n(M;\mathbb{K}) = 1$. $\square$

We are now ready for the final steps of the proof.

STEP 9. *M is simply connected*

PROOF. By the theorem of Bonnet–Myers, $\pi_1(M)$ is finite, so contain an element a of prime period p. Let $[a] \cong \mathbb{Z}_p$ be the subgroup generated by a. Let $\pi : M_a \to M$ be a covering map with $\pi_1(M_a) \cong [a]$ and consider in M_a the covering metric so that $f_a := f\circ\pi : M_a \to \mathbb{R}^{n+2}$ is an isometric immersion. Observe that M_a is compact, with positive curvature and orientable if $p \neq 2$, since, $\pi_1(M_a)$ does not contain subgroups of order two. We may therefore apply the Betti numbers estimate to M_a obtaining $\sum_{i=1}^{n-1}\beta_i(M_a;\mathbb{Z}_p) \le 1$. But $H_1(M_a;\mathbb{Z}_p) \cong H_1(M_a;\mathbb{Z})\otimes\mathbb{Z}_p \cong \mathbb{Z}_p$, by the universal coefficients theorem, and, by Poinaré duality, $H_{n-1}(M_a;\mathbb{Z}_p) \cong \mathbb{Z}_p$ which gives the contradiction $\sum_{i=1}^{n-1}\beta_i(M_a;\mathbb{Z}_p) \ge 2$. $\square$

STEP 10. *The normal bundle νM of the isometric immersion f is trivial.*

PROOF. Since M is simply connected, hence orientable, orientable, the normal bundle νM is an orientable vector bundle; since its fibers are two-dimensional, in order to prove that νM is trivial it suffices to exhibit a continuous never vanishing global section of νM (see Exercise **??**). We then define a section $\xi : M \to \nu M$ by taking $\xi(x)$ to be the middle point of the arc $\mathcal{S}_x$ for all $x \in M$. Although intuitive, the continuity of s has to be proven by a technical argument, which goes as follows. Let $\eta, \eta^\perp$ be an orthonormal frame defined in an an open neighborhood $U \subset M$ of x. Denote by $\theta(y, X)$ the angle between $\eta(y)$ and $\alpha(X, X),\ X \in T_y M$. Since $\mathcal{S}_y \neq \nu_y^1 M,\ \forall y \in U$, it follows that we can choose a continuous determination of θ. We set:

$$\theta_m(y) = inf\{\theta(y, X) : X \in T_y M\}, \quad \theta_M(y) = sup\{\theta(y, X) : X \in T_y M\}.$$

Them θ_m and θ_M are continuous in U, so is $\theta(y) = \frac{1}{2}(\theta_n(y) + \theta_M(y)$. But:

$$\xi(y) = \cos\theta(y)\eta + \sin\theta(y)\eta^\perp,$$

so ξ is continuous. □

We can naw conclude the proof of the theorem: From Step ? we know that M is simply connected. So, by theorem ?? it is sufficient to prove that $\beta_i(M; \mathbb{K}) = 0,\ i = 1, \ldots, n-1$ and for every field $\mathbb{K}$. Suppose this is not the case. Then, by step ?? $\beta_i(M; \mathbb{K}) = 1$ for some $i = 1, \ldots, n-1$ and all the others Betti numbers are zero (in the above range). But this would imply that the Euler characteristic of M is odd, contradicting Theorem ?? since νM is trivial. □

3.5.6. REMARK. If $n \geq 4$ a compact n-dimensional manifold, homotopy equivalent to a sphere, is homeomorphic to a sphere by the positive answer to the generalized Poincaré conjecture.

3.5.7. REMARK. If $n = 2$, the classical Gauss–Bonnet theorem imply that the manifold is diffeomorphic to S^2 or $\mathbb{R}P^2$. We do not know if there exist an immersion of the real projective plane into $\mathbb{R}^4$ such that the induced metric has positive curvature. It is known, however, that if such immersion exist, it can not be an embedding.

It is possible to extend Moore theorem to the case of compact manifolds with non negative curvature. However the proof require complementary techniques an we refer to [**?**] and [**?**] for a proof of the following result:

3.5.8. THEOREM. *Let M be an n-dimensional compact, connected Riemannian manifold with non negative sectional curvature, $n \geq 3$, and $f : M \to \mathbb{R}^{n+2}$ an isometric immersion. Then:*

(1) *If M is simply connected, then either M is a homotopy sphere or it is isometric to a Riemannian product $M_1^{n_1} \times M_2^{n_2}$ and f is the product of two convex embedding $f_i : M_i^{n_i} \to \mathbb{R}^{n_i+1}$.*
(2) *If M is not simply connected, either is covered by S^3 or diffeomorphic to $S^1 \times S^{n-1}$, in the orientable case, or to a generalized Klein bottle [3] in the non orientable case.*

.

[3]The generalized Klein bottle is the non orientable S^{n-1} bundle over S^1.

3.6. Quasi-convex hypersurfaces

Let $f : M^n \to \mathbb{R}^{n+1}$ be an isometric immersion of a compact, connected Riemannian manifold. If the sectional curvature of M^n is non negative, we have seen that f is an embedding and $f(M)$ is the boundary of a convex open set. The main point of the proof was the fact that for a regular value ξ of the Gauss map and $(x, \xi) \in \nu M$, then the eigenvalues of A_ξ have the same sign and, conversely, it is obvious that a "convex embedding" satisfies the above condition. In this section we will considere some important geometric conditions on M^n that imply that f satisfies the following weaker condition:

3.6.1. DEFINITION. The immersion f is *quasi-convex* if all but at most one of the eigenvalues of A_ξ have the same sign.

The above condition is empty if $n \leq 3$ so *for the rest of this section we will assume* $n \geq 4$. From Theorem **??** we have:

3.6.2. THEOREM. *Let M^n be an n-dimensional compact, connected Riemannian manifold $n \geq 4$, and $f : M^n \to \mathbb{R}^{n+1}$ be a quasi-convex immersion. Then M^n has the homotopy type of a CW-complex with no cell in dimension $k, k \in \{2, \dots, n-2\}$. In particular:*

(1) $H_k(M^n; \mathbb{Z}) = \{0\}, k = 2, \dots, n-2$.
(2) $H_1(M^n; \mathbb{Z})$ *is a free* Abelian *group on* β_1 *generators.*
(3) $\pi_1(M^n)$ *is a free group in* β_1 *generators.*

We will discuss now two interesting conditions on the intrinsic geometry of M^n that imply that f is quasi-convex.

3.6.1. Conformally flat hypersurfaces.

Conformally flat manifolds are the analogous, in conformal geometry, of manifolds of constant curvature in Riemannian geometry. We recall that:

3.6.3. DEFINITION. An n-dimensional Riemannian manifold M^n is *(locally) conformally flat*, if $\forall x \in M^n$, there exist an open neighborhood $U \subseteq M^n$ of x and a conformal diffeomorphism of U onto an open set of $\mathbb{R}^n$.

We observe that 2-dimensional Riemannian manifolds are always conformally flat, due to the existence of isothermal coordinates, so we will assume, in what follows, that $n \geq 3$.

Let $\{E_1, \dots, E_n\}$ be an orthonormal basis for T_xM. Recall that the *Ricci tensor* $Q : T_xM \to T_xM$ is defined as:

$$Q(X) = \sum_1^n R(X, E_i)E_i,$$

and, for a unit vector $X \in T_xM$, the *Ricci curvature* is given by $Ricc(X) = < Q(X), X >$. The *scalar curvature* of M^n at x is the trace of the Ricci tensor,

$$S = \sum_1^n < Q(E_i), E_i > = \sum_1^n Ricc(E_i).$$

We define the *Schouten tensor*, $\gamma : T_xM^n \to T_xM^n$ as:

$$\gamma(X) = \frac{1}{n-2}[Q(X) - \frac{SX}{2(n-1)}],$$

and the *Weyl tensor* $W : T_xM^n \times T_xM^n \to End(T_xM^n)$ as:

$$W(X,Y) = R(X,Y) - \gamma(X) \wedge Y - X \wedge \gamma(Y),$$

where $(Z \wedge K)T := < Z, T > K - < K, T > Z$.

The basic (pointwise) characterization of conformally flat manifolds is the following:

3.6.4. THEOREM. *let $n \geq 3$. Then M^n is conformally flat if and only if:*

(1) $W = 0$.
(2) γ *is a Codazzi tensor, i.e.*

$$(\nabla_X \gamma)(Y) = (\nabla_Y \gamma(X), \forall X, Y \in T_xM^n, \forall x \in M^n,\ X, Y \in T_xM^n.$$

Moreover, if $n = 3$ the Weyl tensor always vanishes, and if $n \geq 4$, the vanishing of the Weyl tensor implies that γ is Codazzi.

We will prove the following characterization of conformally flat hypersurfaces, due originally to Cartan:

3.6.5. THEOREM. *Let M^n be a Riemannian manifold, $n \geq 4$, and $f : M^n \to \mathbb{R}^{n+1}$ be an isometric immersion. Then M^n is conformally flat if and only if f is* quasi-umbilic *i.e., the shape operator has an eigenvalue of multiplicity at least $n-1$. In particular, conformally flat hypersurfaces are quasi-convex.*

PROOF. Let $\{E_1, \dots, E_n\}$ be an orthonormal basis of T_xM^n such that $A_\xi E_i = \lambda_i E_i,\ (x, \xi) \in \nu^1 M$. Then, by the Gauss equation, we get:

$$\gamma(E_i) = \frac{1}{n-2}[Ricc(E_i) - \frac{S}{2(n-1)}]E_i.$$

Therefore, the Weyl tensor vanishes if and only if:

$$(n-2)\lambda_i\lambda_j = Ricc(E_i) + Ricc(E_j) - \frac{S}{n-1},\ i, j = 1, \dots, n.$$

Let i, j, k, l be distinct indices. If $W = 0$, the above equation gives:

$$\lambda_i\lambda_j + \lambda_k\lambda_l - \lambda_i\lambda_k - \lambda_l\lambda_j = (\lambda_i - \lambda_l)(\lambda_j - \lambda_k) = 0,$$

The above condition is verified for all four distinct indices if and only if at list $n-1$ of the λ's are equal i.e., if and only if the immersion is quasi-umbilic. Conversely it is obvious that if f is quasi-umbilic, then M^n is conformally flat. $\square$

3.6.6. REMARK. If $f : M^3 \to \mathbb{R}^4$ is a quasi-umbilic immersion, it is easily seen, that M^3 is conformally flat i.e., it's Schouten tensor is Codazzi. However there are example of isometric immersions of conformally flat 3-manifolds with distinct principal curvatures. The classification of such immersions is still an open problem.

3.6.7. REMARK. More is known on the structure of compact conformally flat hypersurfaces of $\mathbb{R}^{n+1}$. In fact is proved in [**?**], that:

3.6.8. THEOREM. *Let $f : M^n \to \mathbb{R}^{n+1}$ be an isometric immersion of a compact, oriented, connected, conformally flat manifold, $n \geq 4$. Then M^n is conformally diffeomorphic to a sphere S^n with "handles" of type $[0,1] \times S^{n-1}$ attached.*

Observe that the above result is quite analogous to the classification of compact orientable surfaces.

For isometric immersion of conformally flat manifolds in higher codimension, we have the following generalization of the Cartan's result due to J. D. Moore [**?**]:

3.6.9. THEOREM. *Let $f : M^n \to \mathbb{R}^{n+p}$ be an isometric immersion of a conformally flat manifold, $p \leq n-3$. Then, $\forall x \in M^n$ there exist a subspace $U \subseteq T_xM^n$ of dimension at least $(n-p)$ and $\xi \in \nu_x^1 M$, such that the second fundamental form, restricted to U, is given by:*

$$\alpha(X,Y) =< X, Y > \xi.$$

Again applying Theorem **??** to the height functions we get:

3.6.10. COROLLARY. *$f : M^n \to \mathbb{R}^{n+p}$ be an isometric immersion of a compact, connected, conformally flat manifold, $p \leq n-3$. Then M^n has the homotopy type of a CW-complex with no cells in dimension $k, p < k < n-p$. In particular the homology vanishes in that range of dimensions.*

PROOF. Let ξ be a regular value of the Gauss map. Then the Hessian of h_ξ has, at a critical point, an eigenvalue of multiplicity at least $(n-p)$. hence the index is smaller or equal to p or greater or equal to $(n-p)$ and the conclusion follows. □

3.6.2. Manifolds with nonnegative isotropic curvature.

One of the reasons why sectional curvature is a basic invariant in Riemannian geometry is that it appears in an important way in the formula of the second variation of the energy functional, giving therefore informations on the stability and, more in general, on the index of geodesics. It is a classical technique to use those information to study the topology of the manifold. If we look at the space of sufficiently smooth maps from a surface Σ to a Riemannian manifold, we have an energy functional:

$$E(\phi) = \int_\Sigma \|d\phi\|^2 d\Sigma,$$

whose critical points are the "harmonic maps". In order to study the topology of the target manifold, we are naturally lead to consider the corresponding index form.

This program was essentially introduced in [**?**]and it turns out that the convenient invariant to study this index is the concepts of *isotropic curvature* that we will describe now.

Let M be a Riemannian manifold. For $x \in M$ we consider the *complexified tangent space* $T_xM^{\mathbb{C}} = T_xM \oplus i\,T_xM$ and we consider the unique extensions of the Riemannian inner product $\langle\cdot,\cdot\rangle$ of T_xM to a complex bilinear form $\langle\cdot,\cdot\rangle_{\mathbb{C}}$ in $T_xM^{\mathbb{C}}$ and to a Hermitian inner product $\langle\cdot,\cdot\rangle_{\underline{\mathbb{C}}}$ in $T_xM^{\mathbb{C}}$; more explicitly:

$$\langle v_1 + iv_2, w_1 + iw_2\rangle_{\mathbb{C}} = \langle v_1, w_1\rangle - \langle v_2, w_2\rangle + i\big(\langle v_2, w_1\rangle + \langle v_1, w_2\rangle\big), \tag{3.6.1}$$

$$\langle v_1 + iv_2, w_1 + iw_2\rangle_{\underline{\mathbb{C}}} = \langle v_1, w_1\rangle + \langle v_2, w_2\rangle + i\big(\langle v_2, w_1\rangle - \langle v_1, w_2\rangle\big), \tag{3.6.2}$$

for all $v_1, v_2, w_1, w_2 \in T_xM$.

3.6.11. DEFINITION. A complex subspace $S \subset T_xM^{\mathbb{C}}$ is called *isotropic* if $\langle v, w\rangle_{\mathbb{C}} = 0$ for all $v, w \in S$.

Obviously $S \subset T_xM^{\mathbb{C}}$ is isotropic if and only if the complex subspaces S and $\overline{S} = \{\bar v : v \in S\}$ are orthogonal with respect to $\langle\cdot,\cdot\rangle_{\underline{\mathbb{C}}}$. In particular, if S is isotropic then $S \cap \overline{S} = \{0\}$ and $\dim_{\mathbb{C}}(S) \le \dim(M)$. The following lemma shows how one can construct isotropic subspaces of $T_xM^{\mathbb{C}}$.

3.6.12. LEMMA. *If $(b_j)_{j=1}^{2r}$ is an orthonormal family in T_xM then the family $\big(\frac{1}{\sqrt{2}}(b_j + ib_{r+j})\big)_{j=1}^{r}$ is a $\langle\cdot,\cdot\rangle_{\underline{\mathbb{C}}}$-orthonormal complex basis for an isotropic subspace S of $T_xM^{\mathbb{C}}$.*

Conversely, if $S \subset T_xM^{\mathbb{C}}$ is an isotropic subspace and if $(Z_j)_{j=1}^{r}$ is a $\langle\cdot,\cdot\rangle_{\underline{\mathbb{C}}}$-orthonormal complex basis for S then $\big(\sqrt{2}\,\Re(Z_j), \sqrt{2}\,\Im(Z_j)\big)_{j=1}^{r}$ is an orthonormal family in T_xM, where $\Re(Z_j), \Im(Z_j) \in T_xM$ denote respectively the real and imaginary parts of $Z_j \in T_xM^{\mathbb{C}}$.

PROOF. It is a straightforward calculation using (3.6.1) and (3.6.2). □

For every $x \in M$ we now consider the unique extension of the trilinear map:

$$T_xM \times T_xM \times T_xM \ni (v_1, v_2, v_3) \longmapsto R_x(v_1, v_2)v_3 \in T_xM$$

to a map $R_x^{\underline{\mathbb{C}}} : T_xM^{\mathbb{C}} \times T_xM^{\mathbb{C}} \times T_xM^{\mathbb{C}} \to T_xM^{\mathbb{C}}$ that is complex linear in the first two variables and conjugate linear in the third. We write $R^{\underline{\mathbb{C}}}(X,Y)Z$ for the value of $R_x^{\underline{\mathbb{C}}}$ on a triple (X, Y, Z) (we will usually omit the point $x \in M$ for simplicity). From the standard symmetries of the curvature tensor, one easily obtains the following identities:

$$R^{\underline{\mathbb{C}}}(X,Y)Z = -R^{\underline{\mathbb{C}}}(Y,X)Z, \quad \big\langle R^{\underline{\mathbb{C}}}(X,Y)Z, T\big\rangle_{\underline{\mathbb{C}}} = -\big\langle R^{\underline{\mathbb{C}}}(X,Y)T, Z\big\rangle_{\underline{\mathbb{C}}},$$
$$\big\langle R^{\underline{\mathbb{C}}}(X,Y)Z, T\big\rangle_{\underline{\mathbb{C}}} = \overline{\big\langle R^{\underline{\mathbb{C}}}(Z,T)X, Y\big\rangle_{\underline{\mathbb{C}}}},$$

for every $X, Y, Z, T \in T_xM^{\mathbb{C}}$. In particular, $\big\langle R^{\underline{\mathbb{C}}}(X,Y)X, Y\big\rangle_{\underline{\mathbb{C}}}$ is a real number.

Given $\mathbb{C}$-linearly independent vectors $Z, W \in T_xM^{\mathbb{C}}$, we define the *complex sectional curvature* of the complex plane spanned by Z and W to be the real number:

$$K^{\mathbb{C}}(Z,W) = -\frac{\big\langle R^{\mathbb{C}}(Z,W)Z, W\big\rangle_{\mathbb{C}}}{\langle Z,Z\rangle_{\mathbb{C}}\langle W,W\rangle_{\mathbb{C}} - \big|\langle Z,W\rangle_{\mathbb{C}}\big|^2} \in \mathbb{R}.$$

It is easy to see that $K^{\mathbb{C}}(Z,W)$ depends only on the complex plane spanned by Z and W and not on the particular basis chosen on such plane (see Exercise 3.16).

We will say that an n-dimensional Riemannian manifold M ($n \ge 4$) has *non negative isotropic curvature* if $K^{\mathbb{C}}(Z,W) \ge 0$ for every $x \in M$ and every $Z, W \in T_xM^{\mathbb{C}}$ that form a basis for an isotropic subspace of $T_xM^{\mathbb{C}}$.

3.6.13. LEMMA. *Assume that M has non negative isotropic curvature. Then, for every $x \in M$ and every orthonormal family (e_1, e_2, e_3, e_4) in T_xM, we have:*

$$K_{12} + K_{14} + K_{23} + K_{34} \ge 0,$$

where K_{ij} denotes the sectional curvature of M in the plane spanned by e_i and e_j.

PROOF. Set $Z = e_1 + ie_3$ and $W = e_2 + ie_4$. It is easy to see that Z and W form a (complex) basis for an isotropic plane in $T_xM^{\mathbb{C}}$. A straightforward computation using the standard symmetries of the curvature tensor R shows that the isotropic curvature corresponding to such plane is given by:

$$K^{\mathbb{C}}(Z,W) = K_{12} + K_{14} + K_{23} + K_{34} + 2\big\langle R(e_3,e_1)e_4, e_2\big\rangle.$$

Similarly, the isotropic curvature corresponding to the complex plane spanned by $\overline{Z} = e_1 - ie_3$ and W is given by:

$$K^{\mathbb{C}}(\overline{Z},W) = K_{12} + K_{14} + K_{23} + K_{34} - 2\big\langle R(e_3,e_1)e_4, e_2\big\rangle.$$

Adding the two (non negative) isotropic curvatures $K^{\mathbb{C}}(Z,W)$ and $K^{\mathbb{C}}(\overline{Z},W)$ we have the desired conclusion. □

3.6.14. THEOREM. *Let M be a compact n-dimensional Riemannian manifold ($n \ge 4$) having non negative isotropic curvature. Then every isometric immersion $f : M^n \to \mathbb{R}^{n+1}$ is quasi-convex.*

3.6.15. REMARK. Using estimates of the index of harmonic spheres in a Riemannian manifold as well as a quite sophisticated Morse Theory for the energy functional on the space of H^1 maps of S^2 into a Riemannian manifold, it was proved in [**?**] the following beautiful result:

3.6.16. THEOREM. *Let M^n, $n \ge 4$ be a compact, simply connected Riemannian manifold with positive isotropic curvature. Then M^n is homeomorphic to the sphere S^n.*

It is an open problem if, in the above hypothesis, M^n is diffeomorphic to a sphere.

3.7. Hypersurfaces of finite geometric type.

Let $f : M = M^n \to R^{n+p}$ be an isometric immersion of an n-dimensional Riemannian manifold. Recall that the *mean curvature vector* H is defined as the trace of the second fundamental form (see exercise?). If $H = 0$, the immersion is called a *minimal immersion*. Minimal immersions are the critical points of the area functional, i.e., if $D \subseteq M$ is a compact domain and f_t is a family of immersions of D with $f_t|\partial D = f|\partial D$, the function:

$$A(t) = \int_D dM_t,$$

where dM_t is the volume density induced by f_t on D, has zero derivative at $t = 0$.

For $n = 2, p = 1$ the theory of minimal surfaces in $\mathbb{R}^3$ is a classical and very extended topic in differential geometry and complex analysis, at least if M is orientable. The main point is that, in this case, the classical Gauss map is an holomorphic function into $S^2 \subset \mathbb{R}^3$ and the immersion can be recovered by complex analytic methods, starting from the Gauss map and the metric (Enneper-Weierstrass representation theorem). In the class of orientable, complete minimal surfaces the subclass of the ones with finite total curvature, i.e. $\int_{M^2} kdM > -\infty$,[4] is a very important one and has quite interesting topological-geometric properties. We list some of them:

- M is conformally diffeomorphic to a compact Riemann surface $\overline{M}$ minus a finite number of points, say $p_1, \dots p_k \in \overline{M}$. The points p_i are called the *ends* of M.
- The (classical) Gauss map $\mathfrak{G} : M \to S^2$ extend to an holomorphic map $\overline{\mathfrak{G}} : \overline{M} \to S^2$. In particular it is singular on a finite set if M is not flat (hence not totally geodesic).
- For each end $p_i \in \overline{M}$ there exist a neighborhood U_i such that the composition of $f|(U_i \setminus p_i)$ with the projection onto $\overline{\mathfrak{G}}(p_i)^\perp$ is a finite covering of order $I(p_i)$ over the complement of a ball in $\overline{\mathfrak{G}}(p_i)^\perp$.

We consider now a class of oriented hypersurface which share the properties of minimal surfaces of finite total curvature. From now on, by the *Gauss map* $\mathfrak{G}$, we will intend the *classical* Gauss map, i.e. the restriction of the Gauus map to one of the components of $\nu^1 M$.

3.7.1. DEFINITION. An immersion $f : M \to \mathbb{R}^{n+1}$ of an n-dimensional, connected, oriented manifold is of *finite geometric type* if:

(1) M is complete in the induced metric.
(2) M is diffeomorphic to $\overline{M} \setminus \{p_1, \dots, p_k\}$ where $\overline{M}$ is compact, and $\mathfrak{G} : M \to S^n$ extend to a smooth map $\overline{\mathfrak{G}} : \overline{M} \to S^n$.
(3) For each end $p_i \in \overline{M}$ there exist a neighborhood U_i such that the composition of $f|(U_i \setminus p_i)$ with the projection onto $\overline{\mathfrak{G}}(p_i)^\perp$ is a finite covering of order $I(p_i)$ over the complement of a ball in $\overline{\mathfrak{G}}(p_i)^\perp$.

[4] Observe that the Gaussian curvature of a minimal surface in $\mathbb{R}^3$ is always nonpositive.

(4) The Gauss-Kronecker curvature $G(x) = det(A_{\mathfrak{G}(x)})$ is zero only on a finite union of connected submanifolds of dimension $\leq n-2$.

3.7.2. REMARK. The integer $I(p_i)$ is still called the *geometric index of the immersion at the end* p_i. If $n \geq 3, I(p_i) = 1$ since the complement of a ball, in those dimensions, is simply connected. If $n = 2, I(p_i)$ is the number of times that $f|(U_i \setminus p_i)$ wings around $\overline{\mathfrak{G}}(p_i)^\perp$. In particular $I(p_i) = 1$ if and only if $f|(U_i \setminus p_i)$ is an embedding. In this case we will say that the *end is embedded*.

3.7.3. REMARK. It is proved in [JM] that condition (3) in the definition above is really a consequence of conditions (1) and (2) . In fact much moore is proved in that paper. In particular the fact that the Gauss map extends to an end p means that M has a “tangent space” at p in the following strong sense: The intersection of $f(M)$ with a sphere of a large radius R, normalized on the unit sphere $S^n(1)$, converges in the C^1 topology, when $R \to \infty$, to the sphere $S^n(1) \cap \overline{\mathfrak{G}}(p)^\perp$. It follows that if f is an embedding, the extended Gauss map assumes, at the ends at most two values, and, in this case, the values are opposite.

The Gauss-Kronecker curvature is well define up to sign, since depend on the chose of the orientation. However, if the dimension is even, it is well defined independently of the orientation. Since we will be essentially interested in the even dimensional case, the choise of the orientation will not be a problem. Olso, in the even dimensional case, the condition on the Gauss-Kroneker curvature imply that the total absolute curvature is two, so f embeds M as the boundary of a convex body. So we will make, from now on, the following:

3.7.4. ASSUMPTION. *M is even dimensional, non compact and f is of finite geometric type.*

3.7.5. REMARK. Since the singular points of the Gauss map do not disconnect M, the sign of the Gauss-Kronecker curvature is constant, and we will denote it by σ.

Let $\xi \in S^n$ be a regular value of the Gauss map, and h_ξ be the height function in the ξ direction. Then h_ξ has only non degenerate critical points and the gradient of h_ξ at $x \in M$ is, up to identifying locally M with $f(M)$, the projection of ξ onto T_xM. So the projection of ξ onto the tangent spaces to M gives a smooth vector field, $X = \nabla h_\xi$ whose singularities are the critical points of h_ξ. Since the index of the gradient of a function at a non degenerate critical point of (Morse) index λ in $(-1)^\lambda$, we get:

3.7.6. LEMMA. *The index of X at a singular point is σ.*

We will study now the behavior of X near the ends.

3.7.7. LEMMA. *The index of X at an end p such that $\overline{\mathfrak{G}}(p) \neq \pm\xi$ is $1 + I(p)$.*

PROOF. We consider, first, the case when the end is embedded, i.e. $I(p) = 1$. Since $\overline{\mathfrak{G}}(p) \neq \pm\xi,\ (df)(X_\xi)$ is an almost constant vector field along f (in a small neighborhood of the end) whose projection on the hyperplane $\overline{\mathfrak{G}}(p)^\perp$ has norm bounded away from zero in a neighborhood of infinity. Therefore, the index of the projection, along a big sphere in $\overline{\mathfrak{G}}(p)^{\perp}$, is zero and the projection extends to a non vanishing vector field on the interior of the sphere. We project the extended field on the unit sphere of R^{n+1} by stereographic projection obtaining a vector field $\tilde{X}_{\xi}$ on the unit sphere with only one singularity, at the south pole. Consequently the index of $\tilde{X}_{\xi}$ is $1 + (-1)^n = 2$. Since the composition of the immersion, projection onto $\overline{\mathfrak{G}}(p)^{\perp}$ and stereographic projection is an orientation preserving diffeomorphism of a small neighborhood of p onto a small neighborhood of the south pole, which send X_{ξ} onto $\tilde{X}_{\xi}$, the conclusion follows.

For the non embedded case (which occurs only for $n = 2$), we recall the tangency formula for computing index of a singularity of a plane vector field:
Let γ be a closed simple curve around a singularity such that the field is non zero along γ and tangent only at a finite number of points. Let n_e the number of points of γ where the integral curve of the vector field is (locally) outside γ and n_i the number of points where the integral curve is (locally) inside γ. Then the index of the vector field is $(2 + n_i - n_e)/2$.
Going back to the case in question, we consider a simple closed curve γ around p. Since the composition of the immersion and the projection onto $\overline{\mathfrak{G}}(p)^{\perp}$ is an $I(p)$-fold covering in a small punctured neighborhood of p, the image α of γ is a closed curve in $R^2 - \{(0, 0)\}$ with winding number $I(p)$. Up to homotopy, we can suppose that γ is an $I(p)$-fold covering of a closed simple curve. We proceed as above and observe that, for each lap, the projected vector field has index $0 = (2 + n_i - n_e)/2$. Therefore, $n_i - n_e = -2$ for each complete lap. We observe that, external (resp. internal) tangency of the flow of the projected field along α corresponds to internal (resp. external) tangency along γ of the flow of X_{ξ}. Therefore, the index of X_{ξ}, along γ, is $(2 + I(p)(n_e - n_i))/2 = 1 + I(p)$. □

From the above, summing the indeces of the vector field, we obtain:

3.7.8. THEOREM. *The Euler characteristic of $\overline{M}$ is*

$$\chi(\overline{M}) = \sum_i (1 + I(p_i)) + 2\sigma m,$$

where m is the degree of the (classical) Gauss map.

We will give naw some applications of the above formula.

Since the Gauss map of a minimal surface is holomorphic, and the tendency of an holomorphic map is to be surjective, most attention has been posed on the problem of determining the “size” of the image of the Gauss map. In this context the best result was obtained by Fujimoto in 1988 who proved that the image of the Gauss map of a *complete*, non flat minimal surface can omits at most four points (and there are many examples where $\mathfrak{G}$ omits exactly four points). In the context of non flat, complete minimal surfaces of finite total curvature, it was proved by Osserman in 1964 that the Gauss map of such a surface omits at most three point. Clearly the catenoid is an example of a surface of the above type whose Gauss map omits two points. It is still an open problem if there are example of complete, non flat minimal surfaces of finite total curvature, whose Gauss map omits exactly three points. Using our arguments we will give naw a proof of Osserman's theorem in the more general context of surfaces of finite geometric type.

3.7.9. THEOREM. *Let $f : M^2 \to \mathbb{R}^3$ be an immersion of finite geometric type. Then the Gauss map omits at most three point.*

PROOF. By hypothesis, the Gauss map $\overline{\mathfrak{G}} : \overline{M} \to S^2(1)$ is a branched covering, branched (possibly) at the flat points and at the ends. At a branch point p, the branching number $\nu(p)$ is the cardinality of the intersection of a small neighborhood of p with the inverse image of a regular value near $\overline{\mathfrak{G}}(p)$. So, the branching number is, always, at least one and bigger than one only at the *effective branch points* which, by our assumptions, are finite in number. In this situation we have the so called Riemann-Hurwitz formula:

$$\chi(\overline{M}) = m\chi(S^2) + \sum(1-\nu(p)) = 2m + \sum(1-\nu(p)) \ . \tag{3.7.1}$$

Let us suppose that the Gauss map omits n points, $\xi_1, \dots, \xi_n$. Let $A_i = \{p \in \overline{M} : \overline{\mathfrak{G}}(p) = \xi_i\}$, $B = \{p \in \overline{M} : \overline{\mathfrak{G}}(p) \neq \xi_i\ \}$ and $C = \{q \in M; \nu(q) > 1\}$. Let ξ be a regular value of $\mathfrak{G}$, $\xi \neq \xi_i$. We write the above formula in the following form:

$$\chi(\overline{M}) = 2m + \sum_{i=1}^{n}\sum_{p\in A_i}(1-\nu(p)) + \sum_{p\in B}(1-\nu(p)) + \sum_{p\in C}(1-\nu(p)) \ . \tag{3.7.2}$$

Observe that $\sum_{p\in A_i}\nu(p) = m$ and $\sum_{i=1}^n |A_i| + |B| = k$. Then:

$$\chi(\overline{M}) = (2-n)m + k - \sum_{p\in B}\nu(p) + \sum_{p\in C}(1-\nu(p)) \ . \tag{3.7.3}$$

Comparing with Equation ?, we obtain:

(3.7.4)

Therefore, $n < 4$, as claimed.

□

A simple analysis of the proof gives the following

3.7.10. COROLLARY. *On the hypothesis of Theorem (**??**), if $n = 3$ then $\chi(\overline{M}) \leq 0$. Moreover, if $\chi(\overline{M}) = 0$, we have:*

(1) $m = k$
(2) $B = \emptyset = C$, *and*
(3) $\sum I(p_i) = k$, *i.e., each end is embedded.*

3.7.11. REMARK. The proof, in the case of complete minimal surfaces of finite total curvature, is very similar to this one, but for the fact that the basic formulas for the Euler characteristic of $\overline{M}$ are obtained via the Weierstrass representation, which, clearly, does not exist outside the minimal case.

The advantage of this point of view is that it extends to higher dimensional hypersurfaces, while the use of complex analysis is restricted to the case $n = 2$. To stress this point we will prove the following theorem that is new even in the case of minimal hypersurfaces of "finite total curvature":

3.7.12. THEOREM. *Let $f : M^{2n} \to R^{2n+1}$ be an immersion of finite geometric type, $n > 1$. Suppose the critical fibers, i.e. the inverse image by the Gauss map of the critical values, form a stratified subset N of dimension less then $n - 1$. Then:*

(1) *M^{2n} is, topologically, a sphere minus two points.*
(2) *If M^{2n} is minimal, it is a catenoid.*

PROOF. First we observe that since $\overline{M}$ is compact, every regular value of $\overline{\mathfrak{G}}$ has a neighborhood that is evenly covered. In particular $\overline{\mathfrak{G}}|(\overline{M} \setminus N)$ is a covering map. Consider a map $\alpha : S^k \to S^{2n}$. By the standard transversality theorem, up to homotopy, we can suppose that α is transversal to $\overline{\mathfrak{G}}|N$, hence disjoint from $\overline{\mathfrak{G}}(N)$ if $k \leq n$. Therefore the inclusion $S^{2n} \setminus \overline{\mathfrak{G}}(N) \to S^{2n}$ induces an epimorphism between the homotopy groups in dimension $\leq n$. Also, if $k \leq n, \alpha$ is homotopic to a constant, hence it extends to a map $\tilde{\alpha} : D^{k+1} \to S^{2n}$. Applying again the transversality argument to $\tilde{\alpha}$, we may assume that the extended map has image disjoint from $\overline{\mathfrak{G}}(N)$. Therefore $S^{2n} \setminus \overline{\mathfrak{G}}(N)$ has vanishing homotopy up to dimension n, in particular is simply connected. It follows that $\overline{\mathfrak{G}}|(\overline{M}^{2n} \setminus N) : (\overline{M}^{2n} \setminus N) \to S^{2n} \setminus \overline{\mathfrak{G}}(N)$ is a diffeomorphism, hence $\overline{\mathfrak{G}}$ is a map of degree one. It also follows that $\overline{M}^{2n} \setminus N$ has vanishing homotopy groups up to dimension n, hence the k-dimensional homology vanishes, if $k \leq n$. By Poincaré duality the homology vanishes in dimension $k = 1, \ldots, 2n-1$. Hence $\overline{M}^{2n}$ is a simply connected homology sphere, hence homotopy equivalent to a sphere and homeomorphic to a sphere by the positive answer to the generalized Poicaré conjecture.

In particular, Equation (**??**) implies that

$$2 = \chi(\overline{M}) = 2(k + \sigma),$$

hence $k = 2$ and $\sigma = -1$. This prove part the first assertion.
the second assertion follows from a theorem of R. Schoen:
The only minimal immersions which are regular at infinity and have two ends are the catenoid and pairs of planes.
We just observe that, minimal hypersurfaces of finite geometric type, are *regular at infinity* in the sense of Schoen, if the ends are embedded, which is the case since $n > 1$. □

Exercises for Chapter 3

The Fundamental equations of an Isometric Immersion.

EXERCISE 3.1. Let E be a vector bundle over a differentiable manifold M with a connection ∇ with zero curvature. Given $x_0 \in M$ and $X \in T_{x_0}M$, show that there exist a local section $\tilde{X}$ of E, defined in a neighborhood U of x_0, and such that $\nabla_Y \tilde{X} = 0,\ \forall Y \in T_y M, y \in U$.

EXERCISE 3.2. Provide details for the following sketched proof of a simplified version of Theorem ?:

3.7.13. THEOREM. *Let Ω be an open, simply connected subset of $\mathbb{R}^n$¿ Let g be a Riemannian metric in Ω with Levi Civita connection ∇, curvature R and $A : \Omega \to End(\mathbb{R}^n)$ be a tensor field valued in the g-symmetric endomorphisms. Suppose that A verifies the equations of Gauss and Codazzi for $c = 0$, i.e.*

- $R(X,Y) = A(X) \wedge A(Y)$,
- $\nabla_X(AY) - A(\nabla_X Y) = \nabla_Y(AX) - A(\nabla_Y X)$.

Then there exist an isometric immersion $f : (\Omega, g) \to \mathbb{R}^{n+1}$.

PROOF. Consider $T\Omega \oplus \epsilon$, where $\epsilon = \Omega \times \mathbb{R}$, with the direct sum (fiber) metric. Let $\xi : \Omega \to \mathbb{R},\ \xi(x) = 1$. Look at ξ as a section of $T\Omega \oplus \epsilon$. Define a connection ∇' by the rules:

- $\nabla'_X Y = \nabla_X Y + < AX, Y > \xi$,
- $\nabla'_X \xi = -AX$.

The equations of Gauss and Codazzi imply that the curvature of ∇' is zero. Let $p \in \Omega, \{x_1, \dots, x_n\}$ be coordinates in Ω such that $\{\frac{\partial}{\partial x_i}(p)\}$ is a g-orthonormal basis at p. Let $E_i = \frac{\partial}{\partial x_i}(p), E_{n+1} = \xi(p)$. Since ∇' is flat and Ω is simply connected, we can extend the above basis to a ∇'-parallel orthonormal frame field $\{\tilde{E}_1, \dots, \tilde{E}_{n+1}\}$. Then:

$$\frac{\partial}{\partial x_i} = \sum_{k=1}^{n} a_{ik} \tilde{E}_k, \quad a_{ik} : \Omega \to \mathbb{R},$$

and $g_{ij} = g(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}) = \sum_{k=1}^{n_1} a_{ik} a_{jk}$. Since the $\tilde{E}_i$'s are parallel,

$$\nabla'_{\frac{\partial}{\partial x_i}} \frac{\partial}{\partial x_j} = \sum_{k=1}^{n+1} \frac{\partial a_{jk}}{\partial x_k} \tilde{E}_j.$$

Using the symmetry of A, we get:

$$\frac{\partial a_{jk}}{\partial x_i} = \frac{\partial a_{ik}}{\partial x_j}.$$

Hence, since Ω is simply connected, there exist functions $f_k : \Omega \to \mathbb{R}, \frac{\partial f_k}{\partial x_j} = a_{jk}$. Then the map $f = (f_1, \dots, f_{n+1})$ gives the desired immersion. □

3.7.14. REMARK. The proof of Theorem ? goes essentially on the same lines.

EXERCISE 3.3. Let $f : M \to \mathbb{R}^{n+p}$ be an isometric immersion of a n-dimensional Riemannian manifold M. Suppose there exist a 1-dimensional subbundle $L \subseteq \nu M$ such that $\alpha_x(X,Y) \in L_x \ \forall x \in M,\ X, Y \in T_x M$. Prove that M admits a (local) isometric immersion into $\mathbb{R}^{n+1}$. In particular if $p = 1$ and N is a q-dimensional totally geodesic submanifold of M, then N admits a local isometric immersion into $\mathbb{R}^{q+1}$. Discuss the example of an helix on a cylinder.

EXERCISE 3.4. Let $f : M \to \overline{M}$ be an isometric immersion with $\dim(M) = n$, $\dim(\overline{M}) = n + p$. Let E denote the vector bundle $f^*T\overline{M}$ over M; as usual, we identify TM with a subbundle of E using $\mathrm{d}f$ and $\iota : TM \to E$ will denote the inclusion. We have a direct sum $E = TM \oplus \nu M$ and E is endowed with the connection $f^*\overline{\nabla}$ which is the pull-back by f of the Levi–Civita connection $\overline{\nabla}$ of $\overline{M}$. The projections of $\overline{\nabla}$ in TM and νM (in the sense of Exercise **??**) are respectively the Levi–Civita connection ∇ of M and the normal connection $\nabla^\perp$ of the immersion f; the second fundamental form of TM in E with respect to νM is the usual second fundamental form $\mathbb{I}$ of the immersion f and the second fundamental form of νM in E with respect to TM is given by:

$$T_xM \times \nu_xM \ni (v, \eta) \longmapsto -A_\eta(v) \in T_xM,$$

for all $x \in M$. The vector bundle E has a Riemannian structure on its fibers induced from the Riemannian metric of $\overline{M}$; such Riemannian structure is parallel with respect to the connection $f^*\overline{\nabla}$. Let $(X_1, \ldots, X_{n+p})$ be a local orthonormal referential of E with $(X_i)_{i=1}^n$ an orthonormal referential of TM, so that we also obtain an orthonormal referential $(X_\alpha)_{\alpha=n+1}^{n+p}$ of νM; denote by $(\overline{\theta}_1, \ldots, \overline{\theta}_{n+p})$ the dual referential of $(X_1, \ldots, X_{n+p})$ and by $(\theta_i)_{i=1}^n$ the dual referential of $(X_i)_{i=1}^n$, so that $\overline{\theta}_i \circ \iota = \theta_i$ for $i = 1, \ldots, n$. Associated to the given orthonormal frame and connections of the vector bundles TM, νM and E we have associated connection and curvature forms ω, Ω, $\omega^\perp$, $\Omega^\perp$, $\overline{\varpi}$ and $\overline{\Omega}$ respectively (recall Exercise 2.22). We will use Latin letters i, j for indices ranging in $1, \ldots, n$ and Greek letters α, β for indices ranging in $n + 1, \ldots, n + p$. Show that:

(a) $\overline{\varpi}_{ij} = \omega_{ij}$, $\overline{\varpi}_{\alpha\beta} = \omega^\perp_{\alpha\beta}$ and $\overline{\varpi}_{\alpha i} = -\overline{\varpi}_{i\alpha} = A_{X_\alpha}(X_i)$ where we identify the vector $A_{X_\alpha}(X_i) \in TM$ with the covector $\langle A_{X_\alpha}(X_i), \cdot\rangle$.
(b) $f^*\overline{\nabla}$ has zero torsion.
(c) Show that equation (2.6.5) for $\overline{\varpi}$ and $\overline{\theta}$ is equivalent to the symmetry of the Weingarten operator.
(d) Show that equation (2.6.4) for $\overline{\varpi}$ and $\overline{\Omega}$ is equivalent to the following:

$$\overline{\Omega}_{ij} = \Omega_{ij} - \sum_\alpha A_{X_\alpha}(X_i) \wedge A_{X_\alpha}(X_j), \qquad \text{(Gauss)}$$

$$\overline{\Omega}_{\alpha i} = \mathrm{d}A_{X_\alpha}(X_i) + \sum_j A_{X_\alpha}(X_j) \wedge \omega_{ji} + \sum_\beta \omega^\perp_{\alpha\beta} \wedge A_{X_\beta}(X_i), \qquad \text{(Codazzi)}$$

$$\overline{\Omega}_{\alpha\beta} = \Omega^\perp_{\alpha\beta} - \sum_i A_{X_\alpha}(X_i) \wedge A_{X_\beta}(X_i). \qquad \text{(Ricci)}$$

EXERCISE 3.5. Let $f : M \to \overline{M}$ be an isometric immersion. Then f (or sometimes M) is said to be totally geodesic if the second fundamental form vanishes identically. Show that f is totally geodesic if and only is, for all geodesic $\gamma : (a, b) \to M$, $f \circ \gamma$ is a geodesic in $\overline{M}$. Determine the totally geodesic submanifolds of the spaces of constant curvature.

EXERCISE 3.6. Let $f : M \to \mathbb{R}^{n+p}$ be an isometric immersion of a n-dimensional Riemannian manifold M. The *mean curvature vector field* $H : M \to$

νM, is the trace of the second fundamental form. More precisely, if $\{E_1, \ldots, E_n\}$ is an orthonormal basis for T_xM and $\{\xi_1, \ldots \xi_p\}$ is an orthonormal basis for $\nu_x M$,

$$H(x) = \frac{1}{n}\sum_{i=1}^{n} \alpha_x(E_i, E_i) = \frac{1}{n}\sum_{i=1}^{p} trace(A_{\xi_i})\xi_i.$$

Suppose that f is *totally umbilical*, i.e. $A_\xi = \lambda(\xi)Id, \ \forall \xi \in \nu M$.

(1) Show that $\alpha(X, Y) =< X, Y > H$, and conclude that $A_\xi = 0$ if $< \xi, H >= 0$.
(2) Show that $\nabla^{\perp}_{E_i} H = 0, \ i = 1, \ldots, n$. Conclude that $\|H\|$ is locally constant.
(3) Let $x \in M$ be a fixed point and $\gamma : [0, \epsilon) \to M$ be a smooth curve with $\gamma(0) = x$. Let $\eta : [0, \epsilon) \to \nu M$ be a normal vector field along γ, parallel in the normal connection, with $< \eta(0), H(x) >= 0$. Show that $< \eta(t), H(\gamma(t)) >= 0$ and use this to prove that $\overline{\nabla}_{\dot\gamma(t)}\eta(t) = 0$.
(4) Show that $< f(\gamma(t)) - f(x), \eta(t) >= 0 \ \forall t \in [0.\epsilon)$ and conclude that $f(\gamma(t))$ belongs to the affine subspace $\mathbb{A}(x)$ passing trough x and spanned by $H(x)$ and T_xM. Observe that this subspace is either n-dimensional or $(n+1)$-dimensional, depending if $H(x) = 0$ or not.
(5) Suppose M connected. Show that if $H = 0, \ f(M) \subseteq \mathbb{A}(x) = T_xM$, and is $H \neq 0$, the function:

$$c(x) = f(x) + \|H\|^{-2}H$$

is constant and therefore $f(M)$ is contained in the sphere of $\mathbb{A}$ centered at c and of radius $\|H\|^{-1}$.

EXERCISE 3.7. Let $f : M \to \mathbb{R}^{n+p}$ be an isometric immersion of a n-dimensional *connected* Riemannian manifold M. Suppose there exist a q-dimensional subbundle $L \subseteq \nu M$ such that L contains the image of the second fundamental form and L is *parallel*, i.e., if $\xi \in \Gamma(L), \ \nabla^{\perp}_X \xi \in \Gamma(L), \ \forall X \in TM$. Observe that in this case, the orthogonal complement of L in the normal bundle is also parallel. Use the ideas of the previous exercise to show that $f(M)$ is contained in the affine subspace trough a point $f(x)$, spanned by T_xM and L_x. Compare this fact with the results of exercise ??.

EXERCISE 3.8. Let $f : M \to \mathbb{R}^{n+1}$ be an isometric immersion of a n-dimensional Riemannian manifold M, and ξ a unit normal field. Let λ be a principal curvature, i.e. an eigenvalue of the shape operator A_ξ and suppose λ has constant multiplicity d in an open set $U \subseteq M$. It is known that the distribution:

$$\mathcal{D}_\lambda = Ker(A_\xi - \lambda Id),$$

is smooth in U.

(1) Prove that $\mathcal{D}_\lambda$ is integrable and if $d \geq 2, \ \lambda$ is constant along the integral leaves of $\mathcal{D}$. (Hint: Use the Codazzi equations).
(2) Show that the leaves of $\mathcal{D}$ are totally umbilical in $\mathbb{R}^{n+1}$. If $\lambda = 0$ they are actually totally geodesic. In this case, show that the affine tangent space is constant along any geodesic of a leaf.

EXERCISE 3.9. Let M be an n-dimensional differentiable manifold and $f : M \to \mathbb{R}^{n+1}$ a smooth immersion. Assume that for every $x \in M$, $f(M)$ is contained in one closed half-space determined by the affine hyper-plane $f(x) + \mathrm{Im}(\mathrm{d}f_x)$. Prove that M is orientable.

EXERCISE 3.10. Consider the surface $M = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^{2n} + x_2^{2m} + x_3^{2k} - 1 = 0\}$, where n, m, k are odd positive integers. Prove that M is compact and use the Gauss Bonnet theorem to compute the integral of the curvature of M (*hint:* Consider $\overline{x}_1 = x_1^n$, ...)

EXERCISE 3.11. Consider the surface $M = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^{2k} - 1 = 0\}$, k any positive integer. Prove that M is compact and all the height functions have exactly two critical points. Conclude that M is the boundary of a convex body.

EXERCISE 3.12. Let $F : \mathbb{R}^{n+1} \to \mathbb{R}$ be a smooth function and $0 \in \mathbb{R}$ a regular value of F. Consider the hypersurface $M^n = \{x \in \mathbb{R}^{n+1} : F(x) = 0\}$. If $\xi = (1, 0, \dots, 0) \in S^n$, the critical points of the height function h_ξ are solutions of the system:

$$\frac{\partial F}{\partial x_i}(x) = 0, \;\; i = 2, \dots, n+1, \quad F(x) = 0.$$

Use the implicit function theorem to prove that the index of h_ξ at a critical point x is the index of the matrix:

$$-(\frac{\partial F}{\partial x_1}(x))^{-1}(\frac{\partial^2 F}{\partial x_i \partial x_j}(x)), \quad i, j = 2, \dots, n+1.$$

EXERCISE 3.13. Let $F(x_1, \dots, x_{n+1}) = (\sum_{i=1}^{n+1} x_i^2 - 5)^2 - 16(1 - \sum_{i=3}^{n+1} x_i^2)$. Consider $M = F^{-1}(0)$. Assume $n \geq 4$.

(1) Let $\xi = (1, 0, \dots, 0) \in S^n$. Show that the height function h_ξ is a Morse function with exactly four critical points. Compute de index of h_ξ at the critical points and use it to compute the homology of M.
(2) Show that M may be obtained by the following geometric construction: Consider $G = SO(2) \times SO(n-1)$ and the product action $G \times (\mathbb{R}^2 \times \mathbb{R}^{n-1}) \to (\mathbb{R}^2 \times \mathbb{R}^{n-1}) = \mathbb{R}^{n+1}$. Consider the circle γ in the $\{x_1, x_3\}$-plane centered at $(2, 0)$ and of radius 1. Then M is the orbit of γ under the action of G. Conclude that M is a manifold of G- cohomogeneity one, i.e. G is a group of isometries of M such that the minimal codimension of the orbits is one.
(3) Show that M is the image of the map:

$$f : \mathbb{R} \times S^1 \times S^{n-1} \to \mathbb{R}^{n+1}, \quad f(t, u, v) = ((sint + 2)u, costv).$$

Conclude that M is a tube of radius one around the circle $(2u, 0)$.
(4) Use the above considerations to compute the second fundamental form of M, at least at points where f is non singular. Conclude that M is conformally flat.

3.7.15. REMARK. The hypersurface above has quite interesting properties. For example it is shown in [MN?] that, up to the choice of the circle γ, i.e. its plane, center and radius, it is the only compact hypersurface of dimension $n \geq 4$ which is conformally flat, of cohomogeneity one (with respect to a closed subgroup of isometries) and *is not an hypersurface of revolution*, i.e. is nor invariant under the action of a subgroup of isometries of the ambient space which leaves a straight line pointwise fixed.

EXERCISE 3.14. Let V be a real vector space. A *complex structure* on V is a linear endomorphism $J : V \to V$ with $J^2 = -\mathrm{Id}$. Given a complex structure J on V then there is a unique way to extend the scalar multiplication of V to $\mathbb{C}$ so that V becomes a complex vector space and $iv = J(v)$ for all $v \in V$; we denote such complex vector space by (V, J).

Let $J^{\mathbb{C}}$ be the unique complex linear extension of J to $V^{\mathbb{C}}$, so that $(J^{\mathbb{C}})^2$ equals minus the identity of $V^{\mathbb{C}}$. Set:

$$V^{\mathfrak{h}} = \big\{ v \in V^{\mathbb{C}} : J^{\mathbb{C}}(v) = iv \big\},$$
$$V^{\mathfrak{a}} = \big\{ v \in V^{\mathbb{C}} : J^{\mathbb{C}}(v) = -iv \big\};$$

$V^{\mathfrak{h}}$ and $V^{\mathfrak{a}}$ are called respectively the *holomorphic* and the *anti-holomorphic* subspaces of $V^{\mathbb{C}}$ corresponding to the complex structure J of V. Show that:

(1) $V^{\mathfrak{h}}$ and $V^{\mathfrak{a}}$ are complex subspaces of $V^{\mathbb{C}}$;
(2) the maps $(V, J) \ni v \mapsto v - iJ(v) \in V^{\mathfrak{h}}$ and $(V, J) \ni v \mapsto v + iJ(v) \in V^{\mathfrak{a}}$ are respectively a complex linear isomorphism and a conjugate linear isomorphism;
(3) $V^{\mathbb{C}} = V^{\mathfrak{h}} \oplus V^{\mathfrak{a}}$;
(4) $V^{\mathfrak{a}}$ is conjugate to $V^{\mathfrak{h}}$;
(5) if S is a complex subspace of $V^{\mathbb{C}}$ such that $V^{\mathbb{C}} = S \oplus \overline{S}$ then there exists a unique complex structure J on V with $V^{\mathfrak{h}} = S$.

EXERCISE 3.15. Let V be a real vector space and $\langle \cdot, \cdot \rangle$ a positive definite inner product in V. Denote by $\langle \cdot, \cdot \rangle_{\mathbb{C}}$ and $\langle \cdot, \cdot \rangle_{\underline{\mathbb{C}}}$ respectively the complex bilinear and the sesqui-linear extensions of $\langle \cdot, \cdot \rangle$ to $V^{\mathbb{C}}$. A complex subspace S of $V^{\mathbb{C}}$ is called *isotropic* if $\langle \cdot, \cdot \rangle_{\mathbb{C}}$ vanishes on S. Show that:

(1) if $S \subset V^{\mathbb{C}}$ is isotropic if and only if S is $\langle \cdot, \cdot \rangle_{\underline{\mathbb{C}}}$-orthogonal to $\overline{S}$;
(2) if $S \subset V^{\mathbb{C}}$ is isotropic then $S \cap \overline{S} = \{0\}$ and there exists a unique real subspace $W \subset V$ such that $W^{\mathbb{C}} = S \oplus \overline{S}$;
(3) if $\dim(V) = n$ then every isotropic subspace $S \subset V^{\mathbb{C}}$ has complex dimension less than or equal to $\frac{n}{2}$;
(4) if $\dim(V) = n$ is even then the isotropic subspaces of $V^{\mathbb{C}}$ having complex dimension $\frac{n}{2}$ are precisely the holomorphic subspaces corresponding to the complex structures J of V that are anti-symmetric with respect to $\langle \cdot, \cdot \rangle$.

EXERCISE 3.16. If (Z, W) and (Z', W') are bases of the same complex subspace of $T_x M^{\mathbb{C}}$, show that $K^{\mathbb{C}}(Z, W) = K^{\mathbb{C}}(Z', W')$ (*hint*: show that $K^{\mathbb{C}}(Z + \lambda W, W) = K^{\mathbb{C}}(Z, W)$ and that $K^{\mathbb{C}}(\lambda Z, W) = K^{\mathbb{C}}(Z, W)$ for complex $\lambda \neq 0$).

EXERCISE 3.17. Let $f : M^2 \to \mathbb{R}^3$ be an *embedding* of finite geometric type. In this case it is known that the Gauss map assumes, at the ends, at most two values, and, if assume two, they are opposite (see Remark ??). Let k be the number of ends and m the degree of the Gauss map. Suppose that the Gaussian curvature of M never vanishes. Prove that:

(1) $k \leq m + 1$.
(2) $\sum_{i=1}^{k} \nu(p_i) = 2m$.
(3) $2m \leq k \leq m + 1$.

Conclude that $\overline{M}^2$ is homeomorphic to S^2 and $k = 2$.

3.7.16. REMARK. It is known that a complete minimal surface with finite total curvature and two ends is a catenoid. In the context of minimal surfaces of finite total curvature, the above result is originally due to L. Rodrigues (see [R]).

EXERCISE 3.18. Consider $M^{2n} = S^n(2^{-\frac{1}{2}}) \times S^n(2^{-\frac{1}{2}}) = \{(x, y) \in \mathbb{R}^{n_1} \oplus \mathbb{R}^{n+1} : \|x\| = \|y\| = 2^{-\frac{1}{2}}, < x, y >= 0\} \subset S^{2n+1} \subset \mathbb{R}^{2n+2}$.

(1) Compute the second fundamental form of M in S^{2n+1} and in $\mathbb{R}^{2n+2}$. Conclude that M is minimal in S^{2n+1}, i.e., the trace of the second fundamental form of M in S^{2n+1} is zero.
(2) Let $p = (p_1, p_2) \in M$ and $\pi : S^{2n+1} \to p^{\perp} \cong \mathbb{R}^{2n+1}$ be the stereographic projection. It can be shown that the Gauss-Kronecker curvature of $\pi(M)$ vanishes only along $\pi(p_1 \times S^n(2^{-\frac{1}{2}}) \cup S^n(2^{-\frac{1}{2}}) \times p_2)$. Conclude that $\pi(M)$ is an hypersurface of finite geometric type. This shows that the hypothesis of Theorem ?? are "almost optimal".

3.7.17. REMARK. If $f : M^n \to \mathbb{R}^{n+1}$ is an immersion of finite geometric type and the ends are embedded, composing with the inverse of the stereographic projection, we obtain an immersion of M into S^{n+1}. Conversely, if M is an hypersurface of S^{n+1} and $p \in M$, the stereographic projection of M from p, gives an hypersurface of $\mathbb{R}^{n+1}$ which is non compact and of finite geometric type if the condition on the Gauss-Kroneker curvature is verified.

CHAPTER 4

Morse Theory on non Compact Manifolds

4.1. What's not working in the case of non compact manifolds?

If we try to extend the results of Morse theory to the case of non compact manifolds in a naive way we immediately find counter-examples to all of the statements given in Sections **??**, 2.5, **??** and **??**. To start with, consider the height function with respect to the axis of an infinite circular cylinder, i.e., consider the smooth map $f : \mathbb{R} \times S^1 \to \mathbb{R}$ given by the projection onto the first coordinate. The map f has no critical points at all, although $\beta_0(M; \mathbb{Q}) = \beta_1(M; \mathbb{Q}) = 1$; thus, the weak Morse inequalities (and hence also the strong ones) do not hold. Even in the case of bounded functions, trivial counter-examples to the Morse inequalities may be obtained by considering the height function on a sphere with a finite number of points removed. Also the non critical neck principle (and its Corollaries 2.3.12 and **??**) do not hold in the non compact case: let for instance M be the sphere with one point in the equator removed and let $f : M \to \mathbb{R}$ be the height function with respect to the axis passing through the poles. Observe that the (non empty) sublevels of f below the equator are contractible, although the sublevels below the north pole containing a neighborhood of the equator have the homotopy type of the circle S^1.

It is easy to single out the main obstruction caused by the lack of compactness in the proof of the non-critical neck principle: the multiple of the gradient of f whose flow was used to move the levels of f may not be a *complete* vector field. If we find a hypothesis that makes such field complete then the non-critical neck principle (and its Corollaries 2.3.12 and **??**) will work! Observe also that compactness is used in the proof of Proposition 2.5.1 only to guarantee the finiteness of critical points at a critical level (and to make the use of the Corollary 2.3.12 of the non-critical neck principle valid).

In order to guarantee that the vector field X used in the proof of the non critical neck principle is complete in the non compact case, one can use the following strategy: if there exists a complete Riemannian metric on M for which $\|\nabla f\|$ stays away from zero on the inverse image by f of a non critical interval $[a, b]$ then X will be bounded with respect to such complete Riemannian metric and will therefore be a complete vector field.

In order to extend the Morse theory to the case of non compact manifolds we will make an assumption concerning the existence of a complete Riemannian metric with respect to which f satisfies the so called *Palais–Smale condition* which implies in particular that $\|\nabla f\|$ stays away from zero on the inverse image by f of a non critical interval $[a, b]$.

One more essential feature of smooth maps on compact manifolds was used in the proof of Theorem 2.5.5. Namely, we constructed a CW-complex Y inductively, by analyzing the contribution of each critical value of f. It was important to know, however, that this construction had a well-defined starting point: the sublevels f^c of f are empty for $c < \min f$. In the non compact case then it will be important to assume that f is bounded from below in order to generalize Theorem 2.5.5.

In this chapter we will extend Morse theory beyond the realm of compact manifolds; more specifically, we extend Morse theory to the case of (possibly infinite-dimensional) Hilbert manifolds. Many readers could wonder at this point why don't we deal with finite-dimensional non compact manifolds. Well, obviously the theory developed in this chapter also works on the finite-dimensional case; if the reader is more interested in such case, he (she) could just ignore the details of functional analysis and read the theory pretending that it is written for finite-dimensional Riemannian manifolds. It happens, however, that one extremely powerful application of Morse theory appears when one considers functionals defined on spaces of maps between finite-dimensional manifolds; the study of critical points for such functionals is what is usually known as *Calculus of Variations*. The prettiest and simplest application of Morse theory to infinite-dimensional manifolds is the one concerning the energy functional in the space of curves connecting two fixed points in a complete finite-dimensional Riemannian manifold; in that case, critical points are precisely the *geodesics* connecting those points so that Morse theory gives us several interesting global results on Riemannian geometry. We develop this application of Morse theory in full detail.

4.2. Review of Functional Analysis

In this section we recall a few selected topics from basic functional analysis as well as some simple aspects of calculus on Banach spaces and on Banach manifolds. In this section (and actually in the whole chapter) *all vector spaces are assumed to be real*, unless otherwise stated. This assumption may seen a little odd for those who may be familiar with functional analysis books that are almost entirely written only for *complex* vector spaces. Given for instance a normed complex vector space, one can always forget about its complex structure and work with the underlying real normed space. From a topological point of view, this change of scalars is irrelevant, although the field of scalars is important from a linear-algebraic point of view. For instance, in the study of spectral theory for linear operators it is almost impossible to work in the real case, since most of the techniques applied involves holomorphic single-variable (Banach space-valued) functions. But we are not interested in spectral theory and actually all the examples in which we will apply the theory of this section will concern only real spaces; so, although many of the results stated in this section would have a complex-analogue, we prefer to work only in the real case for definiteness.

4.2.1. DEFINITION. Let X be a (real) vector space. We call X

- a *topological vector space* if X is endowed with a topology that makes the vector space operations:
$$X \times X \ni (x, y) \longmapsto x + y \in X, \quad \mathbb{R} \times X \ni (c, x) \longmapsto cx \in X,$$
continuous;
- a *Banach space* if X is endowed with a norm $\|\cdot\| : X \to \mathbb{R}$ that induces a complete metric on X (X is automatically a topological vector space with the topology induced from such metric);
- a *Banachable space* if X is a topological vector space for which *there exists* a norm on X that induces the topology of X and that makes X into a Banach space;
- a *Hilbert space* if X is endowed with an inner product $\langle\cdot,\cdot\rangle : X \times X \to \mathbb{R}$ whose corresponding norm makes X into a Banach space;
- a *Hilbertable space* if X is a topological vector space for which *there exists* an inner product on X that induces the topology of X and that makes X into a Hilbert space.

If X is a Banachable space then a norm $\|\cdot\|$ on X that induces the topology of X will be called a *Banach space norm* for X (every such norm makes X into a Banach space — see Exercise 4.4). Similarly, if X is a Hilbertable space then an inner product on X that induces the topology of X will be called a *Hilbert space inner product* for X (any such inner product makes X into a Hilbert space).

Below we recall some classical examples of Banach and Hilbert spaces. All integrals are always understood to be *Lebesgue integrals*; as usual, the expression "for almost all" (or "almost everywhere") means that some property should hold outside a set of measure zero.

4.2.2. EXAMPLE. Let $f : [a, b] \to \mathbb{R}^n$ be a Measurable function. For every real $p \in [1, +\infty[$ we set:
$$\|f\|_{L^p} = \left(\int_a^b \big\|f(t)\big\|^p \, \mathrm{d}t\right)^{\frac{1}{p}} \in [0, +\infty]$$
where $\|\cdot\|$ denotes an arbitrary norm on $\mathbb{R}^n$ (see also Remark 4.2.5 below). We call $\|\cdot\|_{L^p}$ the *L^p-norm* (corresponding to the chosen norm $\|\cdot\|$ on $\mathbb{R}^n$); when a Measurable function $f : [a, b] \to \mathbb{R}^n$ has finite L^p-norm we usually say that the function f *is in L^p* or that f is a *L^p-function* The *Minkowski inequality* states that for every Measurable functions $f, g : [a, b] \to \mathbb{R}^n$ we have:
$$\|f + g\|_{L^p} \le \|f\|_{L^p} + \|g\|_{L^p};$$
moreover, it is easy to see that $\|f\|_{L^p} = 0$ if and only if $f(t) = 0$ for almost all $t \in [a, b]$. Hence the set of all measurable functions $f : [a, b] \to \mathbb{R}^n$ with $\|f\|_{L^p} < +\infty$ is a subspace of the space of all $\mathbb{R}^n$-valued maps on $[a, b]$ and $\|\cdot\|_{L^p}$ is a semi-norm on it. The corresponding normed space (see Exercise 4.1) is denoted by $L^p\big([a, b], \mathbb{R}^n\big)$. An element of $L^p\big([a, b], \mathbb{R}^n\big)$ is an equivalence class of L^p functions, where the equivalence relation $\sim$ is $f \sim g \Leftrightarrow f = g$ almost everywhere.

Nevertheless, in order to simplify the language, one usually pretends that the elements of $L^p\big([a,b],\mathbb{R}^n\big)$ are functions; obviously, one has to be careful with such attitude in verifying that the statements being made about elements of $L^p\big([a,b]\mathbb{R}^n\big)$ do not depend on the representative of the equivalence class (for instance, you cannot evaluate an element of $L^p\big([a,b],\mathbb{R}^n\big)$ at a point of $[a,b]$). It is well known that $L^p\big([a,b],\mathbb{R}^n\big)$ is a Banach space with the norm $\|\cdot\|_{L^p}$; when $[a,b]$ and n are fixed by the context, we may simply talk about *the space* L^p. Observe that the topology on the L^p space does not depend on the norm chosen in $\mathbb{R}^n$. If the norm $\|\cdot\|$ on $\mathbb{R}^n$ is induced by an inner product $\langle\cdot,\cdot\rangle$ and if $p=2$ then the L^p-norm is induced from the L^2-*inner product* given by:

$$\langle f,g\rangle_{L^2}=\int_a^b\big\langle f(t),g(t)\big\rangle\,\mathrm{d}t,$$

so that $L^2\big([a,b],\mathbb{R}^n\big)$ endowed with $\langle\cdot,\cdot\rangle_{L^2}$ is a Hilbert space. All the theory of L^p-spaces may be developed, more in general, for $\mathbb{R}^n$-valued maps on arbitrary Measure spaces, but we won't need that. It is also usual to define the L^p-space for $p=+\infty$ (see Exercise 4.10); again, we won't use that.

4.2.3. EXAMPLE. If A is an arbitrary set and if $\big(X,\|\cdot\|\big)$ is a Banach space then the set of all bounded maps $f:A\to X$ is again a Banach space endowed with the norm:

$$\|f\|_{\sup}=\sup_{a\in A}\big\|f(a)\big\|;$$

we denote such Banach space by $\mathfrak{B}(A,X)$. If A is a topological space then the subspace $C^0(A,X)$ of $\mathfrak{B}(A,X)$ consisting of continuous maps is closed and therefore it is again a Banach space. Sometimes we may prefer using the notation:

$$\|f\|_{C^0}=\|f\|_{\sup}.$$

Observe that a sequence $(f_n)_{n\ge1}$ in $\mathfrak{B}(A,X)$ converges to some f with respect to $\|\cdot\|_{\sup}$ if and only if $f_n\to f$ uniformly on A. Observe also that, although $\|\cdot\|_{\sup}$ depends on the norm $\|\cdot\|$ of X, if one replaces the norm of X by an equivalent one then the norm $\|\cdot\|_{\sup}$ on $\mathfrak{B}(A,X)$ will also be replaced by an equivalent one. We can thus think of $\mathfrak{B}(A,X)$ as a Banachable space if X is a Banachable space.

4.2.4. EXAMPLE. If $f:[a,b]\to\mathbb{R}^n$ is a map of class C^k ($0\le k<\infty$) then we set:

$$\|f\|_{C^k}=\sum_{i=0}^k\big\|f^{(i)}\big\|_{C^0},$$

where $f^{(i)}$ denotes the i-th derivative of f (and $f^{(0)}=f$). The space:

$$C^k\big([a,b],\mathbb{R}^n\big)=\big\{f:[a,b]\to\mathbb{R}^n: f\text{ is of class }C^k\big\},$$

endowed with the norm $\|\cdot\|_{C^k}$ is a Banach space.

4.2.5. REMARK. In Examples 4.2.2 and 4.2.4 we have considered in principle only $\mathbb{R}^n$-valued maps. Obviously there is no harm in replacing $\mathbb{R}^n$ by an arbitrary finite-dimensional vector space and we will indeed do that quite often.

Recall that a linear map $T : X \to Y$ between Banach spaces is continuous if and only if (see Exercise 4.3):

$$\|T\| = \sup_{\|x\|\le 1} \big\|T(x)\big\| < +\infty; \tag{4.2.1}$$

more in general, a multi-linear map $B : X_1 \times \cdots \times X_k \to Y$ is continuous if and only if (see Exercise 4.5):

$$\|B\| = \sup_{\|x_1\|\le 1,\ldots,\|x_k\|\le 1} \big\|B(x_1, \ldots, x_k)\big\| < +\infty. \tag{4.2.2}$$

A linear (respectively, multi-linear) map satisfying condition (4.2.1) (respectively, condition (4.2.2)) is usually called a *bounded* linear (respectively, multi-linear) map. Observe then that boundedness[1] actually is equivalent to continuity for linear (or multi-linear) maps.

The notation introduced on page **??** concerning spaces of multi-linear maps is no longer efficient in the context of functional analysis. We make the following:

4.2.6. CONVENTION. When dealing with topological vector spaces (like Banach spaces or Hilbert spaces) the notations introduced on page **??** should be changed so that the spaces $\mathrm{Lin}(V, W)$, $\mathrm{Lin}(V)$, V^*, $\mathrm{Bil}(V, V'; W)$, etc…, contain only *continuous* linear and multi-linear maps. For instance, if X, Y are Banach spaces then $\mathrm{Lin}(X, Y)$ denotes the space of continuous linear maps from X to Y.

Recall that multi-linear maps defined on finite-dimensional vector spaces are automatically continuous, so that the convention above is compatible with the notation introduced on page **??**. We observe also that if $X_1, X_2, \ldots, X_k, Y$ are Banach spaces then the space of continuous multi-linear maps from $X_1 \times \cdots \times X_k$ to Y is again a Banach space endowed with the norm (4.2.2).

4.2.7. EXAMPLE. If $B : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^p$ is an arbitrary bilinear map then the map:

$$\widehat{B} : C^0\big([a, b], \mathbb{R}^m\big) \times L^2\big([a, b], \mathbb{R}^n\big) \longrightarrow L^2\big([a, b], \mathbb{R}^p\big)$$

defined by $\widehat{B}(f, g)(t) = B\big(f(t), g(t)\big)$, $t \in [a, b]$, is bilinear and continuous. Namely:

$$\big\|\widehat{B}(f, g)\big\|_{L^2}^2 = \int_a^b B\big(f(t), g(t)\big)^2 \,\mathrm{d}t \le \|B\|^2 \|f\|_{C^0}^2 \int_a^b \big\|g(t)\big\|^2 \,\mathrm{d}t,$$

and therefore $\big\|\widehat{B}\big\| \le \|B\|$. We will have particular interest in the continuity of the bilinear map:

$$\widehat{B} : C^0\big([a, b], \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n)\big) \times L^2\big([a, b], \mathbb{R}^m\big) \longrightarrow L^2\big([a, b], \mathbb{R}^n\big) \tag{4.2.3}$$

given by:

$$\widehat{B}(T, f)(t) = T(t) \cdot f(t), \quad t \in [a, b], \tag{4.2.4}$$

[1] One should observe that the term boundedness in the context of multi-linear maps does not have its usual meaning; for instance, non zero linear maps never have bounded image.

for all $T \in C^0\big([a,b], \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n)\big)$, $f \in L^2\big([a,b], \mathbb{R}^m\big)$. Observe that the continuity of the bilinear map (4.2.3) implies by Exercise 4.9 the continuity of the linear map:
(4.2.5)
$$C^0\big([a,b], \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n)\big) \ni T \longmapsto \widehat{B}(T, \cdot) \in \mathrm{Lin}\big(L^2([a,b], \mathbb{R}^m), L^2([a,b], \mathbb{R}^n)\big).$$

In Examples 4.2.2, 4.2.3 and 4.2.4 above, the only space of maps admitting the structure of a Hilbert space was L^2. The problem is that, on one hand, Hilbert spaces are much easier to work with than Banach spaces (in terms of abstract functional analysis) while, on the other hand, differential operators (like the derivative operator $\gamma \mapsto \gamma'$) cannot be bounded (globally defined) linear maps on L^2. We thus need a Hilbert space consisting of maps with higher regularity than L^2. Such problem is solved by the introduction of the *Sobolev spaces*. There are several possible approaches for the general theory of Sobolev spaces, but for our purposes, we need only a very particular aspect of such theory; namely, we will define below the space of $\mathbb{R}^n$-valued H^1 maps on a compact interval. There is a very simple definition for such Sobolev space using the notion of absolutely continuous map:

4.2.8. DEFINITION. A curve $\gamma : [a,b] \to \mathbb{R}^n$ is called *absolutely continuous* if for every $\varepsilon > 0$ there exists $\delta > 0$ such that given disjoint open subintervals (x_1, y_1),..., (x_k, y_k) of $[a,b]$ with $\sum_{i=1}^k y_i - x_i < \delta$ then:
$$\sum_{i=1}^k \big\|\gamma(x_i) - \gamma(y_i)\big\| < \varepsilon.$$

Obviously every absolutely continuous curve is continuous and every Lipschitz continuous curve (and in particular every piecewise C^1 curve) is absolutely continuous. The theorem below gives an equivalent definition of the notion of absolutely continuous curve.

4.2.9. THEOREM. *A curve $\gamma : [a,b] \to \mathbb{R}^n$ is absolutely continuous if and only if the following three conditions are satisfied:*

- *the derivative $\gamma'(t) = \lim_{h \to 0} \frac{\gamma(t+h) - \gamma(t)}{h}$ exists for almost every t in $[a,b]$;*
- *the (almost everywhere defined) map $\gamma' : [a,b] \to \mathbb{R}^n$ is integrable;*
- *$\gamma(t) = \gamma(a) + \int_a^t \gamma'$ for all $t \in [a,b]$.*

Moreover, if $\phi : [a,b] \to \mathbb{R}^n$ is an integrable map then the curve $\gamma : [a,b] \to \mathbb{R}^n$ defined by $\gamma(t) = \int_a^t \phi$ is absolutely continuous and $\gamma' = \phi$ almost everywhere.

PROOF. See for instance [**138**]. □

We can now proceed with the definition of the Sobolev space H^1.

4.2.10. DEFINITION. We say that a curve $\gamma : [a,b] \to \mathbb{R}^n$ is *of Sobolev class* H^1 (shortly, *of class* H^1) if γ is absolutely continuous and the (almost everywhere defined) map $\gamma' : [a,b] \to \mathbb{R}^n$ is in $L^2\big([a,b], \mathbb{R}^n\big)$. We denote by $H^1\big([a,b], \mathbb{R}^n\big)$

the set of all maps $\gamma:[a,b]\to\mathbb{R}^n$ of class H^1 and we define the H^1-*inner product* of $\gamma_1,\gamma_2\in H^1\big([a,b],\mathbb{R}^n\big)$ by:

$$\langle\gamma_1,\gamma_2\rangle_{H^1}=\big\langle\gamma_1(a),\gamma_2(a)\big\rangle+\langle\gamma_1,\gamma_2\rangle_{L^2}. \tag{4.2.6}$$

The norm corresponding to $\langle\cdot,\cdot\rangle_{H^1}$ will be denoted by $\|\cdot\|_{H^1}$ and will be called the H^1-*norm*.

It is easy to see that $H^1\big([a,b],\mathbb{R}^n\big)$ is a vector subspace of $C^0\big([a,b],\mathbb{R}^n\big)$ and that $\langle\cdot,\cdot\rangle_{H^1}$ makes $H^1\big([a,b],\mathbb{R}^n\big)$ into a Hilbert space such that the inclusion of $H^1\big([a,b],\mathbb{R}^n\big)$ in $C^0\big([a,b],\mathbb{R}^n\big)$ is continuous. For more details see Exercise 4.17.

There are several continuous inclusions between the Banach spaces discussed so far. They are listed in Exercise 4.18.

Observe that a continuous bilinear form $B:\mathcal{H}\times\mathcal{H}\to\mathbb{R}$ on a Hilbert space $\mathcal{H}$ is nondegenerate if and only if the linear map:

$$\mathcal{H}\ni x\longmapsto B(x,\cdot)\in\mathcal{H}^* \tag{4.2.7}$$

is injective; equivalently, B is nondegenerate if the linear map that represents B with respect to the Hilbert space inner product of $\mathcal{H}$ is injective. We give the following definition:

4.2.11. DEFINITION. A continuous bilinear form $B:\mathcal{H}\times\mathcal{H}\to\mathbb{R}$ is called *strongly nondegenerate* is the linear map (4.2.7) is an isomorphism; equivalently, B is strongly nondegenerate if the linear map that represents B with respect to the Hilbert space inner product of $\mathcal{H}$ is an isomorphism.

The following gives a characterization of the Hilbert space inner products of a Hilbertable space. Recall that, given a Hilbert space $\big(\mathcal{H},\langle\cdot,\cdot\rangle\big)$, a bounded linear operator $P:\mathcal{H}\to\mathcal{H}$ is called *positive* if P is self-adjoint (i.e., $\langle Px,y\rangle=\langle x,Py\rangle$ for all $x,y\in\mathcal{H}$) and $\langle Px,x\rangle\ge0$ for all $x\in\mathcal{H}$.

4.2.12. PROPOSITION. *Let $(\mathcal{H},\langle\cdot,\cdot\rangle)$ be a Hilbert space and let $B:\mathcal{H}\times\mathcal{H}\to\mathbb{R}$ be a bounded bilinear form. Then B is a (positive definite) Hilbert space inner product for $\mathcal{H}$ if and only if B is represented by a positive isomorphism $P:\mathcal{H}\to\mathcal{H}$ (see Exercise 4.19).*

PROOF. It is easy to see that, given $P\in\mathrm{Lin}(\mathcal{H})$ then $\langle\cdot,\cdot\rangle_1=\langle P\cdot,\cdot\rangle$ is an inner product in $\mathcal{H}$ if and only if P is positive anad injective. We have to show that $\langle\cdot,\cdot\rangle_1$ is a Hilbert space inner product for $\big(\mathcal{H},\langle\cdot,\cdot\rangle\big)$ (i.e., that $\langle\cdot,\cdot\rangle_1$ defines the same topology as $\langle\cdot,\cdot\rangle$) if and only if P is an isomorphism. Observe that we do not know whether $\big(\mathcal{H},\langle\cdot,\cdot\rangle_1\big)$ is a Hilbert space, but it is at least a normed (and a topological) vector space. Since $\langle\cdot,\cdot\rangle_1$ is a bounded bilinear form on $\big(\mathcal{H},\langle\cdot,\cdot\rangle\big)$, it is easy to see that the identity operator:

$$\mathrm{Id}:\big(\mathcal{H},\langle\cdot,\cdot\rangle\big)\longrightarrow\big(\mathcal{H},\langle\cdot,\cdot\rangle_1\big) \tag{4.2.8}$$

is bounded. Obviously, $\langle\cdot,\cdot\rangle_1$ defines the same topology as $\langle\cdot,\cdot\rangle$ if and only if (4.2.8) is a homeomorphism; it thus follows from the Open Mapping Theorem that $\langle\cdot,\cdot\rangle_1$ defines the same topology as $\langle\cdot,\cdot\rangle$ if and only if $\big(\mathcal{H},\langle\cdot,\cdot\rangle_1\big)$ is a Hilbert space. To complete the proof, we will show then that P is bijective if and only if $\big(\mathcal{H}, \langle\cdot,\cdot\rangle_1\big)$ is a Hilbert space. To this aim, consider the commutative diagram:

$$\begin{tikzcd}
 & \big(\mathcal{H}, \langle\cdot,\cdot\rangle_1\big)^* \arrow[dr, "\mathrm{Id}^*"] & \\
\big(\mathcal{H}, \langle\cdot,\cdot\rangle_1\big) \arrow[ur, "\mathfrak{R}_1"] \arrow[dr, "P"'] & & \big(\mathcal{H}, \langle\cdot,\cdot\rangle\big)^* \\
 & \big(\mathcal{H}, \langle\cdot,\cdot\rangle\big) \arrow[ur, "\cong", "\mathfrak{R}"'] &
\end{tikzcd} \tag{4.2.9}$$

where $\mathfrak{R}_1$ is given by $x \mapsto \langle x, \cdot\rangle_1$, $\mathfrak{R}$ is given by $x \mapsto \langle x, \cdot\rangle$, and Id^* denotes the transpose operator of Id. It follows easily from the Cauchy–Schwarz inequality that $\mathfrak{R}$ and $\mathfrak{R}_1$ are isometric immersions; moreover, since $\big(\mathcal{H}, \langle\cdot,\cdot\rangle\big)$ is a Hilbert space, it follows from Riesz's representation theorem that $\mathfrak{R}$ is indeed an isometry. Moreover, Id^* is simply an inclusion map and hence it is injective. Assuming that P is bijective then both arrows in the bottom triangle of diagram (4.2.9) are bijective and therefore $\mathrm{Id}^* \circ \mathfrak{R}_1$ is bijective. Since Id^* is injective, it follows that $\mathfrak{R}_1$ is bijective and it is therefore an isometry; we conclude that $\big(\mathcal{H}, \langle\cdot,\cdot\rangle_1\big)$ is a Hilbert space, since the dual of a normed space is always complete.

Conversely, assume that $\big(\mathcal{H}, \langle\cdot,\cdot\rangle_1\big)$ is a Hilbert space. Then (4.2.8) is a homeomorphism and therefore Id^* is bijective; moreover, by Riesz's representation theorem, the map $\mathfrak{R}_1$ is an isometry. But also $\mathfrak{R}$ is an isometry and hence P is bijective. This concludes the proof. □

In order to developed infinite-dimensional Morse theory we will need a generalization of Sylvester's theorem of Inertia for Hilbert spaces. This task will take a little work. We start by recalling the following tool:

4.2.13. PROPOSITION (continuous functional calculus). *Let $\mathcal{H}$ be a Hilbert space and let $T : \mathcal{H} \to \mathcal{H}$ be a bounded self-adjoint operator. Then there exists a unique continuous homomorphism of algebras with unity*[2]*:*

$$\phi_T : C^0(\sigma(T), \mathbb{R}) \longrightarrow \mathrm{Lin}(\mathcal{H})$$

such that $\phi_T(i) = T$, where $i : \sigma(T) \to \mathbb{R}$ denotes the inclusion. Moreover, $\phi_T(f)$ is a self-adjoint operator for every continuous map $f : \sigma(T) \to \mathbb{R}$ and the homomorphism ϕ_T is an isometry, i.e., the operator norm of $\phi_T(f)$ equals the sup norm of $f \in C^0(\sigma(T), \mathbb{R})$.

PROOF. See [**134**, Chapter VII, Section 1] for the case where $\mathcal{H}$ is a complex Hilbert space. The case of a real Hilbert space can be obtained by a complexification argument[3]. □

[2]This means that ϕ_T is linear, $\phi_T(fg) = \phi_T(f) \circ \phi_T(g)$, for all $f, g \in C^0(\sigma(T), \mathbb{R})$ and that $\phi_T(1) = \mathrm{Id}$.

[3]The *complexification* of a real Hilbert space is the complex Hilbert space $\mathcal{H}^{\mathbb{C}} = \mathcal{H} \oplus \mathcal{H}$, with complex structure $i(x, y) = (-y, x)$ and Hermitean product obtained by extending the inner product

4.2.14. REMARK. We list a few more properties of the operators $\phi_T(f)$ that follow easily from Proposition 4.2.13.

- *If $p(x) = \sum_{k=0}^n a_k x^k$ is a polynomial then $\phi_T(p) = \sum_{k=0}^n a_k T_k$, where* $T^0 = \mathrm{Id}$. Follows directly from the fact that ϕ_T is a homomorphism and from the fact that ϕ_T maps the inclusion $i : \sigma(T) \to \mathbb{R}$ to T.
- *For any $f, g \in C^0(\sigma(T), \mathbb{R})$ the operators $\phi_T(f)$ and $\phi_T(g)$ commute; in particular, each operator $\phi_T(f)$ commutes with T.* This follows directly from the observation that the algebra $C^0(\sigma(T), \mathbb{R})$ is commutative.
- *If $f : \sigma(T) \to \mathbb{R}$ is a non negative function then $f(T)$ is a positive operator.* Choose $g \in C^0(\sigma(T), \mathbb{R})$ with $g^2 = f$ and observe that $\big\langle \phi_T(f)x, x \big\rangle = \big\langle \phi_T(g)x, \phi_T(g)x \big\rangle$ for all $x \in \mathcal{H}$.
- *If $f \in C^0(\sigma(T), \mathbb{R})$ satisfies $f^2 = f$ then $\phi_T(f)$ is an orthogonal projection onto a closed subspace of $\mathcal{H}$.* Observe simply that $\phi_T(f)^2 = \phi_T(f)$ and that $\phi_T(f)$ is self-adjoint (see Exercise 4.28).
- *If a closed subspace V of $\mathcal{H}$ is invariant by T then V is also invariant by $\phi_T(f)$, for all $f \in C^0(\sigma(T), \mathbb{R})$.* This is immediate if f is a polynomial; otherwise, it follows from the continuity of ϕ_T, since any continuous map in a compact subset of $\mathbb{R}$ is a uniform limit of polynomials.

We can now finally prove the following:

4.2.15. LEMMA. *Let $B : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ be a strongly nondegenerate bounded symmetric bilinear form on a Hilbert space $(\mathcal{H}, \langle \cdot, \cdot \rangle)$. Then there exists a direct sum decomposition $\mathcal{H} = \mathcal{H}_+ \oplus \mathcal{H}_-$, where $\mathcal{H}_+$ and $\mathcal{H}_-$ are closed subspaces of $\mathcal{H}$ that are orthogonal with respect to both $\langle \cdot, \cdot \rangle$ and B and such that $B|_{\mathcal{H}_+}$, $-B|_{\mathcal{H}_-}$ are (positive definite) Hilbert space inner products.*

PROOF. Let $T : \mathcal{H} \to \mathcal{H}$ be the bounded self-adjoint operator that represents B. Since T is an isomorphism we have that $0 \notin \sigma(T)$ and therefore we can write $\sigma(T) = \sigma_+ \cup \sigma_-$ where $\sigma_+ = \sigma(T) \cap \left]0, +\infty\right[$ and $\sigma_- = \sigma(T) \cap \left]-\infty, 0\right[$. Denote by $\chi_{\sigma_+}, \chi_{\sigma_-} \in C^0(\sigma(T), \mathbb{R})$ the characteristic maps of σ_+ and σ_- respectively, i.e., χ_{σ_+} (respectively, χ_{σ_-}) equals 1 on σ_+ (respectively, on σ_-) and equals zero otherwise. Observe that χ_{σ_+} and χ_{σ_-} are indeed continuous on $\sigma(T)$, since σ_+ and σ_- are open in $\sigma(T)$. Using the continuous functional calculus (Proposition 4.2.13), we obtain bounded self-adjoint operators $P_+ = \phi_T(\chi_{\sigma_+})$ and $P_- = \phi_T(\chi_{\sigma_-})$ on $\mathcal{H}$. Using the equalities $(\chi_{\sigma_+})^2 = \chi_{\sigma_+}$, $(\chi_{\sigma_-})^2 = \chi_{\sigma_-}$, $\chi_{\sigma_+} + \chi_{\sigma_-} = 1$ and Remark 4.2.14 we obtain that P_+ and P_- are orthogonal projections onto closed subspaces $\mathcal{H}_+$ and $\mathcal{H}_-$ of $\mathcal{H}$ respectively, and that $\mathcal{H} = \mathcal{H}_+ \oplus \mathcal{H}_-$ is a direct sum decomposition that is orthogonal with respect to $\langle \cdot, \cdot \rangle$. Since P_+ and P_- commute with T (see Remark 4.2.14), it follows that both $\mathcal{H}_+$ and $\mathcal{H}_-$ are invariant by T, so that $\mathcal{H}_+$ and $\mathcal{H}_-$ are also B-orthogonal. If $i : \sigma(T) \to \mathbb{R}$ denotes the inclusion then $T \circ P_+ = \phi_T(i\chi_{\sigma_+})$ and, since $i\chi_{\sigma_+}$ is a non negative function, Remark 4.2.14 implies that for every $x \in \mathcal{H}_+$:

$$B(x, x) = \langle Tx, x \rangle = \big\langle (T \circ P_+)x, x \big\rangle \geq 0.$$

---

of $\mathcal{H}$ to a sesqui-linear map. Every bounded self-adjoint operator $T : \mathcal{H} \to \mathcal{H}$ extends uniquely to a (complex linear) bounded self-adjoint operator $T^{\mathbb{C}} : \mathcal{H}^{\mathbb{C}} \to \mathcal{H}^{\mathbb{C}}$

Similarly, by considering the non negative function $-i\chi_{\sigma_-}$ one shows that B is negative semi-definite on $\mathcal{H}_-$. Finally, the fact that $T:\mathcal{H}\to\mathcal{H}$ is a self-adjoint isomorphism implies that its restriction to the invariant subspaces $\mathcal{H}_+$ and $\mathcal{H}_-$ is again an isomorphism (see Exercise 4.30), so that $B|_{\mathcal{H}_+}$ and $-B|_{\mathcal{H}_-}$ are represented by positive isomorphisms of $\mathcal{H}_+$ and $\mathcal{H}_-$ respectively. The conclusion follows from Proposition 4.2.12. □

4.3. Calculus on Banach Spaces and Banach Manifolds

We now make a quick review on the subject of Calculus on Banach spaces. We start with the following:

4.3.1. DEFINITION. Let X, Y be Banach spaces, $U\subset X$ an open subset and $f:U\to Y$ a map. We say that f is *differentiable* at a point $x\in U$ if there exists a continuous linear map $T:X\to Y$ such that the map r defined by the equality:

$$f(x+h)=f(x)+T(h)+r(h),$$

satisfies $\lim_{h\to 0}\frac{r(h)}{\|h\|}=0$.

If f is differentiable at x then it is easy to check that:

$$T(v)=\lim_{t\to 0}\frac{f(x+tv)-f(x)}{t},$$

for all $v\in X$. This implies that T is unique when it exists; we call T the *differential* of f at x and we write $\mathrm{d}f(x)=T$.

4.3.2. REMARK. It is easy to see that the statement "f is differentiable at x and $\mathrm{d}f(x)=T$" is invariant under substitution of the norms in X and Y by equivalent ones. In particular, differentiability is a well-defined notion for Banachable spaces.

If f is differentiable at every point of U, we say that f is *differentiable in* U; in such case, we can consider the map:

$$\mathrm{d}f:U\longrightarrow \mathrm{Lin}(X,Y),$$

defined by $x\mapsto \mathrm{d}f(x)$. Since $\mathrm{Lin}(X,Y)$ is again a Banach space, we can again ask whether $\mathrm{d}f$ is a differentiable map. If it is, we obtain a map:

$$\mathrm{d}^2f=\mathrm{d}(\mathrm{d}f):U\longrightarrow \mathrm{Lin}\big(X,\mathrm{Lin}(X,Y)\big)$$

called the *second order differential* of f. In general, if f can be differentiated k times, we can consider its k-th order differential (defined recursively by $\mathrm{d}^kf=\mathrm{d}\big(\mathrm{d}^{k-1}f\big)$) which is a map of the form:

$$\mathrm{d}^kf:U\longrightarrow \underbrace{\mathrm{Lin}\big(X,\mathrm{Lin}\big(X,\cdots,\mathrm{Lin}}_{k\ \text{Lin's}}(X,Y)\big)\cdots\big).$$

The counter-domain of d^kf may be identified with a nicer space, namely we have an isometry (see Exercise 4.9):

$$\underbrace{\mathrm{Lin}\big(X,\mathrm{Lin}\big(X,\cdots,\mathrm{Lin}}_{k\ \text{Lin's}}(X,Y)\big)\cdots\big)\ni T\longmapsto \widehat{T}\in \mathrm{Multlin}_k(X;Y)$$

defined by:

$$\widehat{T}(v_1, v_2, \dots, v_k) = T(v_1)(v_2)\cdots(v_k),$$

for all $v_1, \dots, v_k \in X$, where $\mathrm{Multlin}_k(X;Y)$ denotes the Banach space of all continuous k-linear maps $B : X \times \cdots \times X \to Y$.

If a map $f : U \subset X \to Y$ is k times differentiable and if its k-th order differential $\mathrm{d}^k f : U \to \mathrm{Multlin}_k(X;Y)$ is continuous then we say that f is a *map of class* C^k. If f is of class C^k for all $k \in \mathbb{N}$, we say that f is a *map of class* C^∞.

From now on, one can develop the theory of differentiable Calculus on Banach spaces just like one does in finite-dimensional spaces. One now can prove the chain rule, the mean value inequality, Schwarz's theorem (on the symmetry of the higher order differentials), the inverse and implicit function theorems and so on. The whole theory goes on like in the finite-dimensional case, with essentially no differences (and in most cases no additional difficulty). The main relevant difference lies on the local form of immersions and submersions. We look at the problem more closely below.

Recall that a closed subspace S of a Banach space X is called *complemented* if there exists a closed subspace $S' \subset X$ with $X = S \oplus S'$.

4.3.3. DEFINITION. Let X, Y be Banach spaces, $U \subset X$ an open subset and $f : U \to Y$ a map. Assume that f is differentiable at some $x \in U$. We say that f is a *submersion* at x if the differential $\mathrm{d}f(x) : X \to Y$ is surjective and if its (automatically closed) kernel $\mathrm{Ker}\big(\mathrm{d}f(x)\big)$ is complemented in X. We say that f is an *immersion* at x if the differential $\mathrm{d}f(x) : X \to Y$ is injective and if its image is closed and complemented in Y.

Our point here is that the standard proofs of the local form of immersions and submersions only work in the Banach space case if one uses the notions of immersion and submersion described above. In finite-dimensional spaces, all subspaces are closed and complemented, so that Definition 4.3.3 reduces to the standard one. We remark also that on Hilbert spaces all closed subspaces are complemented (there is always the orthogonal complement!). Hence, if X is a Hilbert space then $f : U \subset X \to Y$ is a submersion at $x \in U$ iff $\mathrm{d}f(x)$ is surjective; similarly, if Y is a Hilbert space then f is an immersion at x iff $\mathrm{d}f(x)$ is injective and has closed image.

What we need now is a practical method for proving differentiability of maps between Banach spaces in concrete examples. This is the subject of Lemma 4.3.5 below; first we need a definition.

4.3.4. DEFINITION. Let Y be a Banach space. A *separating family* for Y is a set $\mathcal{F}$ of bounded linear operators $\lambda : Y \to Z_\lambda$, with Z_λ a Banach space, such that for each non zero $v \in Y$ there exists $\lambda \in \mathcal{F}$ with $\lambda(y) \neq 0$.

4.3.5. LEMMA (weak differentiation principle). *Let X, Y be Banach spaces, $f : U \to Y$ a map defined on an open subset $U \subset X$ and $\mathcal{F}$ a separating family for Y. Assume that there exists a continuous map $g : U \to \mathrm{Lin}(X,Y)$ such that for every $x \in U$, $v \in X$, $\lambda \in \mathcal{F}$, the directional derivative $\frac{\partial(\lambda\circ f)}{\partial v}(x)$ exists and equals $\lambda\big(g(x)\cdot v\big)$. Then f is of class C^1 and $\mathrm{d}f = g$.*

PROOF. Let $x \in U$ be fixed and define r by:

$$f(x+h) = f(x) + g(x)\cdot h + r(h);$$

all we have to show is that $\lim_{h\to 0}\frac{r(h)}{\|h\|} = 0$. If h is small enough, the closed line segment $[x, x+h]$ is contained in U; moreover, under the hypothesis of the lemma, it is easy to see that for every $\lambda \in \mathcal{F}$ the curve:

$$[0,1] \ni t \longmapsto (\lambda \circ f)(x+th)$$

is differentiable and that is derivative is given by:

$$\frac{\mathrm{d}}{\mathrm{d}t}(\lambda\circ f)(x+th) = \lambda\big(g(x+th)\cdot h\big).$$

We can thus apply the Fundamental Theorem of Calculus[4] to obtain:

$$\lambda\big(r(h)\big) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}(\lambda\circ f)(x+th)\,\mathrm{d}t - \lambda\big(g(x)\cdot h\big)$$
$$= \lambda\left(\int_0^1 g(x+th)\cdot h\,\mathrm{d}t - g(x)\cdot h\right).$$

Since $\mathcal{F}$ separates points in Y we can "cancel" λ on both sides of the equality above obtaining:

$$r(h) = \int_0^1 g(x+th)\cdot h\,\mathrm{d}t - g(x)\cdot h = \left(\int_0^1 \big[g(x+th) - g(x)\big]\,\mathrm{d}t\right)\cdot h;$$

hence:

$$\big\|r(h)\big\| \le \|h\| \sup_{t\in[0,1]} \big\|g(x+th) - g(x)\big\|.$$

The conclusion follows from the continuity of g. □

We now make a quick study on the subject of length of curves in Banach spaces.

A curve $\gamma : I \to X$ defined on an arbitrary interval $I \subset \mathbb{R}$, taking values on a Banach space X is said to be *piecewise* C^1 if there exists a finite subset $\{t_0, t_1, \ldots, t_k\} \subset I$, $t_0 < t_1 < \cdots < t_k$, such that $\gamma|_{[t_i, t_{i+1}]}$ is of class C^1 for $i = 1, \ldots, k-1$ and $\gamma|_{]-\infty, t_0]\cap I}$ and $\gamma|_{[t_k, +\infty[\cap I}$ are of class C^1.

4.3.6. DEFINITION. Let $\big(X, \|\cdot\|\big)$ be a Banach space and let $\gamma : I \to X$ be a piecewise C^1 curve defined in an arbitrary interval $I \subset \mathbb{R}$. The *length* of γ is defined by:

$$L(\gamma) = \int_I \big\|\gamma'(t)\big\|\,\mathrm{d}t \in [0, +\infty].$$

In Exercise 4.12 the reader is asked to show that a line segment is a shortest path connecting two points in a Banach space.

[4]Here we need a theory of integration for Banach space valued curves. One possibility is to use the *Bochner integral* (see [**162**]), but actually one can use simpler approaches in this case. For instance, one can use the notion of *weak integration* (see Exercise 4.20).

4.3.7. LEMMA. *If $\gamma : I \to X$ is a piecewise C^1 curve defined on an arbitrary interval $I \subset \mathbb{R}$ taking values in a Banach space X and if $L(\gamma) < +\infty$, then the image of γ is relatively compact in X.*

PROOF. Given $\varepsilon > 0$, since $\int_I \big\|\gamma'(t)\big\| \,\mathrm{d}t$ is finite, we can find a compact interval $J \subset I$ such that $I \setminus J$ is a disjoint union of two intervals I_1, I_2 and:

$$\int_{I_1} \big\|\gamma'(t)\big\| \,\mathrm{d}t + \int_{I_2} \big\|\gamma'(t)\big\| \,\mathrm{d}t < \varepsilon.$$

For $t, s \in I_1$, $t < s$, using the result of Exercise 4.12 we get: $\big\|\gamma(t) - \gamma(s)\big\| \le L(\gamma|_{[t,s]}) < \varepsilon$, so that the diameter of $\gamma(I_1)$ is less than or equal to ε; similarly, the diameter of $\gamma(I_2)$ is less than or equal to ε. Finally, since $\gamma(J) \subset X$ is compact, it can be covered by a finite number of subsets of X of diameter less than ε. Thus $\gamma(I)$ is totally bounded and hence relatively compact in the complete metric space X. $\square$

We now deal with Banach manifolds. For the basic stuff, there is no big difference between the theory of Banach manifolds and the theory of finite-dimensional manifolds. We just give a few basic definitions for completeness.

Let $\mathcal{M}$ be a set. A *chart* on $\mathcal{M}$ is a bijection $\varphi : U \to \widetilde{U}$, where $\widetilde{U}$ is an open subset of some Banach space X. Given charts $\varphi : U \to \widetilde{U}$, $\psi : V \to \widetilde{V}$ on $\mathcal{M}$, with $\widetilde{U}$ open in the Banach space X and $\widetilde{V}$ open in the Banach space Y, then we say that φ and ψ are *compatible* if either $U \cap V = \emptyset$ or the map:

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

is a smooth diffeomorphism between open sets. An *atlas* $\mathcal{A}$ for $\mathcal{M}$ is a set of pairwise compatible charts on $\mathcal{M}$ whose domains cover $\mathcal{M}$. A *Banach manifold* is a set $\mathcal{M}$ endowed with a maximal atlas $\mathcal{A}$. An atlas $\mathcal{A}$ on $\mathcal{M}$ induces a unique topology on $\mathcal{M}$ for which the domains of the charts in $\mathcal{A}$ are open and the charts in $\mathcal{A}$ are homeomorphisms. Such topology is defined by:

$$Z \subset \mathcal{M} \text{ is open} \iff \begin{array}{c} \varphi(Z \cap U) \text{ is open in } X, \\ \text{for every chart } \varphi : U \to \widetilde{U} \subset X \text{ in} \\ \text{the atlas } \mathcal{A}. \end{array}$$

If a Banach manifold $\mathcal{M}$ admits an atlas consisting only of charts taking values on Hilbert spaces then we call $\mathcal{M}$ a *Hilbert manifold*.

4.3.8. CONVENTION. For the rest of this section and until the end of Section **??** we will not make any assumptions on the topology of the Banach manifolds $\mathcal{M}$ (not even Hausdorff!). In Section **??**, we will usually deal with a finite-dimensional manifold M, for which the conventions of Section **??** apply, i.e., M should be Hausdorff and second countable; at the same time, we will have infinite-dimensional manifolds $\mathcal{M}$ whose points are *curves* on M and we do not want to waist time in proving topological properties of such $\mathcal{M}$. Actually, we will se in Corollary 4.3.22 that a Hilbert manifold admitting a Riemannian metric is automatically T_4.

As in the case of Calculus on finite-dimensional manifolds, one can now define the notion of map of class C^k between Banach manifolds (using local charts) and one can extend all the local theorems of the Calculus on Banach spaces to the context of Banach manifolds. We now make a few remarks concerning the tangent space of a Banach manifold.

Let $\mathcal{M}$ be a Banach manifold and let $x \in \mathcal{M}$ be fixed. As in the finite-dimensional case, the tangent space $T_x\mathcal{M}$ can be defined using equivalence classes of curves in $\mathcal{M}$ passing through x. More explicitly, consider the set A of all smooth curves $\gamma : \left]-\varepsilon, \varepsilon\right[\to \mathcal{M}$ with $\gamma(0) = x$; we define an equivalence relation on A by requiring that $\gamma, \mu \in A$ are equivalent if for some (and hence every) chart φ around x in $\mathcal{M}$ we have $(\varphi \circ \gamma)'(0) = (\varphi \circ \mu)'(0)$. The tangent space $T_x\mathcal{M}$ is defined to be the quotient of A by such equivalence relation. Observe that every chart $\varphi : U \to \widetilde{U} \subset X$ with $x \in U$ induces a bijection $\widehat{\varphi} : T_x\mathcal{M} \to X$ that sends the class of γ to $(\varphi \circ \gamma)'(0)$. If φ and ψ are both charts around x then the bijections $\widehat{\varphi}$ and $\widehat{\psi}$ differ by the differential of the transition map $\psi \circ \varphi^{-1}$ at $\varphi(x)$; such differential is a continuous isomorphism between Banach spaces and therefore all charts induce on $T_x\mathcal{M}$ the same vector space structure and the same topology.

Our point here is that the tangent space $T_x\mathcal{M}$ is a *Banachable space*, not a Banach space, i.e., there is no canonically fixed norm on $T_x\mathcal{M}$. Only the topology of $T_x\mathcal{M}$ is canonical. Observe that if $\mathcal{M}$ is a Hilbert manifold then its tangent spaces are Hilbertable spaces.

One can now, as in the finite-dimensional case, define the differential of a differentiable map between Banach manifolds as being a continuous linear map between the appropriate tangent spaces. Definition 4.3.3 can now be generalized in the obvious way to the context of manifolds.

We now define the notion of a submanifold of a Banach manifold.

4.3.9. DEFINITION. Let $\mathcal{M}$ be a Banach manifold and let $\mathcal{N} \subset \mathcal{M}$ be a subset. A chart $\varphi : U \to \widetilde{U} \subset X$ for $\mathcal{M}$ is called a *submanifold chart* for $\mathcal{N}$ if there exists a closed and complemented subspace $Y \subset X$ such that $\varphi(U \cap \mathcal{N}) = \widetilde{U} \cap Y$. If $\mathcal{N}$ can be covered by the domains of a family of submanifold charts for $\mathcal{N}$ then we say that $\mathcal{N}$ is a *Banach submanifold* of $\mathcal{M}$.

If $\mathcal{N}$ is a Banach submanifold of $\mathcal{M}$ then the submanifold charts can be restricted to form an atlas of $\mathcal{N}$, so that $\mathcal{N}$ also becomes a Banach manifold. The inclusion $i : \mathcal{N} \to \mathcal{M}$ is a smooth *embedding*, i.e., it is an immersion and a homeomorphism onto its image. The differential of the inclusion i can be used to identify, for every $x \in \mathcal{N}$, the tangent space $T_x\mathcal{N}$ with a closed and complemented subspace of the tangent space $T_x\mathcal{M}$.

The following result should come to no surprise:

4.3.10. PROPOSITION. *Let $\mathcal{M}$, $\mathcal{N}$ be Banach manifolds and let $f : \mathcal{M} \to \mathcal{N}$ be a smooth map. If $c \in \mathcal{N}$ is a* regular value *of f, i.e., if f is a submersion at all points of $f^{-1}(c)$ then $f^{-1}(c)$ is a Banach submanifold of $\mathcal{M}$. Moreover, its tangent space is given by:*

$$T_x f^{-1}(c) = \operatorname{Ker}\big(\mathrm{d}f(x)\big),$$

for all $x \in f^{-1}(c)$.

PROOF. It is a simple consequence of the local form of submersions, as in the finite-dimensional case. □

Infinite dimensional Banach manifolds cannot be locally compact. This sometimes brings problems. Some of this problems are solved by Lemma 4.3.12 below. First, we need a definition.

4.3.11. DEFINITION. If $\mathcal{M}$ is a Banach manifold then a chart $\varphi : U \to \widetilde{U} \subset X$ for $\mathcal{M}$ is called *regular* if whenever $F \subset X$ is closed in X and contained in $\widetilde{U}$ then $\varphi^{-1}(F)$ is closed in $\mathcal{M}$.

4.3.12. LEMMA. *If $\mathcal{M}$ is a Banach manifold and $\varphi : U \to \widetilde{U} \subset X$ is a chart for $\mathcal{M}$ then for every open set V in $\mathcal{M}$ with $\overline{V} \subset U$, the chart $\varphi|_V : V \to \varphi(V)$ is regular. In particular, if $\mathcal{M}$ is T_3 then for every chart $\varphi : U \to \widetilde{U} \subset X$ and every $x \in U$ there exists a restriction of φ to an open neighborhood of x that is a regular chart.*

PROOF. We leave it as an exercise to the reader (see Exercise 4.26). □

We now study infinite-dimensional Riemannian manifolds.

4.3.13. DEFINITION. Let $\mathcal{M}$ be a Hilbert manifold. A *Riemannian metric* for $\mathcal{M}$ is a map g that associates to every $x \in \mathcal{M}$ a Hilbert space inner product g_x on the Hilbertable space $T_x\mathcal{M}$ in such a way that for every chart $\varphi : U \to \widetilde{U} \subset \mathcal{H}$ taking values in a Hilbert space $\mathcal{H}$, the map:

$$\widehat{g} : \widetilde{U} \longrightarrow \mathrm{Bil}(\mathcal{H}),$$

defined by:

$$\widehat{g}(x) = g_x\big(\mathrm{d}\varphi(x)^{-1}\cdot, \mathrm{d}\varphi(x)^{-1}\cdot\big),$$

is smooth. A Hilbert manifold $\mathcal{M}$ endowed with a Riemannian metric g will be called a *Riemannian manifold*.

The smoothness of the transition maps between local charts implies easily that in order to check that g is a Riemannian metric one has only to show the smoothness of $\widehat{g}$ for charts φ running through a fixed atlas of $\mathcal{M}$.

We won't need to study much Riemannian geometry in Hilbert manifolds. We just present below a few selected topics that will be used in the later sections.

We start with the definition of arc-length and distance.

4.3.14. DEFINITION. Let $(\mathcal{M}, g)$ be a Riemannian manifold. If $\gamma : I \to \mathcal{M}$ is a piecewise C^1 curve defined on an arbitrary interval $I \subset \mathbb{R}$ then the *length* of γ is the (possibly infinite) non negative real number:

$$L(\gamma) = \int_I \big\|\gamma'(t)\big\| \,\mathrm{d}t \in [0, +\infty].$$

For $x, y \in \mathcal{M}$ we define the *distance* from x to y as the infimum of the lengths of curves in M connecting x and y, i.e., we set:

(4.3.1)
$$\mathrm{dist}(x,y) = \inf\big\{L(\gamma) : \gamma : [a,b] \to \mathcal{M} \text{ piecewise } C^1,\ \gamma(a) = x,\ \gamma(b) = y\big\}.$$

If the set on the righthand side of the equality above is empty (i.e., if x and y are not in the same connected component of $\mathcal{M}$) then we set $\mathrm{dist}(x,y) = +\infty$.

The following properties of the distance function defined above are obvious:

- $\mathrm{dist}(x,x) = 0$ for all $x \in \mathcal{M}$;
- $\mathrm{dist}(x,y) = \mathrm{dist}(y,x)$ for all $x, y \in \mathcal{M}$;
- $\mathrm{dist}(x,z) \le \mathrm{dist}(x,y) + \mathrm{dist}(y,z)$, for all $x, y, z \in \mathcal{M}$.

The triangle inequality above follows from the obvious fact that length of curves is additive by concatenation and from the fact that the concatenation of piecewise C^1 curves is again piecewise C^1.

4.3.15. DEFINITION. Let $(\mathcal{M}, g)$ be a Riemannian manifold. We say that a chart $\varphi : U \to \widetilde{U}$ taking values on an open set $\widetilde{U}$ of a Hilbert space $\big(\mathcal{H}, \langle\cdot,\cdot\rangle\big)$ is *metric-relating* if there exists positive constants $k_{\min}, k_{\max} \in \mathbb{R}$ such that:

$$k_{\min}\big\langle \mathrm{d}\varphi_x(v), \mathrm{d}\varphi_x(v)\big\rangle^{\frac12} \le g_x(v,v)^{\frac12} \le k_{\max}\big\langle \mathrm{d}\varphi_x(v), \mathrm{d}\varphi_x(v)\big\rangle^{\frac12}, \tag{4.3.2}$$

for all $x \in U$, $v \in T_xM$.

Since we assume that g_x is a Hilbert space inner product for T_xM, the constants $k_{\min}$, $k_{\max}$ satisfying (4.3.2) can be chosen for each $x \in U$; saying that φ is metric-relating means that $k_{\min}$ and $k_{\max}$ can be chosen *independently* of $x \in U$. The continuity of the Riemannian metric of $\mathcal{M}$ implies that “small” charts are indeed metric-relating (see Exercise 4.13).

4.3.16. LEMMA. *Let $(\mathcal{M}, g)$ be a Riemannian manifold and assume that $\varphi : U \to \widetilde{U}$ is a metric-relating chart taking values in a Hilbert space $\big(\mathcal{H}, \langle\cdot,\cdot\rangle\big)$. Choose constants $k_{\min}, k_{\max} > 0$ such that* (4.3.2) *holds. If $\widetilde{U}$ is convex then for any $x, y \in U$ we have:*

$$\mathrm{dist}(x,y) \le k_{\max}\big\|\varphi(x) - \varphi(y)\big\|.$$

PROOF. Set $\gamma(t) = \varphi^{-1}\big((1-t)\varphi(x) + t\varphi(y)\big)$ for $t \in [0,1]$ and observe that:

$$\mathrm{dist}(x,y) \le \int_0^1 \big\|\gamma'(t)\big\|\,\mathrm{d}t \le k_{\max}\big\|\varphi(x) - \varphi(y)\big\|. \qquad \square$$

4.3.17. LEMMA. *Let $(\mathcal{M}, g)$ be a Riemannian manifold and assume that $\varphi : U \to \widetilde{U}$ is a metric-relating chart taking values in a Hilbert space $\big(\mathcal{H}, \langle\cdot,\cdot\rangle\big)$. Choose constants $k_{\min}, k_{\max} > 0$ such that* (4.3.2) *holds. Let $F \subset \mathcal{H}$ be a closed subset of $\mathcal{H}$ contained in $\widetilde{U}$ and let $\gamma : [a,b] \to \mathcal{M}$ be a piecewise C^1 curve with $\gamma(a) \in \varphi^{-1}(F)$. If $L(\gamma) < k_{\min} \cdot \mathrm{dist}\big((\varphi\circ\gamma)(a), \partial F\big)$ then*[5] *the image of γ is contained in $\varphi^{-1}(F)$ (and hence in U).*

[5]If the boundary ∂F of F in $\mathcal{H}$ is empty (i.e., if $F = \mathcal{H}$) then the distance $\mathrm{dist}\big((\varphi\circ\gamma)(a), \partial F\big)$ should be interpreted as $+\infty$. In this case the theorem states that *any* piecewise C^1 curve γ starting

PROOF. Consider the set:

$$A = \big\{t \in [a,b] : \gamma\big([a,t]\big) \subset \varphi^{-1}(F)\big\};$$

A is not empty because $a \in A$ and therefore we can consider the supremum $c = \sup A \in [a,b]$. Assume by contradiction that $c < b$. Obviously we have $\gamma\big(\,[a,c[\,\big) \subset \varphi^{-1}(F)$, so that $\varphi \circ \gamma|_{[a,c[}$ is a well-defined piecewise C^1 curve in $\mathcal{H}$. Since $\gamma|_{[a,c[}$ has finite length in the Riemannian manifold M (because $\gamma|_{[a,c[}$ has a piecewise C^1 extension to $[a,c]$) and since φ is metric-relating, it follows that $\varphi \circ \gamma|_{[a,c[}$ is a curve of finite length in the Riemannian manifold $\mathcal{H}$ endowed with the constant Riemannian metric $\langle\cdot,\cdot\rangle$. It follows from Lemma 4.3.7 that $\varphi \circ \gamma|_{[a,c[}$ has relatively compact image in $\mathcal{H}$ and therefore we can find a sequence $(t_n)_{n\ge1}$ in $[a,c[$ with $t_n \to c$ and $(\varphi \circ \gamma)(t_n) \to \tilde{x}$, for some $\tilde{x} \in \mathcal{H}$. Using that F is closed, we obtain that $\tilde{x} \in F \subset \widetilde{U}$ and therefore $\tilde{x} = \varphi(x)$ for some $x \in \varphi^{-1}(F) \subset U$. Since $\varphi : U \to \widetilde{U}$ is a homeomorphism, we conclude that $\gamma(t_n) \to x$ and therefore $x = \gamma(c) \in \varphi^{-1}(F)$. But $(\varphi \circ \gamma)(c)$ cannot belong to the interior of F because (since $c < b$) this would imply that $c + \varepsilon \in A$ for some small $\varepsilon > 0$. We have proven that $(\varphi \circ \gamma)(c) \in \partial F$; now we compute as follows:

$$\begin{aligned} L(\gamma) &\ge \int_a^c g\big(\gamma'(t),\gamma'(t)\big)^{\frac12}\,\mathrm{d}t \ge k_{\min}\int_a^c \big\langle(\varphi\circ\gamma)'(t),(\varphi\circ\gamma)'(t)\big\rangle^{\frac12}\,\mathrm{d}t \\ &\ge k_{\min}\big\|(\varphi\circ\gamma)(c)-(\varphi\circ\gamma)(a)\big\| \ge k_{\min}\cdot\mathrm{dist}\big((\varphi\circ\gamma)(a),\partial F\big), \end{aligned}$$

which is a contradiction. □

4.3.18. COROLLARY. *Let $(\mathcal{M}, g)$ be a Riemannian manifold and assume that $\varphi : U \to \widetilde{U}$ is a metric-relating chart taking values in a Hilbert space $\big(\mathcal{H},\langle\cdot,\cdot\rangle\big)$. Choose constants $k_{\min}, k_{\max} > 0$ such that* (4.3.2) *holds. Assume that F is a closed subset of $\mathcal{H}$ contained in $\widetilde{U}$. If $x \in \varphi^{-1}(F)$, $y \in \mathcal{M}$ satisfy:*

$$\mathrm{dist}(x,y) < k_{\min}\cdot\mathrm{dist}\big(\varphi(x),\partial F\big),$$

then $y \in \varphi^{-1}(F) \subset U$ and:

$$\big\|\varphi(x)-\varphi(y)\big\| \le \frac{1}{k_{\min}}\,\mathrm{dist}(x,y).$$

PROOF. For any $\varepsilon > 0$ we can choose a piecewise C^1 curve $\gamma : [a,b] \to \mathcal{M}$ with $\gamma(a) = x$, $\gamma(b) = y$ and $L(\gamma)$ smaller than both $\mathrm{dist}(x,y) + \varepsilon$ and $k_{\min}\mathrm{dist}\cdot\big(\varphi(x),\partial F\big)$. By Lemma 4.3.17, we have $\mathrm{Im}(\gamma) \subset \varphi^{-1}(F)$ and in particular $y \in \varphi^{-1}(F)$. Moreover:

$$\big\|\varphi(x)-\varphi(y)\big\| \le \int_a^b \big\|(\varphi\circ\gamma)'(t)\big\|\,\mathrm{d}t \le \frac{1}{k_{\min}}\int_a^b \big\|\gamma'(t)\big\|\,\mathrm{d}t < \frac{1}{k_{\min}}\big(\mathrm{dist}(x,y)+\varepsilon\big),$$

where in the first inequality we have used the result of Exercise 4.12. The conclusion now follows by observing that $\varepsilon > 0$ can be taken arbitrarily small. □

at $\varphi^{-1}(F)$ has image contained in $\varphi^{-1}(F)$. In particular, U is actually an arc-connected component of $\mathcal{M}$.

4.3.19. COROLLARY. *Let* $(\mathcal{M}, g)$ *be a Riemannian manifold and assume that* $\varphi : U \to \widetilde{U}$ *is a metric-relating chart taking values in a Hilbert space* $\big(\mathcal{H}, \langle\cdot,\cdot\rangle\big)$. *Let* $(x_n)_{n\ge1}$ *be a Cauchy sequence in* $\mathcal{M}$ *(i.e.,* $\mathrm{dist}(x_n, x_m) \xrightarrow{n,m\to+\infty} 0$*). Assume that we can find a closed subset* F *of* $\mathcal{H}$ *contained in* $\widetilde{U}$ *such that* $x_n \in F \subset U$ *for all* n *and:*

$$\inf_{n\ge1} \mathrm{dist}\big(\varphi(x_n), \partial F\big) > 0. \tag{4.3.3}$$

Then, the sequence $\big(\varphi(x_n)\big)_{n\ge1}$ *is Cauchy (and hence convergent) in the Hilbert space* $\mathcal{H}$.

PROOF. Denote by $c > 0$ the infimum on the left hand side of formula (4.3.3) and choose constants $k_{\mathrm{min}}, k_{\mathrm{max}} > 0$ for which (4.3.2) holds. Since $(x_n)_{n\ge1}$ is Cauchy, we can find $n_0 \in \mathbb{N}$ such that $n, m \ge n_0$ imply $\mathrm{dist}(x_n, x_m) < k_{\mathrm{min}}c$. By Corollary 4.3.18, we have:

$$\big\|\varphi(x_n) - \varphi(x_m)\big\| \le \frac{1}{k_{\mathrm{min}}} \mathrm{dist}(x_n, x_m),$$

for all $n, m \ge n_0$. The conclusion follows. □

4.3.20. COROLLARY. *Let* $(\mathcal{M}, g)$ *be a Riemannian manifold and assume that* $\varphi : U \to \widetilde{U}$ *is a metric-relating chart taking values in a Hilbert space* $\big(\mathcal{H}, \langle\cdot,\cdot\rangle\big)$. *Choose constants* $k_{\mathrm{min}}, k_{\mathrm{max}} > 0$ *such that* (4.3.2) *holds. Assume that the closed ball* $\mathrm{B}[0;r] \subset \mathcal{H}$ *is contained in* $\widetilde{U}$ *for some* $r > 0$ *and choose* $r_0 > 0$ *small enough so that:*

$$r_0 \le \min\Big\{\frac{r}{2}, \frac{k_{\mathrm{min}}}{k_{\mathrm{max}}}\frac{r}{4}\Big\};$$

then, setting $V = \varphi^{-1}\big(\mathrm{B}(0;r_0)\big) \subset U$*, we have:*

$$\big\|\varphi(x) - \varphi(y)\big\| \le \frac{1}{k_{\mathrm{min}}} \mathrm{dist}(x, y),$$

for all $x, y \in V$.

PROOF. Let $x, y \in V$ be fixed. Since $\varphi(V)$ is convex, Lemma 4.3.16 implies that:

$$\mathrm{dist}(x, y) \le k_{\mathrm{max}}\big\|\varphi(x) - \varphi(y)\big\| < 2k_{\mathrm{max}}r_0 \le \frac{k_{\mathrm{min}}r}{2}.$$

Taking $F = \mathrm{B}[0;r]$, since $\varphi(x) \in \mathrm{B}\big(0; \frac{r}{2}\big)$, we have:

$$\mathrm{dist}\big(\varphi(x), \partial F\big) = \mathrm{dist}\big(\varphi(x), S(0;r)\big) > \frac{r}{2}.$$

The conclusion now follows from Corollary 4.3.18. □

We can now prove the following:

4.3.21. PROPOSITION. *If* $(\mathcal{M}, g)$ *is a connected Riemannian manifold then the distance function introduced in Definition 4.3.1 is indeed a (metric space) metric; moreover, the topology induced by such metric coincides with the topology of the manifold* $\mathcal{M}$.

PROOF. In order to prove that dist is a (metric space) metric it suffices to show that $\mathrm{dist}(x,y) > 0$ when $x, y \in M$ are distinct. Let then $x, y \in M$ be distinct and assume by contradiction that $\mathrm{dist}(x,y) = 0$. Choose a metric-relating chart $\varphi : U \to \widetilde{U}$ with $x \in U$ and $\varphi(x) = 0$, where $\widetilde{U}$ is an open subset of a Hilbert space $\big(\mathcal{H}, \langle\cdot,\cdot\rangle\big)$; choose also $k_{\mathrm{min}}, k_{\mathrm{max}} > 0$ satisfying (4.3.2). Since $\mathrm{dist}(x,y) = 0$, for any $r > 0$ with $\mathrm{B}[0;r] \subset \widetilde{U}$ we can find a piecewise C^1 curve connecting x and y with length less than rk_{min}; applying Lemma 4.3.17 with $F = \mathrm{B}[0;r]$ we conclude that $y \in U$ and that $\varphi(y)$ is in $\mathrm{B}[0;r]$. Since $r > 0$ can be taken arbitrarily small we obtain that $\varphi(x) = \varphi(y)$, contradicting the injectivity of the chart φ.

We now prove that if $Z \subset \mathcal{M}$ is open with respect to the manifold topology of $\mathcal{M}$ then Z is open with respect to the topology induced by dist. Choose $x \in Z$ and let $\varphi : U \to \widetilde{U}$, $\mathcal{H}$, k_{min}, k_{max} and r be as above; we can assume also that $U \subset Z$. Applying Lemma 4.3.17 with $F = \mathrm{B}[0;r]$ we conclude that if a piecewise C^1 curve $\gamma : [a,b] \to \mathcal{M}$ satisfies $\gamma(a) = x$ and $L(\gamma) < k_{\mathrm{min}} r$ then $\gamma(b) \in U$. It follows that the open ball of radius $k_{\mathrm{min}} r$ and center x with respect to the metric dist is contained in U (and in Z). Thus, Z is open with respect to the topology induced by dist.

Assume now that Z is open with respect to the topology induced by dist. Choose $x \in Z$ and let $\varphi : U \to \widetilde{U} \subset \mathcal{H}$ be a metric-relating chart with $x \in U$ and $\widetilde{U}$ convex. Since Z is open in $(\mathcal{M}, \mathrm{dist})$, $Z \cap U$ is open in $(U, \mathrm{dist}|_{U\times U})$; Lemma 4.3.16 tells us that $\varphi^{-1} : \widetilde{U} \to (U, \mathrm{dist}|_{U\times U})$ is Lipschitz continuous and therefore $\varphi(Z \cap U)$ is open in $\widetilde{U}$ (and in $\mathcal{H}$). Since φ is a chart of $\mathcal{M}$, it follows that $Z \cap U$ is an open neighborhood of x with respect to the manifold topology of $\mathcal{M}$. Thus Z is open in the manifold topology of $\mathcal{M}$. □

4.3.22. COROLLARY. *If a Hilbert manifold $\mathcal{M}$ admits a Riemannian metric then every connected component of $\mathcal{M}$ is metrizable (we don't make any* a priori *assumptions on the topology of $\mathcal{M}$!). In particular, $\mathcal{M}$ is T_4.* □

The following definition will be essential in the development of infinite-dimensional Morse theory.

4.3.23. DEFINITION. If $\mathcal{M}$ is a Hilbert manifold and g is a Riemannian metric for $\mathcal{M}$ then we say that a subset $F \subset \mathcal{M}$ is *complete* if its intersection with every connected component of $\mathcal{M}$ is a complete metric space (endowed with the metric dist).

Now we can generalize Lemma 4.3.7 to the context of manifolds.

4.3.24. LEMMA. *Let $\mathcal{M}$ be a Riemannian manifold. If $\gamma : I \to \mathcal{M}$ is a piecewise C^1 curve of finite length defined on an arbitrary interval $I \subset \mathbb{R}$ then the image $\gamma(I)$ of γ is totally bounded. In particular, if $\gamma(I)$ is contained in some complete subset of $\mathcal{M}$ then $\gamma(I)$ is relatively compact.*

PROOF. It is identical to the proof of Lemma 4.3.7. □

4.4. Dynamics of the Gradient Flow in the non Compact Case

4.5. The Morse Relations in the non Compact Case

4.6. The CW-Complex Associated to a Morse Function on a non Compact Manifold

4.7. The Morse–Witten Complex in the non Compact Case

Exercises for Chapter 4

Calculus on Banach spaces and Banach manifolds.

EXERCISE 4.1. Let X be a vector space. A map $X \ni x \mapsto \|x\| \in \mathbb{R}$ is called a *semi-norm* if the following conditions hold:

- $\|x\| \geq 0$ for all $x \in X$;
- $\|cx\| = |c|\,\|x\|$ for all $c \in \mathbb{R}$, $x \in X$;
- $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.

A semi-norm $\|\cdot\|$ is called a *norm* if in addition $\|x\| = 0$ implies $x = 0$. If $\|\cdot\|$ is a semi-norm on X show that:

- the set $N = \big\{x \in X : \|x\| = 0\big\}$ is a subspace of X;
- the map:
$$X/N \ni x + N \longmapsto \|x\| \in \mathbb{R}$$
is well-defined and it defines a norm on the quotient space X/N.

EXERCISE 4.2. A *normed vector space* is a vector space X endowed with a norm $\|\cdot\|$. Show that the topology induced from such norm makes X into a topological vector space.

EXERCISE 4.3. Let X, Y be normed vector spaces and let $T : X \to Y$ be a linear map. Show that the following are equivalent:

- T is continuous;
- T is continuous at the origin;
- T is bounded on the unit ball of X;
- $\big\|T(x)\big\| \leq c\|x\|$ for all $x \in X$ and some $c \in \mathbb{R}$;
- T is Lipschitz-continuous.

EXERCISE 4.4. Let X be a vector space and let $\|\cdot\|_1$, $\|\cdot\|_2$ be norms on X. Show that the following conditions are equivalent:

- $\|\cdot\|_1$ and $\|\cdot\|_2$ induce the same topology on X;
- there exists positive constants $k_{\min}$, $k_{\max}$ with:
$$k_{\min}\|x\|_1 \leq \|x\|_2 \leq k_{\max}\|x\|_1,$$
for all $x \in X$.

(*hint*: use the result of Exercise 4.3 with $T = \mathrm{Id} : \big(X, \|\cdot\|_1\big) \to \big(X, \|\cdot\|_2\big)$).

EXERCISE 4.5. Generalize Exercise 4.3 to multi-linear maps; more explicitly, given normed spaces $X_1, \ldots, X_k, Y$ and a multi-linear map $B : X_1 \times \cdots \times X_k \to Y$, show that the following conditions are equivalent:

- B is continuous;
- B is continuous at the origin;
- B is bounded on $\prod_{i=1}^{k} \mathrm{B}[X_i]$.

Observe that continuous multi-linear maps are not Lipschitz continuous in general.

EXERCISE 4.6. Let X be a real vector space, $\big(Y, \|\cdot\|\big)$ a real Banach space and $T : X \to Y$ a linear isomorphism. Show that:

$$\|x\|_T = \big\|T(x)\big\|, \quad x \in X,$$

defines a norm on X that makes it into a Banach space. We call $\|\cdot\|_T$ the norm *induced* by T on X. Observe that $\|\cdot\|_T$ is the *unique* norm on X that makes T into an isometry. Show also that if X is previously endowed with a norm that makes T continuous then such norm is equivalent to $\|\cdot\|_T$ (*hint*: use the open mapping theorem).

EXERCISE 4.7. Let $\big(X, \|\cdot\|\big)$, $\big(Y, \|\cdot\|\big)$ be Banach spaces and let $\mathcal{T}$ be a set of continuous linear isomorphisms $T : X \to Y$. Assuming that:

$$\sup_{T \in \mathcal{T}} \|T\| < +\infty, \quad \sup_{T \in \mathcal{T}} \big\|T^{-1}\big\| < +\infty,$$

show that there exists constants $k_1, k_2 > 0$ (which do not depend on $T \in \mathcal{T}$) such that:

$$k_1\|x\| \le \|x\|_T \le k_2\|x\|,$$

for all $x \in X$ and all $T \in \mathcal{T}$.

EXERCISE 4.8. Given Banach spaces X, Y and a bounded injective linear map $T : X \to Y$, show that $\mathrm{Im}(T)$ is closed in Y if and only if $T : X \to T(X)$ is a homeomorphism when $T(X)$ is regarded with the topology induced from Y. Conclude that the following *principle of reduction of counter-domain* holds. Assume that we are given a commutative diagram:

$$\begin{tikzcd} & Y \\ \mathcal{Z} \arrow[ur, "f"] \arrow[r, "f_0"'] & X \arrow[u, "T"'] \end{tikzcd}$$

where X, Y are Banach spaces, $\mathcal{Z}$ is a topological space and $T : X \to Y$ is a bounded injective linear map with closed image. Then f is continuous if and only if f_0 is continuous.

EXERCISE 4.9. If X, Y, Z are normed vector spaces and $\widehat{T} : X \times Y \to Z$ and $T : X \to \mathrm{Lin}(Y, Z)$ are respectively a bilinear and a linear map related by the equality:

$$\widehat{T}(x, y) = T(x)(y), \quad x \in X,\ y \in Y,$$

show that $\|\widehat{T}\| = \|T\| \in [0, +\infty]$. Conclude that $\widehat{T}$ is continuous if and only if T is continuous. Generalize this result to multi-linear maps by proving that if X_1,

$X_2, \ldots, X_k, Y$ are normed vector spaces then the correspondence $T \leftrightarrow \widehat{T}$ defined by the equality:

$$\widehat{T}(x_1, x_2, \ldots, x_k) = T(x_1)(x_2)\ldots(x_k),$$

defines an isometry between the normed space of bounded multi-linear maps from $X_1 \times \cdots \times X_k$ to Y and the normed space:

$$\mathrm{Lin}\big(X_1, \mathrm{Lin}\big(X_2, \ldots, \mathrm{Lin}(X_k, Y)\big)\cdots\big).$$

EXERCISE 4.10. Let $(\Omega, \mathcal{A}, \mu)$ be a Measure space, i.e., Ω is a set, $\mathcal{A}$ is a σ-algebra on Ω and $\mu : \mathcal{A} \to [0, +\infty]$ is a (σ-additive) measure on $\mathcal{A}$. If you're not very familiar with general measure theory, simply assume that $\Omega = [a, b] \subset \mathbb{R}$, $\mathcal{A} =$ Lebesgue measurable subsets of $[a, b]$ and that $\mu =$ Lebesgue measure. If $f : \Omega \to \mathbb{R}^n$ is a measurable function and if $\|\cdot\|$ is a fixed norm on $\mathbb{R}^n$ we set:

$$\|f\|_{L^\infty} = \sup\big\{c \in \mathbb{R} : f^{-1}\big(\,]c, +\infty[\,\big) \text{ has null measure}\big\} \in [0, +\infty].$$

If $\|f\|_{L^\infty} < +\infty$ then we say that f is *essentially bounded*. Show that the set of essentially bounded measurable $\mathbb{R}^n$-valued maps on Ω is a subspace of the vector space of all $\mathbb{R}^n$-valued maps on Ω; show that $\|\cdot\|_{L^\infty}$ defines a semi-norm on that space and that $\|f\|_{L^\infty} = 0$ iff $f = 0$ almost everywhere. The normed space corresponding to such semi-norm (see Exercise 4.1) is denoted by $L^\infty(\Omega, \mathbb{R}^n)$. Show that $L^\infty(\Omega, \mathbb{R}^n)$ is a Banach space.

EXERCISE 4.11 (Gronwall's inequality). Let $\delta, \phi : [a, b] \to \mathbb{R}$ be non negative maps with δ continuous and ϕ integrable. Assume that:

$$\delta(t) \le c + \int_a^t \phi(s)\delta(s)\,\mathrm{d}s, \tag{4.7.1}$$

for all $t \in [a, b]$ and some fixed $c \in \mathbb{R}$. The goal of this exercise is to prove the inequality:

$$\delta(t) \le c\,\exp\Big(\int_a^t \phi(s)\,\mathrm{d}s\Big), \tag{4.7.2}$$

for all $t \in [a, b]$. Below we give the main steps of the proof.

- Define a sequence of continuous maps $K_n : [a, b] \to \mathbb{R}$ recursively by setting $K_0 \equiv 1$ and:

  $$K_{n+1}(t) = \int_a^t \phi(s)K_n(s)\,\mathrm{d}s, \quad n \ge 0.$$

  Show by induction on n that:

  $$0 \le K_n(t) \le \frac{1}{n!}\Big(\int_a^t \phi(s)\,\mathrm{d}s\Big)^n, \tag{4.7.3}$$

  for all $t \in [a, b]$, $n \ge 0$ (*hint*: observe that, under the induction hypothesis, we have:

  $$\phi(s)K_n(s) \le \frac{\mathrm{d}}{\mathrm{d}s}\frac{1}{(n+1)!}\Big(\int_a^s \phi(u)\,\mathrm{d}u\Big)^{n+1},$$

  for all $s \in [a, b]$.)

- Show by induction on n that:

$$\delta(t) \leq c\sum_{i=0}^{n} K_i(t) + K_n(t)\int_a^t \phi(s)\delta(s)\,\mathrm{d}s, \tag{4.7.4}$$

  for all $t \in [a,b]$, $n \geq 0$ (*hint*: use the induction hypothesis to estimate the integrand on (4.7.1) from above).
- Use (4.7.4) and (4.7.3) to prove (4.7.2) (*hint*: (4.7.3) implies that K_n tends to zero).

EXERCISE 4.12. If X is a Banach space and if $\gamma : [a,b] \to X$ is a piecewise C^1 curve, show that:

$$\big\|\gamma(b) - \gamma(a)\big\| \leq L(\gamma). \tag{4.7.5}$$

(*hint*: choose a linear functional $\lambda \in X^*$ with $\|\lambda\| = 1$ and $\lambda\big(\gamma(b) - \gamma(a)\big) = \big\|\gamma(b) - \gamma)(a)\big\|$. Apply the Fundamental Theorem of Calculus to the map $\lambda \circ \gamma : [a,b] \to \mathbb{R}$).

Observe that if X is a Hilbert space then the equality in (4.7.5) holds if and only if $\gamma'(t)$ is a positive multiple of $\gamma(b) - \gamma(a)$ for (almost) all $t \in [a,b]$. On the other hand if X is not a Hilbert space then there may exists curves connecting two points $p, q \in X$ with length is $\|p - q\|$ but whose image is not contained in the line segment $[p,q]$ (can you find an example in $\mathbb{R}^2$ endowed with the norm $\big\|(x_1,x_2)\big\| = \max\big\{|x_1|,|x_2|\big\}$?).

EXERCISE 4.13. Let $\mathcal{M}$ be a Riemannian manifold and let $\varphi : U \to \widetilde{U}$ be a chart, where $\widetilde{U}$ is open in a Hilbert space $\big(\mathcal{H}, \langle\cdot,\cdot\rangle\big)$. Show that every $x \in U$ has an open neighborhood V in U such that $\varphi|_V : V \to \varphi(V)$ is a metric-relating chart.

EXERCISE 4.14. Let $U \subset \mathbb{R} \times \mathbb{R}^n$ be an open subset. Show that the set:

$$\mathfrak{H}_{C^0}[U] = \big\{\gamma \in C^0\big([a,b],\mathbb{R}^n\big) : \big(t,\gamma(t)\big) \in U, \text{ for all } t \in [a,b]\big\}$$

is open in $C^0\big([a,b],\mathbb{R}^n\big)$. Moreover, given a continuous map $\alpha : U \to \mathbb{R}^n$, show that the map:

$$\mathfrak{H}_{C^0}[\alpha] : \mathfrak{H}_{C^0}[U] \longrightarrow C^0\big([a,b],\mathbb{R}^n\big)$$

defined by:

$$\mathfrak{H}_{C^0}[\alpha](\gamma)(t) = \alpha\big(t,\gamma(t)\big), \quad t \in [a,b],$$

for all $\gamma \in \mathfrak{H}_{C^0}[U]$, is continuous.

EXERCISE 4.15. Prove the following elementary properties of absolutely continuous functions:

- $\gamma : [a,b] \to \mathbb{R}^n$ is absolutely continuous if and only if each of its coordinates $\gamma_i : [a,b] \to \mathbb{R}$, $i = 1,\dots,n$, is absolutely continuous;
- Show that if $\gamma : [a,b] \to \mathbb{R}^n$ is absolutely continuous then $\gamma|_{[c,d]}$ is absolutely continuous for every subinterval $[c,d] \subset [a,b]$.
- if $\gamma : [a,b] \to \mathbb{R}^n$ is a curve and there exists $c \in \,]a,b[$ such that $\gamma|_{[a,c]}$ and $\gamma|_{[c,b]}$ are absolutely continuous then γ is absolutely continuous;

- if $f: X \to \mathbb{R}^n$ is a locally Lipschitz continuous[6] map defined on a subset X of $\mathbb{R}^m$ and if $\gamma : [a, b] \to \mathbb{R}^m$ is an absolutely continuous curve with $\mathrm{Im}(\gamma) \subset X$ then $f \circ \gamma : [a, b] \to \mathbb{R}^n$ is also absolutely continuous;
- absolutely continuous curves $\gamma : [a, b] \to \mathbb{R}^n$ form a vector subspace of $C^0\big([a, b], \mathbb{R}^n\big)$; if $n = 1$, they also form a subalgebra of $C^0\big([a, b], \mathbb{R}\big)$.

EXERCISE 4.16. A *partition* of an interval $[a, b]$ is a finite subset $P \subset [a, b]$ such that $a, b \in P$; we write $P = \{t_0, \ldots, t_k\}$ with $a = t_0 < t_1 < \cdots < t_k = b$. The *variation* of a curve $\gamma : [a, b] \to \mathbb{R}^n$ with respect to a partition P is defined by:

$$\mathrm{Var}(\gamma; P) = \sum_{i=0}^{k-1} \big\|\gamma(t_{i+1}) - \gamma(t_i)\big\|;$$

the *total variation* (or *length*) of γ, denoted by $\mathrm{Var}(\gamma)$, is defined to be the supremum of the variations of γ with respect to all possible partitions P of $[a, b]$. If $\mathrm{Var}(\gamma) < +\infty$ then γ is called a map of *bounded variation* (or a *rectifiable curve*). Denote by $\mathrm{BV}\big([a, b], \mathbb{R}^n\big)$ the set of all maps $\gamma : [a, b] \to \mathbb{R}^n$ of bounded variation.

- Show that “the line is the shortest path between two points”, i.e., for every $\gamma : [a, b] \to \mathbb{R}^n$ we have:
$$\big\|\gamma(b) - \gamma(a)\big\| \le \mathrm{Var}(\gamma).$$
- Show that if $\gamma : [a, b] \to \mathbb{R}^n$ is of bounded variation then for every subinterval $[c, d] \subset [a, b]$ the restriction $\gamma|_{[c,d]}$ is of bounded variation and $\mathrm{Var}(\gamma|_{[c,d]}) \le \mathrm{Var}(\gamma)$.
- Show that $\gamma : [a, b] \to \mathbb{R}^n$ is of bounded variation if and only if each of its coordinates $\gamma_i : [a, b] \to \mathbb{R}$ is.
- Given $\gamma : [a, b] \to \mathbb{R}^n$ and $c \in \,]a, b[$ show that if $\gamma|_{[a,c]}$ and $\gamma|_{[c,b]}$ are of bounded variation then so is γ and:
$$\mathrm{Var}(\gamma) = \mathrm{Var}(\gamma|_{[a,c]}) + \mathrm{Var}(\gamma|_{[c,b]}).$$
- Show that $\mathrm{BV}\big([a, b], \mathbb{R}^n\big)$ is a vector subspace of the space $\mathfrak{B}\big([a, b], \mathbb{R}^n\big)$ of all bounded $\mathbb{R}^n$-valued functions on $[a, b]$.
- Show that if $f : X \to \mathbb{R}^n$ is a locally Lipschitz continuous map defined on a subset $X \subset \mathbb{R}^m$ and if $\gamma : [a, b] \to \mathbb{R}^m$ is a curve of bounded variation with $\mathrm{Im}(\gamma) \subset X$ then $f \circ \gamma$ is of bounded variation.
- Show that if $\sigma : [c, d] \to [a, b]$ is a monotone surjective map then $\gamma : [a, b] \to \mathbb{R}^n$ is of bounded variation if and only if $\gamma \circ \sigma$ is, and that $\mathrm{Var}(\gamma) = \mathrm{Var}(\gamma \circ \sigma)$.
- Show that, for fixed $t_0 \in [a, b]$:
$$\|\gamma| = \big\|\gamma(t_0)\big| + \mathrm{Var}(\gamma)$$
defined a norm on $\mathrm{BV}\big([a, b], \mathbb{R}^n\big)$ that makes it into a Banach space.
- Show that the inclusion of $\mathrm{BV}\big([a, b], \mathbb{R}^n\big)$ in $\mathfrak{B}\big([a, b], \mathbb{R}^n\big)$ is continuous.
- Show that every absolutely continuous curve is of bounded variation.

[6]Observe that this is the case if X is open and f is of class C^1.

EXERCISE 4.17. Show that $H^1\big([a,b],\mathbb{R}^n\big)$ is subspace of the vector space $C^0\big([a,b],\mathbb{R}^n\big)$ and that the map:

$$H^1\big([a,b],\mathbb{R}^n\big) \ni \gamma \longmapsto (\gamma,\gamma') \in C^0\big([a,b],\mathbb{R}^n\big) \oplus L^2\big([a,b],\mathbb{R}^n\big) \tag{4.7.6}$$

is linear injective with closed image. Conclude that $H^1\big([a,b],\mathbb{R}^n\big)$ becomes a Banachable space with the topology induced from (4.7.6); a possible norm for this topology is:

$$\|f\| = \|f\|_\infty + \|f'\|_{L^2}.$$

Consider now the linear maps:

$$H^1\big([a,b],\mathbb{R}^n\big) \ni \gamma \longmapsto (\gamma,\gamma') \in L^2\big([a,b],\mathbb{R}^n\big) \oplus L^2\big([a,b],\mathbb{R}^n\big), \tag{4.7.7}$$

$$H^1\big([a,b],\mathbb{R}^n\big) \ni \gamma \longmapsto \big(\gamma(t_0),\gamma'\big) \in \mathbb{R}^n \oplus L^2\big([a,b],\mathbb{R}^n\big), \tag{4.7.8}$$

where $t_0 \in \mathbb{R}$ is fixed. Assuming that $H^1\big([a,b],\mathbb{R}^n\big)$ is endowed with the topology induced from (4.7.6), show that (4.7.7) is a continuous linear injective map with closed image and that (4.7.8) is a continuous linear isomorphism. Conclude that both the inner products:

$$\langle \gamma_1,\gamma_1\rangle = \langle \gamma_1,\gamma_2\rangle_{L^2} + \langle \gamma_1',\gamma_2'\rangle_{L^2}, \tag{4.7.9}$$

$$\langle \gamma_1,\gamma_1\rangle = \big\langle \gamma_1(t_0),\gamma_2(t_0)\big\rangle + \langle \gamma_1',\gamma_2'\rangle_{L^2}, \tag{4.7.10}$$

induce the same topology on $H^1\big([a,b],\mathbb{R}^n\big)$ that (4.7.6) does (so that the topological vector space $H^1\big([a,b],\mathbb{R}^n\big)$ becomes indeed a Hilbert space with any of the equivalent inner products (4.7.9) and (4.7.10)).

EXERCISE 4.18. Show that the following inclusion maps are (well-defined and) continuous:

(a) $L^q\big([a,b],\mathbb{R}^n\big) \hookrightarrow L^p\big([a,b],\mathbb{R}^n\big)$ for $1 \le p \le q \le +\infty$;
(b) $C^0\big([a,b],\mathbb{R}^n\big) \hookrightarrow L^p\big([a,b],\mathbb{R}^n\big)$, $1 \le p \le +\infty$;
(c) $C^l\big([a,b],\mathbb{R}^n\big) \hookrightarrow C^k\big([a,b],\mathbb{R}^n\big)$, $0 \le k \le l$;
(d) $H^1\big([a,b],\mathbb{R}^n\big) \hookrightarrow C^0\big([a,b],\mathbb{R}^n\big)$;
(e) $C^1\big([a,b],\mathbb{R}^n\big) \hookrightarrow H^1\big([a,b],\mathbb{R}^n\big)$.

hint: for item (a) use the *Hölder inequality*:

$$\int_a^b fg \le \|f\|_{L^p} + \|g\|_{L^{p'}},$$

where $\frac{1}{p} + \frac{1}{p'} = 1$.

EXERCISE 4.19. Let $\big(\mathcal{H},\langle\cdot,\cdot\rangle\big)$ be a Hilbert space. We say that a continuous linear map $T:\mathcal{H}\to\mathcal{H}$ *represents* a continuous bilinear map $B:\mathcal{H}\times\mathcal{H}\to\mathbb{R}$ if:

$$B(x,y) = \big\langle T(x),y\big\rangle,$$

for all $x,y\in\mathcal{H}$. Show that if $B:\mathcal{H}\times\mathcal{H}\to\mathbb{R}$ is a continuous bilinear map then there exists a *unique* continuous linear map $T:\mathcal{H}\to\mathcal{H}$ that represents B.

EXERCISE 4.20. Let X be a Banach space. A map $f : [a,b] \to X$ is called *weakly integrable* if there exists a vector $I \in X$ such that for every continuous linear functional $\lambda \in X^*$ the map $\lambda \circ f : [a,b] \to \mathbb{R}$ is (Lebesgue) integrable and $\int_a^b \lambda \circ f = \lambda(I)$. Show that:

- the vector I above is unique when it exists (*hint*: use Hahn–Banach's theorem); it is called the *weak integral* of f and it is denoted by $\int_a^b f$.
- weakly integrable maps form a subspace of the space of all X-valued maps on $[a,b]$;
- the weak integral is an X-valued linear map on the space of weakly integrable maps $f : [a,b] \to X$;
- if $f : [a,b] \to X$ is bounded and weakly integrable then:

$$\Big\| \int_a^b f \Big\| \le (b-a) \sup_{t\in[a,b]} \big\|f(t)\big\|.$$

  hint: by Hahn–Banach's theorem, there exists $\lambda \in X^*$ with $\|\lambda\| = 1$ and $\lambda \cdot \int_a^b f = \Big\| \int_a^b f \Big\|$).
- the uniform limit of weakly integrable maps is weakly integrable;
- if f is *simple*, i.e., if $\mathrm{Im}(f) = \{x_1, \ldots, x_n\} \subset X$ is finite and if the sets $f^{-1}(x_i) \subset [a,b]$ are measurable then f is weakly integrable and:

$$\int_a^b f = \sum_{i=1}^n x_i \cdot \text{measure}\big(f^{-1}(x_i)\big).$$

- every continuous map $f : [a,b] \to X$ is weakly integrable (*hint*: every continuous map is a uniformly limit of maps as the one in the item above).
- if $f : [a,b] \to X$ is continuous then $F(t) = \int_a^t f$ is of class C^1 and $F' = f$.

EXERCISE 4.21. Show that the continuous isomorphism (4.7.8) maps the subspace $C^\infty\big([a,b],\mathbb{R}^n\big)$ of $H^1\big([a,b],\mathbb{R}^n\big)$ onto $\mathbb{R}^n \oplus C^\infty\big([a,b],\mathbb{R}^n\big)$. Conclude (using the standard fact that $C^\infty\big([a,b],\mathbb{R}^n\big)$ is dense in $L^2\big([a,b],\mathbb{R}^n\big)$) that the space $C^\infty\big([a,b],\mathbb{R}^n\big)$ is dense in $H^1\big([a,b],\mathbb{R}^n\big)$.

EXERCISE 4.22. Let $f, \phi : [a,b] \to \mathbb{R}$ be non negative functions, with f absolutely continuous and ϕ integrable. Show that if:

$$\big|f'(t)\big| \le \phi(t)\sqrt{f(t)},$$

for almost all $t \in [a,b]$ then:

$$\big|\sqrt{f(b)} - \sqrt{f(a)}\,\big| \le \frac{1}{2}\int_a^b \phi(t)\,\mathrm{d}t.$$

(*hint*: if f is positive, use the Fundamental Theorem of Calculus for the absolutely continuous function $\sqrt{f}$; in the general case, replace f by $f + \varepsilon$ and then make $\varepsilon \to 0^+$).

EXERCISE 4.23. Let $\mathcal{M}$ be a Riemannian manifold and $f : \mathcal{M} \to \mathbb{R}$ be a non negative map of class C^1; assume that for some constant $k \geq 0$ we have:

$$\big\|\mathrm{d}f(x)\big\| \leq k\sqrt{f(x)},$$

for all $x \in \mathcal{M}$. Show that for every $x, y \in \mathcal{M}$ we have:

$$\big|\sqrt{f(y)} - \sqrt{f(x)}\,\big| \leq \frac{k}{2}\,\mathrm{dist}(x, y).$$

(*hint*: for every piecewise C^1 curve $\gamma : [a, b] \to \mathcal{M}$ connecting x and y we have $\big|(f \circ \gamma)'(t)\big| \leq \phi(t)\sqrt{(f \circ \gamma)(t)}$, where $\phi(t) = k\big\|\gamma'(t)\big\|$; apply the result of Exercise 4.22 to $f \circ \gamma$ and ϕ).

EXERCISE 4.24. Let $\mathcal{M}, \mathcal{N}$ be Riemannian manifolds and $f : \mathcal{M} \to \mathcal{N}$ a map of class C^1. Assume that for some constant $k > 0$ we have:

$$\big\|\mathrm{d}f(x)\big\| \leq k,$$

for all $x \in \mathcal{M}$. Show that f is Lipschitz continuous with constant k, i.e.:

$$\mathrm{dist}\big(f(x), f(y)\big) \leq k\,\mathrm{dist}(x, y),$$

for all $x, y \in \mathcal{M}$ (*hint*: for every piecewise C^1 curve γ connecting x and y, show that $L(f \circ \gamma) \leq kL(\gamma)$).

EXERCISE 4.25. For every non negative real numbers a, b, show that $(a + b)^2 \leq 2(a^2 + b^2)$.

EXERCISE 4.26. Prove Lemma 4.3.12.

EXERCISE 4.27. Let X be a topological space and assume that X can be written as a disjoint union $X = \bigcup_{i \in I} X_i$ of open subsets $X_i \subset X$ such that each X_i is metrizable. Prove that a subspace $K \subset X$ is compact if and only if K is sequentially compact (*hint*: show that if K is sequentially compact then K intercepts at most a finite number of X_i's). In particular, a subset K of a Riemannian manifold is compact if and only if it is sequentially compact.

EXERCISE 4.28. Let V be a vector space. Given a linear operator $P : V \to V$ show that the following conditions are equivalent:

- P is a projection operator;
- $P(x) = x$ for all $x \in \mathrm{Im}(P)$;
- there exists a subspace $W \subset V$ such that $V = W \oplus \mathrm{Im}(P)$ and such that $P(w + x) = x$ for all $w \in W$, $x \in \mathrm{Im}(P)$, i.e., P is the projection onto the second coordinate corresponding to the direct sum $W \oplus \mathrm{Im}(P)$;
- $V = \mathrm{Ker}(P) \oplus \mathrm{Im}(P)$ and P is the projection onto the second coordinate with respect to the direct sum $\mathrm{Ker}(P) \oplus \mathrm{Im}(P)$.

Now assume that V is real and that V is endowed with an inner product. Given a projection operator $P : V \to V$, show that P is the orthogonal projection operator onto $\mathrm{Im}(P)$ if and only if P is self-adjoint.

EXERCISE 4.29. Let $\mathcal{H}$ be a Hilbert space and let $(x_n)_{n\geq 1}$ be a sequence in $\mathcal{H}$ that converges weakly to $x \in \mathcal{H}$. If $\lim_{n\to+\infty} \|x_n\| = \|x\|$ show that $(x_n)_{n\geq 1}$ converges to x in the norm topology.

EXERCISE 4.30. Let $\mathcal{H}$ be a Hilbert space. Show that a self-adjoint operator $T : \mathcal{H} \to \mathcal{H}$ is injective if and only if its image is dense in $\mathcal{H}$ (*hint*: $\mathrm{Ker}(T)$ is the orthogonal complement of $\mathrm{Im}(T)$). Conclude that if T is a self-adjoint isomorphism and if $V \subset \mathcal{H}$ is a closed invariant subspace then $T|_V : V \to V$ is also an isomorphism.

EXERCISE 4.31. Let $\mathcal{H}$ be a Hilbert space and $T : \mathcal{H} \to \mathcal{H}$ a self-adjoint operator. Given a closed invariant subspace $V \subset \mathcal{H}$, show that:

$$\sigma(T|_V) \subset \sigma(T).$$

(*hint*: use Exercise 4.30).

Infinite dimensional Morse theory.

EXERCISE 4.32. Let $\mathcal{H}$ be a Hilbert space and let $\alpha : \mathcal{H} \to \mathbb{R}$ be a non zero continuous linear functional. Show that the restriction of α to the unit sphere $S(\mathcal{H})$ satisfies the Palais–Smale condition with respect to the Riemannian metric induced from $\mathcal{H}$ (*hint*: use the result of Exercise 4.29).

The Hilbert Manifold Structure of $\mathbf{H^1}\big([\mathbf{a},\mathbf{b}],\mathbf{M}\big)$.

EXERCISE 4.33. Let $n \in \mathbb{N}$ be fixed and let $\mathfrak{M}$ be a set $\mathbb{R}^n$-valued continuous curves defined on compact intervals (different elements of $\mathfrak{M}$ may be defined on different intervals. Assume that the following properties hold:

(a) if $\gamma : [a,b] \to \mathbb{R}^n$ is in $\mathfrak{M}$ and $[c,d] \subset [a,b]$ is a subinterval then $\gamma|_{[c,d]}$ is in $\mathfrak{M}$;
(b) if $\gamma : [a,b] \to \mathbb{R}^n$ is a curve and there exists $c \in \,]a,b[$ such that both $\gamma|_{[a,c]}$ and $\gamma|_{[c,b]}$ are in $\mathfrak{M}$ then γ is in $\mathfrak{M}$;
(c) if $\gamma : [a,b] \to \mathbb{R}^n$ is in $\mathfrak{M}$, $\alpha : U \to V$ is a smooth diffeomorphism between open subsets $U, V \subset \mathbb{R}^n$ and if $\gamma\big([a,b]\big) \subset U$ then $\alpha \circ \gamma$ is in $\mathfrak{M}$.

Given an n-dimensional differentiable manifold M we say that a curve $\gamma : [a,b] \to M$ is *of class* $\mathfrak{M}$ if it is continuous and for every local chart $\varphi : U \to \widetilde{U}$ and every $[c,d] \subset [a,b]$ with $\gamma\big([c,d]\big) \subset U$ we have that $\varphi \circ \gamma|_{[c,d]}$ is in $\mathfrak{M}$. Show that the following conditions are equivalent for a curve $\gamma : [a,b] \to M$:

- γ is of class $\mathfrak{M}$;
- for every $t_0 \in [a,b]$ there exists $\varepsilon > 0$ and a chart $\varphi : U \to \widetilde{U}$ of M such that $\gamma\big([t_0-\varepsilon, t_0+\varepsilon] \cap [a,b]\big) \subset U$ and $\varphi \circ \gamma|_{[t_0-\varepsilon,t_0+\varepsilon]\cap[a,b]}$ is in $\mathfrak{M}$;
- there exists a partition $a = t_0 < t_1 < \cdots < t_k = b$ of $[a,b]$ and a family of charts $\varphi_i : U_i \to \widetilde{U}_i$, $i = 0,\ldots,k-1$, of M such that $\gamma\big([t_i,t_{i+1}]\big) \subset U_i$ and $\varphi_i \circ \gamma|_{[t_i,t_{i+1}]}$ is in $\mathfrak{M}$ for all $i = 0,\ldots,k-1$.

EXERCISE 4.34. Let M be a differentiable manifold and $\gamma : [a,b] \to M$ a curve of class H^1. Given a vector field $v : [a,b] \to TM$ along γ show that the following conditions are equivalent:

- v is of class H^1;
- for every $t_0 \in [a,b]$ there exists $\varepsilon > 0$ and a chart $\varphi : U \to \widetilde{U}$ in M with $\gamma\big([t_0-\varepsilon, t_0+\varepsilon] \cap [a,b]\big) \subset U$ and such that:

$$[t_0-\varepsilon, t_0+\varepsilon] \cap [a,b] \ni t \longmapsto \mathrm{d}\varphi\big(\gamma(t)\big) \cdot v(t) \in \mathbb{R}^n$$

  is of class H^1.
- for every subinterval $[c,d] \subset [a,b]$ and every chart $\varphi : U \to \widetilde{U}$ in M with $\gamma\big([c,d]\big) \subset U$ the map:

$$[c,d] \ni t \longmapsto \mathrm{d}\varphi\big(\gamma(t)\big) \cdot v(t) \in \mathbb{R}^n$$

  is of class H^1.

EXERCISE 4.35. The goal of this exercise is to fill in the details of the final part of the proof of Proposition 5.2.1.

- Show that the map:

$$\zeta : \begin{cases} C^0\big([a,b], \mathrm{Bil}(\mathbb{R}^n)\big) \\ \times \\ \mathrm{Lin}\big(H^1([a,b],\mathbb{R}^n), L^2([a,b],\mathbb{R}^n)\big) \\ \times \\ \mathrm{Lin}\big(H^1([a,b],\mathbb{R}^n), L^2([a,b],\mathbb{R}^n)\big) \\ \times \\ H^1\big([a,b],\mathbb{R}^n\big) \\ \times \\ H^1\big([a,b],\mathbb{R}^n\big) \end{cases} \longrightarrow \mathbb{R}$$

  defined by:

$$\zeta(G, D_1, D_2, u_1, u_2) = \int_a^b G(t)\big[\big(D_1(u_1)\big)(t), \big(D_2(u_2)\big)(t)\big]\,\mathrm{d}t,$$

  is multi-linear and continuous.
- Conclude from the item above that:

$$\left.\begin{matrix} C^0\big([a,b], \mathrm{Bil}(\mathbb{R}^n)\big) \\ \times \\ \mathrm{Lin}\big(H^1([a,b],\mathbb{R}^n), L^2([a,b],\mathbb{R}^n)\big) \\ \times \\ \mathrm{Lin}\big(H^1([a,b],\mathbb{R}^n), L^2([a,b],\mathbb{R}^n)\big) \end{matrix}\right\} \ni (G, D_1, D_2) \stackrel{\hat{\zeta}}{\longmapsto} \zeta(G, D_1, D_2, \cdot, \cdot)$$

  defines a continuous $\mathrm{Bil}\big(H^1([a,b],\mathbb{R}^n)\big)$-valued trilinear map.
- Show that the map:

$$\xi : \mathrm{Bil}(\mathbb{R}^n) \longrightarrow \mathrm{Bil}\big(H^1([a,b],\mathbb{R}^n)\big)$$

  defined by $\xi(B)(u_1,u_2) = B\big(u_1(a), u_2(a)\big)$ is linear and continuous.
- conclude the proof of Proposition 5.2.1 by showing that the map (5.2.12) can be assembled as:

$$\xi \circ \tilde{g} \circ \big(a, \mathrm{Eval}_a\big) + \hat{\zeta} \circ \big(\mathfrak{H}[\tilde{g}], \widetilde{D}, \widetilde{D}\big),$$

where $\mathrm{Eval}_a : H^1\big([a,b],\mathbb{R}^n\big) \to \mathbb{R}^n$ denotes the map $\tilde{\gamma} \mapsto \tilde{\gamma}(a)$ of evaluation at a.

EXERCISE 4.36. Show that the map:

$$F : H^1\big([a,b],M\big) \longrightarrow M \times M$$

given by $F(\gamma) = \big(\gamma(a),\gamma(b)\big)$ is a smooth submersion. Conclude that the subset $H^1_{pq}\big([a,b],M\big)$ of $H^1\big([a,b],M\big)$ consisting of curves connecting p and q is a smooth submanifold of $H^1\big([a,b],M\big)$ and that its tangent space at a point γ consists of the vector fields along γ that vanish at the endpoints (*hint*: use Proposition 4.3.10).

CHAPTER 5

Applications of Morse Theory in the non Compact Case

5.1. Banach Manifolds of Maps

We now show how Lemma 4.3.5 can be applied in practice to prove differentiability of maps between Banach spaces.

If $U \subset \mathbb{R} \times \mathbb{R}^m$ is an open subset, we denote by $\mathfrak{H}[U]$ the set of all curves $\gamma : [a,b] \to \mathbb{R}^m$ of class H^1 whose graph is contained in U, i.e.:

$$\mathfrak{H}[U] = \big\{\gamma \in H^1\big([a,b], \mathbb{R}^m\big) : \big(t, \gamma(t)\big) \in U, \text{ for all } t \in [a,b]\big\}.$$

If $\alpha : U \to \mathbb{R}^n$ is a map of class C^1, we define a map:

$$\mathfrak{H}[\alpha] : \mathfrak{H}[U] \longrightarrow H^1\big([a,b], \mathbb{R}^n\big)$$

by setting:

$$\mathfrak{H}[\alpha](\gamma)(t) = \alpha\big(t, \gamma(t)\big), \quad t \in [a,b], \tag{5.1.1}$$

for all $\gamma \in \mathfrak{H}[U]$. We have the following:

5.1.1. THEOREM. *If $\alpha : U \to \mathbb{R}^n$ is a map of class C^k ($1 \le k \le \infty$) defined on an open subset $U \subset \mathbb{R}^m$ then $\mathfrak{H}[U]$ is open in $H^1\big([a,b], \mathbb{R}^m\big)$ and $\mathfrak{H}[\alpha]$ is of class C^{k-1}. Moreover, if $k \ge 2$ then the differential of $\mathfrak{H}[\alpha]$ is given by:*

$$\mathrm{d}\mathfrak{H}[\alpha]_\gamma(v)(t) = \frac{\partial \alpha}{\partial x}\big(t, \gamma(t)\big) \cdot v(t), \quad t \in [a,b], \tag{5.1.2}$$

for all $\gamma \in \mathfrak{H}[U]$, $v \in H^1\big([a,b], \mathbb{R}^m\big)$.

The proof of Theorem 5.1.1 will be split into several lemmas. We start by proving the continuity of $\mathfrak{H}[\alpha]$.

5.1.2. LEMMA. *If $\alpha : U \to \mathbb{R}^n$ is a map of class C^1 defined on an open subset $U \subset \mathbb{R}^m$ then $\mathfrak{H}[U]$ is open in $H^1\big([a,b], \mathbb{R}^m\big)$ and $\mathfrak{H}[\alpha]$ is continuous.*

PROOF. The fact that $\mathfrak{H}[U]$ is open in $H^1\big([a,b], \mathbb{R}^m\big)$ follows from the fact that $\mathfrak{H}[U]$ is open with respect to the C^0-norm (see Exercise 4.14) and from the fact that the inclusion of H^1 in C^0 is continuous. Using the result of Exercises 4.8 and 4.17 we see that in order to prove the continuity of $\mathfrak{H}[\alpha]$ it suffices to prove the continuity of the composite maps:

$$\mathfrak{H}[U] \xrightarrow{\mathfrak{H}[\alpha]} H^1\big([a,b], \mathbb{R}^m\big) \xrightarrow{\text{inclusion}} C^0\big([a,b], \mathbb{R}^m\big) \tag{5.1.3}$$

$$\mathfrak{H}[U] \xrightarrow{\mathfrak{H}[\alpha]} H^1\big([a,b], \mathbb{R}^m\big) \xrightarrow{\text{derivation}} L^2\big([a,b], \mathbb{R}^m\big) \tag{5.1.4}$$

The continuity of (5.1.3) follows from Exercise 4.14 and from the continuity of the inclusion of H^1 in C^0. In order to prove the continuity of (5.1.4) we evaluate it explicitly on $\gamma \in \mathfrak{H}[U]$ obtaining:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathfrak{H}[\alpha](\gamma)(t) = \frac{\partial \alpha}{\partial t}\big(t, \gamma(t)\big) + \frac{\partial \alpha}{\partial x}\big(t, \gamma(t)\big) \cdot \gamma'(t).$$

It follows that (5.1.4) is the sum of the restriction of $\mathfrak{H}_{C^0}[\frac{\partial\alpha}{\partial t}]$ to $\mathfrak{H}[U]$ (see Exercise 4.14) and of the map $\mathfrak{H}[U] \to L^2\big([a,b], \mathbb{R}^m\big)$ described by the following picture:

$$\begin{tikzcd}
\mathfrak{H}[U] \arrow[r, "\mathfrak{H}_{C^0}[\frac{\partial\alpha}{\partial x}]"] & C^0\big([a,b], \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n)\big) & \\
 & \bigoplus \arrow[r, "\text{the map (4.2.3)}"] & L^2\big([a,b], \mathbb{R}^n\big) \\
\mathfrak{H}[U] \arrow[r, "\text{derivation}"'] & L^2\big([a,b], \mathbb{R}^m\big) &
\end{tikzcd}$$

This conclude the proof. $\square$

5.1.3. LEMMA. *If $\alpha : U \to \mathbb{R}^n$ is a map of class C^2 defined on an open subset $U \subset \mathbb{R}^m$ then $\mathfrak{H}[\alpha]$ is of class C^1 and formula* (5.1.2) *holds.*

PROOF. This is a simple application of Lemma 4.3.5. The separating family $\mathcal{F}$ for $H^1\big([a,b], \mathbb{R}^n\big)$ is the family of *evaluation maps*, i.e., for every $t \in [a,b]$ we set:

$$\lambda_t : H^1\big([a,b], \mathbb{R}^n\big) \ni \gamma \longmapsto \gamma(t) \in \mathbb{R}^n,$$

and then we take $\mathcal{F} = \big\{\lambda_t : t \in [a,b]\big\}$. Now take g to be what it is supposed to be, i.e., define:

$$g : \mathfrak{H}[U] \longrightarrow \mathrm{Lin}\big(H^1([a,b], \mathbb{R}^m), H^1([a,b], \mathbb{R}^n)\big)$$

by setting:

$$g(\gamma)(v)(t) = \frac{\partial \alpha}{\partial x}\big(t, \gamma(t)\big) \cdot v(t), \quad t \in [a,b],$$

for all $\gamma \in \mathfrak{H}[U]$, $v \in H^1\big([a,b], \mathbb{R}^m\big)$. Obviously:

$$\frac{\partial\big(\lambda_t \circ \mathfrak{H}[\alpha]\big)}{\partial v}(\gamma) = \frac{\mathrm{d}}{\mathrm{d}s}\alpha\big(t, \gamma(t) + sv(t)\big)\Big|_{s=0} = g(\gamma)(v)(t).$$

The only non trivial part of the proof is the continuity of g which follows from the continuity of $\mathfrak{H}[\frac{\partial\alpha}{\partial x}] : \mathfrak{H}[U] \to H^1\big([a,b], \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n)\big)$ (see Lemma 5.1.2) and from Lemma 5.1.4 below. $\square$

5.1.4. LEMMA. *The map:*

$$\mathcal{O} : H^1\big([a,b], \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n)\big) \longrightarrow \mathrm{Lin}\big(H^1([a,b], \mathbb{R}^m), H^1([a,b], \mathbb{R}^n)\big)$$

defined by:

$$\mathcal{O}(T)(v)(t) = T(t) \cdot v(t),$$

for all $t \in [a,b]$, $T \in H^1\big([a,b], \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n)\big)$, $v \in H^1\big([a,b], \mathbb{R}^m\big)$ is linear and continuous.

PROOF. By Exercise 4.9, it suffices to show that the bilinear map:

$$\widehat{B}: H^1\big([a,b], \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n)\big) \times H^1\big([a,b], \mathbb{R}^m\big) \longrightarrow H^1\big([a,b], \mathbb{R}^n\big)$$

defined by (4.2.4) is continuous. But using the identification:

$$H^1\big([a,b], \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n)\big) \times H^1\big([a,b], \mathbb{R}^m\big) \cong H^1\big([a,b], \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n) \oplus \mathbb{R}^m\big)$$

the map $\widehat{B}$ is precisely $\mathfrak{H}[B]$, where $B : \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n) \times \mathbb{R}^m \to \mathbb{R}^n$ is defined by $B(T, v) = T(v)$. The conclusion follows from Lemma 5.1.2. □

PROOF OF THEOREM 5.1.1. It follows from Lemmas 5.1.2, 5.1.3 and 5.1.4, using induction on k and the fact that $\mathrm{d}\mathfrak{H}[\alpha]$ equals the composite of $\mathfrak{H}[\frac{\partial \alpha}{\partial x}]$ with the continuous linear map $\mathcal{O}$ defined in the statement of Lemma 5.1.4. □

If α is a smooth map then it is not true in general that "left composition with α" defines a smooth map on L^p spaces; in fact, such map may not even be well-defined, i.e., it may happen that f is in L^p, α is smooth but $\alpha \circ f$ is not in L^p. However, "left composition with α" is smooth on L^p when α is linear; the following proposition is a mixture of this observation with Theorem 5.1.1.

5.1.5. PROPOSITION. *Let $\alpha : U \times \mathbb{R}^m \to \mathbb{R}^n$ be a map of class C^k ($1 \le k \le \infty$) where U is open in $\mathbb{R} \times \mathbb{R}^p$; assume that $\alpha(t, x, \cdot)$ is linear on $\mathbb{R}^m$ for every $(t, x) \in U$. Then the map:*

$$\begin{array}{rl} \mathfrak{H}_{H^1\text{-}L^2}[\alpha] : H^1\big([a,b], \mathbb{R}^p\big) \times L^2\big([a,b], \mathbb{R}^m\big) & \\ \cup \qquad\qquad & \\ \mathfrak{H}[U] \times L^2\big([a,b], \mathbb{R}^m\big) & \longrightarrow L^2\big([a,b], \mathbb{R}^n\big) \end{array}$$

defined by:

$$\mathfrak{H}_{H^1\text{-}L^2}[\alpha](\gamma, v)(t) = \alpha\big(t, \gamma(t), v(t)\big), \quad t \in [a,b],$$

for all $\gamma \in \mathfrak{H}[U]$, $v \in L^2\big([a,b], \mathbb{R}^m\big)$, is of class C^{k-1}.

PROOF. Consider the map $\bar{\alpha}$ of class C^k defined by:

$$\bar{\alpha} : U \ni (t, x) \longmapsto \alpha(t, x, \cdot) \in \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n);$$

it follows from Theorem 5.1.1 that $\mathfrak{H}[\bar{\alpha}]$ is of class C^{k-1}. The conclusion follows by observing that $\mathfrak{H}_{H^1\text{-}L^2}[\alpha]$ is the composite of the map:

$$\mathfrak{H}[\bar{\alpha}] \times \mathrm{Id} : \mathfrak{H}[U] \times L^2\big([a,b], \mathbb{R}^m\big) \to H^1\big([a,b], \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n)\big) \times L^2\big([a,b], \mathbb{R}^m\big)$$

with (the restriction to $H^1 \times L^2$ of) the continuous bilinear map (4.2.3). □

5.1.6. DEFINITION. A curve $\gamma : [a,b] \to M$ on a differentiable manifold M is called *of Sobolev class* H^1 (shortly, *of class* H^1) if it is continuous and for every local chart $\varphi : U \to \widetilde{U}$ of M and for every interval $[c,d] \subset [a,b]$ with $\gamma\big([c,d]\big) \subset U$ we have that $\varphi \circ \gamma|_{[c,d]} : [c,d] \to \mathbb{R}^n$ is of class H^1. We denote by $H^1\big([a,b], M\big)$ the set of all curves $\gamma : [a,b] \to M$ of class H^1.

The definition above is not very practical if one wishes to show that a particular curve $\gamma:[a,b]\to M$ is of class H^1. For nicer statements of the definition above see Exercise 4.33 (where we consider a more general context than H^1 that would be suitable also for other purposes).

5.1.7. DEFINITION. A *one parameter family of charts* on an n-dimensional differentiable manifold M is a smooth map $\varphi:U\to\mathbb{R}^n$ defined on an open subset U of $\mathbb{R}\times M$ such that the map:

$$\varphi^\diamond : U \ni (t,x) \longmapsto \big(t,\varphi(t,x)\big) \in \mathbb{R}\times\mathbb{R}^n$$

is a diffeomorphism onto an open subset $\widetilde{U}$ of $\mathbb{R}\times\mathbb{R}^n$. For $t\in\mathbb{R}$ we denote by U_t the (possibly empty) open subset of M defined by:

$$U_t = \big\{x\in M : (t,x)\in U\big\};$$

by $\varphi_t : U_t\to\mathbb{R}^n$ we denote the map $\varphi_t(x)=\varphi(t,x)$ and we set:

$$\widetilde{U}_t = \mathrm{Im}(\varphi_t) = \big\{v\in\mathbb{R}^n : (t,v)\in\widetilde{U}\big\}.$$

We will write $\varphi=(\varphi_t,U_t,\widetilde{U}_t)$ to indicate that φ is a one parameter family of charts and that φ_t, U_t and $\widetilde{U}_t$ are defined as above.

Obviously, if $\varphi=(\varphi_t,U_t,\widetilde{U}_t)$ is a one parameter family of charts then $\varphi_t : U_t\to\widetilde{U}_t$ is a local chart on M for every $t\in\mathbb{R}$. Conversely, it follows from the inverse function theorem that if φ is smooth and each φ_t is a local chart then φ is a one parameter family of charts.

If U is an open subset of $\mathbb{R}\times M$ we denote by $\mathfrak{H}[U]$ the set of curves $\gamma:[a,b]\to M$ of class H^1 whose graph is contained in U, i.e.:

$$\mathfrak{H}[U] = \big\{\gamma\in H^1\big([a,b],M\big) : \big(t,\gamma(t)\big)\in U,\ \text{for all } t\in[a,b]\big\}.$$

If N is a differentiable manifold and $\alpha:U\to N$ is smooth, we define a map:

$$\mathfrak{H}[\alpha] : \mathfrak{H}[U] \longrightarrow H^1\big([a,b],N\big)$$

by formula (5.1.1). Observe that if $\varphi=(\varphi_t,U_t,\widetilde{U}_t)$ is a one parameter family of charts in M then $\mathfrak{H}[\varphi]$ gives a bijection from $\mathfrak{H}[U]$ to $\mathfrak{H}[\widetilde{U}]$.

Let $\varphi=(\varphi_t,U_t,\widetilde{U}_t)$, $\psi=(\psi_t,V_t,\widetilde{V}_t)$ be one parameter families of charts. If $U\cap V\ne\emptyset$ then the *transition function* from φ to ψ is the map

$$\alpha : \underbrace{\bigcup_{t\in\mathbb{R}}\{t\}\times\varphi_t(U_t\cap V_t)}_{\varphi^\diamond(U\cap V)} \longrightarrow \underbrace{\bigcup_{t\in\mathbb{R}}\{t\}\times\psi_t(U_t\cap V_t)}_{\psi^\diamond(U\cap V)}$$

defined by:

$$\alpha(t,v) = \big(t,(\psi_t\circ\varphi_t^{-1})(v)\big),$$

for all $(t,v)\in\mathbb{R}\times\mathbb{R}^n$ with $v\in\varphi_t(U_t\cap V_t)$. Obviously $\alpha=\psi^\diamond\circ\big(\varphi^\diamond\big)^{-1}$ is a smooth diffeomorphism between open subsets of $\mathbb{R}\times\mathbb{R}^n$. It follows from Theorem 5.1.1 that:

$$\mathfrak{H}[\psi]\circ\big(\mathfrak{H}[\varphi]\big)^{-1} = \mathfrak{H}[\alpha] : \mathfrak{H}[\varphi^\diamond(U\cap V)] \longrightarrow \mathfrak{H}[\psi^\diamond(U\cap V)]$$

is a smooth diffeomorphism between open subsets of $H^1\big([a,b],\mathbb{R}^n\big)$.

We have so far proven that for every one parameter family of charts $\varphi = (\varphi_t, U_t, \widetilde{U}_t)_{t\in\mathbb{R}}$, the map $\mathfrak{H}[\varphi]$ is a chart on the set $H^1\big([a,b],M\big)$ and that the charts of the form $\mathfrak{H}[\varphi]$ are pairwise compatible. In order to obtain a differentiable atlas for $H^1\big([a,b],M\big)$ we now need to show that the domains of the charts $\mathfrak{H}[\varphi]$ cover $H^1\big([a,b],M\big)$. This will be a consequence of the following:

5.1.8. PROPOSITION. *Given a continuous curve $\gamma : [a,b] \to M$ on a differentiable manifold M then there exists a one parameter family of charts $\varphi = (\varphi_t, U_t, \widetilde{U}_t)$ on M such that U contains the graph of φ.*

PROOF. Choose an arbitrary Riemannian metric on M. Recall that a positive number $r > 0$ is called a *normal radius* for a point $x \in M$ if the geodesical exponential map $\exp$ maps the ball $\mathrm{B}(0;r)$ of T_xM diffeomorphically onto an open subset of M. We call $r > 0$ a *totally normal* radius for $x \in M$ if r is a normal radius for x and for all the points in the open set $\exp\big(\mathrm{B}(0;r)\big)$. It is an standard argument in Riemannian geometry (see, for instance, [**24**]) that for every compact subset $K \subset M$ we can find a number $r > 0$ that is a totally normal radius for all points of K.

Consider an arbitrary continuous extension of γ to a curve defined in the whole line $\mathbb{R}$. Let $r > 0$ be a totally normal radius for all points of the compact set $K = \gamma\big([a-1,b+1]\big)$. By standard approximation arguments (see [**73**]) we can find a smooth curve $\mu : \,]a-1,b+1[\, \to M$ such that $\mathrm{dist}\big(\gamma(t),\mu(t)\big) < r$ for all $t \in \,]a-1,b+1[$, where dist denotes the distance function corresponding to the Riemannian metric of M. Choose an arbitrary parallel referential along μ, so that we obtain an isomorphism $\sigma_t : T_{\mu(t)}M \to \mathbb{R}^n$ for all $t \in \,]a-1,b+1[$. The conclusion is now obtained by taking U_t to be the exponential of the ball $\mathrm{B}(0;r)$ on $T_{\mu(t)}M$ and by taking φ_t to be the composition of the inverse of the diffeomorphism:

$$\exp : T_{\mu(t)}M \supset \mathrm{B}(0;r) \longmapsto U_t$$

with the isomorphism σ_t, for all $t \in \,]a-1,b+1[$. $\square$

5.1.9. COROLLARY. *If M is a differentiable manifold then the set $\big\{\mathfrak{H}[\varphi]\big\}_{\varphi}$, where φ runs over all possible one parameter families of charts on M, is a differentiable atlas for $H^1\big([a,b],M\big)$.* $\square$

We have endowed $H^1\big([a,b],M\big)$ with the structure of an infinite dimensional Hilbert manifold. As in the case of any Hilbert manifold, the tangent space of $H^1\big([a,b],M\big)$ at a point (i.e., a curve) $\gamma \in H^1\big([a,b],M\big)$ is a Hilbertable space that can be constructed using for instance equivalence classes of curves or any other general construction for tangent spaces of Hilbert manifolds. Nevertheless, such general construction is not useful for practical purposes; we need a more concrete description of $T_\gamma H^1\big([a,b],M\big)$.

For $t_0 \in [a,b]$ we denote by:

$$\mathrm{Eval}_{t_0} : H^1\big([a,b],M\big) \longrightarrow M$$

the *evaluation map at* t_0, i.e., $\mathrm{Eval}_{t_0}(\gamma) = \gamma(t_0)$ for all $\gamma \in H^1\big([a,b],M\big)$. If $\varphi = (\varphi_t, U_t, \widetilde{U}_t)$ is a one parameter family of charts in M then we have a commutative diagram:

$$\begin{tikzcd} H^1\big([a,b],M\big) \supset \mathfrak{H}[U] \arrow[r, "\mathrm{Eval}_{t_0}"] \arrow[d, "\mathfrak{H}[\varphi]"'] & U_t \subset M \arrow[d, "\varphi_{t_0}"] \\ H^1\big([a,b],\mathbb{R}^n\big) \supset \mathfrak{H}[\widetilde{U}] \arrow[r, "\mathrm{Eval}_{t_0}"'] & \widetilde{U}_t \subset \mathbb{R}^n \end{tikzcd} \tag{5.1.5}$$

that says that Eval_{t_0} is represented in the local charts $\mathfrak{H}[\varphi]$ and φ_{t_0} by the map $\mathrm{Eval}_{t_0} : H^1\big([a,b],\mathbb{R}^n\big) \to \mathbb{R}^n$ of evaluation at t_0. This implies that $\mathrm{Eval}_{t_0} : H^1\big([a,b],M\big) \to M$ is smooth for every $t_0 \in [a,b]$.

5.1.10. PROPOSITION. *Let M be a differentiable manifold. For every $\gamma \in H^1\big([a,b],M\big)$, $\mathfrak{v} \in T_\gamma H^1\big([a,b],M\big)$, set:*

$$v(t) = \mathrm{d}(\mathrm{Eval}_t)(\gamma)\cdot\mathfrak{v},$$

for all $t \in [a,b]$, so that $v : [a,b] \to TM$ is a vector field along γ. The curve $v : [a,b] \to TM$ is of class H^1 and the map:

$$TH^1\big([a,b],M\big) \ni \mathfrak{v} \longmapsto v \in H^1\big([a,b],TM\big) \tag{5.1.6}$$

is a smooth diffeomorphism of Hilbert manifolds.

PROOF. Let $\varphi = (\varphi_t, U_t, \widetilde{U}_t)_{t\in\mathbb{R}}$ be a one parameter family of charts in M. For every $t \in \mathbb{R}$, we have that $\mathrm{d}\varphi_t : TU_t \to \widetilde{U}_t \times \mathbb{R}^n$ is a local chart in TM defined on the open subset $TU_t \subset TM$; moreover, it is easy to see that $\varphi = (\mathrm{d}\varphi_t, TU_t, \widetilde{U}_t \times \mathbb{R}^n)_{t\in\mathbb{R}}$ is a one parameter family of charts in TM. Now the differential of $\mathfrak{H}[\varphi]$ gives a local chart:

$$\mathrm{d}\mathfrak{H}[\varphi] : T\mathfrak{H}[U] \longrightarrow \mathfrak{H}[\widetilde{U}] \times H^1\big([a,b],\mathbb{R}^n\big) \subset H^1\big([a,b],\mathbb{R}^n\big) \times H^1\big([a,b],\mathbb{R}^n\big)$$

on the tangent bundle $TH^1\big([a,b],M\big)$. Moreover,

$$\mathfrak{H}[\varphi] : \mathfrak{H}[TU] \longrightarrow \mathfrak{H}[\widetilde{U} \times \mathbb{R}^n] \cong \mathfrak{H}[\widetilde{U}] \times H^1\big([a,b],\mathbb{R}^n\big),$$

is a local chart on $H^1\big([a,b],TM\big)$. Differentiating (5.1.5) one obtains easily the following commutative diagram:

$$\begin{tikzcd} T\mathfrak{H}[U] \arrow[rr, "\text{the map (5.1.6)}"] \arrow[dr, "\mathrm{d}\mathfrak{H}[\varphi]"'] & & \mathfrak{H}[TU] \arrow[dl, "\mathfrak{H}[\varphi]"] \\ & \mathfrak{H}[\widetilde{U}] \times H^1\big([a,b],\mathbb{R}^n\big) & \end{tikzcd}$$

that says that (5.1.6) is represented by the identity with respect to suitable local coordinates. The conclusion follows. □

5.1.11. DEFINITION. If $\gamma : [a,b] \to M$ is a curve of class H^1 then a vector field v along γ is *of class* H^1 if $v : [a,b] \to TM$ is a curve of class H^1 in the differentiable manifold TM.

See Exercise 4.34 for equivalent definitions of vector field of class H^1 along curves.

From now on we will always identify the tangent bundle of $H^1\big([a,b],M\big)$ with $H^1\big([a,b],TM\big)$ via the diffeomorphism (5.1.6). In particular, for every curve $\gamma:[a,b]\to M$ of class H^1, the tangent space $T_\gamma H^1\big([a,b],M\big)$ is identified with the vector space of vector fields of class H^1 along γ.

5.1.12. PROPOSITION. *Given differentiable manifolds M, N and a smooth map $\alpha: U\to N$ defined on an open subset $U\subset\mathbb{R}\times M$ then $\mathfrak{H}[U]$ is open in $H^1\big([a,b],M\big)$ and $\mathfrak{H}[\alpha]:\mathfrak{H}[U]\to H^1\big([a,b],N\big)$ is smooth. Moreover, for every $\gamma\in\mathfrak{H}[U]$ and every $v\in T_\gamma H^1\big([a,b],M\big)$ we have:*

$$\mathrm{d}\mathfrak{H}[\alpha]_\gamma(v)(t)=\frac{\partial\alpha}{\partial x}\big(t,\gamma(t)\big)\cdot v(t),$$

for all $t\in[a,b]$.

PROOF. Follows easily from Theorem 5.1.1 using local charts of the form $\mathfrak{H}[\varphi]$. $\square$

5.1.13. COROLLARY. *Let M, N be finite-dimensional differentiable manifolds and let $H: N\times[a,b]\to M$ be a smooth map (in the sense that H admits a smooth extension to an open neighborhood of $N\times[a,b]$ in $N\times\mathbb{R}$). Then the map:*

$$\widehat{H}: N\ni x\longmapsto H(x,\cdot)\in H^1\big([a,b],M\big)$$

is smooth and its differential is given by:

$$\big[\mathrm{d}\widehat{H}(x)\cdot v\big](t)=\frac{\partial H}{\partial x}(x,t)\cdot v,$$

for all $x\in N$, $v\in T_xN$, $t\in[a,b]$.

PROOF. Consider a smooth extension of H to an open neighborhood of $N\times[a,b]$ in $N\times\mathbb{R}$. Denote by $\mathfrak{c}$ the map:

$$\mathfrak{c}: N\longrightarrow H^1\big([a,b],N\big)$$

that associates to every $x\in N$ the constant curve in N with constant value x; it is easy to see that $\mathfrak{c}$ is smooth. The conclusion now follows from Proposition 5.1.12 by observing that $\widehat{H}=\mathfrak{H}[H]\circ\mathfrak{c}$. $\square$

We set[1]:

$$C^\infty\big([a,b],M\big)=\big\{\gamma:[a,b]\to M:\gamma\text{ is smooth}\big\}.$$

5.1.14. PROPOSITION. *The set $C^\infty\big([a,b],M\big)$ is dense in the Hilbert manifold $H^1\big([a,b],M\big)$.*

[1]A curve $\gamma:[a,b]\to M$ will be called *smooth* if it admits a smooth extension to some open interval containing $[a,b]$.

PROOF. Let $\varphi = (\varphi_t, U_t, \widetilde{U}_t)$ be a one parameter family of charts in M. Since $\mathfrak{H}[\widetilde{U}]$ is open in $H^1\big([a,b], R^n\big)$, it follows from Exercise 4.21 that the intersection $C^\infty\big([a,b],\mathbb{R}^n\big) \cap \mathfrak{H}[\widetilde{U}]$ is dense in $\widetilde{U}$. Since $\mathfrak{H}[\varphi]^{-1} : \mathfrak{H}[\widetilde{U}] \to \mathfrak{H}[U]$ is a continuous map that takes smooth curves to smooth curves, it follows that the closure of $C^\infty\big([a,b],M\big)$ in $H^1\big([a,b],M\big)$ contains $\mathfrak{H}[U]$. The conclusion now follows from Corollary 5.1.9. $\square$

5.2. The Riemannian Metric of $\mathbf{H^1}\big([\mathbf{a},\mathbf{b}],\mathbf{M}\big)$

We will now define a Riemannian metric on the Hilbert manifold $H^1\big([a,b],M\big)$.

5.2.1. PROPOSITION. *Let (M,g) be a finite dimensional Riemannian manifold and let ∇ be an arbitrary connection on M. For every $\gamma \in H^1\big([a,b],M\big)$ the formula:*

(5.2.1)

$$\langle v,w\rangle_\gamma = g\big(v(a),w(a)\big) + \int_a^b g\big(\tfrac{\mathrm{D}v}{\mathrm{d}t},\tfrac{\mathrm{D}w}{\mathrm{d}t}\big)\,\mathrm{d}t,\quad v,w\in T_\gamma H^1\big([a,b],M\big),$$

gives a well-defined Hilbert space inner product on the space $T_\gamma H^1\big([a,b],M\big)$. Moreover, the family:

$$H^1\big([a,b],M\big) \ni \gamma \longmapsto \langle\cdot,\cdot\rangle_\gamma,$$

defines a Riemannian metric on $H^1\big([a,b],M\big)$.

PROOF. Let $\varphi = (\varphi_t, U_t, \widetilde{U}_t)$ be a one parameter family of charts in M. We define smooth maps:

$$\mathfrak{b} : \widetilde{U} \longrightarrow \mathbb{R}^n,\quad A : \widetilde{U} \longrightarrow \mathrm{Lin}(\mathbb{R}^n),$$
$$\Gamma : \widetilde{U} \longrightarrow \mathrm{Bil}(\mathbb{R}^n;\mathbb{R}^n),\quad \tilde{g} : \widetilde{U} \to \mathrm{Bil}(\mathbb{R}^n),$$

by setting:

$$\mathfrak{b}(t,\tilde{x}) = \frac{\partial\varphi}{\partial t}(t,x),\quad A_{(t,\tilde{x})}(e_i) = \mathrm{d}\varphi_t(x)\cdot\frac{\partial X_i}{\partial t}(t,x),$$
$$\Gamma_{(t,\tilde{x})}(e_i,e_j) = \mathrm{d}\varphi_t(x)\cdot\big[\nabla_{X_i}X_j(t,x)\big],\quad \tilde{g}_{(t,\tilde{x})} = g_x\big(\mathrm{d}\varphi_t(x)^{-1}\cdot,\mathrm{d}\varphi_t(x)^{-1}\cdot\big),$$

for all $(t,\tilde{x}) \in \widetilde{U}$, $i,j = 1,\ldots,n$, where $x = \varphi_t^{-1}(\tilde{x})$, $(e_i)_{i=1}^n$ is the canonical basis of $\mathbb{R}^n$ and $X_i(t,x) = \mathrm{d}\varphi_t(x)^{-1}\cdot e_i$, $i = 1,\ldots,n$. In the formulas above we have denoted by $\nabla_{X_i}X_j(t,x)$ the covariant derivative of the vector field $x \mapsto X_j(t,x)$ in the direction $X_i(t,x)$ and by $\frac{\partial X_i}{\partial t}(t,x)$ the standard derivative of the curve $t \mapsto X_i(t,x)$ in T_xM. The objects $\mathfrak{b}$, A, Γ and $\tilde{g}$ encode all the relevant information we need to describe (5.2.1) in the chart $\mathfrak{H}[\varphi]$ of $H^1\big([a,b],M\big)$. Let $\gamma \in \mathfrak{H}[U]$ be given and set $\tilde{\gamma} = \mathfrak{H}[\varphi](\gamma)$, so that:

$$\tilde{\gamma}(t) = \varphi\big(t,\gamma(t)\big),\quad t\in[a,b]. \tag{5.2.2}$$

We denote by $\tilde{\mathrm{d}}(\tilde\gamma):[a,b]\to\mathbb{R}^n$ the "coordinate[2] representation" of γ', i.e., we set:

$$\tilde{\mathrm{d}}(\tilde\gamma)(t)=\mathrm{d}\varphi_t\big(\gamma(t)\big)\cdot\gamma'(t),\quad t\in[a,b];\tag{5.2.3}$$

differentiating (5.2.2) we obtain:

$$\tilde{\mathrm{d}}(\tilde\gamma)(t)=\tilde\gamma'(t)-\mathfrak{b}\big(t,\tilde\gamma(t)\big),\quad t\in[a,b].\tag{5.2.4}$$

Now pick $v\in T_\gamma H^1\big([a,b],M\big)$ and set $\tilde v=\mathrm{d}\mathfrak{H}[\varphi]_\gamma(v)$, so that:

$$\tilde v(t)=\mathrm{d}\varphi_t\big(\gamma(t)\big)\cdot v(t),\quad t\in[a,b];\tag{5.2.5}$$

using the time-dependent referential $(X_i)_{i=1}^n$ we can rewrite (5.2.5) as:

$$v(t)=\sum_{i=1}^n\tilde v_i(t)X_i\big(t,\gamma(t)\big),\quad t\in[a,b].\tag{5.2.6}$$

We denote by $\widetilde{\mathrm{D}}_{\tilde\gamma}(\tilde v):[a,b]\to\mathbb{R}^n$ the "coordinate representation" of $\frac{\mathrm{D}v}{\mathrm{d}t}$, i.e., we set:

$$\widetilde{\mathrm{D}}_{\tilde\gamma}(\tilde v)(t)=\mathrm{d}\varphi_t\big(\gamma(t)\big)\cdot\frac{\mathrm{D}v}{\mathrm{d}t}(t),\quad t\in[a,b];$$

taking the covariant derivative of (5.2.6) with respect to t we get:

$$\widetilde{\mathrm{D}}_{\tilde\gamma}(\tilde v)(t)=\tilde v'(t)+A_{(t,\tilde\gamma(t))}\big(\tilde v(t)\big)+\Gamma_{(t,\tilde\gamma(t))}\big(\tilde{\mathrm{d}}(\tilde\gamma)(t),\tilde v(t)\big),\quad t\in[a,b].\tag{5.2.7}$$

Finally, we can write the representation of (5.2.1) with respect to the local chart $\mathfrak{H}[\varphi]$ as:

$$\langle\!\langle\tilde v,\tilde w\rangle\!\rangle_{\tilde\gamma}=\tilde g_{(a,\tilde\gamma(a))}\big(\tilde v(a),\tilde w(a)\big)+\int_a^b\tilde g_{(t,\tilde\gamma(t))}\big(\widetilde{\mathrm{D}}_{\tilde\gamma}(\tilde v)(t),\widetilde{\mathrm{D}}_{\tilde\gamma}(\tilde w)(t)\big)\,\mathrm{d}t,\tag{5.2.8}$$

for all $\tilde v,\tilde w\in H^1\big([a,b],\mathbb{R}^n\big)$, $\tilde\gamma\in\mathfrak{H}[\widetilde U]$. It is easy to see that $\tilde{\mathrm{d}}(\tilde\gamma)$ and $\widetilde{\mathrm{D}}_{\tilde\gamma}(\tilde v)$ are in $L^2\big([a,b],\mathbb{R}^n\big)$, so that (5.2.8) is a well-defined positive semi-definite symmetric bilinear form on $H^1\big([a,b],\mathbb{R}^n\big)$ for every fixed $\tilde\gamma\in\mathfrak{H}[\widetilde U]$; we claim that (5.2.8) is indeed positive definite and that it is a Hilbert space inner product in $H^1\big([a,b],\mathbb{R}^n\big)$, i.e., it defines the standard topology of $H^1\big([a,b],\mathbb{R}^n\big)$. Keeping in mind the inequalities:

$$0<\inf_{\substack{t\in[a,b]\\ \|z\|=1}}\tilde g_{(t,\tilde\gamma(t))}(z,z)\le\sup_{\substack{t\in[a,b]\\ \|z\|=1}}\tilde g_{(t,\tilde\gamma(t))}(z,z)<+\infty,$$

we see that the claim will be proved once we establish that:

$$H^1\big([a,b],\mathbb{R}^n\big)\ni\tilde v\longmapsto\Big[\big\|\tilde v(a)\big\|^2+\big\|\widetilde{\mathrm{D}}_{\tilde\gamma}(\tilde v)\big\|_{L^2}^2\Big]^{\frac12}\in\mathbb{R}\tag{5.2.9}$$

[2]It is indeed possible to give a Hilbert manifold structure to the set of all L^2-vector fields along H^1-curves in M, so that γ' would be a point of this Hilbert manifold and $\tilde\gamma\mapsto\tilde{\mathrm{d}}(\tilde\gamma)$ would actually be the coordinate representation of the operator $\gamma\mapsto\gamma'$. In order to simplify the exposition we avoid such construction so that formula (5.2.3) should be simply understood as the definition of the term "coordinate representation of γ'".

defines a norm in $H^1\big([a,b],\mathbb{R}^n\big)$ and that such norm induces the standard topology of $H^1\big([a,b],\mathbb{R}^n\big)$. We define a linear map:

$$T_{\tilde\gamma} : H^1\big([a,b],\mathbb{R}^n\big) \longrightarrow \mathbb{R}^n \oplus L^2\big([a,b],\mathbb{R}^n\big)$$

by setting:

$$T_{\tilde\gamma}(\tilde v) = \big(\tilde v(a), \widetilde{\mathrm{D}}_{\tilde\gamma}(\tilde v)\big). \tag{5.2.10}$$

It is easy to see that $T_{\tilde\gamma}$ is a continuous linear map; moreover, it follows from the standard theorem on existence and uniqueness of solutions of linear ODE's with initial data that $T_{\tilde\gamma}$ is bijective. If we endow the counter-domain of $T_{\tilde\gamma}$ with the norm:

$$\mathbb{R}^n \oplus L^2\big([a,b],\mathbb{R}^n\big) \ni (v_0,u) \longrightarrow \Big[\|v_0\|^2 + \|u\|^2_{L^2}\Big]^{\frac12} \in \mathbb{R}, \tag{5.2.11}$$

then (5.2.9) is simply the norm on $H^1\big([a,b],\mathbb{R}^n\big)$ induced by $T_{\tilde\gamma}$ from (5.2.11) (see Exercise 4.6). This proves the claim.

We now have to check that (5.2.8) defines a smooth map:

$$H^1\big([a,b],\mathbb{R}^n\big) \supset \mathfrak{H}[\widetilde U] \ni \tilde\gamma \longmapsto \langle\cdot,\cdot\rangle_{\tilde\gamma} \in \mathrm{Bil}\big(H^1([a,b],R^n)\big). \tag{5.2.12}$$

To this aim, we check the smoothness of all the objects we have introduced. First, we observe that $\tilde{\mathrm{d}}$ defines a smooth map:

$$\tilde{\mathrm{d}} : H^1\big([a,b],\mathbb{R}^n\big) \supset \mathfrak{H}[\widetilde U] \longrightarrow L^2\big([a,b],\mathbb{R}^n\big);$$

namely, this follows from formula (5.2.4), Theorem 5.1.1, the continuity of the inclusion of H^1 in L^2 and from the continuity of the *derivation operator* from H^1 to L^2. Now we must show that (5.2.7) defines a smooth map:

$$\widetilde{\mathrm{D}} : H^1\big([a,b],\mathbb{R}^n\big) \supset \mathfrak{H}[\widetilde U] \ni \tilde\gamma \longmapsto \widetilde{\mathrm{D}}_{\tilde\gamma} \in \mathrm{Lin}\big(H^1([a,b],\mathbb{R}^n), L^2([a,b],\mathbb{R}^n)\big).$$

This can be obtained by using the smoothness of $\tilde{\mathrm{d}}$, the continuity of the linear map (4.2.5) and by applying Proposition 5.1.5 to the map:

$$\widetilde U \times \mathbb{R}^n \ni (t,\tilde x, z) \longmapsto A_{(t,\tilde x)} + \Gamma_{(t,\tilde x)}(z,\cdot) \in \mathrm{Lin}(\mathbb{R}^n)$$

observing that such map is *linear with respect to* z. The final conclusion (see Exercise 4.35) can now be obtained using the smoothness of $\widetilde{\mathrm{D}}$, $\tilde{\mathrm{d}}$ and of the map:

$$\mathfrak{H}[\tilde g] : \mathfrak{H}[\widetilde U] \longrightarrow H^1\big([a,b],\mathrm{Bil}(\mathbb{R}^n)\big). \qquad \square$$

5.2.2. LEMMA. *Let (M,g) be a finite dimensional Riemannian manifold and let ∇ be an arbitrary connection on M; assume that $H^1\big([a,b],M\big)$ is endowed with the Riemannian metric defined in* (5.2.1). *Let $\varphi = (\varphi_t, U_t, \widetilde U_t)$ be a one parameter family of charts on M and let $\widetilde V \subset \widetilde U$ be an open subset of $\mathbb{R}\times\mathbb{R}^n$ such that the closure of $\widetilde V \cap \big([a,b]\times\mathbb{R}^n\big)$ is contained in $\widetilde U$ and it is compact. For a given positive real number $r>0$, we set:*

$$\widetilde{\mathcal U} = \widetilde{\mathcal U}(r,\widetilde V) = \big\{\tilde\gamma \in \mathfrak{H}[\widetilde V] : \|\tilde\gamma'\|_{L^2} < r\big\} \subset \mathfrak{H}[\widetilde U],$$

and $\mathcal U = \mathfrak{H}[\varphi]^{-1}(\widetilde{\mathcal U}) \subset \mathfrak{H}[U]$. Then $\mathcal U$ is open in $\mathfrak{H}[U]$, $\widetilde{\mathcal U}$ is open in $\mathfrak{H}[\widetilde U]$ and the chart $\mathfrak{H}[\varphi]|_{\mathcal U} : \mathcal U \to \widetilde{\mathcal U}$ is metric relating.

PROOF. We define the objects $\mathfrak{b}$, A, Γ, $\tilde{g}$, $\tilde{\mathrm{d}}$ and $\widetilde{\mathrm{D}}$ as in the proof of Proposition 5.2.1, so that (5.2.8) is the representation in the chart $\mathfrak{H}[\varphi]$ of the Riemannian metric (5.2.1). Since $\widetilde{V}\cap\big([a,b]\times\mathbb{R}^n\big)\subset\widetilde{U}$ is compact, we have:

$$0<\inf_{\substack{(t,\tilde{x})\in\widetilde{V}\cap([a,b]\times\mathbb{R}^n)\\ \|z\|=1}}\tilde{g}_{(t,\tilde{x})}(z,z)\le\sup_{\substack{(t,\tilde{x})\in\widetilde{V}\cap([a,b]\times\mathbb{R}^n)\\ \|z\|=1}}\tilde{g}_{(t,\tilde{x})}(z,z)<+\infty. \tag{5.2.13}$$

Keeping in mind the inequalities above, we see that in order to prove that $\mathfrak{H}[\varphi]|_{\mathcal{U}}$ is metric-relating, it suffices to find constants that do not depend on $\tilde{\gamma}\in\widetilde{\mathcal{U}}$ and that relate the norm defined by (5.2.9) and the norm defined by the inner product (4.2.6) (or any of the usual H^1-norms discussed in Exercise 4.17); more explicitly, we have to find $k_1,k_2>0$ with:

$$k_1\|\tilde{v}\|_{H^1}\le\Big[\big\|\tilde{v}(a)\big\|^2+\big\|\widetilde{\mathrm{D}}_{\tilde{\gamma}}(\tilde{v})\big\|_{L^2}^2\Big]^{\frac{1}{2}}\le k_2\|\tilde{v}\|_{H^1},$$

for all $\tilde{v}\in H^1\big([a,b],\mathbb{R}^n\big)$ and all $\tilde{\gamma}\in\widetilde{\mathcal{U}}$. Recalling that (5.2.9) is the norm induced from (5.2.11) by the linear isomorphism (5.2.10) we see (using Exercise 4.7) that the proof will be completed once we show that:

$$\sup_{\tilde{\gamma}\in\widetilde{\mathcal{U}}}\|T_{\tilde{\gamma}}\|<+\infty, \tag{5.2.14}$$

$$\sup_{\tilde{\gamma}\in\widetilde{\mathcal{U}}}\big\|T_{\tilde{\gamma}}^{-1}\big\|<+\infty. \tag{5.2.15}$$

The compactness of $\widetilde{V}\cap\big([a,b]\times\mathbb{R}^n\big)\subset\widetilde{U}$ yields:

$$\sup_{\substack{(t,\tilde{x})\in\widetilde{V}\\ t\in[a,b]}}\|A_{(t,\tilde{x})}\|<+\infty,\quad\sup_{\substack{(t,\tilde{x})\in\widetilde{V}\\ t\in[a,b]}}\|\mathfrak{b}_{(t,\tilde{x})}\|<+\infty,\quad\sup_{\substack{(t,\tilde{x})\in\widetilde{V}\\ t\in[a,b]}}\|\Gamma_{(t,\tilde{x})}\|<+\infty; \tag{5.2.16}$$

using (5.2.16), (5.2.4) and keeping in mind that $\|\tilde{\gamma}'\|_{L^2}$ is bounded for $\tilde{\gamma}\in\widetilde{\mathcal{U}}$ we obtain:

$$\sup_{\tilde{\gamma}\in\widetilde{\mathcal{U}}}\big\|\tilde{\mathrm{d}}(\tilde{\gamma})\big\|_{L^2}<+\infty. \tag{5.2.17}$$

For $\tilde{\gamma}\in\mathfrak{H}[\widetilde{U}]$ we define:

$$K_{\tilde{\gamma}}:[a,b]\longrightarrow\mathrm{Lin}(\mathbb{R}^n)$$

by setting:

$$K_{\tilde{\gamma}}(t)=A_{(t,\tilde{\gamma}(t))}+\Gamma_{(t,\tilde{\gamma}(t))}\big(\tilde{\mathrm{d}}(\tilde{\gamma})(t),\cdot\big),$$

for all $t\in[a,b]$; observe that (recall (5.2.7)):

$$\widetilde{\mathrm{D}}_{\tilde{\gamma}}(\tilde{v})(t)=\tilde{v}'(t)+K_{\tilde{\gamma}}(t)\cdot\tilde{v}(t),$$

for all $\tilde{v}\in H^1\big([a,b],\mathbb{R}^n\big)$ and $t\in[a,b]$. Using (5.2.16) and (5.2.17) we obtain:

$$\sup_{\tilde{\gamma}\in\widetilde{\mathcal{U}}}\|K_{\tilde{\gamma}}\|_{L^2}<+\infty. \tag{5.2.18}$$

Inequality (5.2.14) is now a direct consequence of (5.2.18). The proof of inequality (5.2.15) is a bit more involved and it requires some basic results from the theory of linear differential ODE's.

Pick $v_0 \in \mathbb{R}^n$, $u \in L^2\big([a,b],\mathbb{R}^n\big)$ with $\|v_0\| \le 1$, $\|u\|_{L^2} \le 1$ and set $\tilde v = T_{\tilde\gamma}(v_0,u)$; this means that $\tilde v$ is a solution of the linear differential equation:

$$\tilde v'(t) = -K_{\tilde\gamma}(t)\cdot \tilde v(t) + u(t), \quad t \in [a,b], \tag{5.2.19}$$

satisfying the initial condition $\tilde v(a) = v_0$. We have to find an upper bound for $\|\tilde v\|_{H^1}$ which does not depend on $\tilde\gamma \in \widetilde{\mathcal{U}}$. The nonhomogeneous equation (5.2.19) can be solved using the method of variation of constants which yields:

$$\tilde v(t) = \Phi_{\tilde\gamma}(t)\Big[v_0 + \int_a^t \Phi_{\tilde\gamma}(s)^{-1}\cdot u(s)\,\mathrm{d}s\Big], \quad t \in [a,b], \tag{5.2.20}$$

where $\Phi_{\tilde\gamma} : [a,b] \to \mathrm{Lin}(\mathbb{R}^n)$ is defined by the matrix differential equation:

$$\Phi_{\tilde\gamma}'(t) = -K_{\tilde\gamma}(t)\Phi_{\tilde\gamma}(t), \quad t \in [a,b], \tag{5.2.21}$$

and by the initial condition $\Phi_{\tilde\gamma}(a) = \mathrm{Id}$. Since $\big\|\tilde v(a)\big\| \le 1$, in order to find an upper bound for $\|\tilde v\|_{H^1}$ it is sufficient to find an upper bound for $\|\tilde v'\|_{L^2}$; using (5.2.19), (5.2.18) and the fact that $\|u\|_{L^2} \le 1$, we see that an upper bound for $\|\tilde v'\|_{L^2}$ is easily obtained from an upper bound for $\|\tilde v\|_{C^0}$. Now (5.2.20) implies:

$$\big\|\tilde v(t)\big\| \le \|\Phi_{\tilde\gamma}\|_{C^0}\Big[1 + \big\|\Phi_{\tilde\gamma}^{-1}\big\|_{C^0}\sqrt{b-a}\Big];$$

the proof will then be completed once we show that:

$$\sup_{\tilde\gamma\in\widetilde{\mathcal{U}}} \|\Phi_{\tilde\gamma}\|_{C^0} < +\infty, \tag{5.2.22}$$

$$\sup_{\tilde\gamma\in\widetilde{\mathcal{U}}} \big\|\Phi_{\tilde\gamma}^{-1}\big\|_{C^0} < +\infty. \tag{5.2.23}$$

The proof of (5.2.22) will be obtained now using Gronwall's inequality (see Exercise 4.11); the proof of (5.2.23) can be obtained with a similar argument observing that $\Phi_{\tilde\gamma}^{-1}$ satisfies the linear ODE:

$$\big(\Phi_{\tilde\gamma}^{-1}\big)'(t) = K_{\tilde\gamma}(t)\Phi_{\tilde\gamma}(t)^{-1}, \quad t \in [a,b].$$

We start by rewriting (5.2.21) in integral form obtaining:

$$\Phi_{\tilde\gamma}(t) = \mathrm{Id} - \int_a^t K_{\tilde\gamma}(s)\Phi_{\tilde\gamma}(s)\,\mathrm{d}s, \quad t \in [a,b];$$

hence:

$$\big\|\Phi_{\tilde\gamma}(t)\big\| \le 1 + \int_a^t \big\|K_{\tilde\gamma}(s)\big\|\,\big\|\Phi_{\tilde\gamma}(s)\big\|\,\mathrm{d}s, \quad t \in [a,b].$$

Using (4.7.2) with $\delta(t) = \big\|\Phi_{\tilde\gamma}(t)\big\|$, $\phi(t) = \big\|K_{\tilde\gamma}(t)\big\|$ and $c = 1$ we obtain:

$$\|\Phi_{\tilde\gamma}\|_{C^0} \le \exp\big(\|K_{\tilde\gamma}\|_{L^1}\big);$$

since $\|K_{\tilde\gamma}\|_{L^1} \le \sqrt{b-a}\,\|K_{\tilde\gamma}\|_{L^2}$, the conclusion follows from (5.2.18). $\square$

5.2.3. DEFINITION. For a finite dimensional Riemannian manifold (M,g), the *energy functional* $E : H^1\big([a,b],M\big) \to \mathbb{R}$ is defined by:

$$E(\gamma) = \frac{1}{2}\int_a^b \big\|\gamma'(t)\big\|^2\,\mathrm{d}t.$$

5.2.4. LEMMA. *The energy functional E is smooth and its differential is given by:*

$$\mathrm{d}E_\gamma(v) = \int_a^b \big\langle \gamma'(t), \tfrac{\mathrm{D}v}{\mathrm{d}t}(t)\big\rangle\,\mathrm{d}t. \tag{5.2.24}$$

PROOF. In the notation of the proof of Proposition 5.2.1 we see that the representation of E with respect to the chart $\mathfrak{H}[\varphi]$ is given by:

$$\widetilde{E}(\tilde\gamma) = \frac{1}{2}\int_a^b \tilde g_{(t,\tilde\gamma(t))}\big(\tilde{\mathrm{d}}(\tilde\gamma)(t),\tilde{\mathrm{d}}(\tilde\gamma)(t)\big)\,\mathrm{d}t, \tag{5.2.25}$$

for every $\tilde\gamma \in \mathfrak{H}[\widetilde{U}]$, where $\widetilde{E} = E\circ\mathfrak{H}[\varphi]^{-1}$. The smoothness of $\widetilde{E}$ (and hence of E) now follows from the smoothness of $\tilde d$ (established in the proof of Proposition 5.2.1) and from the smoothness of $\mathfrak{H}[\tilde g]$ (see Theorem 5.1.1), using arguments similar to (actually simpler than) the ones used in Exercise 4.35.

Since $\mathrm{d}E : TH^1\big([a,b],M\big) \to \mathbb{R}$ is continuous (actually, it is smooth) and (by arguments similar to those used to establish the smoothness of E above) the righthand side of (5.2.24) defines a continuous (actually smooth) map on the tangent bundle $TH^1\big([a,b],M\big)$, it follows from Proposition 5.1.14 that it suffices to check equality (5.2.24) when v (and hence γ) is smooth. Let then $\gamma : [a,b] \to M$ be a smooth curve and $v : [a,b] \to TM$ a smooth vector field along γ. There exists a smooth map $]-\varepsilon,\varepsilon[\times[a,b] \ni (s,t) \mapsto H(s,t) \in M$ such that $H(0,t) = \gamma(t)$ and $\frac{\partial H}{\partial s}(0,t) = v(t)$ for all $t \in [a,b]$. Writing $\gamma_s = H(s,\cdot)$ then $]-\varepsilon,\varepsilon[\ni s \mapsto \gamma_s \in H^1\big([a,b],M\big)$ is a smooth curve with $\frac{\mathrm{d}}{\mathrm{d}s}\gamma_s\big|_{s=0} = v$ (see Corollary 5.1.13). We have:

$$\mathrm{d}E_\gamma(v) = \frac{\mathrm{d}}{\mathrm{d}s}E(\gamma_s)\bigg|_{s=0};$$

now a simple computation shows that the righthand side of the formula above equals the righthand side of (5.2.24). □

5.2.5. COROLLARY. *For every $\gamma,\mu \in H^1\big([a,b],M\big)$, we have:*

$$\big|\sqrt{E(\gamma)} - \sqrt{E(\mu)}\,\big| \le \frac{1}{\sqrt{2}}\,\mathrm{dist}(\gamma,\mu),$$

where dist *denotes the distance function on $H^1\big([a,b],M\big)$ corresponding to the Riemannian metric* (5.2.1).

PROOF. From (5.2.24) it follows that:

$$\begin{aligned}\big\|\mathrm{d}E_\gamma(v)\big\| &\le \int_a^b \big\|\gamma'(t)\big\|\,\big\|\tfrac{\mathrm{D}v}{\mathrm{d}t}(t)\big\|\,\mathrm{d}t \le \Big(\int_a^b \big\|\gamma'(t)\big\|^2\,\mathrm{d}t\Big)^{\frac12}\Big(\int_a^b \big\|\tfrac{\mathrm{D}v}{\mathrm{d}t}(t)\big\|^2\,\mathrm{d}t\Big)^{\frac12}\\ &\le \sqrt{2E(\gamma)}\,\langle v,v\rangle_\gamma^{\frac12};\end{aligned}$$

the conclusion is obtained from the result of Exercise 4.23. □

Given continuous curves $\gamma, \mu : [a,b] \to M$ on the Riemannian manifold (M,g) we set:

$$\mathrm{dist}_{C^0}(\gamma,\mu) = \sup_{t\in[a,b]} \mathrm{dist}\big(\gamma(t),\mu(t)\big).$$

5.2.6. LEMMA. *For every $\gamma, \mu \in H^1\big([a,b],M\big)$ we have:*

$$\mathrm{dist}_{C^0}(\gamma,\mu) \le \sqrt{2}\max\big\{1,\sqrt{b-a}\,\big\}\mathrm{dist}(\gamma,\mu),$$

where dist *denotes the distance function on $H^1\big([a,b],M\big)$ corresponding to the Riemannian metric* (5.2.1).

PROOF. We have to show that for every fixed $t \in [a,b]$ we have:

$$\mathrm{dist}\big(\gamma(t),\mu(t)\big) \le \sqrt{2}\max\big\{1,\sqrt{b-a}\,\big\}\,\mathrm{dist}(\gamma,\mu).$$

This will follow from the result of Exercise 4.24 once we show that:

$$\big\|\mathrm{dEval}_t(\gamma)\big\| \le \sqrt{2}\max\big\{1,\sqrt{b-a}\,\big\},$$

for all $\gamma \in H^1\big([a,b],M\big)$, where $\mathrm{Eval}_t : H^1\big([a,b],M\big) \to M$ denotes the map $\gamma \mapsto \gamma(t)$ of evaluation at the instant t. Let $\gamma \in H^1\big([a,b],M\big)$ and $v \in T_\gamma H^1\big([a,b],M\big)$ be fixed; we want to show that:

$$\big\|v(t)\big\|^2 \le 2\max\{1,b-a\}\,\langle v,v\rangle_\gamma.$$

To this aim, let $X : [a,b] \to TM$ be a parallel vector field along γ with $X(t) = v(t)$; since the metric of M is parallel, we have $\big\|X(s)\big\| = \big\|v(t)\big\|$ for all $s \in [a,b]$. Moreover:

$$\big\|v(t)\big\|^2 = \big\langle v(t),X(t)\big\rangle = \big\langle v(a),X(a)\big\rangle + \int_a^t \big\langle \tfrac{\mathrm{D}v}{\mathrm{d}s}(s),X(s)\big\rangle\,\mathrm{d}s;$$

now we compute:

$$\big\|v(t)\big\|^2 \le \big\|v(a)\big\|\,\big\|v(t)\big\| + \big\|v(t)\big\|\int_a^b \big\|\tfrac{\mathrm{D}v}{\mathrm{d}s}(s)\big\|\,\mathrm{d}s.$$

Therefore:

$$\big\|v(t)\big\| \le \|v(a)\| + \int_a^b \big\|\tfrac{\mathrm{D}v}{\mathrm{d}s}(s)\big\|\,\mathrm{d}s,$$

which implies (see Exercise 4.25):

$$\big\|v(t)\big\|^2 \le 2\big\|v(a)\big\|^2 + 2(b-a)\int_a^b \big\|\tfrac{\mathrm{D}v}{\mathrm{d}s}(s)\big\|^2\,\mathrm{d}s \le 2\max\{1,b-a\}\,\langle v,v\rangle_\gamma.$$

This concludes the proof. □

5.2.7. THEOREM. *Let (M,g) be a finite dimensional Riemannian manifold and consider the Hilbert manifold $H^1\big([a,b],M\big)$ endowed with the Riemannian metric* (5.2.1)*, where $\frac{\mathrm{D}}{\mathrm{d}t}$ denotes covariant derivative with respect to the Levi–Civita connection. If M is complete then $H^1\big([a,b],M\big)$ is also complete.*

PROOF. Let $(\gamma_k)_{k\ge1}$ be a Cauchy sequence in $H^1\big([a,b],M\big)$. Lemma 5.2.6 implies that $(\gamma_k)_{k\ge1}$ is also a Cauchy sequence for the metric dist_{C^0} on the space of continuous curves in M; since dist_{C^0} is complete, we conclude that $(\gamma_k)_{k\ge1}$ converges with respect to dist_{C^0} (i.e., converges uniformly) to some continuous curve $\gamma:[a,b]\to M$. Observe that Corollary 5.2.5 also implies that:

$$\sup_{k\ge1}\big|E(\gamma_k)\big|<+\infty. \tag{5.2.26}$$

By Proposition 5.1.8 we can find a one parameter family of charts $\varphi=(\varphi_t,U_t,\widetilde U_t)$ such that U contains the graph of γ; define $\tilde\gamma:[a,b]\to\mathbb{R}^n$ by:

$$\tilde\gamma(t)=\varphi\big(t,\gamma(t)\big),\quad t\in[a,b].$$

Since $(\gamma_k)_{\ge1}$ converges to γ with respect to the metric dist_{C^0}, it follows that the graph of γ_k is contained in U for k sufficiently large; for such k we can set $\tilde\gamma_k=\mathfrak{H}[\varphi](\gamma_k)$, so that $(\tilde\gamma_k)_{k\ge1}$ converges uniformly to $\tilde\gamma$. Choose $R>0$ such that $\mathrm{B}\big[\tilde\gamma(t);R\big]\subset\widetilde U$ for all $t\in[a,b]$ and set:

$$\widetilde V=\big\{(t,\tilde x)\in\mathbb{R}\times\mathbb{R}^n:\big\|\tilde\gamma(t)-\tilde x\big\|<R\big\}\cap\widetilde U;$$

in the formula above we have considered an arbitrary continuous extension of $\tilde\gamma$ to the whole line $\mathbb{R}$. It is easy to see that $\widetilde V$ is an open subset of $\widetilde U$ and that the closure of $\widetilde V\cap\big([a,b],\times\mathbb{R}^n\big)$ is compact and contained in $\widetilde U$. Since $\tilde\gamma_k\to\tilde\gamma$ uniformly, it follows that $\|\tilde\gamma_k-\tilde\gamma\|_{C^0}<R$ for k sufficiently large, so that the graph of $\tilde\gamma_k$ is contained in $\widetilde V$ for such k. For the rest of the proof we assume that some initial portion of the original sequence $(\gamma_k)_{k\ge1}$ was deleted, so that $\tilde\gamma_k$ is (well-defined and) has its graph contained in $\widetilde V$ for all k. Keeping in mind formulas (5.2.25), (5.2.4), (5.2.13) and (5.2.16), it follows from (5.2.26) that:

$$\sup_{k\ge1}\big\|\tilde\gamma_k'\big\|_{L^2}=r<+\infty. \tag{5.2.27}$$

By Lemma 5.2.2, if we set:

$$\widetilde{\mathcal U}=\widetilde{\mathcal U}(3r,\widetilde V),\quad \mathcal U=\mathfrak{H}[\varphi]^{-1}(\widetilde{\mathcal U})$$

then the chart $\mathfrak{H}[\varphi]|_{\mathcal U}:\mathcal U\to\widetilde{\mathcal U}$ is metric relating. Consider the closed subset $F\subset H^1\big([a,b]\mathbb{R}^n\big)$ defined by:

$$F=\big\{\tilde\mu\in H^1\big([a,b],\mathbb{R}^n\big):\|\tilde\mu-\tilde\gamma\|_{C^0}\le\tfrac{R}{2},\quad\|\tilde\mu'\|_{L^2}\le2r\big\};$$

obviously, $F\subset\widetilde{\mathcal U}$ and $\tilde\gamma_k\in F$ for all k sufficiently large. By Corollary 4.3.19, the proof will be concluded if we manage to find $k_0\in\mathbb{N}$ such that:

$$\inf_{k\ge k_0}\mathrm{dist}\big(\tilde\gamma_k,\partial F\big)>0. \tag{5.2.28}$$

Observe that (5.2.28) is equivalent to:

$$\inf_{\substack{\tilde\mu\in\partial F\\k\ge k_0}}\|\tilde\gamma_k-\tilde\mu\|_{C^0}+\|\tilde\gamma_k'-\tilde\mu'\|_{L^2}>0. \tag{5.2.29}$$

Finally, (5.2.29) follows from $\|\tilde\gamma_k-\tilde\gamma\|_{C^0}\to0$ and (5.2.27) by observing that $\tilde\mu\in\partial F$ implies either $\|\tilde\gamma-\tilde\mu\|_{C^0}=\frac{R}{2}$ or $\|\tilde\mu'\|_{L^2}=2r$. $\square$

5.3. Morse Theory for Riemannian Geodesics

5.3.1. LEMMA. *For every $\gamma \in H^1\big([a,b],M\big)$ we have:*

$$L(\gamma) \le \sqrt{2(b-a)E(\gamma)}.$$

PROOF. It is an immediate consequence of the Cauchy-Schwarz inequality. □

Let now $p,q \in M$ be fixed and consider the set:

$$H^1_{pq}\big([a,b],M\big) = \big\{\gamma \in H^1\big([a,b],M\big) : \gamma(a) = p,\ \gamma(b) = q\big\};$$

it follows from the result of Exercise 4.36 that $H^1\big([a,b],M\big)$ is a smooth Hilbert submanifold of $H^1\big([a,b],M\big)$ and that its tangent space is given by:

$$T_\gamma H^1_{pq}\big([a,b],M\big) = \big\{v \in T_\gamma H^1\big([a,b],M\big) : v(a) = v(b) = 0\big\},$$

for all $\gamma \in H^1_{pq}\big([a,b],M\big)$. Obviously the Riemannian metric (5.2.1) restricts to a Riemannian metric in $H^1_{pq}\big([a,b],M\big)$ given by:

$$\langle v,w\rangle_\gamma = \int_a^b g\big(\tfrac{\mathrm{D}v}{\mathrm{d}t},\tfrac{\mathrm{D}w}{\mathrm{d}t}\big)\,\mathrm{d}t, \quad v,w \in T_\gamma H^1\big([a,b],M\big), \tag{5.3.1}$$

for all $\gamma \in H^1_{pq}\big([a,b],M\big)$.

5.3.2. COROLLARY. *Let $(\gamma_k)_{k\ge1}$ be a sequence in $H^1\big([a,b],M\big)$ on which E is bounded. If for some $t_0 \in [a,b]$ the set $\big\{\gamma_k(t_0) : k \ge 1\big\}$ is bounded in M then the set $\big\{\gamma_k(t) : k \ge 1\big\}$ is bounded in M for all $t \in [a,b]$.*

PROOF. Since $\sup_{k\ge1} E(\gamma_k) < +\infty$, Lemma 5.3.1 implies that:

$$\sup_{k\ge1} L(\gamma_k) < +\infty;$$

the conclusion follows by observing that:

$$\mathrm{dist}\big(\gamma_k(t),\gamma_k(t_0)\big) \le L(\gamma_k),$$

for all k. □

5.3.3. COROLLARY. *If $(\gamma_k)_{k\ge1}$ is a sequence in $H^1\big([a,b],M\big)$ on which E is bounded then the set $\big\{\gamma_k : k \ge 1\big\}$ is equicontinuous.*

PROOF. This follows by observing that, for all $t,s \in [a,b]$ and all $k \ge 1$:

$$\mathrm{dist}\big(\gamma_k(t),\gamma_k(s)\big) \le L\big(\gamma_k|_{[t,s]}\big) \le \sqrt{2|t-s|E(\gamma_k)}.$$

□

5.3.4. PROPOSITION. *Let (M,g) be a finite dimensional Riemannian manifold and consider the Hilbert manifold $H^1_{pq}\big([a,b],M\big)$ endowed with the Riemannian metric* (5.3.1)*, where $\frac{\mathrm{D}}{\mathrm{d}t}$ denotes covariant derivative with respect to the Levi–Civita connection. If M is complete then the energy functional $E : H^1\big([a,b],M\big) \to \mathbb{R}$ satisfies the Palais–Smale condition.*

PROOF. Let $(\gamma_k)_{k\ge1}$ be a Palais–Smale sequence for E. Since $\gamma_k(a)=p$ for all k, it follows from Corollary 5.3.2 that the set $\big\{\gamma_k(t):k\ge1\big\}$ is bounded in M for all $t\in[a,b]$; since M is complete, $\big\{\gamma_k(t):k\ge1\big\}$ is relatively compact. Moreover, by Corollary 5.3.3, the set $\big\{\gamma_k:k\ge1\big\}$ is equicontinuous, so that we can apply Arzelá-Ascoli's theorem to conclude that (up to a subsequence) $(\gamma_k)_{k\ge1}$ converges uniformly to some continuous curve $\gamma:[a,b]\to M$. Let $\varphi=(\varphi_t,U_t,\widetilde U_t)$ be a one parameter family of charts such that the graph of γ is contained in U (see Proposition 5.1.8). For k sufficiently large, the graph of γ_k will be contained in U, so that it makes sense to define $\tilde\gamma_k=\mathfrak{H}[\varphi](\gamma_k)$; we define also $\tilde\gamma:[a,b]\to\mathbb{R}^n$ by:

$$\tilde\gamma(t)=\varphi\big(t,\gamma(t)\big),\quad t\in[a,b],$$

so that $(\tilde\gamma_k)_{k\ge1}$ converges uniformly to $\tilde\gamma$ (we delete some initial part of the sequence $(\gamma_k)_{k\ge1}$, if necessary). Let V be an open neighborhood of the graph of γ whose closure is compact and contained in U; set $\varphi^\diamond(V)=\widetilde V$. We now consider the objects A, Γ, $\mathfrak{b}$, $\tilde g$, $\tilde{\mathrm{d}}$ and $\widetilde{\mathrm{D}}$ defined in the proof of Proposition 5.2.1; since V is relatively compact in U, we have the estimates (5.2.13) and (5.2.16). As in the proof of Theorem 5.2.7, the fact that E is bounded on $(\gamma_k)_{k\ge1}$ implies that $(\tilde\gamma_k')_{k\ge1}$ is bounded in L^2 (see (5.2.27)); thus, since $(\tilde\gamma_k)_{k\ge1}$ is bounded in C^0, the sequence $(\tilde\gamma_k)_{k\ge1}$ is bounded in H^1. By passing to a subsequence, we may assume that $(\tilde\gamma_k)_{k\ge1}$ converges weakly in H^1 (necessarily to $\tilde\gamma$); in particular, $\tilde\gamma\in H^1\big([a,b],\mathbb{R}^n\big)$. For each k, we set $\tilde v_k=\tilde\gamma_k-\tilde\gamma$ and $v_k=\mathrm{d}\mathfrak{H}[\varphi]_{\gamma_k}^{-1}(\tilde v_k)$; we have $\tilde v_k(a)=\tilde v_k(b)=0$, which implies $v_k\in T_{\gamma_k}H^1_{pq}\big([a,b],M\big)$. Since $(\tilde v_k)_{k\ge1}$ converges uniformly to zero, it follows from (5.2.13) that:

$$\langle v_k,v_k\rangle_{\gamma_k}=\int_a^b\tilde g_{(t,\tilde\gamma_k(t))}\big(\tilde v_k(t),\tilde v_k(t)\big)\,\mathrm{d}t\xrightarrow{k\to+\infty}0.$$

Since also $\big\|\mathrm{d}E(\gamma_k)\big\|\to0$, we have:

$$\mathrm{d}E_{\gamma_k}\cdot v_k=\int_a^b\tilde g_{(t,\tilde\gamma_k(t))}\big(\tilde{\mathrm{d}}(\tilde\gamma_k)(t),\widetilde{\mathrm{D}}_{\tilde\gamma_k}(\tilde\gamma_k-\tilde\gamma)(t)\big)\,\mathrm{d}t\xrightarrow{k\to+\infty}0. \tag{5.3.2}$$

Using (5.2.16), the fact that $(\tilde\gamma_k')_{k\ge1}$ is bounded in L^2 and the fact that $(\tilde\gamma_k)_{k\ge1}$ converges uniformly to $\tilde\gamma$ we get that:

$$\sup_{k\ge1}\big\|\tilde{\mathrm{d}}(\tilde\gamma_k)\big\|_{L^2}<+\infty,$$

and that $\widetilde{\mathrm{D}}_{\tilde\gamma_k}(\tilde\gamma_k-\tilde\gamma)$ can be written as:

$$\widetilde{\mathrm{D}}_{\tilde\gamma_k}(\tilde\gamma_k-\tilde\gamma)=\tilde\gamma_k'-\tilde\gamma'+u_k, \tag{5.3.3}$$

where:

$$u_k\xrightarrow{k\to+\infty}0\ \text{ in }\ L^2\big([a,b],\mathbb{R}^n\big). \tag{5.3.4}$$

From (5.3.2), (5.3.3), (5.3.4) and (5.2.13), we get:

$$\int_a^b\tilde g_{(t,\tilde\gamma_k(t))}\big(\tilde{\mathrm{d}}(\tilde\gamma_k)(t),\tilde\gamma_k'(t)-\tilde\gamma'(t)\big)\,\mathrm{d}t\xrightarrow{k\to+\infty}0. \tag{5.3.5}$$

For each k we define a linear functional $\alpha_k \in L^2\big([a,b],\mathbb{R}^n\big)^*$ by setting:

$$\alpha_k(z)(t) = \int_a^b \tilde{g}_{(t,\tilde{\gamma}_k(t))}\big[\mathfrak{b}\big(t,\tilde{\gamma}_k(t)\big), z(t)\big]\,\mathrm{d}t, \quad z \in L^2\big([a,b],\mathbb{R}^n\big);$$

since $(\tilde{\gamma}_k)_{k\geq 1}$ converges uniformly to $\tilde{\gamma}$, the sequence of linear functionals $(\alpha_k)_{k\geq 1}$ converges in $L^2\big([a,b],\mathbb{R}^n\big)^*$ to the linear functional:

$$\alpha(z)(t) = \int_a^b \tilde{g}_{(t,\tilde{\gamma}(t))}\big[\mathfrak{b}\big(t,\tilde{\gamma}(t)\big), z(t)\big]\,\mathrm{d}t, \quad z \in L^2\big([a,b],\mathbb{R}^n\big).$$

Thus $\alpha_k \to \alpha$ in $L^2\big([a,b],\mathbb{R}^n\big)^*$ and $\tilde{\gamma}_k' - \tilde{\gamma}' \to 0$ weakly in $L^2\big([a,b],\mathbb{R}^n\big)$; this implies:

$$\int_a^b \tilde{g}_{(t,\tilde{\gamma}_k(t))}\big[\mathfrak{b}\big(t,\tilde{\gamma}_k(t)\big), \tilde{\gamma}_k'(t) - \tilde{\gamma}'(t)\big]\,\mathrm{d}t \xrightarrow{k\to+\infty} 0. \tag{5.3.6}$$

From (5.3.5) and (5.3.6) we get:

$$\int_a^b \tilde{g}_{(t,\tilde{\gamma}_k(t))}\big(\tilde{\gamma}_k'(t), \tilde{\gamma}_k'(t) - \tilde{\gamma}'(t)\big)\,\mathrm{d}t \xrightarrow{k\to+\infty} 0. \tag{5.3.7}$$

Since $t \mapsto \tilde{g}_{(t,\tilde{\gamma}_k(t))}$ converges uniformly to $t \mapsto \tilde{g}(t,\tilde{\gamma}(t))$ and $(\tilde{\gamma}_k')_{k\geq 1}$ is bounded in L^2, it follows from (5.3.7) that:

$$\int_a^b \tilde{g}_{(t,\tilde{\gamma}(t))}\big(\tilde{\gamma}_k'(t), \tilde{\gamma}_k'(t) - \tilde{\gamma}'(t)\big)\,\mathrm{d}t \xrightarrow{k\to+\infty} 0. \tag{5.3.8}$$

Since $\tilde{\gamma}_k' - \tilde{\gamma}' \to 0$ weakly in $L^2\big([a,b],\mathbb{R}^n\big)$ and

$$L^2\big([a,b],\mathbb{R}^n\big) \ni z \longmapsto \int_a^b \tilde{g}_{(t,\tilde{\gamma}(t))}\big(\tilde{\gamma}'(t), z(t)\big)\,\mathrm{d}t \in \mathbb{R}$$

is a continuous linear functional, it follows that:

$$\int_a^b \tilde{g}_{(t,\tilde{\gamma}(t))}\big(\tilde{\gamma}'(t), \tilde{\gamma}_k'(t) - \tilde{\gamma}'(t)\big)\,\mathrm{d}t \xrightarrow{k\to+\infty} 0. \tag{5.3.9}$$

Finally, from (5.3.8), (5.3.9) and (5.2.13) we get:

$$\|\tilde{\gamma}_k' - \tilde{\gamma}'\|_{L^2} \xrightarrow{k\to+\infty} 0;$$

thus $\tilde{\gamma}_k \to \tilde{\gamma}$ in $H^1\big([a,b],\mathbb{R}^n\big)$ and the proof is completed. $\square$

We now recall the statement of the Morse Index Theorem. We will first need a few definitions.

5.3.5. DEFINITION. Let (M,g) be a Riemannian manifold. If $I \subset \mathbb{R}$ is an interval then a *geodesic* $\gamma : I \to M$ is a smooth curve satisfying the equation:

$$\frac{\mathrm{D}}{\mathrm{d}t}\gamma'(t) = 0, \quad t \in I.$$

A smooth vector field J along γ is called a *Jacobi field* if it satisfies the equation:

$$\frac{\mathrm{D}^2}{\mathrm{d}t^2}J(t) = R\big(\gamma'(t), J(t)\big)\gamma'(t), \quad t \in I.$$

Given a geodesic $\gamma : [a,b] \to M$ then an instant $t \in \,]a,b]$ is called *conjugate* for γ if there exists a non zero Jacobi field J along γ with $J(a) = J(t) = 0$; the *multiplicity* of t as a conjugate instant along γ, denoted $\mathrm{mul}(t)$, equals the dimension of the space of all Jacobi fields J along γ with $J(a) = J(t) = 0$. The *geometric index* of a geodesic γ is defined as the sum of the multiplicities of the conjugate instants $t \in \,]a,b[$ along γ:

$$\text{geometric index of } \gamma = \sum_{t \in]a,b[} \mathrm{mul}(t).$$

Two points $p, q \in M$ are called *conjugate* if there exists a geodesic $\gamma : [a,b] \to M$ with $\gamma(a) = p$, $\gamma(b) = q$ and such that $t = b$ is a conjugate instant along γ.

We can now state the following:

5.3.6. THEOREM (Morse index theorem). *If (M,g) is a finite-dimensional Riemannian manifold and $p, q \in M$ are fixed points then the critical points of the energy functional $E : H^1\big([a,b], M\big) \to \mathbb{R}$ are precisely the geodesics $\gamma : [a,b] \to M$ with $\gamma(a) = p$, $\gamma(b) = q$. The Hessian of E at γ is given by the so called* index form*:*

$$\mathrm{Hess}E_\gamma(v,w) = \int_a^b g\big(\tfrac{\mathrm{D}v}{\mathrm{d}t}, \tfrac{\mathrm{D}w}{\mathrm{d}t}\big) + g\big(R\big(\gamma'(t), v(t)\big)\gamma'(t), w(t)\big)\,\mathrm{d}t,$$

for all $v, w \in T_\gamma H^1_{pq}\big([a,b], M\big)$. The kernel of the index form equals the space of Jacobi fields J along γ with $J(a) = J(b) = 0$; in particular, γ is a nondegenerate critical point of E iff $t = b$ is not a conjugate instant along γ. Moreover, all nondegenerate critical points of E are strongly nondegenerate and the Morse index of a critical point γ equals the geometric index of the geodesic γ.

PROOF. See [**98**]. □

In order to present some applications of Morse theory for counting Riemannian geodesics connecting to fixed points, we state without proof the following results.

Recall that the *loop space* of a topological space X at the base point $x_0 \in X$ is defined by:

$$\Omega(X; x_0) = \big\{\gamma : [0,1] \to X : \gamma \text{ is continuous and } \gamma(0) = \gamma(1) = x_0\big\};$$

the space $\Omega(X; x_0)$ is always assumed to be endowed with the compact-open topology.

5.3.7. THEOREM. *If M is a connected finite-dimensional differentiable manifold and if $p, q \in M$ are arbitrary fixed points then $H^1_{pq}\big([0,1], M\big)$ has the same homotopy type as $\Omega(M; x_0)$, for any $x_0 \in M$.*

PROOF. See [**98**, §17]. □

The following is a very deep result relating the singular homology of a space with the singular homology of its loop space.

5.3.8. THEOREM. *Let X be a simply-connected and arc-connected topological space and let $\mathbb{K}$ be a field. Assume that a point $x_0 \in X$ is fixed. If for some $n \geq 2$ we have $H_n(X;\mathbb{K}) \neq 0$ and $H_i(X;\mathbb{K}) = 0$ for all $i > n$ then the singular homology of $\Omega(X;x_0)$ satisfies the following property: for every integer $i \geq 0$ there exists an integer j, $0 < j < n$, such that $H_{i+j}\big(\Omega(X;x_0);\mathbb{K}\big) \neq 0$.*

PROOF. See [**141**, Proposition 11, pg. 483]. □

5.3.9. COROLLARY. *Under the assumptions of Theorem 5.3.8, the loop space $\Omega(X;x_0)$ has infinitely many non zero Betti numbers with respect to the field $\mathbb{K}$.* □

Now using Theorem 5.3.7 and the theory developed in Sections **??** and **??** we obtain readily the following:

5.3.10. THEOREM (Morse relations for Riemannian geodesics). *Let (M,g) be a complete connected finite-dimensional Riemannian manifold. Assume that the points $p,q \in M$ are non conjugate. For every integer $k \geq 0$, denote by κ_k the number of geodesics $\gamma : [0,1] \to M$ from p to q having geometric index equal to k. If $\mathbb{K}$ is an arbitrary field and if $\mathfrak{P}_\lambda\big(\Omega(M;x_0);\mathbb{K}\big)$ denotes the Poincaré polynomial of the loop space $\Omega(M;x_0)$ (with an arbitrary base point $x_0 \in M$) with respect to the field $\mathbb{K}$ then there exists a formal power series $Q(\lambda)$ with coefficients in $\mathbb{N} \cup \{+\infty\}$ such that:*

$$\sum_{k=0}^{+\infty} \kappa_k \lambda^k = \mathfrak{P}_\lambda\big(\Omega(M;x_0);\mathbb{K}\big) + (1+\lambda)Q(\lambda). \tag{5.3.10}$$

□

5.3.11. COROLLARY. *If (M,g) is a complete contractible finite-dimensional Riemannian manifold then the number of geodesics connecting two non conjugate points of M is either odd or infinite.*

PROOF. It can be shown that if M is contractible then also $\Omega(M;x_0)$ is contractible; hence:

$$P_\lambda\big(\Omega(M;x_0);\mathbb{K}\big) = 1.$$

The conclusion follows by using equality (5.3.10) with $\lambda = 1$; namely, if $\lambda = 1$ then the lefthand side of (5.3.10) becomes the total number of geodesics from p to q and the righthand side of (5.3.10) becomes $2Q(1)+1$ (which is either infinite or odd). This concludes the proof. □

5.3.12. COROLLARY. *If (M,g) is a compact Riemannian manifold then the number of geodesics connecting two non conjugate points of M is always infinite.*

PROOF. Let $p,q \in M$ be two fixed non conjugate points. It follows from (5.3.10) that the number of geodesics of index k in M from p to q is greater than or equal to the k-th Betti number of the loop space of M with coefficients in the (arbitrarily fixed) field $\mathbb{K}$. Assume that M is simply-connected. If n denotes the dimension of M then it is well-known[3] that $H_n(M;\mathbb{K}) \neq 0$ and that $H_i(M;\mathbb{K}) = 0$ for $i > n$. It follows from Corollary 5.3.9 that the loop space of M has infinitely many non zero Betti numbers with respect to the field $\mathbb{K}$ and therefore there are infinitely many geodesics connecting p and q.

We now prove the general case (with M not necessarily simply-connected). Let $\pi : \widetilde{M} \to M$ denote the universal covering of M and consider $\widetilde{M}$ endowed with the pull-back of the Riemannian metric of M by π. Choose $\tilde{p}, \tilde{q} \in \widetilde{M}$ with $\pi(\tilde{p}) = p$ and $\pi(\tilde{q}) = q$. The Riemannian manifold $\widetilde{M}$ is again complete. Moreover, since π is a local isometry, it is easy to see that $\tilde{p}$ and $\tilde{q}$ are non conjugate in $\widetilde{M}$. If the fundamental group of M is infinite then the set $\pi^{-1}(q) \subset \widetilde{M}$ is also infinite and therefore we obtain infinitely many geodesics in M connecting p and q by taking projections of the geodesics in $\widetilde{M}$ connecting $\tilde{p}$ and points of $\pi^{-1}(q)$. On the other hand, if the fundamental group of M is finite then $\widetilde{M}$ is again compact and by the first part of the proof we can find infinitely many geodesics in $\widetilde{M}$ connecting $\tilde{p}$ and $\tilde{q}$; their projections in M will provide us with an infinite set of geodesics in M connecting p and q. $\square$

[3] If M is an n-dimensional topological manifold then $H_i(M;G) = 0$ for every abelian group G and every $i > n$. Moreover, if M is orientable (which in our case follows from the simply-connectedness of M) and connected then $H_n(M;G) \cong G$. See, for instance, [**39**, Chapter VIII, §3, §4] for a proof of such results.

APPENDIX A

Thom Class and Thom Isomorphism

APPENDIX B

Hyperbolic Singularities of a Vector Field

In this appendix we present an elementary introduction to the theory of dynamical systems. We introduce the notion of hyperbolic singularity of a vector field on a manifold and we study the stable and unstable manifolds of such singularities. We also prove the theorem of Hartman–Grobman that gives a topological characterization of the flow of a vector field near a hyperbolic singularity.

Let us fix some conventions that will be used throughout the appendix. Let V be a real finite-dimensional vector space. If $A : V \to V$ is a linear endomorphism, we denote by $\sigma(A)$ the set of complex roots of the characteristic polynomial of A; this means that $\sigma(A)$ equals the set of eigenvalues of the complexification of A, which is the unique complex linear extension $A^{\mathbb{C}}$ of A to the complexification $V^{\mathbb{C}}$ of V. For $\lambda \in \sigma(A) \cap \mathbb{R}$ we write:

$$V_\lambda(A) = \bigcup_{k \geq 1} \operatorname{Ker}(A - \lambda)^k,$$

and for $\lambda \in \sigma(A) \setminus \mathbb{R}$ we write:

$$V_\lambda(A) = \Big(\bigcup_{k \geq 1} \operatorname{Ker}(A^{\mathbb{C}} - \lambda)^k \oplus \bigcup_{k \geq 1} \operatorname{Ker}(A^{\mathbb{C}} - \bar{\lambda})^k \Big) \cap V,$$

so that $V_\lambda(A) = V_{\bar{\lambda}}(A)$. The primary decomposition of A is therefore written as:

$$V = \bigoplus_{\substack{\lambda \in \sigma(A) \\ \Im(\lambda) > 0}} V_\lambda(A),$$

where $\Im(\lambda)$ denotes the imaginary part of λ. Observe that if A is symmetric with respect to some inner product of V then $\sigma(A) \subset \mathbb{R}$ and $V_\lambda(A) = \operatorname{Ker}(A - \lambda)$ is simply the λ-eigenspace of A.

We give a basic definition.

B.1. DEFINITION. Let V be a real finite-dimensional vector space. A linear endomorphism $A : V \to V$ is called *hyperbolic* if $\sigma(A)$ contains no purely imaginary complex numbers. The *positive* and the *negative* eigenspaces of A are defined respectively by:

$$V_+(A) = \sum_{\substack{\lambda \in \sigma(A) \\ \Re(\lambda) > 0}} V_\lambda(A), \quad V_-(A) = \sum_{\substack{\lambda \in \sigma(A) \\ \Re(\lambda) < 0}} V_\lambda(A),$$

where $\Re(\lambda)$ denotes the real part of λ.

Obviously if $A : V \to V$ is hyperbolic we obtain the following direct sum decomposition of V in A-invariant subspaces:

$$V = V_+(A) \oplus V_-(A).$$

We now prove a lemma concerning some estimates on the norm of the exponential of a linear map. We consider the spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ endowed with their standard Euclidean norms. The norm of a linear map $A : \mathbb{R}^n \to \mathbb{R}^n$ is defined by:

$$\|A\| = \sup_{\substack{x \in \mathbb{R}^n \\ \|x\| \le 1}} \big\|A(x)\big\|,$$

and the norm of a complex linear map $A : \mathbb{C}^n \to \mathbb{C}^n$ is defined by:

$$\|A\| = \sup_{\substack{x \in \mathbb{C}^n \\ \|x\| \le 1}} \big\|A(x)\big\|.$$

We observe that if $A^{\mathbb{C}} : \mathbb{C}^n \to \mathbb{C}^n$ denotes the complexification of $A : \mathbb{R}^n \to \mathbb{R}^n$ then $\|A\| = \|A^{\mathbb{C}}\|$ (see Exercise B.2).

B.2. LEMMA. *Let* $A : \mathbb{R}^n \to \mathbb{R}^n$ *be a linear map and choose* $\lambda_0, \lambda_1 \in \mathbb{R}$ *such that:*

$$\lambda_0 < \min_{\lambda \in \sigma(A)} \Re(\lambda) \le \max_{\lambda \in \sigma(A)} \Re(\lambda) < \lambda_1.$$

Then there exists a constant $C > 0$ *such that:*

$$\big\|e^{tA}\big\| \le C e^{t\lambda_1}, \quad \big\|e^{-tA}\big\| \le C e^{-t\lambda_0},$$

for all $t \ge 0$*.*

PROOF. Let $A = S + N$ denote the Jordan decomposition of A, i.e., S is semi-simple, N is nilpotent and $SN = NS$. Then:

$$e^{tA} = e^{tS} e^{tN}, \tag{B.1}$$

and:

$$e^{tN} = I + tN + \frac{t^2 N^2}{2} + \cdots + \frac{t^n N^n}{n!},$$

for all $t \in \mathbb{R}$. Thus:

$$\big\|e^{tN}\big\| \le 1 + t\|N\| + \frac{t^2 \|N\|^2}{2} + \cdots + \frac{t^n \|N\|^n}{n!},$$

for all $t \ge 0$ and therefore for every $\varepsilon > 0$ we can find a constant $C_0 > 0$ such that:

$$\big\|e^{tN}\big\| \le C_0 e^{\varepsilon t}, \tag{B.2}$$

for all $t \ge 0$. Now $\sigma(S) = \sigma(A)$ and $S^{\mathbb{C}} : \mathbb{C}^n \to \mathbb{C}^n$ is diagonalizable, so that we can find a complex linear isomorphism $B : \mathbb{C}^n \to \mathbb{C}^n$ such that $D = BS^{\mathbb{C}}B^{-1}$ is a diagonal matrix whose diagonal elements belong to $\sigma(A)$. We have:

$$e^{tS^{\mathbb{C}}} = B^{-1} e^{tD} B,$$

and therefore, for all $t \in \mathbb{R}$:

$$\big\|e^{tS}\big\| = \big\|e^{tS^{\mathbb{C}}}\big\| \le \|B\| \, \|B^{-1}\| \, \big\|e^{tD}\big\|. \tag{B.3}$$

Obviously:

$$\|e^{tD}\| = \max_{\lambda\in\sigma(A)} e^{t\,\Re(\lambda)}, \tag{B.4}$$

for all $t \in \mathbb{R}$. By choosing $\varepsilon > 0$ with $\Re(\lambda) + \varepsilon \le \lambda_1$ for all $\lambda \in \sigma(A)$, formulas (B.1), (B.2), (B.3) and (B.4) imply:

$$\|e^{tA}\| \le C_0\,\|B\|\,\|B^{-1}\|\,e^{t\lambda_1},$$

for all $t \ge 0$. This proves the desired estimate on $\|e^{tA}\|$. The estimate on $\|e^{-tA}\|$ is obtained by replacing A with $-A$. □

We now prove a preparatory lemma concerning linear ODE's whose coefficient matrix is hyperbolic.

B.3. LEMMA. *Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be a hyperbolic linear map and denote by $\pi_+ : \mathbb{R}^n \to \mathbb{R}^n_+(A)$, $\pi_- : \mathbb{R}^n \to \mathbb{R}^n_-(A)$ the projections corresponding to the direct sum decomposition $\mathbb{R}^n = \mathbb{R}^n_+(A) \oplus \mathbb{R}^n_-(A)$. Then for every $x_0 \in \mathbb{R}^n_-(A)$ and every continuous map $u : [0,+\infty[\; \to \mathbb{R}^n$ with $\lim_{t\to+\infty} u(t) = 0$ there exists a unique solution $x : [0,+\infty[\; \to \mathbb{R}^n$ of the nonhomogeneous linear ODE:*

$$x' = Ax + u, \tag{B.5}$$

with $\pi_-(x) = x_0$ and $\lim_{t\to+\infty} x(t) = 0$.

PROOF. Denote by A_+ and A_- respectively the endomorphisms of $\mathbb{R}^n_+(A)$ and $\mathbb{R}^n_-(A)$ given by restrictions of A and choose $\lambda_+, \lambda_- \in \mathbb{R}$ such that:

$$\max_{\lambda\in\sigma(A_-)} \Re(\lambda) < \lambda_- < 0 < \lambda_+ < \min_{\lambda\in\sigma(A_+)} \Re(\lambda).$$

For all $t \ge 0$ we set:

$$x(t) = e^{tA}\Big(x_0 + \int_0^t e^{-sA}\pi_-\big(u(s)\big)\,\mathrm{d}s - \int_t^{+\infty} e^{-sA}\pi_+\big(u(s)\big)\,\mathrm{d}s\Big); \tag{B.6}$$

the convergence of the second integral in (B.6) follows by observing that u is bounded and that, by Lemma B.2:

$$\big\|e^{-sA}\pi_+\big(u(s)\big)\big\| = \big\|e^{-sA_+}\pi_+\big(u(s)\big)\big\| \le Ce^{-s\lambda_+}\big\|u(s)\big\|,$$

for all $s \ge 0$ and some constant $C \ge 0$. A straightforward computation shows that x is a solution of (B.5) with $\pi_-\big(x(0)\big) = x_0$. In order to compute $\lim_{t\to+\infty} x(t)$ we rewrite (B.6) as:

$$x(t) = e^{tA}x_0 + \int_{-\infty}^{+\infty} e^{(t-s)A}\Big[\chi(t-s)\pi_-\big(u(s)\big) - \chi(s-t)\pi_+\big(u(s)\big)\Big]\,\mathrm{d}s$$

$$= e^{tA}x_0 + \int_{-\infty}^{+\infty} e^{sA}\Big[\chi(s)\pi_-\big(u(t-s)\big) - \chi(-s)\pi_+\big(u(t-s)\big)\Big]\,\mathrm{d}s, \tag{B.7}$$

where $\chi : \mathbb{R} \to \mathbb{R}$ denotes the characteristic function of $[0,+\infty[$ and we write $u(s) = 0$ for $s < 0$. Since $x_0 \in \mathbb{R}^n_-(A)$, we have:

$$\|e^{tA}x_0\| = \|e^{tA_-}x_0\| \le Ce^{t\lambda_-}\,\|x_0\|,$$

for all $t \geq 0$ and some constant $C \geq 0$. This shows that $\lim_{t\to+\infty} e^{tA}x_0 = 0$. We now compute the limit as $t \to +\infty$ of the integral in (B.7) using Lebesgue's dominated convergence theorem. Since $\lim_{t\to+\infty} u(t) = 0$, we have:

$$\lim_{t\to+\infty} e^{sA}\Big[\chi(s)\pi_-\big(u(t-s)\big) - \chi(-s)\pi_+\big(u(t-s)\big)\Big] = 0,$$

for fixed $s \in \mathbb{R}$. Moreover, for all $t \geq 0$ and $s \in \mathbb{R}$ we have:

$$\begin{aligned}\Big\|e^{sA}\Big[\chi(s)\pi_-\big(u(t-s)\big) - \chi(-s)\pi_+\big(u(t-s)\big)\Big]\Big\| \\ \leq \big\|\chi(s)e^{sA_-}\pi_-\big(u(t-s)\big)\big\| + \big\|\chi(-s)e^{sA_+}\pi_+\big(u(t-s)\big)\big\| \\ \leq C\big(\chi(s)e^{s\lambda_-} + \chi(-s)e^{s\lambda_+}\big)\|u\|_{C^0},\end{aligned}$$

for some constant $C \geq 0$, where $\|u\|_{C^0} = \sup_{s\geq 0}\big\|u(s)\big\|$. Obviously:

$$\int_{-\infty}^{+\infty} \chi(s)e^{s\lambda_-} + \chi(-s)e^{s\lambda_+}\,\mathrm{d}s < +\infty,$$

which completes the proof that $\lim_{t\to+\infty} x(t) = 0$. Now assume that x_1 and x_2 are solutions of (B.5) with:

$$\pi_-\big(x_1(0)\big) = \pi_-\big(x_2(0)\big) = x_0$$

and $\lim_{t\to+\infty} x_1(t) = \lim_{t\to+\infty} x_2(t) = 0$. Then $x = x_1 - x_2$ is a solution of the homogeneous ODE $x' = Ax$ with $x(0) \in \mathbb{R}^n_+(A)$ and $\lim_{t\to+\infty} x(t) = 0$. Thus $x(t) = e^{tA}x(0) \in \mathbb{R}^n_+(A)$ for all $t \geq 0$ and:

$$\big\|x(0)\big\| = \big\|e^{-tA}x(t)\big\| = \big\|e^{-tA_+}x(t)\big\| \leq e^{-t\lambda_+}\big\|x(t)\big\|;$$

but $\lim_{t\to+\infty} e^{-t\lambda_+}\big\|x(t)\big\| = 0$, which proves that $x(0) = 0$ and hence $x = x_1 - x_2 \equiv 0$. □

Now we study singularities of vector fields on manifolds. If $X : M \to TM$ is a smooth vector field on a manifold M then a *singularity* of X is a point $p \in X$ with $X(p) = 0$. At a singularity p of a vector field X, there exists a natural way of defining a "differential" of X at p, which is a linear endomorphism of T_pM that we will denote by $\nabla X(p)$. If M is an open subset of $\mathbb{R}^n$ then $\nabla X(p)$ is simply the standard differential $\mathrm{d}X(p) : \mathbb{R}^n \to \mathbb{R}^n$. In the case of an arbitrary manifold, $\nabla X(p)$ can be defined for instance as the covariant derivative of X with respect to an arbitrary connection; the fact that p is a singularity implies that such covariant derivative does not depend on the choice of the connection. The linear map $\nabla X(p)$ can also be defined more directly using a local chart around p or, more abstractly, looking at the double tangent bundle TTM. These different possibilities are discussed in Exercise B.3. We remark also that if $f : M \to \mathbb{R}$ is a smooth map on a Riemannian manifold (M, g) then the singularities $p \in M$ of the gradient $X = \nabla f$ of f are precisely the critical points of f and that $\nabla X(p)$ is precisely the linear map that represents the Hessian of f at p with respect to the inner product g_p.

B.4. DEFINITION. Let $X : M \to TM$ be a smooth vector field on a manifold M. A singularity $p \in M$ of X is called *hyperbolic* if the linear endomorphism $\nabla X(p)$ of T_pM is hyperbolic.

We remark that if X is the gradient of a smooth map f on a Riemannian manifold then the hyperbolic singularities of X are precisely the nondegenerate critical points of f.

Now let $X : M \to TM$ be a fixed smooth vector field on a manifold M and let $p \in M$ be a fixed hyperbolic singularity of X. We denote by $F : A \to M$ the flow of X, so that A is open in $\mathbb{R} \times M$, F is smooth and $t \mapsto F(t,x)$ is the maximal integral curve of X with $F(0,x) = x$, for all $x \in M$. For $t \in \mathbb{R}$ we denote by A_t the open subset of M defined by:

$$A_t = \big\{x \in M : (t,x) \in A\big\},$$

and by $F_t : A_t \to M$ the map $F_t = F(t,\cdot)$. Then $F_t : A_t \to A_{-t}$ is a smooth diffeomorphism for all $t \in \mathbb{R}$. The *stable and the unstable manifolds* of p with respect to X are defined respectively by:

$$W_{\mathrm{s}}(p,X) = \big\{x \in M : x \in A_t, \text{ for all } t \geq 0 \text{ and } \lim_{t\to+\infty} F_t(x) = p\big\},$$
$$W_{\mathrm{u}}(p,X) = \big\{x \in M : x \in A_t, \text{ for all } t \leq 0 \text{ and } \lim_{t\to-\infty} F_t(x) = p\big\}.$$

Obviously:

$$W_{\mathrm{s}}(p,X) = W_{\mathrm{u}}(p,-X). \tag{B.8}$$

At this point, there is no evidence that either $W_{\mathrm{s}}(p,X)$ or $W_{\mathrm{u}}(p,X)$ is a manifold of some sort, but this matter will be clarified later in Proposition B.10, where it will be established that the stable and unstable manifolds are immersed submanifolds of M.

Due to (B.8), we will from now on only state results concerning the stable manifold. Analogous results for the unstable manifold can then be obtained by replacing X with $-X$.

If $U \subset M$ is an open neighborhood of p, we will often need to consider the stable manifold $W_{\mathrm{s}}(p, X|_U)$ of p with respect to the vector field X restricted to U. For shortness, we will now write:

$$W_{\mathrm{s}}(p) = W_{\mathrm{s}}(p,X) \quad \text{and} \quad W_{\mathrm{s}}(p;U) = W_{\mathrm{s}}(p, X|_U).$$

Observe that $W_{\mathrm{s}}(p;U)$ is in general *not* the same as $W_{\mathrm{s}}(p) \cap U$; namely, we have:

$$W_{\mathrm{s}}(p;U) = \big\{x \in W_{\mathrm{s}}(p) : F_t(x) \in U, \text{ for all } t \geq 0\big\} \subset W_{\mathrm{s}}(p) \cap U.$$

We now make a few simple remarks concerning the stable manifold and the flow of X that will be used in the proofs of the results presented later on.

B.5. REMARK. For any $t \in \mathbb{R}$, the smooth diffeomorphism $F_t : A_t \to A_{-t}$ restricts to a homeomorphism from $W_{\mathrm{s}}(p) \cap A_t$ onto $W_{\mathrm{s}}(p) \cap A_{-t}$; thus, if Z is open in $W_{\mathrm{s}}(p)$ then both $F_t(Z \cap A_t)$ and $F_t^{-1}(Z) = F_{-t}(Z \cap A_{-t})$ are open in $W_{\mathrm{s}}(p)$.

B.6. REMARK. The stable manifold $W_{\mathrm{s}}(p)$ is arc-connected (with respect to the topology induced by M). Namely, given $x \in W_{\mathrm{s}}(p)$ then

$$\gamma(t) = \begin{cases} F\big(\frac{t}{1-t^2}, x\big), & t \in [0,1[\,, \\ p, & t = 1, \end{cases}$$

defines a continuous curve $\gamma : [0,1] \to W_{\mathrm{s}}(p)$ connecting x to p.

B.7. REMARK. If U, $V \subset M$ are open neighborhoods of p then obviously:

$$W_{\mathrm{s}}(p;U) \cap W_{\mathrm{s}}(p;V) = W_{\mathrm{s}}(p;U \cap V). \tag{B.9}$$

It is also easy to check that for any open neighborhood $U \subset M$ of p and for any $t \in \mathbb{R}$ we have:

$$F_t\big(W_{\mathrm{s}}(p;U) \cap A_t\big) = W_{\mathrm{s}}\big(p; F_t(U \cap A_t)\big), \tag{B.10}$$

$$F_t^{-1}\big(W_{\mathrm{s}}(p;U)\big) = W_{\mathrm{s}}\big(p; F_t^{-1}(U)\big). \tag{B.11}$$

Moreover:

$$W_{\mathrm{s}}(p) = \bigcup_{t \ge 0} F_t^{-1}\big(W_{\mathrm{s}}(p;U)\big) = \bigcup_{t \ge 0} W_{\mathrm{s}}\big(p; F_t^{-1}(U)\big); \tag{B.12}$$

namely, if $x \in W_{\mathrm{s}}(p)$ then, by definition, there exists $t_0 \ge 0$ with $F_t(x) \in U$ for all $t \ge t_0$. This means that $F_{t_0}(x) \in W_{\mathrm{s}}(p;U)$. Observe that the union in (B.12) is monotone, i.e., for $0 \le t \le s$, we have:

$$F_t^{-1}\big(W_{\mathrm{s}}(p;U)\big) \subset F_s^{-1}\big(W_{\mathrm{s}}(p;U)\big).$$

Now we look at the local structure of the stable manifold.

B.8. LEMMA. *There exists an open neighborhood $U \subset M$ of p such that:*

- *$W_{\mathrm{s}}(p;U)$ is an embedded submanifold of M whose tangent space at p is equal to the negative eigenspace of $\nabla X(p)$;*
- *if $V \subset M$ is an open neighborhood of p contained in U then $W_{\mathrm{s}}(p;V)$ is open in $W_{\mathrm{s}}(p;U)$ (and in particular $W_{\mathrm{s}}(p;V)$ is also an embedded submanifold of M).*

PROOF. By choosing a local chart around p, we may assume without loss of generality that M is an open neighborhood of the origin in $\mathbb{R}^n$ and that $p = 0$. For shortness, we denote by $\mathbb{R}^n_+$ and $\mathbb{R}^n_-$ respectively the positive and the negative eigenspaces of $\mathrm{d}X(0) : \mathbb{R}^n \to \mathbb{R}^n$ and by π_+ and π_- the respective projections with respect to the direct sum decomposition $\mathbb{R}^n = \mathbb{R}^n_+ \oplus \mathbb{R}^n_-$.

The strategy is to use the implicit function theorem for maps on Banach spaces. We denote by E^0 the Banach space of continuous maps $\gamma : [0,+\infty[\to \mathbb{R}^n$ such that $\lim_{t\to+\infty} \gamma(t) = 0$, endowed with the norm $\|\gamma\|_{C^0} = \sup_{t\ge 0} \big\|\gamma(t)\big\|$; by E^1 we denote the Banach space of C^1 maps $\gamma : [0,+\infty[\to \mathbb{R}^n$ such that:

$$\lim_{t\to+\infty} \gamma(t) = \lim_{t\to+\infty} \gamma'(t) = 0,$$

endowed with the norm $\|\gamma\|_{C^1} = \|\gamma\|_{C^0} + \|\gamma'\|_{C^0}$. If $U \subset \mathbb{R}^n$ is an open neighborhood of the origin we write:

$$E^k(U) = \big\{\gamma \in E^k : \mathrm{Im}(\gamma) \subset U\big\},$$

so that $E^k(U)$ is an open subset of E^k, for $k = 0, 1$.

Consider the map:

$$\phi : \mathbb{R}^n_- \times E^1(M) \ni (x_0, \gamma) \longmapsto \big(\pi_-(\gamma(0)) - x_0, \gamma' - X \circ \gamma\big) \in \mathbb{R}^n_- \times E^0.$$

The map ϕ is smooth and the partial derivative $\frac{\partial \phi}{\partial \gamma}$ at the origin is given by:

$$\frac{\partial \phi}{\partial \gamma}(0,0)v = \big(\pi_-(v(0)), v' - \mathrm{d}X(0) \circ v\big),$$

for all $v \in E^1$. Lemma B.3 implies that $\frac{\partial \phi}{\partial \gamma}(0,0)$ is an isomorphism and thus, by the implicit function theorem, we can find $r_1, r_2 > 0$ and a smooth map:

$$\sigma : \mathrm{B}(0, r_1; \mathbb{R}^n_-) \longrightarrow \mathrm{B}(0, r_2; E^1)$$

such that $\mathrm{B}(0, r_1; \mathbb{R}^n_-) \times \mathrm{B}(0, r_2; E^1) \subset \mathbb{R}^n_- \times E^1(M)$ and:

$$\big(\mathrm{B}(0, r_1; \mathbb{R}^n_-) \times \mathrm{B}(0, r_2; E^1)\big) \cap \phi^{-1}(0) = \mathrm{Gr}(\sigma),$$

where $\mathrm{B}(0, r; \mathcal{X})$ denotes the open ball of center 0 and radius r of a normed space $\mathcal{X}$ and $\mathrm{Gr}(\sigma)$ denotes the graph of σ. Observe that $\sigma(0) = 0$ and that for all $h \in \mathbb{R}^n_-$:

$$\text{(B.13)} \qquad \mathrm{d}\sigma(0)h = -\Big[\frac{\partial \phi}{\partial \gamma}(0,0)^{-1} \circ \frac{\partial \phi}{\partial x_0}(0,0)\Big] h = v,$$

where $v : [0, +\infty[\, \to \mathbb{R}^n$ is the unique solution of the ODE $v' = \mathrm{d}X(0) \circ v$ with $\pi_-\big(v(0)\big) = h$ and $\lim_{t \to +\infty} v(t) = 0$. From the proof of Lemma B.3 (see (B.6)) it is clear that v is given by:

$$\text{(B.14)} \qquad v(t) = e^{t\, \mathrm{d}X(0)} h.$$

We now choose $U \subset M$ to be an open neighborhood of the origin such that:

$$\sup_{x \in U} \|x\| < \frac{r_2}{2}, \quad \sup_{x \in U} \big\|X(x)\big\| < \frac{r_2}{2}, \quad \sup_{x \in U} \big\|\pi_-(x)\big\| < r_1;$$

observe that if $\gamma : [0, +\infty[\, \to \mathbb{R}^n$ is an integral curve of X with $\mathrm{Im}(\gamma) \subset U$ and $\lim_{t \to +\infty} \gamma(t) = 0$ then $\pi_-\big(\gamma(0)\big) \in \mathrm{B}(0, r_1; \mathbb{R}^n_-)$ and $\gamma \in \mathrm{B}(0, r_2; E^1)$. This means that:

$$\text{(B.15)} \qquad W_{\mathrm{s}}(0; U) = \big\{\gamma(0) : \gamma \in E^1(U) \text{ and } \big(\pi_-(\gamma(0)), \gamma\big) \in \mathrm{Gr}(\sigma)\big\}.$$

We may thus write $W_{\mathrm{s}}(0; U)$ as the graph of a smooth map; more specifically, let $\eta : \mathrm{B}(0, r_1; \mathbb{R}^n_-) \to \mathbb{R}^n_+$ be the smooth map defined by:

$$\eta(x_0) = \pi_+\big(\gamma(0)\big),$$

where $\gamma = \sigma(x_0)$. From (B.13) and (B.14) we see that $\mathrm{d}\eta(0) = 0$. Moreover, from (B.15) we get:

$$W_{\mathrm{s}}(0; U) = \mathrm{Gr}(\eta|_{\sigma^{-1}(E^1(U))}),$$

where $\sigma^{-1}\big(E^1(U)\big)$ is an open subset of $\mathrm{B}(0,r_1;\mathbb{R}^n_-)$. This proves that $W_{\mathrm{s}}(0;U)$ is an embedded submanifold of $\mathbb{R}^n$ whose tangent space at 0 is $\mathrm{Gr}\big(\mathrm{d}\eta(0)\big)=\mathbb{R}^n_-$. Moreover, if $V\ni 0$ is an open subset of U then:

$$W_{\mathrm{s}}(0;V)=\mathrm{Gr}(\eta|_{\sigma^{-1}(E^1(V))}),$$

which proves that $W_{\mathrm{s}}(0;V)$ is open in $W_{\mathrm{s}}(0;U)$, because $\sigma^{-1}\big(E^1(V)\big)$ is open in $\mathrm{B}(0,r_1;\mathbb{R}^n_-)$. □

B.9. REMARK. Choose U as in the statement of Lemma B.8. Given $t\in\mathbb{R}$ and an open neighborhood $Z\subset M$ of p contained in $F_t^{-1}(U)$ then $W_{\mathrm{s}}(p;Z)$ is open in $W_{\mathrm{s}}\big(p;F_t^{-1}(U)\big)=F_t^{-1}\big(W_{\mathrm{s}}(p;U)\big)$ (recall (B.11)). Namely, we have $Z=F_t^{-1}\big(F_t(Z)\big)$ and thus (B.11) implies:

$$W_{\mathrm{s}}(p;Z)=F_t^{-1}\big(W_{\mathrm{s}}(p;F_t(Z))\big).$$

But $F_t(Z)$ is an open neighborhood of p contained in U and thus $W_{\mathrm{s}}\big(p;F_t(Z)\big)$ is open in $W_{\mathrm{s}}(p;U)$; finally, the continuity of F_t implies that $F_t^{-1}\big(W_{\mathrm{s}}(p;F_t(Z))\big)$ is open in $F_t^{-1}\big(W_{\mathrm{s}}(p;U)\big)$.

We can now prove that $W_{\mathrm{s}}(p)$ is an immersed submanifold of M. We adopt the following terminology; if N is an immersed submanifold of M, then by the *manifold topology* of N we mean the topology induced by the atlas of N. Such topology is finer than the *induced topology* of N, which is the topology N inherits from M.

B.10. PROPOSITION. *There exists a unique manifold structure on $W_{\mathrm{s}}(p)$ such that $W_{\mathrm{s}}(p)$ is an immersed submanifold of M and such that, for every open neighborhood $V\subset M$ of p, $W_{\mathrm{s}}(p;V)$ is open in $W_{\mathrm{s}}(p)$ with respect to the manifold topology. Moreover, the following statements hold:*

(1) *the tangent space of $W_{\mathrm{s}}(p)$ at p is equal to the negative eigenspace of $\nabla X(p)$;*
(2) *the vector field X restricts to a smooth vector field on $W_{\mathrm{s}}(p)$;*
(3) *for $x\in W_{\mathrm{s}}(p)$, the maximal integral curve of X passing through x equals the maximal integral curve of $X|_{W_{\mathrm{s}}(p)}$ passing through x;*
(4) *p is a hyperbolic singularity of $X|_{W_{\mathrm{s}}(p)}$ whose stable manifold is equal to $W_{\mathrm{s}}(p)$;*
(5) *$W_{\mathrm{s}}(p)$ is arc-connected with respect to the manifold topology.*

PROOF. Choose $U\subset M$ as in the statement of Lemma B.8. Then $W_{\mathrm{s}}(p;U)$ is an embedded submanifold of M and for all $t\ge 0$, since F_t is a smooth diffeomorphism between open subsets of M, it follows that also $F_t^{-1}\big(W_{\mathrm{s}}(p;U)\big)$ is an embedded submanifold of M. By (B.12) we see then that $W_{\mathrm{s}}(p)$ is a monotone union of embedded submanifolds of M. Our strategy now is to use the result of Exercise B.1 to construct the manifold structure of $W_{\mathrm{s}}(p)$. First, observe that the union in (B.12) can be taken over $t\in\mathbb{N}$, i.e., it can be replaced by a countable union. Now we show that, for $0\le t\le s$, the set $F_t^{-1}\big(W_{\mathrm{s}}(p;U)\big)$ is open in

$F_s^{-1}\big(W_{\rm s}(p;U)\big)$. Using (B.9) and (B.11) we compute:

$$\begin{aligned} F_t^{-1}\big(W_{\rm s}(p;U)\big) &= F_t^{-1}\big(W_{\rm s}(p;U)\big)\cap F_s^{-1}\big(W_{\rm s}(p;U)\big)\\ &= W_{\rm s}\big(p;F_t^{-1}(U)\big)\cap W_{\rm s}\big(p;F_s^{-1}(U)\big) = W_{\rm s}\big(p;F_t^{-1}(U)\cap F_s^{-1}(U)\big). \end{aligned}$$

Since $F_t^{-1}(U)\cap F_s^{-1}(U)$ is an open neighborhood of p in M contained in $F_s^{-1}(U)$, by Remark B.9, we know that $W_{\rm s}\big(p;F_t^{-1}(U)\cap F_s^{-1}(U)\big) = F_t^{-1}\big(W_{\rm s}(p;U)\big)$ is open in $F_s^{-1}\big(W_{\rm s}(p;U)\big)$. So far, we have proven the following facts:

- there exists a manifold structure on $W_{\rm s}(p)$ such that $W_{\rm s}(p)$ is an immersed submanifold of M and such that, for all $t\ge 0$, $W_{\rm s}\big(p;F_t^{-1}(U)\big)$ is open in $W_{\rm s}(p)$ with respect to the manifold topology;
- there exists at most one manifold structure on $W_{\rm s}(p)$ such that $W_{\rm s}(p)$ is an immersed submanifold of M and such that, for every open neighborhood $V\subset M$ of p, $W_{\rm s}(p;V)$ is open in $W_{\rm s}(p)$ with respect to the manifold topology.

To complete the proof of the first part of the statement of the proposition, we consider an open neighborhood $V\subset M$ of p and we show that $W_{\rm s}(p;V)$ is open in $W_{\rm s}(p)$ with respect to the manifold topology. From (B.9) and (B.12) we obtain:

$$W_{\rm s}(p;V) = \bigcup_{t\ge 0} W_{\rm s}(p;V)\cap W_{\rm s}\big(p;F_t^{-1}(U)\big) = \bigcup_{t\ge 0} W_{\rm s}\big(p;V\cap F_t^{-1}(U)\big). \tag{B.16}$$

Since $V\cap F_t^{-1}(U)$ is an open neighborhood of p contained in $F_t^{-1}(U)$, by Remark B.9 we obtain that $W_{\rm s}\big(p;V\cap F_t^{-1}(U)\big)$ is open in $W_{\rm s}\big(p;F_t^{-1}(U)\big)$. But $W_{\rm s}\big(p;F_t^{-1}(U)\big)$ is open in $W_{\rm s}(p)$ with respect to the manifold topology; moreover, $W_{\rm s}\big(p;F_t^{-1}(U)\big)$ inherits the same topology from M and from the manifold topology of $W_{\rm s}(p)$. This shows that $W_{\rm s}\big(p;V\cap F_t^{-1}(U)\big)$ is open in $W_{\rm s}(p)$ with respect to the manifold topology and hence, by (B.16), $W_{\rm s}(p;V)$ is also open in $W_{\rm s}(p)$ with respect to the manifold topology. This completes the proof of the first part of the statement of the proposition. Moreover, since $W_{\rm s}(p;U)$ is an open submanifold of $W_{\rm s}(p)$, statement (1) follows directly from Lemma B.8. We now prove statements (2)—(5).

Let $x\in W_{\rm s}(p)$ be fixed and denote by $\gamma: I\to M$ the maximal integral curve of X with $\gamma(0)=x$. Obviously $\gamma(I)$ is contained in $W_{\rm s}(p)$, but it is not clear in principle that $\gamma: I\to W_{\rm s}(p)$ is smooth. We argue as follows; since $\lim_{t\to+\infty}\gamma(t)=p$, we have $\gamma(t)\in U$ for t sufficiently large and thus, given a bounded subinterval $I'\subset I$, we can find $t\ge 0$ large enough so that $\gamma(I')$ is contained in $F_t^{-1}\big(W_{\rm s}(p;U)\big)$. Since $F_t^{-1}\big(W_{\rm s}(p;U)\big)$ is an embedded submanifold of M and an open submanifold of $W_{\rm s}(p)$, we get that $\gamma|_{I'}: I'\to W_{\rm s}(p)$ is smooth. Since $I'\subset I$ is an arbitrary bounded subinterval, we obtain that $\gamma: I\to W_{\rm s}(p)$ is smooth. Observe that, in particular, $\gamma'(0)=X(x)$ is in $T_xW_{\rm s}(p)$. We have thus proven statements (2) and (3) (see Exercise B.4).

To prove statement (4), observe first that p is obviously a singularity of $X|_{W_{\rm s}(p)}$ and that $\nabla(X|_{W_{\rm s}(p)})(p)$ is equal to the restriction of $\nabla X(p)$ to $T_pW_{\rm s}(p)$, which is equal to the negative eigenspace of $\nabla X(p)$. Thus p is a hyperbolic singularity of

$X|_{W_s(p)}$. To prove that $W_s(p)$ is the stable manifold of p with respect to $X|_{W_s(p)}$, we have to show that for every $x \in W_s(p)$, we have $\lim_{t\to+\infty} F_t(x) = p$ *with respect to the manifold topology* of $W_s(p)$. We know that $\lim_{t\to+\infty} F_t(x) = p$ with respect to the topology of M. But $F_t(x) \in W_s(p;U)$ for t sufficiently large and $W_s(p;U)$ inherits the same topology from M and from the manifold topology of $W_s(p)$. Thus $\lim_{t\to+\infty} F_t(x) = p$ with respect to the manifold topology of $W_s(p)$. This proves statement (4). Finally, now that the continuity of $t \mapsto F_t(x)$ and the limit $\lim_{t\to+\infty} F_t(x) = p$ have been established with respect to the manifold topology of $W_s(p)$, statement (5) follows using the same argument used in Remark B.6. $\square$

We now study conditions under which the stable manifold is embedded in M.

B.11. PROPOSITION. *The following conditions are equivalent:*

(1) $W_s(p)$ *is an embedded submanifold of* M*;*
(2) *for every open neighborhood* $V \subset M$ *of* p, $W_s(p;V)$ *is open in* $W_s(p)$ *with respect to the induced topology;*
(3) *every open neighborhood* $V \subset M$ *of* p *contains an open neighborhood* $Z \subset M$ *of* p *such that* $W_s(p;Z)$ *is open in* $W_s(p)$ *with respect to the induced topology;*
(4) *every open neighborhood* $V \subset M$ *of* p *contains an open neighborhood* $Z \subset M$ *of* p *such that* $W_s(p;Z) = W_s(p) \cap Z$*;*
(5) *there exists an open neighborhood* $V \subset M$ *of* p *such that* $W_s(p) \cap V$ *is an embedded submanifold of* M.

PROOF.
(1)⇒(2). Since $W_s(p)$ is an embedded submanifold of M, the manifold structure given to $W_s(p)$ by Proposition B.10 must coincide with the one that makes $W_s(p)$ embedded in M. This proves (2).

(2)⇒(3). Trivial.

(3)⇒(4). Let $V \subset M$ be an open neighborhood of p; by (3), there exists an open neighborhood $V_0 \subset V$ of p such that $W_s(p;V_0)$ is open in $W_s(p)$ with respect to the induced topology. Thus, there exists an open subset $V_1 \subset M$ such that $W_s(p;V_0) = W_s(p) \cap V_1$. Then $V_1 \subset M$ is an open neighborhood of p and it is easy to see that $W_s(p;V_0) \subset W_s(p;V_1)$; setting $Z = V_1 \cap V$, we obtain:

$$\begin{aligned} W_s(p;Z) \subset W_s(p) \cap Z = W_s(p) \cap V_1 \cap V = W_s(p;V_0) \cap V \\ = W_s(p;V_0) \subset W_s(p;V_1) \subset W_s(p;Z), \end{aligned}$$

which proves that $W_s(p;Z) = W_s(p) \cap Z$.

(4)⇒(5). Choose U as in the statement of Lemma B.8. By (4), we can find an open neighborhood $Z \subset U$ of p such that $W_s(p;Z) = W_s(p) \cap Z$. But $Z \subset U$ implies that $W_s(p;Z)$ is an embedded submanifold of M.

(5)⇒(1). Let $V \subset M$ be an open neighborhood of p such that $W_{\mathrm{s}}(p) \cap V$ is an embedded submanifold of M. Obviously:

$$W_{\mathrm{s}}(p) = \bigcup_{t \geq 0} F_t^{-1}\big(W_{\mathrm{s}}(p) \cap V\big),$$

and since F_t is a smooth diffeomorphism between open subsets of M, we have that $F_t^{-1}\big(W_{\mathrm{s}}(p) \cap V\big)$ is an embedded submanifold of M for all $t \geq 0$. But $F_t^{-1}\big(W_{\mathrm{s}}(p) \cap V\big)$ is open in $W_{\mathrm{s}}(p)$ with respect to the *induced topology* (recall Remark B.5), which proves that $W_{\mathrm{s}}(p)$ is embedded in M. □

We now study the case that (M, g) is a Riemannian manifold and that $X = \nabla f$ is the gradient of a smooth map $f : M \to \mathbb{R}$. As we have already observed, the fact that p is a hyperbolic singularity of X means that p is a nondegenerate critical point of f; moreover, $g\big(\nabla X(p), \cdot\big) = \mathrm{Hess} f_p$. Our goal is to show that $W_{\mathrm{s}}(p)$ is always embedded in M, if X is a gradient. We start with the following preparatory lemma.

B.12. LEMMA. *If $X = \nabla f$ and $p \in M$ is a nondegenerate critical point of $f : M \to \mathbb{R}$ with $f(p) = c \in \mathbb{R}$ then, given an open neighborhood $V \subset M$ of p, we can find a smooth chart $\varphi : Z \to \mathrm{B}(0, r; \mathbb{R}^k)$ on the manifold $W_{\mathrm{s}}(p)$ with $p \in Z \subset V$ and*

$$f(x) = c - \big\|\varphi(x)\big\|^2,$$

for all $x \in Z$. Moreover, for $\varepsilon \in \left]0, r^2\right[$ we have:

$$f^{-1}\big(\left[c - \varepsilon, +\infty\right[\big) \cap W_{\mathrm{s}}(p) \subset Z \subset V.$$

In particular, φ restricts to a homeomorphism between $f^{-1}\big(\left[c - \varepsilon, +\infty\right[\big) \cap W_{\mathrm{s}}(p)$ and $\mathrm{B}[0, \sqrt{\varepsilon}; \mathbb{R}^k]$ that carries $f^{-1}(c - \varepsilon) \cap W_{\mathrm{s}}(p)$ to the sphere $S[0, \sqrt{\varepsilon}; \mathbb{R}^k]$ (see Remark B.13 below).

PROOF. We already know (by Proposition B.10) that $W_{\mathrm{s}}(p)$ is an immersed submanifold of M; thus, f restricts to a smooth map $f|_{W_{\mathrm{s}}(p)}$ on the manifold $W_{\mathrm{s}}(p)$. Since $T_p W_{\mathrm{s}}(p)$ is the negative eigenspace of $\mathrm{Hess} f_p$, the existence of the chart φ in $W_{\mathrm{s}}(p)$ follows directly from the Morse Lemma.

Now choose $\varepsilon > 0$ with $\varepsilon < r^2$. We will show that the set:

$$f^{-1}\big(\left[c - \varepsilon, +\infty\right[\big) \cap W_{\mathrm{s}}(p) = f^{-1}\big([c - \varepsilon, c]\big) \cap W_{\mathrm{s}}(p)$$

is contained in Z (and hence in V). Choose $x \in W_{\mathrm{s}}(p)$ with $f(x) \geq c - \varepsilon$ and assume by contradiction that $x \notin Z$. Let $\gamma : I \to M$ denote the maximal integral curve of X such that $\gamma(0) = x$. From Proposition B.10 we know that the map $\gamma : I \to W_{\mathrm{s}}(p)$ is continuous when $W_{\mathrm{s}}(p)$ is endowed with the manifold topology and that $\lim_{t \to +\infty} \gamma(t) = p$ also with respect to the manifold topology of $W_{\mathrm{s}}(p)$. Thus, for t sufficiently large, we have $\gamma(t) \in Z$ and $f\big(\gamma(t)\big) > c - \varepsilon$. But $\gamma(t) \in Z$ and $f\big(\gamma(t)\big) > c - \varepsilon$ imply $\gamma(t) \in \varphi^{-1}\big(\mathrm{B}(0, \sqrt{\varepsilon}; \mathbb{R}^k)\big)$. Since $\gamma(0) = x$ is not in $\varphi^{-1}\big(\mathrm{B}[0, \sqrt{\varepsilon}; \mathbb{R}^k]\big)$, it follows from the result of Exercise B.5 that there must exist $t > 0$ with $\gamma(t) \in Z$ and $\big\|\varphi\big(\gamma(t)\big)\big\| = \varepsilon$. Thus:

$$f\big(\gamma(t)\big) = c - \varepsilon \leq f\big(\gamma(0)\big),$$

contradicting the fact that $f \circ \gamma$ is strictly increasing. $\square$

B.13. REMARK. From the proof of Lemma B.12 we know that Z is open in $W_{\mathrm{s}}(p)$ with respect to the *manifold topology* and that $\varphi : Z \to \mathrm{B}(0, r; \mathbb{R}^k)$ is a homeomorphism if Z is endowed with the *manifold topology* of $W_{\mathrm{s}}(p)$. However, in Theorem B.14 below we will see that $W_{\mathrm{s}}(p)$ is embedded in M and thus the manifold topology of $W_{\mathrm{s}}(p)$ coincides with the induced topology.

We can now prove that the stable manifold is embedded in the case of gradient vector fields.

B.14. THEOREM. *Let (M, g) be a Riemannian manifold, $f : M \to \mathbb{R}$ be a smooth map and $p \in M$ be a nondegenerate critical point of f. Then $W_{\mathrm{s}}(p, \nabla f)$ is a connected embedded submanifold of M whose tangent space at p equals the negative eigenspace of the linear endomorphism of T_pM that represents $\mathrm{Hess} f_p$ with respect to the inner product g_p.*

PROOF. We will prove that condition (4) in the statement of Proposition B.11 holds. Let $V \subset M$ be an open neighborhood of p. Set $c = f(p)$ and choose $\varepsilon > 0$ as in the statement of Lemma B.12. Setting $Z = f^{-1}\big(\,]c - \varepsilon, +\infty[\,\big) \cap V$ then, since f is increasing in the flow lines of ∇f, it is easy to see that $W_{\mathrm{s}}(p; Z) = W_{\mathrm{s}}(p) \cap Z$. Thus, $W_{\mathrm{s}}(p)$ is an embedded submanifold of M. The other claims in the statement of the theorem follow from Proposition B.10. $\square$

Our goal now is to give a topological characterization of the flow of a vector field near a hyperbolic singularity. We need some definitions.

B.15. DEFINITION. Let $X : M \to TM$, $Y : N \to TN$ be smooth vector fields on manifolds M, N, and let $f : M \to N$ be a continuous map. We say that Y is *f-related* to X if f carries the flow of X to the flow of Y, i.e., if for every integral curve $\gamma : I \to M$ of X, $f \circ \gamma$ is an integral curve of Y. If there exists a homeomorphism $f : M \to N$ such that Y is f-related to X, we say that X and Y are *topologically conjugated*.

A few simple facts concerning the definition above are discussed in Exercise B.6.

Our goal is to prove the following:

B.16. THEOREM (Hartman–Grobman). *Let $X : M \to TM$ be a smooth vector field on a manifold M and let $p \in M$ be a hyperbolic singularity of X. Then there exists an open neighborhood $U \subset M$ of p and an open neighborhood $\widetilde{U} \subset T_pM$ of the origin such that $X|_U$ is topologically conjugated to the (restricted) linear vector field $\nabla X(p)|_{\widetilde{U}}$.*

The proof of Theorem B.16 will take some work. We need several preliminary lemmas. To keep the reader motivated throughout the process, we present below an outline of the proof.

Sketch of the proof of Theorem B.16. Using a local chart around p, one can obviously assume that M is an open subset of $\mathbb{R}^n$ and that $p = 0$. The flow at time t of the linear vector field $A = \mathrm{d}X(0) : \mathbb{R}^n \to \mathbb{R}^n$, is given by the linear isomorphism e^{tA}. Since A has no purely imaginary eigenvalues, the isomorphism e^{tA} has no eigenvalues on the unit circle for $t \neq 0$; isomorphisms with such property will be called *hyperbolic isomorphisms*. Using the implicit function theorem on Banach spaces in a suitable way, we prove that small C^1-perturbations of a hyperbolic isomorphism $L : \mathbb{R}^n \to \mathbb{R}^n$ are *topologically conjugated* in the sense that given two such perturbations $L + \phi_1$, $L + \phi_2$ we can find a homeomorphism $h : \mathbb{R}^n \to \mathbb{R}^n$ such that $h(L+\phi_2)h^{-1} = L+\phi_1$; actually, h is shown to be unique is a small C^0-neighborhood of the identity map. We then replace the vector field X by a global vector field $\widehat{X} : \mathbb{R}^n \to \mathbb{R}^n$ which equals X in a small neighborhood of the origin and equals A far from the origin; $\widehat{X}$ is also chosen so that its flow-at-time-one-map $\widehat{F}_1 : \mathbb{R}^n \to \mathbb{R}^n$ is C^1-close to $L = e^A$. Thus, we know that $\widehat{F}_1$ is topologically conjugated to L by a homeomorphism $h : \mathbb{R}^n \to \mathbb{R}^n$; we then show that h actually carries the entire flow of $\widehat{X}$ to the flow $t \mapsto e^{tA}$ of A. Since $\widehat{X}$ equals X near the origin, the proof is completed.

We now give the details of the proof. First, a formal definition.

B.17. DEFINITION. Let V be a real finite-dimensional vector space. A linear isomorphism $L : V \to V$ is called a *hyperbolic isomorphism* if there is no $\lambda \in \sigma(L)$ with $|\lambda| = 1$. If $L : V \to V$ is a hyperbolic isomorphism, we have the following direct sum decomposition of V into L-invariant subspaces:

$$V = V_{\mathrm{u}}(L) \oplus V_{\mathrm{s}}(L),$$

where:

$$V_{\mathrm{u}}(L) = \sum_{\substack{\lambda\in\sigma(L)\\ |\lambda|>1}} V_\lambda(L), \quad V_{\mathrm{s}}(L) = \sum_{\substack{\lambda\in\sigma(L)\\ |\lambda|<1}} V_\lambda(L).$$

We have thus the following:

B.18. LEMMA. *If $A : \mathbb{R}^n \to \mathbb{R}^n$ is a hyperbolic linear map then $L = e^A$ is a hyperbolic isomorphism.*

PROOF. Using the Jordan decomposition of A (see the proof of Lemma B.2) it is easy to see that:

$$\sigma(L) = \big\{e^\lambda : \lambda \in \sigma(A)\big\}.$$

The conclusion follows. □

As explained in the sketch of the proof of Theorem B.16, we will use the implicit function theorem on Banach spaces to prove that small C^1-perturbations of a hyperbolic isomorphism are topologically conjugated; when verifying the hypothesis of the implicit function theorem, we will need some tools from linear algebra that are given below.

B.19. LEMMA. *Given a linear map $L : \mathbb{R}^n \to \mathbb{R}^n$ and a real number a with $a > \max_{\lambda\in\sigma(L)} |\lambda|$ then there exists a norm $\|\cdot\|'$ on $\mathbb{R}^n$ such that:*

$$\|L\|' = \sup_{\|v\|'\le 1} \big\|L(v)\big\|' \le a.$$

PROOF. Let $L = S + N$ be the Jordan decomposition of L, so that S is semi-simple, N is nilpotent and $SN = NS$. Let $B : \mathbb{C}^n \to \mathbb{C}^n$ be a complex linear isomorphism such that $D = BS^{\mathbb{C}}B^{-1} : \mathbb{C}^n \to \mathbb{C}^n$ is diagonal (with diagonal elements in $\sigma(L)$). Define a norm $\|\cdot\|^{\diamondsuit}$ on $\mathbb{C}^n$ by:

$$\|x\|^{\diamondsuit} = \big\|B(x)\big\|, \quad x \in \mathbb{C}^n,$$

where, $\|\cdot\|$ denotes the usual Euclidean norm on $\mathbb{C}^n$. We denote also by $\|\cdot\|^{\diamondsuit}$ the norm on $\mathbb{R}^n$ obtained by taking the restriction of $\|\cdot\|^{\diamondsuit}$. One checks easily that:

$$\|S\|^{\diamondsuit} \le \|S^{\mathbb{C}}\|^{\diamondsuit} = \|D\| = \sup_{\lambda \in \sigma(L)} |\lambda|.$$

Now choose $\varepsilon > 0$ with $\sup_{\lambda\in\sigma(L)} |\lambda| + \varepsilon \le a$ and define the norm $\|\cdot\|'$ on $\mathbb{R}^n$ by setting:

$$\|x\|' = \sum_{k=0}^{+\infty} \frac{1}{\varepsilon^k}\big\|N^k(x)\big\|^{\diamondsuit} = \sum_{k=0}^{n-1} \frac{1}{\varepsilon^k}\big\|N^k(x)\big\|^{\diamondsuit}, \quad x \in \mathbb{R}^n.$$

If $x \in \mathbb{R}^n$, $\|x\|' \le 1$ we compute:

$$\big\|N(x)\big\|' = \varepsilon \sum_{k=1}^{n-1} \frac{1}{\varepsilon^k}\big\|N^k(x)\big\|^{\diamondsuit} \le \varepsilon,$$

$$\big\|S(x)\big\|' = \sum_{k=0}^{n-1} \frac{1}{\varepsilon^k}\big\|SN^k(x)\big\|^{\diamondsuit} \le \sup_{\lambda\in\sigma(L)} |\lambda|.$$

Hence:

$$\|L\|' = \|S + N\|' \le \|S\|' + \|N\|' \le \sup_{\lambda\in\sigma(L)} |\lambda| + \varepsilon \le a. \qquad \square$$

Let us introduce some notation. Denote by $C^0_{\mathrm{bu}}(\mathbb{R}^n)$ the Banach space of all uniformly continuous bounded maps $u : \mathbb{R}^n \to \mathbb{R}^n$ endowed with the norm:

$$\|u\|_{C^0} = \sup_{x\in\mathbb{R}^n} \big\|u(x)\big\|;$$

by $C^1_{\mathrm{bu}}(\mathbb{R}^n)$ we denote the Banach space of all bounded maps $u : \mathbb{R}^n \to \mathbb{R}^n$ of class C^1 such that $\mathrm{d}u : \mathbb{R}^n \to \mathrm{Lin}(\mathbb{R}^n)$ is bounded and uniformly continuous. The space $C^1_{\mathrm{bu}}(\mathbb{R}^n)$ is endowed with the norm:

$$\|u\|_{C^1} = \sup_{x\in\mathbb{R}^n} \big\|u(x)\big\| + \sup_{x\in\mathbb{R}^n} \big\|\mathrm{d}u(x)\big\|.$$

Observe that each $u \in C^1_{\mathrm{bu}}(\mathbb{R}^n)$ is Lipschitz (and hence uniformly continuous) because $\mathrm{d}u$ is bounded.

B.20. LEMMA. *Let $L : \mathbb{R}^n \to \mathbb{R}^n$ be a hyperbolic isomorphism. The linear map:*

$$C^0_{\mathrm{bu}}(\mathbb{R}^n) \ni u \longmapsto u \circ L - L \circ u \in C^0_{\mathrm{bu}}(\mathbb{R}^n) \tag{B.17}$$

is an isomorphism.

PROOF. For shortness, we set $\mathbb{R}^n_{\rm u} = \mathbb{R}^n_{\rm u}(L)$, $R^n_{\rm s} = \mathbb{R}^n_{\rm s}(L)$ and we denote by $L_{\rm u}$, $L_{\rm s}$ respectively the linear isomorphisms of $\mathbb{R}^n_{\rm u}$ and $\mathbb{R}^n_{\rm s}$ obtained by taking restrictions of L. We can write the Banach space $C^0_{\rm bu}(\mathbb{R}^n)$ as the direct sum of two closed subspaces as follows:

$$C^0_{\rm bu}(\mathbb{R}^n) = C^0_{\rm bu}(\mathbb{R}^n, \mathbb{R}^n_{\rm u}) \oplus C^0_{\rm bu}(\mathbb{R}^n, \mathbb{R}^n_{\rm s}),$$

where:

$$C^0_{\rm bu}(\mathbb{R}^n, \mathbb{R}^n_{\rm u}) = \big\{u \in C^0_{\rm bu}(\mathbb{R}^n) : \mathrm{Im}(u) \subset \mathbb{R}^n_{\rm u}\big\},$$
$$C^0_{\rm bu}(\mathbb{R}^n, \mathbb{R}^n_{\rm s}) = \big\{u \in C^0_{\rm bu}(\mathbb{R}^n) : \mathrm{Im}(u) \subset \mathbb{R}^n_{\rm s}\big\}.$$

The subspaces $C^0_{\rm bu}(\mathbb{R}^n, \mathbb{R}^n_{\rm u})$, $C^0_{\rm bu}(\mathbb{R}^n, \mathbb{R}^n_{\rm s})$ are invariant by (B.17); thus, the proof of the lemma will be completed once we show that the maps:

$$C^0_{\rm bu}(\mathbb{R}^n, \mathbb{R}^n_{\rm u}) \ni u \longmapsto u \circ L - L_{\rm u} \circ u \in C^0_{\rm bu}(\mathbb{R}^n, \mathbb{R}^n_{\rm u}), \tag{B.18}$$

$$C^0_{\rm bu}(\mathbb{R}^n, \mathbb{R}^n_{\rm s}) \ni u \longmapsto u \circ L - L_{\rm s} \circ u \in C^0_{\rm bu}(\mathbb{R}^n, \mathbb{R}^n_{\rm s}), \tag{B.19}$$

obtained by taking restrictions of (B.17) are isomorphisms. We have:

$$u \circ L - L_{\rm u} \circ u = L_{\rm u} \circ (L_{\rm u}^{-1} \circ u \circ L - u)$$

and thus, to prove that (B.18) is an isomorphism, it suffices to prove that

$$C^0_{\rm bu}(\mathbb{R}^n, \mathbb{R}^n_{\rm u}) \ni u \longmapsto L_{\rm u}^{-1} \circ u \circ L - u \in C^0_{\rm bu}(\mathbb{R}^n, \mathbb{R}^n_{\rm u}) \tag{B.20}$$

is an isomorphism. But (B.20) is a perturbation of the identity by the map:

$$C^0_{\rm bu}(\mathbb{R}^n, \mathbb{R}^n_{\rm u}) \ni u \longmapsto L_{\rm u}^{-1} \circ u \circ L \in C^0_{\rm bu}(\mathbb{R}^n, \mathbb{R}^n_{\rm u}); \tag{B.21}$$

the idea is to find a norm on $C^0_{\rm bu}(\mathbb{R}^n, \mathbb{R}^n_{\rm u})$, equivalent to $\|\cdot\|_{C^0}$, such that the norm of (B.21) is smaller than 1. This will prove that (B.20) (and hence (B.18)) is an isomorphism. Since $\sigma(L_{\rm u}) \subset \mathbb{C}$ is contained in the complement of the unit closed ball, $\sigma(L_{\rm u}^{-1})$ is contained in the open unit ball and thus, by Lemma B.19, we can find a norm $\|\cdot\|'$ on $\mathbb{R}^n_{\rm u}$ such that $\|L_{\rm u}^{-1}\|' < 1$. Then, the norm:

$$\|u\|'_{C^0} = \sup_{x \in \mathbb{R}^n} \big\|u(x)\big\|',$$

on $C^0_{\rm bu}(\mathbb{R}^n, \mathbb{R}^n_{\rm u})$ is equivalent to $\|\cdot\|_{C^0}$ and it is easy to see that, with respect to $\|\cdot\|'_{C^0}$, the operator (B.21) has norm smaller than 1. The proof that (B.19) is an isomorphism is similar; one writes:

$$u \circ L - L_{\rm s} \circ u = (u - L_{\rm s} \circ u \circ L^{-1}) \circ L,$$

and then it is possible to choose a norm $\|\cdot\|'$ on $\mathbb{R}^n_{\rm s}$ such that $\|L_{\rm s}\|' < 1$. The conclusion follows. □

We now prove that C^1-small perturbations of a hyperbolic isomorphisms are topologically conjugated.

B.21. LEMMA. *Given a hyperbolic isomorphism $L : \mathbb{R}^n \to \mathbb{R}^n$, there exists $\varepsilon > 0$ and $\delta > 0$ with the following property; given $\phi \in C^1_{\rm bu}(\mathbb{R}^n)$ with $\|\phi\|_{C^1} < \varepsilon$ there exists a unique $u \in C^0_{\rm bu}(\mathbb{R}^n)$ with $\|u\|_{C^0} < \delta$ and:*

$$(\mathrm{Id} + u) \circ (L + \phi) = L \circ (\mathrm{Id} + u). \tag{B.22}$$

Moreover, for such u, the map $\mathrm{Id} + u : \mathbb{R}^n \to \mathbb{R}^n$ is a homeomorphism.

PROOF. Consider the map:

$$F : C^1_{\mathrm{bu}}(R^n) \times C^1_{\mathrm{bu}}(\mathbb{R}^n) \times C^0_{\mathrm{bu}}(\mathbb{R}^n) \longrightarrow C^0_{\mathrm{bu}}(\mathbb{R}^n),$$

given by:

$$F(\phi_1, \phi_2, u) = (\mathrm{Id} + u) \circ (L + \phi_1) - (L + \phi_2) \circ (\mathrm{Id} + u).$$

Then (B.22) is equivalent to $F(\phi, 0, u) = 0$. We can rewrite F as:

$$F(\phi_1, \phi_2, u) = \phi_1 + u \circ (L + \phi_1) - L \circ u - \phi_2 \circ (\mathrm{Id} + u).$$

One can check that F is continuous and that the differential $\frac{\partial F}{\partial u}$ exists and it is also continuous; actually, $\frac{\partial F}{\partial u}$ is given by:

$$\big[\tfrac{\partial F}{\partial u}(\phi_1, \phi_2, u)v\big](x) = \big[v \circ (L + \phi_1)\big](x) - (L \circ v)(x) - \mathrm{d}\phi_2\big(x + u(x)\big)v(x),$$

for all $v \in C^0_{\mathrm{bu}}(\mathbb{R}^n)$, $x \in \mathbb{R}^n$. In particular:

$$\tfrac{\partial F}{\partial u}(0, 0, 0)v = v \circ L - L \circ v,$$

and thus, by Lemma B.20, $\frac{\partial F}{\partial u}(0, 0, 0)$ is an isomorphism. By the version of the implicit function theorem given in Exercise B.8, we can find $\varepsilon_1, \delta_1 > 0$ such that for every $\phi_1, \phi_2 \in C^1_{\mathrm{bu}}(\mathbb{R}^n)$ with $\|\phi_1\|_{C^1} < \varepsilon_1$, $\|\phi_2\|_{C^1} < \varepsilon_1$, there exists a unique $u \in C^0_{\mathrm{bu}}(\mathbb{R}^n)$ with $\|u\|_{C^0} < \delta_1$ and $F(\phi_1, \phi_2, u)$=0; moreover, the map $(\phi_1, \phi_2) \mapsto u$ is continuous. Thus, setting $\delta = \frac{\delta_1}{2}$, we can choose $\varepsilon > 0$, $\varepsilon < \varepsilon_1$, such that if $\|\phi_1\|_{C^1} < \varepsilon$, $\|\phi_2\|_{C^1} < \varepsilon$, then the corresponding map u satisfies $\|u\|_{C^0} < \delta$. Obviously, given $\phi \in C^1_{\mathrm{bu}}(\mathbb{R}^n)$ with $\|\phi\|_{C^1} < \varepsilon$, there exists a unique $u \in C^0_{\mathrm{bu}}(\mathbb{R}^n)$ such that $\|u\|_{C^0} < \delta$ and $F(\phi, 0, u) = 0$, i.e., such that (B.22) holds. It remains to prove that $\mathrm{Id} + u$ is a homeomorphism of $\mathbb{R}^n$. Let $v \in C^0_{\mathrm{bu}}(\mathbb{R}^n)$ be the unique map with $\|v\|_{C^0} < \delta$ and $F(0, \phi, v) = 0$. We have:

$$(\mathrm{Id} + v) \circ L = (L + \phi) \circ (\mathrm{Id} + v). \tag{B.23}$$

From (B.22) and (B.23) we get:

$$(\mathrm{Id} + v) \circ (\mathrm{Id} + u) \circ (L + \phi) = (L + \phi) \circ (\mathrm{Id} + v) \circ (\mathrm{Id} + u),$$
$$(\mathrm{Id} + u) \circ (\mathrm{Id} + v) \circ L = L \circ (\mathrm{Id} + u) \circ (\mathrm{Id} + v);$$

in other words:

$$F\big(\phi, \phi, (\mathrm{Id} + v) \circ (\mathrm{Id} + u) - \mathrm{Id}\big) = 0, \quad F\big(0, 0, (\mathrm{Id} + u) \circ (\mathrm{Id} + v) - \mathrm{Id}\big) = 0.$$

Observe now that $(\mathrm{Id} + v) \circ (\mathrm{Id} + u) - \mathrm{Id} \in C^0_{\mathrm{bu}}(\mathbb{R}^n)$ and that:

$$\big\|(\mathrm{Id} + v) \circ (\mathrm{Id} + u) - \mathrm{Id}\big\|_{C^0} = \big\|u + v \circ (\mathrm{Id} + u)\big\|_{C^0} < 2\delta = \delta_1;$$

thus, $(\mathrm{Id} + v) \circ (\mathrm{Id} + u) - \mathrm{Id} = 0$. Similarly, $(\mathrm{Id} + u) \circ (\mathrm{Id} + v) - \mathrm{Id} = 0$ and thus $\mathrm{Id} + u$ and $\mathrm{Id} + v$ are mutually inverse homeomorphisms. $\square$

B.22. LEMMA. *Let $X : M \to TM$ be a smooth vector field, $p \in M$ be a hyperbolic singularity of X and let $Z \subset M$ be an open set containing the whole unstable manifold $W_{\mathrm{u}}(p)$, except possibly for p itself (i.e., $W_{\mathrm{u}}(p) \setminus \{p\} \subset Z$). Then there exists a neighborhood $V \subset M$ of p such that for every $x \in V$, either $x \in W_{\mathrm{s}}(p)$ or $t \cdot x \in Z$ for some $t > 0$.*

PROOF. By the Theorem of Hartman–Grobman (Theorem B.16) we can find an open neighborhood $U \subset M$ of p and an open neighborhood $\widetilde{U} \subset T_pM$ of the origin such that $X|_U$ is topologically conjugated to $\mathrm{d}X(p)|_{\widetilde{U}}$. Let $\varphi : U \to \widetilde{U}$ be a homeomorphism such that $\mathrm{d}X(p)|_{\widetilde{U}}$ is φ-related to $X|_U$. We set $A = \mathrm{d}X(p)$, $(T_pM)_+ = (T_pM)_+(A)$, $(T_pM)_- = (T_pM)_-(A)$ and we denote by A_+, A_- respectively the endomorphisms of $(T_pM)_+$, $(T_pM)_-$ obtained by taking restrictions of A. The flow line of A passing through $(v_+, v_-) \in T_pM = (T_pM)_+ \oplus (T_pM)_-$ at $t = 0$ is given by:

$$\mathbb{R} \ni t \longmapsto \big(e^{tA_+}v_+, e^{tA_-}v_-\big) \in (T_pM)_+ \oplus (T_pM)_-. \tag{B.24}$$

Choose an arbitrary norm on T_pM and constants $\lambda_+, \lambda_- \in \mathbb{R}$ such that:

$$\max_{\lambda \in \sigma(A_-)} \Re(\lambda) < \lambda_- < 0 < \lambda_+ < \min_{\lambda \in \sigma(A_+)} \Re(\lambda).$$

By Lemma B.2 we can find a constant $C \geq 1$ such that:

$$\big\|e^{-tA_+}\big\| \leq Ce^{-t\lambda_+}, \quad \big\|e^{tA_-}\big\| \leq Ce^{t\lambda_-},$$

for all $t \geq 0$. Thus, for $v_+ \in (T_pM)_+$, $v_- \in (T_pM)_-$, we have:

$$\big\|e^{-tA_+}v_+\big\| \leq Ce^{-t\lambda_+}\|v_+\| \leq C\|v_+\|, \tag{B.25}$$

$$\big\|e^{tA_-}v_-\big\| \leq Ce^{t\lambda_-}\|v_-\| \leq C\|v_-\|, \tag{B.26}$$

for all $t \geq 0$. Choose $r > 0$ such that:

$$\mathrm{B}\big(0, r; (T_pM)_+\big) \times \mathrm{B}\big(0, r; (T_pM)_-\big) \subset \widetilde{U}. \tag{B.27}$$

Choosing $r' > 0$ with $r' < \frac{r}{C}$ then inequalities (B.25) and (B.26) show that the closed ball $\mathrm{B}\big[0, r'; (T_pM)_+\big]$ (resp., the closed ball $\mathrm{B}\big[0, r'; (T_pM)_-\big]$) is contained in the unstable manifold (resp., in the stable manifold) of the origin with respect to the vector field $A|_{\widetilde{U}}$. Thus, by the result of Exercise B.7, we have:

$$\varphi^{-1}\Big(\mathrm{B}\big[0, r'; (T_pM)_+\big]\Big) \subset W_{\mathrm{u}}(p), \quad \varphi^{-1}\Big(\mathrm{B}\big[0, r'; (T_pM)_-\big]\Big) \subset W_{\mathrm{s}}(p).$$

It follows that the open set $\varphi(Z \cap U) \subset T_pM$ contains $\mathrm{B}\big[0, r'; (T_pM)_+\big] \setminus \{0\}$. Since the sphere $S\big[0, r'; (T_pM)_+\big]$ is compact, we can find $\varepsilon > 0$ such that:

$$S\big[0, r'; (T_pM)_+\big] \times \mathrm{B}\big(0, \varepsilon; (T_pM)_-\big) \subset \varphi(Z \cap U). \tag{B.28}$$

Now choose $\varepsilon' > 0$ with $\varepsilon' \leq r'$ and $\varepsilon' \leq \frac{\varepsilon}{C}$. We claim that:

$$V = \varphi^{-1}\Big(\mathrm{B}\big(0, r'; (T_pM)_+\big) \times \mathrm{B}\big(0, \varepsilon'; (T_pM)_-\big)\Big)$$

is the neighborhood of p we are looking for. Namely, choose $x \in V$ and write $\varphi(x) = (v_+, v_-) \in T_pM$, so that $\|v_+\| < r'$ and $\|v_-\| < \varepsilon'$. If $v_+ = 0$ then

$\varphi(x)$ is in $\mathrm{B}\big[0, r'; (T_pM)_-\big]$, so that $x \in W_{\mathrm{s}}(p)$. Now assume that $v_+ \neq 0$. By the choice of ε' and by inequality (B.26) we have:

$$\big\|e^{tA_-} v_-\big\| < r, \tag{B.29}$$

$$\big\|e^{tA_-} v_-\big\| < \varepsilon, \tag{B.30}$$

for all $t \geq 0$. Also, for $t \geq 0$, we have:

$$\big\|e^{tA_+} v_+\big\| \geq \big\|e^{-tA_+}\big\|^{-1} \|v_+\| \geq \frac{1}{C} e^{t\lambda_+} \|v_+\|;$$

since $v_+ \neq 0$, we have $\lim_{t \to +\infty} \big\|e^{tA_+} v_+\big\| = +\infty$. Moreover, since $\|v_+\| < r'$, there exists $t > 0$ such that $\big\|e^{tA_+} v_+\big\| = r'$; denote by $t_0 > 0$ the least such t. Then:

$$\big\|e^{tA_+} v_+\big\| \leq r' < r,$$

for $t \in [0, t_0]$; by (B.27) and (B.29), the flow line (B.24) stays in $\widetilde{U}$ for $t \in [0, t_0]$. Thus:

$$[0, t_0] \ni t \longmapsto \varphi^{-1}\big(e^{tA_+} v_+, e^{tA_-} v_-\big) \in U \subset M$$

is a flow line of X that passes through x at $t = 0$. Moreover, by (B.28) and (B.30), we have:

$$\big(e^{t_0 A_+} v_+, e^{t_0 A_-} v_-\big) \in S\big[0, r'; (T_pM)_+\big] \times \mathrm{B}\big(0, \varepsilon; (T_pM)_-\big) \subset \varphi(Z \cap U),$$

so that:

$$t_0 \cdot x = \varphi^{-1}\big(e^{t_0 A_+} v_+, e^{t_0 A_-} v_-\big) \in Z.$$

This concludes the proof. $\square$

EXERCISE B.1. Let M be a manifold and let $(N_i)_{i \in I}$ be a family of embedded submanifolds of M. Assume that for all $i, j \in I$, the set $N_i \cap N_j$ is open in N_i and in N_j. Assume also that there exists a countable subset $J \subset I$ such that $\bigcup_{i \in I} N_i = \bigcup_{i \in J} N_i$. Then there exists a unique manifold structure on $N = \bigcup_{i \in I} N_i$ such that N is an immersed submanifold of M and such that, for all $i \in I$, N_i is open in N with respect to the manifold topology. Moreover, for all $i \in I$, the manifold structure that N_i inherits as an open subset of N is equal to the manifold structure that makes N_i an embedded submanifold of M.

EXERCISE B.2. Let V be a real vector space endowed with an inner product $\langle \cdot, \cdot \rangle$ and denote by $\|\cdot\|$ the corresponding norm. Assume that the complexification $V^{\mathbb{C}}$ of V is endowed with the unique Hermitean product that extends $\langle \cdot, \cdot \rangle$ and with the corresponding norm given by $\|x + iy\| = \big(\|x\|^2 + \|y\|^2\big)^{\frac{1}{2}}$, for all $x, y \in V$. Show that if $A : V \to V$ is a linear map and $A^{\mathbb{C}} : V^{\mathbb{C}} \to V^{\mathbb{C}}$ denotes the complexification of V then $\|A\| = \|A^{\mathbb{C}}\|$, i.e.:

$$\sup_{\substack{x \in V \\ \|x\| \leq 1}} \big\|A(x)\big\| = \sup_{\substack{x, y \in V \\ \|x + iy\| \leq 1}} \big\|A^{\mathbb{C}}(x + iy)\big\|.$$

EXERCISE B.3. Let $X : M \to TM$ be a smooth vector field on a manifold M and let $p \in M$ be a singularity of X. Show that the covariant derivative of X at p defines a linear endomorphism $\nabla X(p)$ of T_pM that does not depend on the choice of the connection. Given a local chart $\varphi : U \to \widetilde{U} \subset \mathbb{R}^n$ on M with $p \in U$, show that the linear map:

$$\nabla X(p) = \mathrm{d}\varphi(p)^{-1} \circ \mathrm{d}\widetilde{X}\big(\varphi(p)\big) \circ \mathrm{d}\varphi(p) : T_pM \longrightarrow T_pM,$$

does not depend on the choice of the chart φ, where $\widetilde{X}\big(\varphi(x)\big) = \mathrm{d}\varphi_x\big(X(x)\big)$ denotes the coordinate representation of X. Show that both definitions of $\nabla X(p)$ given above coincide. Denote by 0_p the zero vector of T_pM and consider the direct sum decomposition:

$$T_{0_p}(TM) = H \oplus V, \tag{B.31}$$

where V is the vertical space at 0_p and H is the tangent space to the zero section of TM at 0_p. If $\pi_V : T_{0_p}(TM) \to V \cong T_pM$ denotes the projection with respect to the decomposition (B.31), show that the linear map $\nabla X(p)$ equals the composite of π_V with $\mathrm{d}X(p) : T_pM \to T_{0_p}M$.

EXERCISE B.4. Let M be a manifold, $X : M \to TM$ be a smooth vector field and $N \subset M$ be an immersed submanifold. If $X(x) \in T_xN$ for all $x \in N$, show that $X|_N : N \to TN$ is a smooth vector field on N.

EXERCISE B.5. Let M be a topological manifold and $\varphi : U \to \widetilde{U}$ be a local chart on M, with U open in M and $\widetilde{U}$ open in $\mathbb{R}^n$. Let B be a subset of $\widetilde{U}$ such that $\varphi^{-1}(B)$ is closed in M. Let $\gamma : [a, b] \to M$ be a continuous curve such that $\gamma(a) \notin \varphi^{-1}(B)$, $\gamma(b) \in U$ and such that $\varphi\big(\gamma(b)\big)$ is in the interior of B in $\mathbb{R}^n$. Show that there exists $t \in \,]a, b[$ such that $\gamma(t) \in U$ and $\varphi\big(\gamma(t)\big)$ is in the boundary of B in $\mathbb{R}^n$.

EXERCISE B.6. Let M, N be manifolds, $X : M \to TM$, $Y : N \to TN$ be smooth vector fields and $f : M \to N$ be a continuous map.

- if f is of class C^1 then Y is f-related to X if and only if $\mathrm{d}f_x\big(X(x)\big) = Y\big(f(x)\big)$ for all $x \in M$;
- if f is a homeomorphism and Y is f-related to X then X is f^{-1}-related to Y. Moreover, if $\gamma : I \to M$ is a *maximal* integral curve of X then $f \circ \gamma$ is also a maximal integral curve of Y;
- "topological conjugacy" is an equivalence relation on the class of smooth vector fields on manifolds.

EXERCISE B.7. Let M, N be manifolds, $X : M \to TM$, $Y : N \to TN$ be smooth vector fields and $f : M \to N$ be a continuous map such that Y is f-related to X. Show that:

- f carries singularities of X to singularities of Y;
- if $p \in M$ is a hyperbolic singularity of X and $f(p)$ is a hyperbolic singularity of Y then $f\big(W_{\mathrm{s}}(p, X)\big) \subset W_{\mathrm{s}}\big(f(p), Y\big)$ and $f\big(W_{\mathrm{u}}(p, X)\big) \subset W_{\mathrm{u}}\big(f(p), Y\big)$. If f is a homeomorphism conclude that $f\big(W_{\mathrm{s}}(p, X)\big) = W_{\mathrm{s}}\big(f(p), Y\big)$ and $f\big(W_{\mathrm{u}}(p, X)\big) = W_{\mathrm{u}}\big(f(p), Y\big)$.

EXERCISE B.8. Prove the following version of the implicit function theorem. Let E_i, $i = 1, 2, 3$, be Banach spaces and let $f : U_1 \times U_2 \to E_3$ be a continuous map defined on an open subset $U_1 \times U_2 \subset E_1 \times E_2$. Let $(x_0, y_0) \in U_1 \times U_2$ be fixed and set $z_0 = f(x_0, y_0)$. Assume that for every $x \in U_1$, the map $f(x, \cdot) : U_2 \to E_3$ is differentiable and that $\frac{\partial f}{\partial y} : U_1 \times U_2 \to \mathrm{Lin}(E_2, E_3)$ is continuous. Then, if $\frac{\partial f}{\partial y}(x_0, y_0) : E_2 \to E_3$ is an isomorphism, we can find $r_1, r_2 > 0$ and a continuous map:

$$\sigma : \mathrm{B}(x_0, r_1; E_1) \longrightarrow \mathrm{B}(y_0, r_2; E_2)$$

such that $\mathrm{B}(x_0, r_1; E_1) \subset U_1$, $\mathrm{B}(y_0, r_2; E_2) \subset U_2$ and:

$$\big(\mathrm{B}(x_0, r_1; E_1) \times \mathrm{B}(y_0, r_2; E_2)\big) \cap f^{-1}(z_0) = \mathrm{Gr}(\sigma).$$

In other words, for every $x \in U_1$ with $\|x - x_0\| < r_1$, there exists a unique $y \in U_2$ with $\|y - y_0\| < r_2$ and $f(x, y) = z_0$; moreover, the map $x \mapsto y = \sigma(x)$ is continuous.

APPENDIX C

The Morse–Smale Condition

APPENDIX D

Floer Homology

APPENDIX E

Tightness & Tautness

An important step in the development of the theory initiated with the Chern-Lashof theorem (see Section 3.4 in this book or [**35**]) was the reformulation of the point of view in terms of critical point theory by Kuiper ([**79**]). He showed that for a given compact smooth[1] manifold M, the infimum of the total absolute curvature $\tau(f)$ over all immersions of M into all Euclidean spaces is the Morse number $\gamma(M)$, which is defined as the minimum number of critical points which any Morse function can possess (see also [**142, 160**]). Moreover, this lower bound is attained if and only if every Morse height function in the ambient space has $\gamma(M)$ critical points on M. Such an immersion f is said to have *minimum total absolute curvature*.

Further development and reformulation came with the introduction by Kuiper in [**83**] of a concept of generalized convexity in terms of intersections with half-spaces and injectivity of induced maps on homology. Note that the designation "tight" in this context was first used by Banchoff in [**7**] in conjunction with his introduction of the two-piece property. An immersion f of a compact manifold M into an Euclidean space is said to be *tight with respect to the field of coefficients* F (or, for short, F*-tight*) if the induced homomorphism

$$H_*(f^{-1}\mathcal{H}; F) \to H_*(M; F) \tag{E.1}$$

in singular homology is injective for almost every closed half-space $\mathcal{H}$ in the ambient Euclidean space, whereas f is said to have the *two-piece property* (TPP) if $f^{-1}\mathcal{H}$ is connected for every closed half-space $\mathcal{H}$ in the ambient Euclidean space. It can be easily shown that in both of these definitions we need only to consider half-spaces $\mathcal{H}$ which are defined by height functions that restrict to Morse functions on M. Plainly, then, we see that every tight immersion has the TPP. It is also interesting to notice that these properties are invariant under projective transformations, in the sense that one adds a hyperplane at infinity and considers images of submanifolds under projective transformations that do not meet the hyperplane at infinity.

An equivalent definition of F-tightness for an immersion $f : M \to \mathbb{R}^m$ is requiring that every height function $h_\xi(x) = \langle f(x), \xi \rangle$, $x \in M$, which is a Morse function has the property that its number of critical points is equal to the sum of the Betti numbers of M relative to F, i. e. h_ξ is F*-perfect* Likewise, the TPP for f is

[1] In spite of the fact that there is an interesting theory of topological and polyhedral tight & taut immersions (see [**78, 85**]), we shall restrict our discussion to smooth immersions.

equivalent to the requirement that every height function h_ξ which is a Morse function has exactly one maximum and one minimum on M.) It follows that a F-tight immersion of a compact manifold has minimum total absolute curvature, since total absolute curvature is the mean number of critical points of height functions on M. Note that in this case we also have that the Morse number $\gamma(M)$ equals the sum of the Betti numbers of M relative to F. Hence, we see that the concepts of F-tightness and minimum total absolute curvature are equivalent for immersions of manifolds satisfying the condition that $\gamma(M)$ equals the sum of the Betti numbers of M relative to F, but there are examples of manifolds immersed with minimal total absolute curvature where this condition does not hold (see [**86**] for the case where $F = \mathbb{Z}_2$).

An important observation of Kuiper regarding the codimension of substantial[2] tight immersions into Euclidean spaces appeared already in his first papers [**79, 82**] on the subject: a substantial immersion f of a compact n-dimensional manifold that satisfies the TPP admits a point where the second osculating space coincides with the ambient space. Here the second osculating space of f at p is the affine space spanned by the first and second partial derivatives of f at p. Counting these derivatives shows that the dimension of the second osculating space can be at most $\frac{1}{2}n(n+3)$. Therefore the codimension of the immersion can be at most $\frac{1}{2}n(n+1)$. The Veronese embedding of the real projective space $\mathbb{R}P^n$ is tight in $\mathbb{R}^{\frac{1}{2}n(n+3)}$ showing that this estimate is optimal.

In the case of surfaces, tightness, minimum total absolute curvature and the TPP are all equivalent concepts. Therefore, as observed in [**36**], the tightly embedded compact oriented surfaces in $\mathbb{R}^3$ are precisely the oriented surfaces in $\mathbb{R}^3$ with the property that points of positive Gauss curvature lie on the boundary of the convex hull of the surface. As to non orientable surfaces, all these admit tight immersions into $\mathbb{R}^3$, but the projective plane, the Klein bottle and the projective plane with one handle which are prohibited; these results were proved by Kuiper in [**80, 81**], except for the case of the projective plane with one handle which was solved much later by Haab (see [**65**]).

The easiest examples of tight surfaces in $\mathbb{R}^4$ are the tori given by products of two convex curves and the stereographic projection of the Veronese embedding of the projective plane in S^4 to $\mathbb{R}^4$. Otherwise, all compact orientable surfaces admit substantial tight immersions into $\mathbb{R}^4$ (see [**84**]) but the two-sphere, which is prohibited by the Chern-Lashof theorem. The non orientable surfaces with the exception of the Klein bottle were claimed also to admit such immersions by Kuiper in [**84, 85**], although he did not give concrete examples (see also [**33**], pp. 80–81). The case of the Klein bottle is still open.

The highest dimension of an Euclidean space into which a surface can be substantially and tightly embedded is five. Kuiper proved in [**83**] one of the most remarkable facts in the theory, namely that a substantial tight immersion of a surface in $\mathbb{R}^5$ is projectively equivalent to the Veronese embedding of the real projective

[2]An immersion $f : M \to \mathbb{R}^m$ is called *substantial* if its image does not lie in any affine hyperplane of $\mathbb{R}^m$.

plane. This result was generalized by Little and Pohl in [**89**] who proved that a substantial tight immersion of a compact n-dimensional manifold into $\mathbb{R}^{\frac{1}{2}n(n+3)}$ that satisfies the TPP is projectively equivalent to the Veronese embedding of the real projective space $\mathbb{R}P^n$. We remark that the standard embeddings of the other projective spaces (complex, quaternionic and octonionic) are also tight, namely FP^n embeds substantially and tightly into $\mathbb{R}^m$ for $m = n + d\frac{n(n+1)}{2}$, where $F = \mathbb{R}$, $\mathbb{C}$, $\mathbf{H}$, $\mathbf{O}$ and $d = \dim_{\mathbb{R}} F$ (recall that $\mathbf{O}P^n$ is defined only for $n = 2$). Further generalizations of these results were given for tight immersions of compact $2k$-dimensional manifolds that are $(k-1)$-connected but not k-connected (*highly connected manifolds*), see for example [**84, 85, 153**]. More examples of tight immersions will be given below when we discuss taut immersions.

The beginnings of the study of taut immersions can be traced back to Banchoff's paper [**8**] where he attempted to classify tight surfaces which lie in a Euclidean sphere $S^m \subset \mathbb{R}^{m+1}$. Since a hyperplane in $\mathbb{R}^{m+1}$ intersects S^m in a great or small $(m-1)$-sphere, the usual TPP is equivalent to the TPP with respect to hyperspheres in S^m for a spherical immersion. This problem is in turn equivalent via stereographic projection to the study of surfaces in $\mathbb{R}^m$ which have the TPP with respect to hyperspheres and hyperplanes, i. e. the *spherical two-piece property* (STPP). It turns out that a compact surface in $\mathbb{R}^3$ that satisfies the STPP is either a round sphere or a cyclide of Dupin (see [**8**]); the latter can all be constructed as the image of a torus of revolution under a Möbius transformation, where one has to permit that the axis of revolution can meet the generating circle[3].

It is easily seen that the STPP for an immersion $f : M \to \mathbb{R}^m$ is equivalent to the requirement that every Morse distance function $L_q(x) = |f(x) - q|^2$, $q \in \mathbb{R}^m$, have exactly one maximum and one minimum on M. Carter and West generalized the STPP in [**26**] and defined an immersion f of a compact manifold to be *taut with respect to the field* F (or F*-taut*, for short) if every Morse distance function L_q has the minimum number of critical points allowed by the Morse inequalities with respect to F. It follows from this definition that a taut immersion f must be an embedding, for if q is was double point in the image then the distance function L_q would have two minima and one could then perturb q, if necessary, in order to obtain a Morse distance function with two local minima. Moreover, as was done for tightness, one sees that a submanifold $M \subset \mathbb{R}^m$ is F-taut if and only if the induced homomorphism

$$H_*(M \cap B; F) \to H_*(M; F) \tag{E.2}$$

in singular homology is injective for almost every closed ball B in $\mathbb{R}^m$. It is then clear that tautness is conformally invariant. Since any intersection of a closed ball in $\mathbb{R}^m$ with S^{m-1} can also be given as the intersection of some closed half-space with S^{m-1}, it also follows that a tight spherical immersion is taut. Furthermore, one sees that a taut submanifold M is tight, because for any half-space $\mathcal{H}$ defined

[3]The cyclides were introduced by Dupin in [**41**] as the envelope of the family of spheres tangent to three fixed spheres. The characterization of the cyclides quoted above is due to Mannheim, see the account in [**88**].

by a Morse height function one can construct a closed ball B such that $M \cap \mathcal{H}$ is a strong deformation retract of $M \cap B$.

We next give some examples of tautly embedded submanifolds. The Clifford tori $S^{n_1}(r_1) \times \cdots \times S^{n_k}(r_k) \subset S^{n_1+\ldots+n_k}(1)$ where $r_1^2 + \ldots + r_k^2 = 1$, and the standard embeddings of the projective spaces FP^n, $F = \mathbb{R}, \mathbb{C}, \mathbb{H}$, **O** are taut, since these are tight spherical embeddings. In the case of spheres, a substantial taut embedding of a sphere must be spherical and of codimension one. In fact, such an $f : S^n \to \mathbb{R}^m$ is tight, whence $m = n + 1$ and $f(S^n)$ is a convex hypersurface by the Chern-Lashof theorem. Now stereographic projection maps $f(S^n)$ into a taut submanifold of $\mathbb{R}^{n+2}$ which cannot be substantial, again by the Chern-Lashof theorem. Therefore, we see that $f(S^n)$ is spherical. If M is an n-dimensional taut hypersurface in $\mathbb{R}^{n+1}$ which has the same integral homology as $S^k \times S^{n-k}$, then Cecil and Ryan proved in [**31**] that M has precisely two principal curvatures at each point and that the principal curvatures are constant along the corresponding curvature distributions. They called such hypersurfaces, compact or not, cyclides of Dupin. This generalizes the two-dimensional cyclides.

A very rich class of examples of tautly embedded submanifolds is given by the generalized (real) flag manifolds. Bott and Samelson introduced in [**21**] (see also [**16**]) the concept of variational completeness for isometric group actions.

Roughly speaking, the action of a compact connected Lie group on a complete Riemannian manifold is *variationally complete* if it produces enough Jacobi fields along geodesics to determine the multiplicities of focal points to the orbits. They proved that the orbits of variationally complete linear representations are tautly embedded with respect to $\mathbb{Z}_2$ coefficients, and that the isotropy representations of symmetric spaces are variationally complete. As a consequence, the orbits of the isotropy representations of the symmetric spaces, which are called *generalized flag manifolds* (or *R-spaces*, although this now seems to be an older terminology), are all $\mathbb{Z}_2$-taut submanifolds. It is very interesting to remark, perhaps owing to the fact that Bott and Samelson neither state their result in our terminology nor mention total absolute curvature, how far their result may have gone unnoticed in this context, as Takeuchi and Kobayashi reproved it later independently in [**147**]. On another note, a characterization of the *symmetric* generalized flag manifolds was given by Ferus in [**45**] who showed these to be the only compact extrinsically symmetric submanifolds of Euclidean space. He also used this characterization to give another, elegant proof of the tautness of these submanifolds.

The generalized flag manifolds are homogeneous examples of submanifolds which belong to another very important, more general class of submanifolds, called isoparametric submanifolds. The theory of isoparametric hypersurfaces has a long story that goes way back but that can be said to have in É. Cartan the first one of its main contributors (see the survey [**157**]). An *isoparametric hypersurface* in a simply-connected real space form is a hypersurface with constant principal curvatures. In the course of his work on the subject, Cartan noticed that isoparametric hypersurfaces in spheres are a much more rich and difficult object of study than its counterparts in Euclidean and hyperbolic spaces. In fact, until today there is no complete classification of them. The subject seems to have been forgotten for over thirty years after Cartan, when Münzner (see also [**109**]) wrote the two very influential papers [**105, 106**]. Using his results, Cecil and Ryan observed in [**32**] that isoparametric hypersurfaces and their focal manifolds are taut.

In the eighties, some generalizations of the concept of isoparametric hypersurface to higher codimensions were proposed, but the today commonly accepted one seems to have been first given by Harle in [**67**] (see also Carter and West [**27, 28**] and Terng [**148**]). An *isoparametric submanifold* of a simply-connected space form is a submanifold whose normal bundle is flat and such that, for every locally defined parallel normal vector field, the eigenvalues of the corresponding Weingarten operator are constant. Examples of inhomogeneous isoparametric hypersurfaces in spheres were constructed in [**114, 115**] and, more systematically, in [**46**]. In contrast, Thorbergsson proved in [**155**] the striking result that a compact irreducible isoparametric submanifold of substantial codimension greater than or equal to 3 in an Euclidean space is homogeneous (see [**70, 110**] for other proofs of this fact), and then it follows from [**118**] that it must be a generalized flag manifold. Hsiang, Palais and Terng studied in [**74**] the topology of isoparametric submanifolds and proved, among other things, that they and their focal submanifolds are taut. This result also follows from the work of Thorbergsson in [**152**]. Both in [**74**] and [**152**] the method to prove tautness is to use curvature spheres to construct explicit cycles that represent a basis for the $\mathbb{Z}_2$-homology, which can be viewed as a generalization of the method of Bott and Samelson to show that the generalized flag manifolds are taut.

Another related class of submanifolds are the Dupin hypersurfaces. Pinkall introduced this class in [**130**] (see also [**131**]) as a simultaneous generalization of the cyclides of Dupin referred to above and of isoparametric hypersurfaces. Let M be an immersed hypersurface in a real space form. A *curvature surface* of M is a smooth submanifold S such that for every $x \in M$ the tangent space T_xS is a maximal eigenspace of the Weingarten operator of M at x. We say that M is a *Dupin hypersurface* if a continuous principal curvature function on M is constant along the corresponding curvature surfaces of M. If in addition the multiplicities of the principal curvatures are constant on M, we say that M is a *proper Dupin hypersurface*. ike tautness, the Dupin and proper Dupin conditions are invariant under Möbius transformations and under stereographic projection.

It follows from the Codazzi equation that if the dimension of a curvature surface S of an arbitrary hypersurface is greater than one, then the corresponding principal curvature is constant on S and S is an open subset of an umbilical submanifold of the space form of dimension equal to the multiplicity of the principal curvature. Since the definition of Dupin does not insist on the *existence* of curvature surfaces, one has only to check whether each principal curvature is constant along each of its lines of curvature in order to verify the Dupin condition.

The natural framework for the study of Dupin hypersurfaces is Lie sphere geometry (see [**29, 131**]), which is a contact geometry and was introduced by Lie. One reason for introducing Lie sphere geometry is that a parallel hypersurface to a Dupin hypersurface is also Dupin in some sense, even if it may develop singularities. This situation is similar to the singularities of the cyclides. It turns out that the Dupin and proper Dupin conditions are invariant under the group of Lie sphere transformations, which is generated by the Möbius transformations and the parallel transformations. Obviously, the image of an isoparametric hypersurface in S^m under stereographic projection from a point not in the hypersurface is a compact proper Dupin hypersurface embedded in $\mathbb{R}^m$. Similarly, the image of isoparametric hypersurface in S^m under a Lie sphere transformation of S^m is a compact proper Dupin hypersurface embedded in S^m, but not all compact proper Dupin hypersurfaces embedded in S^m are obtained this way as the examples in [**133, 100**] show.

Thorbergsson showed in [**152**] that a complete proper Dupin hypersurface embedded in $\mathbb{R}^n$ is taut. Pinkall [**132**] and Miyaoka [**99**] then independently showed that a taut hypersurface is Dupin (not necessarily proper). More generally, a tube around a taut submanifold is Dupin.

There is a very interesting theorem by Ozawa ([**113**]) which states that the set of critical points of a distance function of a taut submanifold decomposes into critical submanifolds which are nondegenerate in the sense of Bott. As a first application, we see that the injectivity of the homomorphism (E.2) holds for *every* closed ball B in the ambient space if the manifold M is tautly embedded. As another application, one sees rather easily that the subclass of Dupin hypersurfaces given by the the taut hypersurfaces admit curvature surfaces through any given point and any given maximal eigenspace of the Weingarten operator at that point. To this day it remains a difficult problem to establish whether a compact Dupin hypersurface admitting existence of curvature surfaces as above needs to be taut.

Most of the examples of taut embeddings known are homogeneous spaces. In [**154**] Thorbergsson posed some questions regarding the problem of which homogeneous spaces admit taut embeddings and derived some necessary topological conditions for the existence of a taut embedding which allowed him to conclude that certain homogeneous spaces cannot be tautly embedded (see also [**69**]), among others the lens spaces distinct from the real projective space. Olmos showed in [**111**] that a compact homogeneous submanifold embedded in Euclidean space with a flat normal bundle is a generalized flag manifold. Many proofs have been given of the tautness of special cases of generalized flag manifolds where the arguments are easier. No new examples of taut embeddings of homogeneous spaces besides the generalized flag manifolds were known until Gorodski and Thorbergsson classified in [**61**] the irreducible representations of compact Lie groups all of whose orbits are tautly embedded. It turns out that the classification includes three new representations which are not isotropy representations of symmetric spaces, thereby supplying many new examples of ($\mathbb{Z}_2$-) tautly embedded homogeneous spaces. In [**62**] Gorodski and Thorbergsson provided another proof of the $\mathbb{Z}_2$-tautness of those orbits by adapting the construction of the cycles of Bott and Samelson to that case. It is interesting to remark that those three representations coincide precisely with the representations of cohomogeneity three of the compact Lie groups which are not orbit-equivalent to the isotropy representation of a symmetric space.

Tautness was generalized to immersions into arbitrary complete Riemannian manifolds by Grove and Halperin ([**64**]) and, independently, by Terng and Thorbergsson ([**150**]), and it follows from the work of Bott and Samelson that orbits of variationally complete actions are taut. Let N be a complete Riemannian manifold. A proper immersion $\phi : M \to N$ is said to be *taut* if the energy functional $E_p : P(N, \phi \times p) \to \mathbb{R}$ is a perfect Morse function for every $p \in N$ that is not a focal point of M, where $P(N, \phi \times p)$ denotes the space of pairs (q, γ) such that $q \in M$ and γ a H^1-path $\gamma : [0, 1] \to N$ such that $(\gamma(0), \gamma(1)) = (\phi(q), p)$. In [**150**] it is proved that a taut immersion is an embedding if the range is simply-connected, and an analogue of Ozawa's theorem is stated and proved. The question of the tautness of a distance sphere is discussed and shown to be equivalent to the tautness of its center. All points in a compact symmetric space are easily seen to be taut, but the question of the existence of other simply-connected compact Riemannian manifolds all of whose points are taut is still open and turns out to be more general than the Blaschke conjecture (see [**15**] for a discussion of this conjecture).

This short account about the development of the the theory of tight & taut immersions has presented only a partial selection of topics. For a wider discussion and more details the reader is referred to the excellent surveys [**30, 156, 157**] and monograph [**33**], and to the references therein.

Bibliography

[1] A. Abbondandolo, *A New Cohomology for the Morse Theory of Strongly Indefinite Functionals on Hilbert Spaces*, Top. Math. Nonlin. Anal. **9** (1997), 325–382.

[2] A. Abbondandolo, *Morse Theory for Asymptotically Linear Hamiltonian Systems*, Nonlinear Anal. TMA **39** (2000), 997–1049.

[3] A. Abbondandolo, *Morse theory for Hamiltonian systems*, Pitman Research Notes in Mathematics 425, CRC, London, 2001.

[4] A. Abbondandolo, P. Mejer, *Morse Homology on Hilbert Spaces*, Commun. Pure Appl. Math. **54** (2001), 689–760.

[5] A. A. Agrachev, A. V. Sarychev, *Abnormal sub-Riemannian Geodesics: Morse Index and Rigidity*, Ann. Inst. Henri Poincaré **13**, n. 6 (1996), 635–690.

[6] V. I. Arnol'd, *Characteristic Class Entering in Quantization Conditions*, Funct. Anal. Appl. **1** (1967), 1–13.

[7] T. F. Banchoff. Tightly embedded 2-dimensional polyhedral manifolds. *Amer. J. Math.*, 87:462–472, 1965.

[8] T. F. Banchoff. The spherical two-piece property and tight surfaces in spheres. *J. Differential Geom.*, 4:193–205, 1970.

[9] J. K. Beem, *Conformal Changes and Geodesic Completeness*, Comm. Math. Phys. **49** (1976), 179–186.

[10] J. K. Beem, H. Buseman, *Axiom for Indefinite Metrics*, Cir. Mat. Palermo **15** (1966), 223–246.

[11] J. K. Beem, P. E. Ehrlich, K. L. Easley, *Global Lorentzian Geometry*, 2nd Edition, Marcel Dekker, Inc., New York and Basel, 1996.

[12] J. K. Beem, P. E. Parker, *Pseudoconvexity and Geodesic Connectedness*, Ann. Mat. Pura e Applicata **155,** No. 4, 1989, 137–142.

[13] V. Benci, *A New Approach to Morse–Conley Theory and Some Applications*, Ann. Mat. Pura e Appl. **158** (1991), 231–305.

[14] V. Benci, D. Fortunato, F. Giannoni, *On the Existence of Infinitely Many Geodesics on Space-Time Manifolds*, Ann. Inst. Henri Poincaré, Analyse Nonlinéaire, **8** (1991), 79–102.

[15] A. Besse. *Manifolds all of whose geodesics are closed*, volume 93 of *Ergeb. Math. Ihrer Grez.* Springer, Berlin, 1978.

[16] R. Bott. An application of the Morse theory to the topology of Lie groups. *Bull. Soc. Math. France*, 84:251–281, 1956.

[17] R. Bott, *Lectures on Morse Theory, Old and New*, Bull. Amer. Math. Soc. **7**, No. 2 (1982), 331–358.

[18] R. Bott, *The Periodicity Theorem for the Classical Groups*, Ann. of Math. (2) **70** 2 (1959), 179–203.

[19] R. Bott, *On the Iteration of Closed Geodesics and the Sturm Intersection Theorem*, Commun. Pure Appl. Marth. **9** (1956), 171–206.

[20] R. Bott, *Morse Theory Indomitable*, Inst. Hautes Études Sci. Publ. Math. **68** (1988), 99–114.

[21] R. Bott and H. Samelson. Applications of the theory of Morse to symmetric spaces. *Amer. J. Math.*, 80:964–1029, 1958. Correction in Amer. J. Math. **83** (1961), 207–208.

[22] H. Brezis, *Analyse Fonctionelle*, Masson, Paris, 1983.

[23] A. M. Candela, M. Sánchez, *Geodesic Connectedness in Gödel Type Space-Times*, Diff. Geom. Appl. **12** (2000), 105–120.

[24] M. do Carmo, *Riemannian Geometry*, Birkhäuser, Boston, 1992.

[25] M. do Carmo, E. L. Lima, *Immersions of Manifolds with Non-Negative Sectional Curvature*, Bol. Soc. Bras. Mat. **2**, vol. 2 (1971), 9–22.

[26] S. Carter and A. West. Tight and taut immersions. *Proc. London. Math. Soc.*, 25:701–720, 1972.

[27] S. Carter and A. West. Generalised Cartan polynomials. *J. London. Math. Soc.*, 32:305–316, 1985.

[28] S. Carter and A. West. Isoparametric systems and transnormality. *Proc. London. Math. Soc.*, 51:520–542, 1985.

[29] T. E. Cecil. *Lie sphere geometry*. Universitext. Springer, New York, 1992.

[30] T. E. Cecil. Taut and Dupin submanifolds. In T. E. Ryan and S.-S. Chern, editors, *Tight and Taut Submanifolds*, Math. Sci. Res. Inst. Publ. 32, pages 135–180. Cambridge University Press, 1997.

[31] T. E. Cecil and P. J. Ryan. Focal sets, taut embeddings and the cyclides of Dupin. *Math. Ann.*, 236:177–190, 1978.

[32] T. E. Cecil and P. J. Ryan. *Tight spherical embeddings*, pages 94–104. Number 838 in Lecture Notes in Math. Springer-Verlag, 1981.

[33] T. E. Cecil and P. J. Ryan. *Tight and Taut Immersions of Manifolds*. Number 107 in Research Notes in Mathematics. Pitman, 1985.

[34] K. C. Chang, *Infinite Dimensional Morse Theory and Multiple Solutions Problems*, Birkäuser, Basel, 1993.

[35] S. S. Chern and R. Lashof. On the total curvature of immersed manifolds. *Amer. J. Math.*, 79:306–318, 1957.

[36] S. S. Chern and R. Lashof. On the total curvature of immersed manifolds II. *Michigan Math. J.*, 5:5–12, 1958.

[37] E. A. Coddington, N. Levinson, *Theory of Ordinary Differential Equations*, McGraw–Hill Book Company, New York, Toronto, London, 1955.

[38] C. Conley, E. Zehnder, *The Birkhoff-Lewis Fixed Point Theorem and a Conjecture of V. I. Arnold*, Invent. Math. **73** (1983), 33–49.

[39] A. Dold, *Lectures on Algebraic Topology*, Second Edition, Springer–Verlag, Berlin, Heidelberg, New York, 1980.

[40] J. J. Duistermaat, *On the Morse Index in Variational Calculus*, Adv. in Math. **21** (1976), 173–195.

[41] C. Dupin. *Applications de géométrie et de méchanique*. Paris, 1822.

[42] H. M. Edwards, *A Generalized Sturm Theorem*, Ann. of Math. **80** (1964), 22–57.

[43] P. E. Ehrlich, S. Kim, *A Focal Index Theorem for Null Geodesics*, J. Geom. Phys. **6**, n. 4 (1989), 657–670.

[44] I. Ekeland, *An Index Theory for Periodic Solutions of Convex Hamiltonian Systems*, Proc. of Symposia in Pure Math. vol. 45 (1986), Part I, 395–423.

[45] D. Ferus. The tightness of extrinsic symmetric submanifolds. *Math. Z.*, 181:563–565, 1982.

[46] D. Ferus, H. Karcher, and H. F. Muenzner. Cliffordalgebren and neue isoparametrische Hyperflaechen. *Math. Z.*, 177:479–502, 1981.

[47] A. Floer, *An Instanton Invariant for* 3*-Manifolds*, Commun. Math. Phys. **118** (1988), 215–240.

[48] A. Floer, *Symplectic Fixed Points and Holomorphic Spheres*, Commun. Math. Phys. **120** (1989), 215–240.

[49] A. Floer, *Witten's Complex and Infinite Dimensional Morse Theory*, J. Diff. Geom. **30** (1989), 207–221.

[50] A. Floer, *A Relative Morse Index for the Symplectic Action*, Commun. Pure Appl. Math. **41** (1988), 393–407.

[51] A. Floer, *The Unregularized Gradient Flow of the Symplectic Action*, Commun. Pure Appl. Math. **41** (1988), 775–813.

[52] D. B. Fuks, *Maslov–Arnol'd Characteristic Classes*, Soviet Math. Dokl. **9**, n. 1 (1968), 96–99.

[53] R. Giambó, F. Giannoni, P. Piccione, *Existence, Multiplicity and Regularity for sub-Riemannian Geodesics by Variational Methods*, preprint 1999.

[54] F. Giannoni, A. Masiello, P. Piccione, D. Tausk, *A Generalized Index Theorem for Morse–Sturm Systems and Applications to semi-Riemannian Geometry*, preprint 1999, (`LANL math.DG/9908056`), to appear in the Asian Journal of Mathematics.

[55] F. Giannoni, A. Masiello, P. Piccione, *A Morse Theory for Massive Particles and Photons by Fermat Principles in General Relativity*, Journal of Geometry and Physics **35**, vol. 1 (2000), 1–35.

[56] F. Giannoni, A. Masiello, P. Piccione, *A Morse Theory for Light Rays in Stably Causal Lorentzian Manifolds*, Annales de l'Institut H. Poincaré - Physique Theorique **69**, no. 4 (1998), p. 359–412.

[57] F. Giannoni, P. Piccione, *An Intrinsic Approach to the Geodesical Connectedness of Stationary Lorentzian Manifolds*, Communications in Analysis and Geometry **7**, n. 1 (1999), p. 157–197.

[58] F. Giannoni, P. Piccione, *An Existence Theory for Relativistic Brachistochrones in Stationary Spacetimes*, J. Math. Phys. **39** (1998), vol. 11, 6137–6152.

[59] F. Giannoni, P. Piccione, R. Sampalmieri, *On the Geodesical Connectedness for a Class of Semi-Riemannian Manifolds*, Journal of Mathematical Analysis and Applications **252** (2000), no. 1, 444–476.

[60] F. Giannoni, P. Piccione, J. A. Verderesi, *An Approach to the Relativistic Brachistochrone Problem by sub–Riemannian Geometry*, J. Math. Phys. **38**, n. 12 (1997), 6367–6381.

[61] C. Gorodski and G. Thorbergsson. Representations of compact Lie groups and the osculating spaces of their orbits. preprint, Univ. of Cologne, 2000.

[62] C. Gorodski and G. Thorbergsson. Cycles of Bott-Samelson type for taut representations. preprint, Univ. of Cologne, 2001.

[63] D. Gromoll, W. Meyer, *Periodic geodesics on compact riemannian manifolds*, J. Differential Geometry **3** (1969), 493–510.

[64] K. Grove and S. Halperin. Elliptic isometries, condition (C) and proper maps. *Arch. Math. (Basel)*, 56:288–299, 1991.

[65] F. Haab. Immersions tendues de surfaces dans E^3. *Comment. Math. Helv.*, 67:182–202, 1992.

[66] U. Hamenstädt, *Some Regularity Theorems for Carnot–Carathéodory Metrics*, J. Diff. Geom. **32** (1990), 819–850.

[67] C. E. Harle. Isoparametric families of submanifolds. *Bol. Soc. Brasil. Mat.*, 13:35–48, 1982.

[68] S. W. Hawking, G. F. Ellis, *The Large Scale Structure of Spacetime*, Cambridge Univ. Press, London, New York, 1973.

[69] J. Hebda. The possible cohomology ring of certain types of taut submanifolds. *Nagoya Math. J.*, 111:85–97, 1988.

[70] E. Heintze and X. Liu. Homogeneity of infinite-dimensional isoparametric submanifolds. *Ann. Math.*, 149:149–181, 1999.

[71] A. D. Helfer, *Conjugate Points on Spacelike Geodesics or Pseudo-Self-Adjoint Morse-Sturm-Liouville Systems*, Pacific J. Math. **164**, n. 2 (1994), 321–340.

[72] A. D. Helfer, *Conjugate Points and Higher Arnol'd–Maslov Classes*, Contemporary Mathematics vol. 170 (1994), 135–147.

[73] M. W. Hirsch, *Differential Topology*, Springer–Verlag, 1976.

[74] W.-Y. Hsiang, R. S. Palais, and C.-L. Terng. The topology of isoparametric submanifolds. *J. Differential Geom.*, 27:423–460, 1988.

[75] N. Jacobson, Basic Algebra I, second edition, W. H. Freeman and Co., New York, 1985.

[76] I. Kishimoto, *The Morse Index Theorem for Carnot–Carathéodory Spaces*, J. Math. Kyoto Univ. **38**, 2 (1998), 287–293.

[77] S. Kobayashi, K. Nomizu, *Foundations of Differential Geometry*, vol. 2, Interscience, J. Wiley, 1969.

[78] W. Kuehnel. *Tight polyhedral submanifolds and tight triangulations*, volume 1612 of *Lecture Notes in Math.* Springer, 1995.

[79] N. H. Kuiper. Immersions with minimal total absolute curvature. In *Coll. de géométrie diff.*, pages 75–88. Centre Belge de Recherches Math., Bruxelles, 1958.

[80] N. H. Kuiper. On surfaces in euclidean three-space. *Bull. Soc. Math. Belg.*, 12:5–22, 1960.

[81] N. H. Kuiper. Convex immersions of closed surfaces in E^3. Nonorientable closed surfaces in E^3 with total minimal absolute Gauss-curvature. *Comment. Math. Helv.*, 35:85–92, 1961.

[82] N. H. Kuiper. Sur les immersions à courbure totale minimale. In *Séminaire de Topologie et Géometrie Différentielle C. Ereshmann, Paris*, volume II. 1961. Recueil d'exposés faits en 1958-1959-1960.

[83] N. H. Kuiper. On convex maps. *Nieuw Archief voor Wisk.*, 10:147–164, 1962.

[84] N. H. Kuiper. Minimal total absolute curvature for immersions. *Invent. Math.*, 10:209–238, 1970.

[85] N. H. Kuiper. Tight embeddings and maps. submanifolds of geometrical class three in E^n. In *The Chern Symposium 1979*, pages 97–145. Springer-Verlag, 1980.

[86] N. H. Kuiper and W. Meeks III. Total curvature for knotted surfaces. *Invent. Math.*, 77:25–69, 1984.

[87] S. Lang, *Differential Manifolds*, Springer-Verlag, Berlin, 1985.

[88] R. v. Lilienthal. *Besondere Flaechen*, volume III of *Enzyklopaedie der Math. Wissenschaften*, chapter 3. Geometrie. Teubner, Leipzig, 1902–1927.

[89] J. A. Little and W. F. Pohl. On tight immersions of maximal codimension. *Invent. Math.*, 13:179–204, 1971.

[90] Chun-gen Liu, Y. Long, *Iteration inequalities of the Maslov-type index theory with applications*, J. Differential Equations **165** (2000), no. 2, 355–376.

[91] W. Liu, H. J. Sussmann, *Shortest Paths for Sub-Riemannian Metrics on Rank–*2 *Distribution*, Memoirs AMS **564**, vol. 118, 1995.

[92] Y. Long, *Precise iteration formulae of the Maslov-type index theory and ellipticity of closed characteristics*, Adv. Math. **154** (2000), no. 1, 76–131.

[93] Y. Long, X. Xu, *Periodic solutions for a class of nonautonomous Hamiltonian systems*, Nonlinear Anal. **41** (2000), no. 3-4, Ser. A: Theory Methods, 455–463.

[94] A. Marino, G. Prodi, *Metodi Perturbativi nella Teoria di Morse*, Boll. Un. Mat. Ital. (4) 11 (1975), no. 3, suppl., 1–32.

[95] A. Masiello, *Variational Methods in Lorentzian Geometry*, Pitman Res. Notes in Math. **309**, London, 1994.

[96] W. S. Massey, *A basic course in algebraic topology*, Graduate Texts in Mathematics, 127. Springer-Verlag, New York, 1991.

[97] J. Mawhin, M. Willem, *Critical Point Theory and Hamiltonian Systems*, Springer-Verlag, Berlin, 1989.

[98] J. Milnor, *Morse Theory*, Annals of Mathematics Studies n. 51, Princeton Univ. Press, 1969.

[99] R. Miyaoka. Taut embeddings and Dupin hypersurfaces. In *Differential geometry of submanifolds, Kyoto, 1984*, number 1090 in Lecture Notes in Math., pages 15–23. Springer, 1984.

[100] R. Miyaoka and T. Ozawa. Construction of taut embeddings and Cecil-Ryan conjecture. In *Geometry of manifolds, Matsumoto, 1988*, number 8 in Perpect. Math., pages 181–189. Academic Press, 1989.

[101] F. Mercuri, P. Piccione, D. Tausk, *Stability of the Focal and the Geometric Index in semi-Riemannian Geometry via the Maslov Index*, Technical Report RT-MAT 99-08, Mathematics Department, University of São Paulo, Brazil, 1999. (`LANL math.DG/9905096`)

[102] R. Montgomery, *Singular Extremals on Lie Groups*, Math. Control Signals **7** (1994), 217–234.

[103] R. Montgomery, *Abnormal minimizers*, SIAM J. Control Optimization **32**, No. 6 (1994), 1605–1620.

[104] R. Montgomery, *A Survey of Singular Curves in Sub-Riemannian Geometry*, J. Dyn. Contr. Sys. **1** (1995), 49–90.

[105] H. F. Muenzner. Isoparametrische Hyperflaechen in Sphaeren, I. *Math. Ann.*, 251:57–71, 1980.

[106] H. F. Muenzner. Isoparametrische Hyperflaechen in Sphaeren, II. *Math. Ann.*, 256:215–232, 1981.

[107] J. R. Munkres, *Elements of Algebraic Topology*, The Benjamin/Cummning Publishing Co., Inc., 1984.

[108] R. C. N. Marques, D. V. Tausk, *The Morse Index Theorem for Periodic Geodesics in Stationary Lorentzian Manifolds*, preprint 2000.

[109] K. Nomizu. Some results in É. Cartan's theory of isoparametric families of hypersurfaces. *Bull. Amer. Math. Soc.*, 79:1184–1188, 1973.

[110] C. Olmos. Isoparametric submanifolds and their homogeneous structures. *J. Differential Geom.*, 38:225–234, 1993.

[111] C. Olmos. Homogeneous submanifolds of higher rank and parallel mean curvature. *J. Differential Geom.*, 39:605–627, 1994.

[112] B. O'Neill, *Semi–Riemannian Geometry with Applications to Relativity*, Academic Press, New York, 1983.

[113] T. Ozawa. On the critical sets of distance functions to a taut submanifold. *Math. Ann.*, 276:91–96, 1986.

[114] H. Ozeki and M. Takeuchi. On some types of isoparametric hypersurfaces in spheres, I. *Tôhoku Math. J.*, 27:515–559, 1975.

[115] H. Ozeki and M. Takeuchi. On some types of isoparametric hypersurfaces in spheres, II. *Tôhoku Math. J.*, 28:7–55, 1976.

[116] R. Palais, *Foundations of Global Nonlinear Analysis*, W. A. Benjamin, 1968.

[117] R. Palais, *Morse Theory on Hilbert Manifolds*, Topology **2** (1963), 299–340.

[118] R. S. Palais and C.-L. Terng. A general theory of canonical forms. *Trans. Amer. Math. Soc.*, 300:771–789, 1987.

[119] R. S. Palais, Chuu-lian Terng, *Critical Point Theory and Submanifolds Geometry*, Lecture Notes in Mathematics n. 1353, Springer-Verlag, 1988.

[120] P. Piccione, R. Sampalmieri, *Geodesical Connectedness of Compact Lorentzian Manifolds*, Dynam. Sys. Appl. **5** (1996), 479–502.

[121] P. Piccione, D. Tausk, *A Note on the Morse Index Theorem for Geodesics between Submanifolds in semi-Riemannian Geometry*, J. Math. Phys. **40**, vol. 12 (1999), 6682–6688.

[122] P. Piccione, D. Tausk, *The Maslov Index and a Generalized Morse Index Theorem for Non Positive Definite Metrics*, Comptes Rendus de l'Académie de Sciences de Paris, vol. 331, issue 5 (2000), p. 385–389.

[123] P. Piccione, D. Tausk, *An Index Theorem for non Periodic Solutions of Hamiltonian Systems*, Proceedings of the London Mathematical Society (3) 83 (2001), 351–389.

[124] P. Piccione, D. Tausk, *On the Geometry of Grassmannians and the Symplectic Group: the Maslov Index and Its Applications*, notes of a short course given at the "XI Escola de Geometria Diferencial", Universidade Federal Fluminense, Niteroi, RJ, Brazil, August 2000.

[125] P. Piccione, D. Tausk, *On the Banach Differentiable Structure for Sets of Maps with Non-Compact Domains*, to appear in Journal of Nonlinear Analysis: Series A Methods and Applications (2001).

[126] P. Piccione, D. Tausk, *On the Distribution of Conjugate Points along semi-Riemannian Geodesics*, to appear in Communications in Analysis and Geometry 2001. (`LANL math.DG/0011038`)

[127] P. Piccione, D. V. Tausk, *On the Maslov and the Morse Index for Constrained Variational Problems*, to appear in Journal des Mathématiques Pures et Appliquées.

[128] P. Piccione, D. Tausk, *The Morse Index Theorem in semi-Riemannian Geometry*, preprint 2000.

[129] P. Piccione, D. Tausk, *An Index Theory for Paths that are Solutions of a Class of Strongly Indefinite Variational Problems*, preprint 2001.

[130] U. Pinkall. Dupin'sche Hyperflaechen. Doctoral Dissertation, Univ. Freiburg, 1981.

[131] U. Pinkall. Dupin hypersurfaces. *Math. Ann.*, 270:427–440, 1985.

[132] U. Pinkall. Curvature properties of taut submanifolds. *Geom. Dedicata*, 20:79–83, 1986.

[133] U. Pinkall and G. Thorbergsson. Deformations of Dupin hypersurfaces. *Proc. Amer. Math. Soc.*, 107:1037–1043, 1989.

[134] M. Reed, B. Simon, *Methods of Modern Mathematical Analysis I: Functional Analysis*, Academic Press, New York, 1980.

[135] A. Romero, M. Sánchez, *On Completeness of Certain Families of Semi-Riemannian Manifolds*, Geometriae Dedicata **53** (1994), 103–117.

[136] A. Romero, M. Sánchez, *New Properties and Examples of Incomplete Lorentzian Tori*, J. Math. Phys. **34**, 4 (1994), 1992–1997.

[137] A. Romero, M. Sánchez, *Completeness of Compact Lorentz Manifolds Admitting a Timelike Conformal Killing Vector Field*, Proc. AMS **123**, 9 (1995), 2831–2833.

[138] W. Rudin, *Real and complex analysis*, Third edition, McGraw-Hill Book Co., New York, 1987.

[139] D. Salamon, E. Zehnder, *Morse Theory for Periodic Solutions of Hamiltonian Systems and the Maslov Index*, Commun. Pure Appl. Math. **45** (1992), 1303–1360.

[140] M. Schwarz, *Morse Homology*, Birkäuser, Basel, 1993.

[141] J.–P. Serre, *Homologie Singulière des Espaces Fibrés*, Ann. of Math. **54**, No. 3 (1951), 425–505.

[142] R. W. Sharpe. A proof of the Chern-Lashof conjecture in dimensions greater than five. *Comment. Math. Helv.*, 64:221–235, 1989.

[143] S. Smale, *On Gradient Dynamical Systems*, Ann. of Math. **74** (1961), 199–206.

[144] S. Smale, *An Infinite Dimensional Version of Sard's Theorem*, Amer. J. Math. **87** (1965), 861–866.

[145] S. Smale, *Differentiable Dynamical Systems*, Bull. Amer. Math. Soc. **73** (1967), 747–817.

[146] A. Szulkin, *Cohomology and Morse Theory for Strongly Indefinite Functionals*, Math. Z. **209** (1992), 375–418.

[147] M. Takeuchi and S. Kobayashi. Minimal embeddings of R-spaces. *J. Differential Geom.*, 2:203–215, 1968.

[148] C.-L. Terng. Isoparametric submanifolds and their Coxeter groups. *J. Differential Geom.*, 21:79–107, 1985.

[149] C.-L. Terng, *Recent progress in submanifold geometry*. Differential geometry: partial differential equations on manifolds (Los Angeles, CA, 1990), 439–484, Proc. Sympos. Pure Math., 54, Part 1, Amer. Math. Soc., Providence, RI, 1993.

[150] C.-L. Terng and G. Thorbergsson. Taut immersions into complete Riemannian manifolds. In T. E. Ryan and S.-S. Chern, editors, *Tight and Taut Submanifolds*, Math. Sci. Res. Inst. Publ. 32, pages 181–228. Cambridge University Press, 1997.

[151] R. Thom, *Sur une partition em cellules associée á une fonction sur une varieté*, C. R. Acad. Sci. Paris Sér. A-B **228** (1949), 973–975.

[152] G. Thorbergsson. Dupin hypersurfaces. *Bull. London Math. Soc.*, 15:493–498, 1983.

[153] G. Thorbergsson. Tight immersions of highly connected manifolds. *Comment. Math. Helv.*, 61:102–121, 1986.

[154] G. Thorbergsson. Homogeneous spaces without taut embeddings. *Duke Math. J.*, 57:347–355, 1988.

[155] G. Thorbergsson. Isoparametric foliations and their buildings. *Ann. of Math.*, (2) 133:429–446, 1991.

[156] G. Thorbergsson. Smooth tight immersions. *Jber. d. Dt. Math.-Verein.*, 100:23–35, 1998.

[157] G. Thorbergsson. *A survey on isoparametric hypersurfaces and their generalizations*, volume I of *Handbook of Differential Geometry*, chapter 10. Elsevier Science, 2000.

[158] A. J. Tromba, *A General Approach to Morse Theory*, J. Diff. Geom. **12** (1977), 47–85.

[159] A. Weinstein, *Symplectic Manifolds and Their Lagrangian Submanifolds*, Adv. in Math. **6** (1971), 329–349.

[160] J. P. Wilson. The total absolute curvature of an immersed manifold. *J. London Math. Soc.*, 40:362–366, 1965.

[161] E. Witten, *Supersymmetry and Morse Theory*, J. Diff. Geom. **17** (1982), 661–692.

[162] K. Yosida, *Functional Analysis*, Springer-Verlag, 1966.

List of Symbols

Index

Z