



Analysis on Manifolds



uOttawa

Benoit Dionne
University of Ottawa

© Benoit Dionne, 2025 (University of Ottawa)

Introduction to the analysis on manifolds / modern differential geometry for an advanced undergraduate or graduate course.

This document is available on the following sites.

uO Research: <http://hdl.handle.net/10393/50918>

GitHub: https://github.com/BenoitDionne/Analysis_on_Manifolds



Unless otherwise stated, these lecture notes are made available under the terms of the license [Creative Commons Attribution - Non Commercial-Share Alike 4.0 International](https://creativecommons.org/licenses/by-nc-sa/4.0/) (CC BY-NC-SA 4.0)

Cover Page:

Casa Milà (popularly known as La Pedrera) by the architect Antoni Gaudi, Barcelona, photo by Louise Oegema.

Facade of the Roman library of Celsus in Ephesus, Anatolia (Asia Minor) in the Izmir province, photo by Jean Dionne.

Contents

Preface	1
I Integration on Manifold	5
Chapter 1 Inverse and Implicit Function Theorems	7
1.1. Review	7
1.2. Local Immersion Theorem	8
Chapter 2 The Riemann Integral of Real Valued Functions of One Variable	11
2.1. Definition and Properties	11
2.2. Darboux Theorem	17
2.3. Change of Variable	20
2.4. Sequence of Functions	23
2.5. Exercises	27
2.5.1. Supplementary Topics	47
Chapter 3 The Riemann Integral of Real Valued Functions of Several Variables	61
3.1. Definition and Properties	61
3.2. Bounded Curves in $\mathbb{R}^n$	71
3.3. Iterated Integrals	73
3.4. Partition of Unity	79
3.5. General Definition of the Integral	83
3.6. Change of Variables	88

3.7.	Sard's Theorem	103
3.8.	Exercises	106
3.8.1.	Supplementary Topics	113
Chapter 4	Differential Forms	149
4.1.	Multilinear Algebra	149
4.2.	Orientation	161
4.3.	Tangent Spaces	163
4.4.	Vector Fields	166
4.5.	Differential Forms	167
4.6.	Domains of Integration and their Boundary	178
4.7.	Integration on Chains	182
Chapter 5	Manifolds	187
5.1.	Preliminaries	187
5.2.	Local Charts, Atlases and Manifolds	189
5.2.1.	Manifolds in $\mathbb{R}^n$	195
5.3.	Tangent Spaces and Oriented Manifolds	199
5.3.1.	Tangent Spaces as Equivalence Classes of Curves	206
5.3.2.	Tangent Spaces as Spaces of differential Linear Operators	208
5.4.	Vector Fields	211
5.5.	Differential Forms	212
5.6.	Integration on Manifolds	217
5.6.1.	Generalization of Stokes' Theorem	225
5.7.	The Volume Element	243
5.7.1.	Polar Coordinates	249
Chapter 6	Classical Vector Calculus	255
6.1.	Gradient, Curl and Divergence	255
6.2.	Line Integral	256
6.2.1.	A Little Note on Arc Length	262
6.3.	Surface Integrals	265
6.3.1.	A Little Note on Surface Area	268
6.4.	The Classical trio	269

6.5.	Applications	280
6.5.1.	Tangent Vector, Curvature and Torsion	280
6.5.2.	Physical Interpretations	285
6.6.	Exercises	291
6.6.1.	Line and Surface Integrals	291
6.6.2.	Green's, Stokes' and Divergence Theorems	308
6.6.3.	Vector Fields	334
II	Differential Geometry	341
Chapter 7	Cohomology	343
7.1.	Preliminaries	343
7.2.	Cohomology Modules	351
7.2.1.	Cohomology Modules with Compact Supports	351
7.2.2.	Cohomology Modules (General Case)	360
7.2.3.	Mayer-Vietoris Argument	371
7.3.	Applications	377
Chapter 8	Algebraic Topology	383
8.1.	Fundamental Groups	384
8.1.1.	Covering Spaces	389
8.2.	Simplicial Homology	397
8.2.1.	Simplicial Complexes	397
8.2.2.	Barycentric Subdivisions	400
8.2.3.	Simplicial Approximation	403
8.2.4.	Oriented Simplicial Homology	407
8.2.5.	Reduced and Relative Homology	412
8.3.	Simplicial Cohomology	414
8.3.1.	Cohomology Modules	414
8.3.2.	Relative Cohomology	416
8.3.3.	Cup Product	417
8.3.4.	Relation Between Simplicial Cohomology and de Rham Cohomology	418
8.4.	Singular Homology	427

8.4.1.	Homotopy Modules	427
8.4.2.	Exact Sequences	433
8.4.3.	Chain Complexes	436
8.4.4.	Reduced Homology	449
8.4.5.	Relation Between $\pi_1(X, x_0)$ and $H_1(X; R)$	450
8.4.6.	Relative Homology	456
8.4.7.	Excision	464
8.4.8.	Mayer-Vietoris Sequence	478
8.4.9.	Further Studies	484
8.5.	Singular Cohomology	485
8.5.1.	Cohomotopy Modules	485
8.5.2.	Cochain Complexes	487
8.5.3.	Relative Cohomology	489
8.5.4.	Cup Product	491
8.5.5.	Relation Between Singular Cohomology and de Rham Cohomology	493
8.6.	Relation Between Simplicial and Singular Theory	500
8.6.1.	Relation Between Simplicial Homology and Singular Homology	501
8.6.2.	Relation Between Simplicial Cohomology and Singular Cohomology	513
Chapter 9	Riemannian Geometry	515
9.1.	Riemannian Manifold	515
9.2.	Preliminaries	516
9.2.1.	Group of Rotations	516
9.2.2.	The Tangent Bundle $T(TS)$	518
9.2.3.	The Tangent Bundle TS_c	520
9.3.	Parallel Translation and Connection	526
9.4.	Structural Equations and Curvature	539
9.4.1.	First (Cartan) Structural Equation	543
9.4.2.	Second (Cartan) Structural Equation	546
9.5.	Geodesic	548
9.6.	Gauss-Bonnet Theorem	550
9.7.	Lie Bracket	560
9.7.1.	An Equivalent Definition of the Lie Bracket	563

9.8.	Geodesic Coordinate Systems	566
9.9.	Euclidean and Non-Euclidean Geometries	582
9.9.1.	Orientation Preserving Isometries	583
9.9.2.	Group of Orientation Preserving Isometries	584
9.9.3.	Manifolds of Constant Curvatures	587
Bibliography		593
Index		595

Preface

The first part of these lecture notes provides a clear, detailed and complete presentation of Riemann integrals for functions of several variables including an introduction to manifolds and integration on manifolds.

The first part of these lecture notes is a complement to and an expansion of Spivak's **Calculus on Manifolds** [26]. We review some of the topics in the book, adding some additional information and results. We also reproduce (in some cases with major differences) some of the proofs that were sketchy. After having learned calculus of functions of several variables in their first year at the university, students study basic analysis. This usually includes some concepts of metric spaces and topology, the derivative of functions of several variables including the inverse and implicit function theorems, and sometime a rigorous introduction to Riemann integration of functions of one variable. The rigorous presentation of Riemann integration for functions of several variables including integration on manifolds is completely ignored. The next time that students may study integration of functions is at the graduate level when they learn about Lebesgue integrals and so never cover rigorously vector calculus and integration on manifolds. The first part of these lecture notes fills this gap that we often find in undergraduate mathematics programs.

A large number of problems, many with detailed solutions, are included in Chapters 2, 3 and 6. These problems come from several textbooks listed in the reference, in particular from Folland's **Advanced Calculus** [7]. For many of them, they are standard problems that one can find in all advanced calculus textbooks. These lecture notes are therefore also ideal for self study and as a reference for students.

The second part of these lecture notes provides a modern introduction to differential geometry. There are several unusual features to the presentation of this topic in these lecture notes. One of them is that the more geometric presentation of the tangent space to a manifold at a point is used almost everywhere in these notes. The only time that the definition of the tangent space of manifold at a point as a space of linear differential operators is used is to define the Lie Bracket where it is really the easiest way to define the Lie Bracket. Another unusual feature is the inclusion of a chapter on algebraic geometry. We feel that this is necessary because algebraic geometry provides a very nice and efficient technique to reach conclusions about manifolds that would otherwise require very technical and convoluted proofs. The second part begins with an introduction to de Rham cohomology. We then provide a solid introduction to algebraic topology including the simplicial and singular approaches to homology and cohomology. We prove the relation between de Rham

cohomology and the general concept of cohomology.

The last chapter of these lecture notes and its second part is an introduction to Riemannian geometry. The presentation of this topic is very geometric and hopefully provides the reader with a good visual understanding of the properties associated to Riemannian geometry. In particular, this chapter includes a proof of the beautiful Gauss-Bonnet theorem. We have tried to stay away from formal computations frequently associated with this subject.

Some Terminology

Metric and Topological Spaces

The **k-norm** of $\mathbf{x} \in \mathbb{R}^n$ is defined by

$$\|\mathbf{x}\|_k = \begin{cases} \left(\sum_{i=1}^n |x_i|^k \right)^{1/k} & \text{if } k > 0 \\ \max_{1 \leq i \leq n} |x_i| & \text{if } k = \infty \end{cases}$$

We adopt the convention that $\|\cdot\|$ and $\|\cdot\|_2$ will denote the **2-norm** or **Euclidean norm**.

The **diameter** of a set $W \subset \mathbb{R}^n$ is the number $\text{diam } W = \inf\{\|\mathbf{x} - \mathbf{y}\| : \mathbf{x}, \mathbf{y} \in W\}$.

The **distance** between two sets $V, W \subset \mathbb{R}^n$ is the number $\text{dist}(V, W) = \inf\{\|\mathbf{x} - \mathbf{y}\| : \mathbf{x} \in V \text{ and } \mathbf{y} \in W\}$. The distance between a point $\mathbf{x} \in \mathbb{R}^n$ and a set $W \subset \mathbb{R}^n$ is defined as $\text{dist}(\mathbf{x}, W) = \text{dist}(\{\mathbf{x}\}, W)$.

The **open ball** of radius r centred at $\mathbf{y} \in \mathbb{R}^n$ is denoted $B_r(\mathbf{y})$; namely,

$$B_r(\mathbf{y}) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{y}\| < r\}.$$

We will also use this notation to denote open balls in other metric spaces.

Let W be a subset in $\mathbb{R}^n$. The **closure** of W , denoted $\overline{W}$, is the smallest closed set containing W . We have $\overline{W} = \bigcap_{\substack{W \subset V \\ V \text{ closed}}} V$. One of the most important case in these lecture

notes is the closure of a rectangle $R = \prod_{i=1}^n I_i$ where the I_i are bounded intervals. We have

$\overline{R} = \prod_{i=1}^n \overline{I}_i$ where $\overline{I}_i$ is the closed interval obtained from the interval I_i by adding the end points.

Let W be a set in $\mathbb{R}^n$. The **interior** of W , denoted W° , is the largest open subset of W . We have $W^\circ = \bigcup_{\substack{V \subset W \\ V \text{ open}}} V$. One of the most important case in these lecture notes is the interior

of a rectangle $R = \prod_{i=1}^n I_i$ where the I_i are bounded intervals. We have $R^\circ = \prod_{i=1}^n I_i^\circ$ where I_i° is the open interval obtained from the interval I_i by removing the end points.

The definitions of the interior and closure of a set $W \subset V \subset \mathbb{R}^n$ relative to the set V are identical to the previous definitions except that $\mathbb{R}^n$ is replaced by V where V has the induced topology from $\mathbb{R}^n$.

Let W be a set in $\mathbb{R}^n$. The **boundary** of W , denoted ∂W , is the set of all points $\mathbf{x} \in \mathbb{R}^n$ such that $B_r(\mathbf{x}) \cap W \neq \emptyset$ and $B_r(\mathbf{x}) \cap (\mathbb{R}^n \setminus W) \neq \emptyset$ for all $r > 0$. One of the most important case in these lecture notes is the boundary of a rectangle $R = \prod_{i=1}^n I_i$ where $I_i = [a_i, b_i]$, $[a_i, b_i[$, $]a_i, b_i]$ or $]a_i, b_i[$ for all i . We have

$$\partial R = \bigcup_{i=1}^n \left((I_1 \times \dots \times I_{i-1} \times \{a_i\} \times I_{i+1} \times \dots \times I_n) \right. \\ \left. \bigcup (I_1 \times \dots \times I_{i-1} \times \{b_i\} \times I_{i+1} \times \dots \times I_n) \right)$$

where $I_1 \times \dots \times I_{i-1}$ is ignored when $i = 1$ and $I_{i+1} \times \dots \times I_n$ is ignored when $i = n$. The reader may want to write down the previous formula for the case $n = 2$ and $n = 3$ to verify its accuracy.

Function Spaces

A function $f : [a, b] \rightarrow \mathbb{R}^n$ is **piecewise continuous** if f is continuous at all but a finite number of points in $[a, b]$ and, at a point $c \in [a, b]$ where f is not continuous, we have $\lim_{x \rightarrow c^-} f(x)$ and $\lim_{x \rightarrow c^+} f(x)$ exist.

Similarly, given $m \in \mathbb{N}$, a function $f : [a, b] \rightarrow \mathbb{R}^n$ is **piecewise C^m continuously differentiable** if f is m -time differentiable at all but a finite number of points in $[a, b]$ and, at a point $c \in [a, b]$ where f is not m -time differentiable, we have $\lim_{x \rightarrow c^-} f^{(j)}(x)$ and $\lim_{x \rightarrow c^+} f^{(j)}(x)$ exist for $0 \leq j \leq m$. Obviously, f is piecewise C^0 continuously differentiable is another way to say that f is piecewise continuously differentiable.

Let V be a subset of $\mathbb{R}^n$ and $f : V \rightarrow \mathbb{R}$. The **supremum norm** of f on V is defined by $\|f\|_{\infty, V} = \sup_{\mathbf{x} \in V} |f(\mathbf{x})|$. We will simply write $\|f\|_{\infty}$ when the set V is clear from the context.

Let V be a subset of $\mathbb{R}^n$. The **support** of a function $f : V \rightarrow \mathbb{R}$ is the set $\text{supp } f = \overline{\{\mathbf{x} \in V : f(\mathbf{x}) \neq 0\}}$.

Let V be an open subset of $\mathbb{R}^n$ and $f : V \rightarrow \mathbb{R}$ be a sufficiently differentiable function, we define the **differential operator** D^α for $\alpha \in \mathbb{N}^n$ by

$$D^\alpha f(\mathbf{x}) = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_n^{\alpha_n}}(\mathbf{x})$$

for $\mathbf{x} \in V$, where $|\alpha| = \sum_{j=1}^n \alpha_j$.

Let V be an open subset of $\mathbb{R}^n$. For $m \in \mathbb{N}$, $C^m(V)$ is the space of all function $f : V \rightarrow \mathbb{R}$ such that all $D^\alpha f : V \rightarrow \mathbb{R}$ exists and is continuous for all $\alpha \in \mathbb{N}^n$ with $|\alpha| \leq m$. Likewise, $C^\infty(V)$ is the space of all function $f : V \rightarrow \mathbb{R}$ such that all $D^\alpha f : V \rightarrow \mathbb{R}$ exists and is continuous for all $\alpha \in \mathbb{N}^n$.

For $m \in \mathbb{N}$, we say that $f \in C^m(\overline{V})$ if $f \in C(\overline{V})$, and there exist an open set $W \supset \overline{V}$ and a function $F \in C^m(W)$ such that $F = f$ on V . Likewise, we say that $f \in C^\infty(\overline{V})$ if $f \in C(\overline{V})$, and there exist an open set $W \supset \overline{V}$ and a function $F \in C^\infty(W)$ such that $F = f$ on V .

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $g : \mathbb{R}^n \rightarrow \mathbb{R}$ be two functions. We write $f(\mathbf{x}) = O(g(\mathbf{x}))$ near the origin if there exists a positive constant K such that $\|f(\mathbf{x})\| < K|g(\mathbf{x})|$ for $\mathbf{x}$ in a neighbourhood of the origin.

Note

In these lecture notes, $\mathbb{N} = \{0, 1, 2, 3, \dots\}$ and $\mathbb{N}^+ = \{1, 2, 3, \dots\}$.

Note

By tradition, $C(V)$ means the space of continuous real valued functions on V . Namely, $C(V) = C^0(V)$. If we really want to talk about the space of continuously differentiable real valued functions on V , then we will write $C^1(V)$.

Part I

Integration on Manifold

Chapter 1

Inverse and Implicit Function Theorems

1.1 Review

We review two of the main results about differentiation that will be very useful later. The proofs of these results can be found in any good textbooks in analysis including [26].

Theorem 1.1.1 (Inverse Function Theorem)

Let V be an open subset of $\mathbb{R}^n$ and $f : V \rightarrow \mathbb{R}^n$ be a continuously differentiable function. Suppose that $Df(\mathbf{a}) \neq 0$ for some $\mathbf{a} \in V$. Then, there exist an open neighbourhood $U \subset V$ of $\mathbf{a}$ and an open neighbourhood W of $f(\mathbf{a})$ such that $f : U \rightarrow W$ is one-to-one and onto, the function $f^{-1} : W \rightarrow U$ is continuously differentiable and

$$Df^{-1}(\mathbf{y}) = (Df(f^{-1}(\mathbf{y})))^{-1}$$

for all $\mathbf{y} \in W$.

A consequence of the Inverse Function Theorem is the following theorem.

Theorem 1.1.2 (Implicit Function Theorem)

Let $f : U \times V \rightarrow \mathbb{R}^m$ be a continuously differentiable function where $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ are open sets. Suppose that $f(\mathbf{a}, \mathbf{b}) = \mathbf{0}$ for some $(\mathbf{a}, \mathbf{b}) \in U \times V$. Let

$$A = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{a}, \mathbf{b}) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{a}, \mathbf{b}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{a}, \mathbf{b}) & \cdots & \frac{\partial f_m}{\partial x_n}(\mathbf{a}, \mathbf{b}) \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} \frac{\partial f_1}{\partial x_{n+1}}(\mathbf{a}, \mathbf{b}) & \cdots & \frac{\partial f_1}{\partial x_{n+m}}(\mathbf{a}, \mathbf{b}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_{n+1}}(\mathbf{a}, \mathbf{b}) & \cdots & \frac{\partial f_m}{\partial x_{n+m}}(\mathbf{a}, \mathbf{b}) \end{pmatrix}.$$

If $\det B \neq 0$, then there exist open neighbourhoods $U_1 \subset U$ of $\mathbf{a}$ and $V_1 \subset V$ of $\mathbf{b}$

with the property that for each $\mathbf{x} \in U_1$ there exists a unique $g(\mathbf{x}) \in V_1$ such that $f(\mathbf{x}, g(\mathbf{x})) = 0$. Moreover, the function $g : U_1 \rightarrow V_1$ is continuously differentiable and $Dg(\mathbf{a}) = -B^{-1}A$.

The Implicit Function Theorem is in fact equivalent to the Inverse Function Theorem. One is a consequence of the other.

1.2 Local Immersion Theorem

Theorem 1.2.1 (Local Immersion Theorem)

Let V be an open subset of $\mathbb{R}^n$ and $f : V \rightarrow \mathbb{R}^m$ be a continuously differentiable function with $m \leq n$. Suppose that there exists $\mathbf{a} \in V$ such that $f(\mathbf{a}) = 0$ and $Df(\mathbf{a})$ has rank m . Then there exist an open set $W \subset \mathbb{R}^n$, an open neighbourhood $U \subset V$ of $\mathbf{a}$ and a diffeomorphism $h : W \rightarrow U$ such that $f(h(\mathbf{x})) = (x_1 \ x_2 \ \dots \ x_m)^\top$ for $\mathbf{x} \in W$.

In the previous theorem, the function h is a local change of coordinates near $\mathbf{a}$.

Proof.

Since $Df(\mathbf{a})$ has rank m , there exists an $(m \times m)$ -submatrix of $Df(\mathbf{a})$ with non-zero determinant. We choose a permutation of the coordinates $P : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that the first m columns of $Df(\mathbf{a})P$ formed an $m \times m$ -matrix with non-zero determinant. Since P is a permutation, it is its own inverse.

Let $\mathbf{b} = P(\mathbf{a})$. The first m columns of $D(f \circ P)(\mathbf{b})$ form an $m \times m$ -matrix with non-zero determinant because $D(f \circ P)(\mathbf{b}) = Df(P(\mathbf{b}))P = Df(\mathbf{a})P$.

Let

$$M = \begin{pmatrix} \frac{\partial(f_1 \circ P)}{\partial x_1}(\mathbf{b}) & \dots & \frac{\partial(f_1 \circ P)}{\partial x_m}(\mathbf{b}) \\ \vdots & \ddots & \vdots \\ \frac{\partial(f_m \circ P)}{\partial x_1}(\mathbf{b}) & \dots & \frac{\partial(f_m \circ P)}{\partial x_m}(\mathbf{b}) \end{pmatrix}.$$

It is the $m \times m$ -submatrix of $D(f \circ P)(\mathbf{b})$ formed by the first m columns. Consider the map

$$H : P(V) \rightarrow \mathbb{R}^n$$

$$\mathbf{x} \mapsto ((f_1 \circ P)(\mathbf{x}) \ \dots \ (f_m \circ P)(\mathbf{x}) \ x_{m+1} \ x_{m+2} \ \dots \ x_n)^\top$$

We have $H(\mathbf{b}) = (0 \ \dots \ 0 \ b_{m+1} \ b_{m+2} \ \dots \ b_n)^\top$ because $(f \circ P)(\mathbf{b}) = f(P(\mathbf{b})) = f(\mathbf{a}) = \mathbf{0}$, and

$$\det D H(\mathbf{b}) = \det \begin{pmatrix} M & * \\ 0 & \text{Id}_{n-m} \end{pmatrix} = \det M \neq 0.$$

It follows from the Inverse Function Theorem that there exist an open neighbourhood $U_1 \subset P(V) \subset \mathbb{R}^n$ of $\mathbf{b}$ and an open neighbourhood $W_1 \subset \mathbb{R}^n$ of $(0 \ \dots \ 0 \ b_{m+1} \ b_{m+2} \ \dots \ b_n)^\top$ such that $H : U_1 \rightarrow W_1$ is a diffeomorphism. Let $\pi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be the projection on the first m components. We have

$$(f \circ P)(H^{-1}(\mathbf{x})) = (\pi \circ H)(H^{-1}(\mathbf{x})) = \pi(\mathbf{x}) = (x_1 \ x_2 \ \dots \ x_m)^\top$$

for $\mathbf{x} \in W_1$. We have the following commutative diagram.

$$\begin{array}{ccc} V & \xrightarrow{f} & \mathbb{R}^m \\ P \uparrow & & \uparrow \pi \\ U_1 & \xleftarrow{H^{-1}} & W_1 \end{array}$$

The conclusion of the theorem is given by $U = P(U_1)$, $W = W_1$ and $h = P \circ H^{-1}$. ■

Chapter 2

The Riemann Integral of Real Valued Functions of One Variable

2.1 Definition and Properties

We begin with several basic definitions.

Definition 2.1.1

A **partition** of an interval $[a, b]$ is a subset $P = \{p_i : 0 \leq i \leq N\}$ of $[a, b]$ such that $a = p_0 < p_1 < p_2 < p_3 < \dots < p_N = b$ with $N \in \mathbb{N}$.

Definition 2.1.2

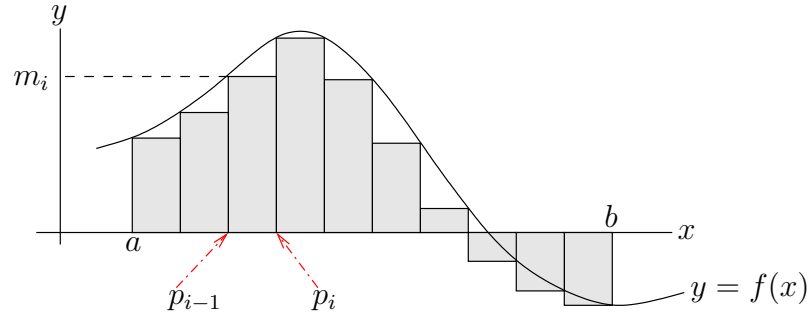
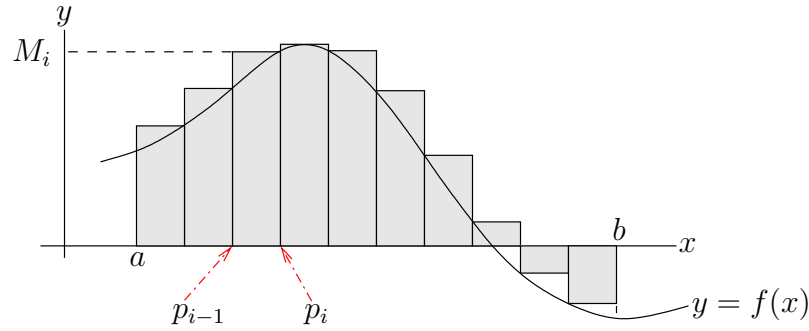
Consider a bounded function $f : [a, b] \rightarrow \mathbb{R}$ and a partition $P = \{p_i : 0 \leq i \leq N\}$ of $[a, b]$. Let $m_i = \inf\{f(x) : x \in [p_{i-1}, p_i]\}$ and $M_i = \sup\{f(x) : x \in [p_{i-1}, p_i]\}$ for $1 \leq i \leq N$. The **lower Riemann sum** of f with respect to the partition P is

$$\mathcal{L}_P(f) = \sum_{j=1}^N m_i(p_i - p_{i-1}) .$$

The **upper Riemann sum** of f with respect to the partition P is

$$\mathcal{U}_P(f) = \sum_{j=1}^N M_i(p_i - p_{i-1})$$

(Figures [2.1](#) and [2.2](#)).

Figure 2.1: A lower Riemann sum for $f : [a, b] \rightarrow \mathbb{R}$ Figure 2.2: An upper Riemann sum for $f : [a, b] \rightarrow \mathbb{R}$ **Definition 2.1.3**

A **refinement** for a partition P of $[a, b]$ is another partition Q of $[a, b]$ such that $Q \supset P$.

Proposition 2.1.4

Consider a bounded function $f : [a, b] \rightarrow \mathbb{R}$ and a partition P of $[a, b]$. If Q is a refinement of a partition P of $[a, b]$, then $\mathcal{L}_P(f) \leq \mathcal{L}_Q(f)$ and $\mathcal{U}_Q(f) \leq \mathcal{U}_P(f)$.

Proof.

Suppose that $P = \{p_i : 0 \leq i \leq N\}$ and that Q contains one more point than P . Let $q \in Q \setminus P$ be this point.

As usual, let $m_i = \inf\{f(x) : x \in [p_{i-1}, p_i]\}$ for $1 \leq i \leq N$. There exists j such that $p_{j-1} < q < p_j$. If $\tilde{m}_j = \inf\{f(x) : x \in [p_{j-1}, q]\}$ and $\check{m}_j = \inf\{f(x) : x \in [q, p_j]\}$, then $\tilde{m}_j \geq m_j$ and $\check{m}_j \geq m_j$. Hence

$$\mathcal{L}_Q(f) - \mathcal{L}_P(f) = \tilde{m}_j(q - p_{j-1}) + \check{m}_j(p_j - q) - m_j(p_j - p_{j-1})$$

$$= (\tilde{m}_j - m_j)(q - p_{j-1}) + (\check{m}_j - m_j)(p_j - q) \geq 0 .$$

Thus $\mathcal{L}_P(f) \leq \mathcal{L}_Q(f)$. If Q contains k more points than P , then k applications of the previous reasoning show that $\mathcal{L}_P(f) \leq \mathcal{L}_Q(f)$.

The proof that $\mathcal{U}_Q(f) \leq \mathcal{U}_P(f)$ is similar. ■

Corollary 2.1.5

Consider a bounded function $f : [a, b] \rightarrow \mathbb{R}$. If P and Q are two partitions of $[a, b]$, then $\mathcal{L}_P(f) \leq \mathcal{U}_Q(f)$.

Proof.

Let $R = P \cup Q$. R is a refinement of both P and Q . Hence $\mathcal{L}_P(f) \leq \mathcal{L}_R(f) \leq \mathcal{U}_R(f) \leq \mathcal{U}_Q(f)$ where the first and last inequalities comes from the previous proposition and the middle inequality is a consequence of the definition of lower and upper Riemann sums. ■

Definition 2.1.6

Consider a bounded function $f : [a, b] \rightarrow \mathbb{R}$. The **lower Riemann integral** of f on $[a, b]$ is defined as

$$\mathcal{L}(f) = \sup\{\mathcal{L}_P(f) : P \text{ is a partition of } [a, b]\} .$$

The **upper Riemann integral** of f on $[a, b]$ is defined as

$$\mathcal{U}(f) = \inf\{\mathcal{U}_P(f) : P \text{ is partition of } [a, b]\} .$$

It follows from Corollary 2.1.5 that $\mathcal{L}(f) \leq \mathcal{U}(f)$. To prove this claim, it suffices to note that, given a partition Q of $[a, b]$, we have that $\mathcal{L}_P(f) \leq \mathcal{U}_Q(f)$ for all partition P of $[a, b]$ implies that $\mathcal{L}(f) \leq \mathcal{U}_Q(f)$. Since Q is an arbitrary partition of $[a, b]$, we get $\mathcal{L}(f) \leq \mathcal{U}(f)$.

Definition 2.1.7

Consider a bounded function $f : [a, b] \rightarrow \mathbb{R}$. If $\mathcal{U}(f) = \mathcal{L}(f)$, then we say that **f is (Riemann) integrable** on the interval $[a, b]$ and the **integral of f** on $[a, b]$ is defined as

$$\int_a^b f(x) \, dx = \mathcal{L}(f) = \mathcal{U}(f) .$$

Theorem 2.1.8

Consider a bounded function $f : [a, b] \rightarrow \mathbb{R}$. The function f is integrable on $[a, b]$ if and only if for every $\epsilon > 0$ there exists a partition P of $[a, b]$ such that $\mathcal{U}_P(f) - \mathcal{L}_P(f) < \epsilon$.

Proof.

$\Leftarrow$) Since $\mathcal{L}_P(f) \leq \mathcal{L}(f) \leq \mathcal{U}(f) \leq \mathcal{U}_P(f)$ for all partition P of $[a, b]$, we have $\mathcal{U}(f) - \mathcal{L}(f) < \epsilon$ for all $\epsilon > 0$. Therefore $\mathcal{U}(f) = \mathcal{L}(f)$ and f is integrable.

$\Rightarrow$) Choose two partitions Q and R of $[a, b]$ such that $\mathcal{L}(f) - \mathcal{L}_Q(f) < \epsilon/2$ and $\mathcal{U}_R(f) - \mathcal{U}(f) < \epsilon/2$. Since f is integrable, we have $\mathcal{U}(f) = \mathcal{L}(f)$. Therefore $\mathcal{U}_R(f) - \mathcal{L}_Q(f) < \epsilon$. If we use the partition $P = Q \cup R$, a refinement of both Q and R , then we get $\mathcal{U}_P(f) - \mathcal{L}_P(f) \leq \mathcal{U}_R(f) - \mathcal{L}_Q(f) < \epsilon$ because $\mathcal{L}_Q(f) \leq \mathcal{L}_R(f) \leq \mathcal{U}_R(f) \leq \mathcal{U}_P(f)$ according to Proposition 2.1.4. ■

Proposition 2.1.9

If $f : [a, b] \rightarrow \mathbb{R}$ is bounded and monotone, then f is integrable on $[a, b]$.

Proof.

Suppose that f is a decreasing function. We consider the partition $P_N = \{p_i = a + i(b-a)/N : 0 \leq i \leq N\}$ of $[a, b]$. Since f is decreasing, we have $m_i = \inf\{f(x) : x \in [p_{i-1}, p_i]\} = f(p_i)$ and $M_i = \sup\{f(x) : x \in [p_{i-1}, p_i]\} = f(p_{i-1})$ for all i . Hence

$$\begin{aligned} \mathcal{U}_{P_N}(f) - \mathcal{L}_{P_N}(f) &= \sum_{i=1}^N M_i(p_i - p_{i-1}) - \sum_{i=1}^N m_i(p_i - p_{i-1}) = \frac{b-a}{N} \sum_{i=1}^N (f(p_{i-1}) - f(p_i)) \\ &= \frac{(b-a)(f(a) - f(b))}{N}. \end{aligned}$$

Given $\epsilon > 0$, we can choose N large enough such that the partition P_N satisfies $\mathcal{U}_{P_N}(f) - \mathcal{L}_{P_N}(f) < \epsilon$. It follows from Theorem 2.1.8 that f is integrable.

The proof for f increasing is similar. ■

Proposition 2.1.10

If $f : [a, b] \rightarrow \mathbb{R}$ and $g : [a, b] \rightarrow \mathbb{R}$ are two integrable functions such that $f(x) \leq g(x)$ for all $x \in [a, b]$, then $\int_a^b f(x) dx \leq \int_a^b g(x) dx$

Proof.

Given a partition $P = \{p_i : 0 \leq i \leq N\}$ of $[a, b]$, let $M_i = \sup\{f(x) : p_{i-1} \leq x \leq p_i\}$ and $\tilde{M}_i = \sup\{g(x) : p_{i-1} \leq x \leq p_i\}$. We have $M_i \leq \tilde{M}_i$ for all i . Hence

$$\mathcal{U}(f) \leq \mathcal{U}_P(f) = \sum_{i=1}^N M_i(p_i - p_{i-1}) \leq \sum_{i=1}^N \tilde{M}_i(p_i - p_{i-1}) = \mathcal{U}_P(g).$$

Since the partition P is arbitrary,

$$\mathcal{U}(f) \leq \inf\{\mathcal{U}_P(g) : P \text{ is a partition of } [a, b]\} = \mathcal{U}(g).$$

Since $f : [a, b] \rightarrow \mathbb{R}$ and $g : [a, b] \rightarrow \mathbb{R}$ are integrable, we get

$$\int_a^b f(x) dx = \mathcal{U}(f) \leq \mathcal{U}(g) = \int_a^b g(x) dx. \quad \blacksquare$$

If $f : [a, b] \rightarrow \mathbb{R}$ is integrable, then $f : [a, b] \rightarrow \mathbb{R}$ is bounded by definition of the Riemann integral. Therefore, there exist two real numbers m and M such that $m \leq f(x) \leq M$ for all $x \in [a, b]$.

Proposition 2.1.11

Suppose that $f : [a, b] \rightarrow [m, M]$ is an integrable function. Suppose that $g : [m, M] \rightarrow \mathbb{R}$ is continuous. Then $h = g \circ f : [a, b] \rightarrow \mathbb{R}$ is integrable on $[a, b]$.

Proof.

Let $K = \sup\{|g(x)| : x \in [m, M]\}$. Since $\tilde{g} : [m, M] \rightarrow [0, \infty[$ defined by $\tilde{g}(x) = |g(x)|$ for $x \in [a, b]$ is continuous and $[m, M]$ is a compact set, we have that $\tilde{g}$ reaches its maximum at some point of $[m, M]$. In particular, $K < \infty$.

Given $\epsilon > 0$, let $\tilde{\epsilon} = \frac{\epsilon}{b - a + 2K}$. Since g is a continuous function on the compact set $[m, M]$, it is uniformly continuous on $[m, M]$. Therefore, there exists $\delta > 0$ such that $\delta < \tilde{\epsilon}$ and $|g(x) - g(y)| < \tilde{\epsilon}$ for $x, y \in [m, M]$ with $|x - y| < \delta$.

Since f is integrable on $[a, b]$, there exists a partition $P = \{p_i : 0 \leq i \leq N\}$ of $[a, b]$ such that $\mathcal{U}_P(f) - \mathcal{L}_P(f) < \delta^2$. Let $m_i = \inf\{f(x) : p_{i-1} \leq x \leq p_i\}$, $M_i = \sup\{f(x) : p_{i-1} \leq x \leq p_i\}$, $\tilde{m}_i = \inf\{h(x) : p_{i-1} \leq x \leq p_i\}$ and $\tilde{M}_i = \sup\{h(x) : p_{i-1} \leq x \leq p_i\}$. Moreover, let $A = \{i \in \{1, 2, \dots, N\} : M_i - m_i < \delta\}$ and $B = \{1, 2, \dots, N\} \setminus A$.

For $i \in A$, we have $|f(x) - f(y)| \leq M_i - m_i < \delta$ for all $x, y \in [p_{i-1}, p_i]$. Hence $|h(x) - h(y)| = |g(f(x)) - g(f(y))| < \tilde{\epsilon}$ for all $x, y \in [p_{i-1}, p_i]$. This implies that

$$\tilde{M}_i - \tilde{m}_i \leq \tilde{\epsilon} \quad \text{for } i \in A. \quad (2.1.1)$$

For $i \in B$, we use the definition of K to get

$$\tilde{M}_i - \tilde{m}_i \leq 2K \quad \text{for } i \in B. \quad (2.1.2)$$

Moreover,

$$\delta \sum_{i \in B} (p_i - p_{i-1}) \leq \sum_{i \in B} (M_i - m_i)(p_i - p_{i-1}) \leq \mathcal{U}_P(f) - \mathcal{L}_P(f) \leq \delta^2$$

yields

$$\sum_{i \in B} (p_i - p_{i-1}) \leq \delta. \quad (2.1.3)$$

It follows from (2.1.1), (2.1.2) and (2.1.3) that

$$\begin{aligned} \mathcal{U}_P(h) - \mathcal{L}_P(h) &= \sum_{i \in A} (\tilde{M}_i - \tilde{m}_i)(p_i - p_{i-1}) + \sum_{i \in B} (\tilde{M}_i - \tilde{m}_i)(p_i - p_{i-1}) \\ &\leq \tilde{\epsilon} \sum_{i \in A} (p_i - p_{i-1}) + 2K \sum_{i \in B} (p_i - p_{i-1}) \leq \tilde{\epsilon}(b - a) + 2K\delta \\ &< \tilde{\epsilon}(b - a + 2K) = \epsilon. \end{aligned}$$

Since ϵ is arbitrary, it follows from Theorem 2.1.8 that h is integrable on $[a, b]$. ■

Proposition 2.1.12

Suppose that $f : [a, b] \rightarrow \mathbb{R}$ is integrable. Then $|f| : [a, b] \rightarrow \mathbb{R}$ is integrable and

$$\left| \int_a^b f(x) \, dx \right| \leq \int_a^b |f(x)| \, dx.$$
Proof.

The function f is a bounded function because it is integrable. Thus, its image is a bounded subset of $\mathbb{R}$. Choose an interval $[m, M]$ containing $f([a, b])$. It follows from the previous proposition with $g(x) = |x|$ for $x \in [m, M]$ that $x \mapsto g(f(x)) = |f(x)|$ is integrable.

Take $\nu = 1$ or -1 such that $\nu \int_a^b f(x) \, dx \geq 0$. It follows from Proposition 2.1.10 and Question 2.2 that

$$0 \leq \left| \int_a^b f(x) \, dx \right| = \nu \int_a^b f(x) \, dx = \int_a^b \nu f(x) \, dx \leq \int_a^b |\nu f(x)| \, dx = \int_a^b |f(x)| \, dx. \quad \blacksquare$$

Remark 2.1.13

Note that $|f|$ integrable does not implies that f is integrable. For instance, the function $f : [0, 1] \rightarrow \mathbb{R}$ defined by $f(x) = 1$ for $x \in \mathbb{Q} \cap [0, 1]$ and $f(x) = -1$ for $x \in (\mathbb{R} \setminus \mathbb{Q}) \cap [0, 1]$ is not integrable (see Question 2.1) but $|f|$ is integrable because $|f(x)| = 1$ for all $x \in [0, 1]$. It is interesting to note that the situation is different for the Lebesgue integral that the reader may have the chance to study in the future. If $|f|$ is Lebesgue integrable, then f is Lebesgue integrable. ♠

Definition 2.1.14

Let $|I|$ denote the length of the interval I .

A set $S \in \mathbb{R}$ has **zero content** if for every $\epsilon > 0$ there exists a finite collection of closed intervals $\{I_i\}_{i=1}^M$ such that $S \subset \bigcup_{i=1}^M I_i$ and $\sum_{i=1}^M |I_i| < \epsilon$.

A set $S \in \mathbb{R}$ has **measure zero** if for every $\epsilon > 0$ there exists a countable collection of closed intervals $\{I_i\}_{i=1}^\infty$ such that $S \subset \bigcup_{i=1}^\infty I_i$ and $\sum_{i=1}^\infty |I_i| < \epsilon$.

Remark 2.1.15

For those who will study the theory of Lebesgue integration in the future, they should note that our definition of sets of measure zero is not as general as the definition of sets of measure zero according to the Lebesgue measure. Sets of measure zero according to the Lebesgue measure are sets of measure zero according to our definition but the converse is not always true. For instance, the set $\mathbb{Q}$ of rational numbers is a set of measure zero according to the Lebesgue measure but it is not a set of measure zero according to our definition. ♠

Definition 2.1.16

A set $S \subset \mathbb{R}$ is **(Jordan) measurable** if S is bounded and ∂S has measure zero.

Theorem 2.1.17

Consider a bounded function $f : [a, b] \rightarrow \mathbb{R}$. The set of all points in $[a, b]$ where f is discontinuous has measure zero if and only if f is integrable on $[a, b]$.

The proof of this theorem is given in Chapter 3 on the integration of real valued functions of several variables.

Given $S \subset \mathbb{R}$, let $\mathcal{X}_S : \mathbb{R} \rightarrow \mathbb{R}$ be the function defined by

$$\mathcal{X}(x) = \begin{cases} 1 & \text{if } x \in S \\ 0 & \text{if } x \in \mathbb{R} \setminus S \end{cases}$$

The only points where $\mathcal{X}_S$ is discontinuous are the points on ∂S .

Definition 2.1.18

Given a Jordan measurable set $S \subset \mathbb{R}$ and a function $f : S \rightarrow \mathbb{R}$, we define the **integral of f on S** , denoted $\int_S f$ or $\int_S f(x) dx$, as the integral $\int_a^b \mathcal{X}_S f$ for $[a, b] \supset S$ if this integral exists ¹.

The definition of the integral of $f : S \rightarrow \mathbb{R}$ is independent of the interval $[a, b]$ containing S because $\mathcal{X}_S(x)f(x) = 0$ for $x \in [a, b] \setminus S$.

We leave it to the reader to prove that the countable union of sets of measure zero is a set of measure zero and that a subset of a set of measure zero is a set of measure zero. Since the set of points where $\mathcal{X}_S f : [a, b] \rightarrow \mathbb{R}$ is discontinuous is a subset of the union of ∂S with the set of points where f is discontinuous, and ∂S has measure zero because it is Jordan measurable, we get that the set of points where $\mathcal{X}_S f : [a, b] \rightarrow \mathbb{R}$ is discontinuous has measure zero if the set of point where $f : S \rightarrow \mathbb{R}$ is discontinuous has measure zero. The converse is also true because the set of points where $f : [a, b] \rightarrow \mathbb{R}$ is discontinuous is a subset of the union of ∂S with the set of points where $\mathcal{X}_S f : [a, b] \rightarrow \mathbb{R}$ is discontinuous.

2.2 Darboux Theorem

Consider an integrable function $f : [a, b] \rightarrow \mathbb{R}$. Given a partition $P = \{p_i : 0 \leq i \leq N\}$ of the interval $[a, b]$ and a set of points $P^* = \{p_i^* : 1 \leq i \leq N\}$ such that $p_{i-1} \leq p_i^* \leq p_i$ for all

¹It is understood that f is extended to $[a, b]$. For instance, we may set $f(x) = 0$ for $x \in [a, b] \setminus S$.

i , the **Riemann sum** associated to P and P^* is the sum

$$R(f, P, P^*) = \sum_{i=1}^N f(p_i^*)(p_i - p_{i-1}) .$$

Theorem 2.2.1 (Darboux Theorem)

Suppose that $f : [a, b] \rightarrow \mathbb{R}$ is an integrable function. Given $\epsilon > 0$, there exists $\delta > 0$ such that

$$\left| R(f, P, P^*) - \int_a^b f(x) dx \right| < \epsilon$$

if the partition $P = \{p_i : 0 \leq i \leq N\}$ of $[a, b]$ is such that $|p_i - p_{i-1}| < \delta$ for all i and all sets P^* as defined above.

This theorem is the reason why, in elementary calculus courses, the Riemann integral $\int_a^b f(x) dx$ of a continuous function $f : [a, b] \rightarrow \mathbb{R}$ is defined as the limit of Riemann sums $R(f, P_j, P_j^*)$ where the partition $P_j = \{p_{j,i} : 0 \leq i \leq N_j\}$ of the interval $[a, b]$ is such that $\max_{1 \leq i \leq N_j} (p_{j,i} - p_{j,i-1}) \rightarrow 0$ as $j \rightarrow \infty$. The points in $P_j^* = \{p_{j,i}^* : 1 \leq i \leq N_j\}$ are often chosen to be the right or left end points of the intervals $[p_{j,i-1}, p_{j,i}]$, or the midpoints of these intervals. The previous theorem ensures that the limit is independent of the partitions P_j and the sets P_j^* chosen.

To prove this theorem, we need the following lemma.

Lemma 2.2.2

Suppose that $f : [a, b] \rightarrow \mathbb{R}$ is an integrable function on $[a, b]$. Let $C = \sup\{|f(x)| : a \leq x \leq b\}$ ². Suppose that $P = \{p_i : 0 \leq i \leq N_P\}$ and $Q = \{q_i : 0 \leq i \leq N_Q\}$ are two partitions of $[a, b]$, and P is such that $p_i - p_{i-1} < \delta$ for all i . Then,

$$\mathcal{U}_P(f) < \mathcal{U}_{P \cup Q}(f) + 2CN_Q\delta \quad \text{and} \quad \mathcal{L}_P(f) > \mathcal{L}_{P \cup Q}(f) - 2CN_Q\delta .$$

Proof.

Let $P \cup Q = \{z_j : 0 \leq j \leq N\}$, $M_i = \sup\{f(x) : p_{i-1} \leq x \leq p_i\}$ for $1 \leq i \leq N_P$ and $\tilde{M}_j = \sup\{f(x) : z_{j-1} \leq x \leq z_j\}$ for $1 \leq j \leq N$.

Given $1 \leq i \leq N_P$, there exist integers j_1 and j_2 such that $0 \leq j_1 < j_2 \leq N$, $z_{j_1} = p_{i-1}$ and $z_{j_2} = p_i$. If $j_2 = j_1 + 1$, then

$$\left| M_i(p_i - p_{i-1}) - \tilde{M}_{j_2}(z_{j_2} - z_{j_1}) \right| = 0 .$$

²The constant C is finite because f is Riemann integrable by assumption and therefore f is bounded on $[a, b]$.

Suppose that $j_2 > j_1 + 1$; namely, there is at least one point of Q between p_{i-1} and p_i . Then

$$\begin{aligned} \left| M_i(p_i - p_{i-1}) - \sum_{j=j_1+1}^{j_2} \tilde{M}_j(z_j - z_{j-1}) \right| &= \left| \sum_{j=j_1+1}^{j_2} M_i(z_j - z_{j-1}) \sum_{j=j_1+1}^{j_2} \tilde{M}_j(z_j - z_{j-1}) \right| \\ &\leq \sum_{j=j_1+1}^{j_2} |M_i - \tilde{M}_j| (z_j - z_{j-1}) \leq \sum_{j=j_1+1}^{j_2} 2C(z_j - z_{j-1}) = 2C(p_i - p_{i-1}) < 2C\delta. \end{aligned}$$

The worse scenario is if each interval $]p_{i-1}, p_i[$ contains at most one point of Q . Note that at least two of the points of Q are in P because $p_0 = q_0 = a$ and $p_{N_P} = q_{N_Q} = b$ by definition of a partition. Thus there are at most $N_Q - 2$ cases where $]p_{i-1}, p_i[$ contains only one point of Q . Hence

$$\mathcal{U}_P(f) - \mathcal{U}_{P \cup Q}(f) < 2C\delta(N_Q - 2) < 2C\delta N_Q.$$

The second inequality in the conclusion of the lemma is a consequence of the first one if we replace f by $-f$. ■

Proof (of Theorem 2.2.1).

Choose $\epsilon > 0$. Since $\mathcal{L}_P(f) \leq R(f, P, P^*) \leq \mathcal{U}_P(f)$ for all partitions P of $[a, b]$ and sets P^* as defined above, it is enough to find δ such that

$$\left| \mathcal{L}_P(f) - \int_a^b f(x) dx \right| < \frac{\epsilon}{2} \quad \text{and} \quad \left| \mathcal{U}_P(f) - \int_a^b f(x) dx \right| < \frac{\epsilon}{2} \quad (2.2.1)$$

for all partitions $P = \{p_i : 0 \leq i \leq N_P\}$ of $[a, b]$ satisfying $p_i - p_{i-1} < \delta$. To be more precise, (2.2.1) implies that $\mathcal{U}_P(f) - \mathcal{L}_P(f) < \epsilon$. Since $\int_a^b f(x) dx$ and $R(f, P, P^*)$ are in the interval $[\mathcal{L}_P(f), \mathcal{U}_P(f)]$, we get the conclusion of the theorem.

Since f is integrable on $[a, b]$, there exists a partition $Q = \{q_i : 0 \leq i \leq N_Q\}$ of $[a, b]$ such that

$$\left| \mathcal{L}_Q(f) - \int_a^b f(x) dx \right| < \frac{\epsilon}{4} \quad \text{and} \quad \left| \mathcal{U}_Q(f) - \int_a^b f(x) dx \right| < \frac{\epsilon}{4}. \quad (2.2.2)$$

Let $C = \sup\{|f(x)|; a \leq x \leq b\}$ and $\delta = \epsilon/(8N_Q C)$. Suppose that $P = \{p_i : 0 \leq i \leq N_P\}$ is a partition of $[a, b]$ such that $p_i - p_{i-1} < \delta$ for all i . From the previous lemma and (2.2.2), we have

$$\int_a^b f(x) dx \leq \mathcal{U}_P(f) < \mathcal{U}_{P \cup Q}(f) + 2N_Q C \delta = \mathcal{U}_{P \cup Q}(f) + \frac{\epsilon}{4} \leq \mathcal{U}_Q(f) + \frac{\epsilon}{4} < \int_a^b f(x) dx + \frac{\epsilon}{2}$$

and

$$\int_a^b f(x) dx \geq \mathcal{L}_P(f) > \mathcal{L}_{P \cup Q}(f) - 2N_Q C \delta = \mathcal{L}_{P \cup Q}(f) - \frac{\epsilon}{4} \geq \mathcal{L}_Q(f) - \frac{\epsilon}{4} > \int_a^b f(x) dx - \frac{\epsilon}{2}$$

where we have used Proposition 2.1.4 to obtain several of the inequalities. These two relations imply (2.2.1). ■

2.3 Change of Variable

Theorem 2.3.1

Suppose that $\phi : [\alpha, \beta] \rightarrow \mathbb{R}$ is a differentiable function such that $\phi'(x) > 0$ for all $x \in [\alpha, \beta]$ and such that ϕ' is integrable on $[\alpha, \beta]$. Moreover, suppose that f is integrable on $[a, b] = \phi([\alpha, \beta])$. Then, $(f \circ \phi)\phi'$ is integrable on $[\alpha, \beta]$ and

$$\int_{\alpha}^{\beta} f(\phi(x))\phi'(x) dx = \int_a^b f(x) dx .$$

Proof.

i) Choose $\epsilon > 0$. Since ϕ' is integrable on $[\alpha, \beta]$, there exists a partition $\tilde{P} = \{\tilde{p}_i : 0 \leq i \leq N\}$ of $[\alpha, \beta]$ such that

$$\mathcal{U}_{\tilde{P}}(\phi') - \mathcal{L}_{\tilde{P}}(\phi') = \sum_{i=1}^N (\tilde{M}_i - \tilde{m}_i)(\tilde{p}_i - \tilde{p}_{i-1}) < \epsilon \quad (2.3.1)$$

where $\tilde{M}_i = \sup\{\phi'(y) : \tilde{p}_{i-1} \leq y \leq \tilde{p}_i\}$ and $\tilde{m}_i = \inf\{\phi'(y) : \tilde{p}_{i-1} \leq y \leq \tilde{p}_i\}$ for all i .

Let $p_i = \phi(\tilde{p}_i)$ for $0 \leq i \leq N$. The set $P = \{p_i : 0 \leq i \leq N\} = \phi(\tilde{P})$ is a partition of $[a, b]$ because $\phi : [\alpha, \beta] \rightarrow [a, b]$ is strictly increasing and onto. In fact, every partition of $[a, b]$ is the image by ϕ of a partition of $[\alpha, \beta]$.

It follows from the Mean Value Theorem that there exists $z_i \in [\tilde{p}_{i-1}, \tilde{p}_i]$ such that $\phi(\tilde{p}_i) - \phi(\tilde{p}_{i-1}) = \phi'(z_i)(\tilde{p}_i - \tilde{p}_{i-1})$ for each i .

Choose $u_i \in [p_{i-1}, p_i]$ for $0 \leq i \leq N$. Since $\phi : [\alpha, \beta] \rightarrow [a, b]$ is strictly increasing and onto, there exists (a unique) $w_i \in [\tilde{p}_{i-1}, \tilde{p}_i]$ such that $u_i = \phi(w_i)$ for each i . We have

$$\sum_{i=1}^N f(u_i)(p_i - p_{i-1}) = \sum_{i=1}^N f(\phi(w_i))(\phi(\tilde{p}_i) - \phi(\tilde{p}_{i-1})) = \sum_{i=1}^N f(\phi(w_i))\phi'(z_i)(\tilde{p}_i - \tilde{p}_{i-1}) . \quad (2.3.2)$$

If $C = \sup\{|f(x)| : a \leq x \leq b\}$, then we get from (2.3.1) that

$$\begin{aligned} & \left| \sum_{i=1}^N f(\phi(w_i))\phi'(w_i)(\tilde{p}_i - \tilde{p}_{i-1}) - \sum_{i=1}^N f(\phi(w_i))\phi'(z_i)(\tilde{p}_i - \tilde{p}_{i-1}) \right| \\ & \leq \sum_{i=1}^N |f(\phi(w_i))| |\phi'(w_i) - \phi'(z_i)| (\tilde{p}_i - \tilde{p}_{i-1}) \\ & \leq C \sum_{i=1}^N |\phi'(w_i) - \phi'(z_i)| (\tilde{p}_i - \tilde{p}_{i-1}) \leq C \sum_{i=1}^N (\tilde{M}_i - \tilde{m}_i)(\tilde{p}_i - \tilde{p}_{i-1}) < \epsilon C . \end{aligned} \quad (2.3.3)$$

Let $\check{M}_i = \sup\{f(\phi(y))\phi'(y) : \tilde{p}_{i-1} \leq y \leq \tilde{p}_i\}$ and $M_i = \sup\{f(x) : p_{i-1} \leq x \leq p_i\}$ for all i .

It follows from (2.3.2) and (2.3.3) that

$$\begin{aligned} \sum_{i=1}^N f(u_i)(p_i - p_{i-1}) &\leq \sum_{i=1}^N f(\phi(w_i))\phi'(w_i)(\tilde{p}_i - \tilde{p}_{i-1}) + \epsilon C \\ &\leq \sum_{i=1}^N \check{M}_i(\tilde{p}_i - \tilde{p}_{i-1}) + \epsilon C = \mathcal{U}_{\tilde{P}}((f \circ \phi)\phi') + \epsilon C . \end{aligned}$$

Since this inequality is true for all $u_i \in [p_{i-1}, p_i]$, we get

$$\mathcal{U}_P(f) = \sum_{i=1}^N M_i(p_i - p_{i-1}) \leq \mathcal{U}_{\tilde{P}}((f \circ \phi)\phi') + \epsilon C . \quad (2.3.4)$$

It also follows from (2.3.2) and (2.3.3) that

$$\sum_{i=1}^N f(\phi(w_i))\phi'(w_i)(\tilde{p}_i - \tilde{p}_{i-1}) - \epsilon C \leq \sum_{i=1}^N f(u_i)(p_i - p_{i-1}) \leq \sum_{i=1}^N M_i(p_i - p_{i-1}) = \mathcal{U}_P(f) .$$

Since this inequality is true for all $w_i \in [\tilde{p}_{i-1}, \tilde{p}_i]$, we get

$$\mathcal{U}_{\tilde{P}}((f \circ \phi)\phi') - \epsilon C = \sum_{i=1}^N \check{M}_i(\tilde{p}_i - \tilde{p}_{i-1}) - \epsilon C \leq \mathcal{U}_P(f) . \quad (2.3.5)$$

Finally, we get from (2.3.4) and (2.3.5) that

$$|\mathcal{U}_P(f) - \mathcal{U}_{\tilde{P}}((f \circ \phi)\phi')| \leq \epsilon C . \quad (2.3.6)$$

Let $\check{m}_i = \inf\{f(\phi(y))\phi'(y) : \tilde{p}_{i-1} \leq y \leq \tilde{p}_i\}$ and $m_i = \inf\{f(x) : p_{i-1} \leq x \leq p_i\}$ for all i .

It follows from (2.3.2) and (2.3.3) that

$$\begin{aligned} \sum_{i=1}^N f(u_i)(p_i - p_{i-1}) &\geq \sum_{i=1}^N f(\phi(w_i))\phi'(w_i)(\tilde{p}_i - \tilde{p}_{i-1}) - \epsilon C \\ &\geq \sum_{i=1}^N \check{m}_i(\tilde{p}_i - \tilde{p}_{i-1}) - \epsilon C = \mathcal{L}_{\tilde{P}}((f \circ \phi)\phi') - \epsilon C . \end{aligned}$$

Since this inequality is true for all $u_i \in [p_{i-1}, p_i]$, we get

$$\mathcal{L}_P(f) = \sum_{i=1}^N m_i(p_i - p_{i-1}) \geq \mathcal{L}_{\tilde{P}}((f \circ \phi)\phi') - \epsilon C . \quad (2.3.7)$$

It also follows from (2.3.2) and (2.3.3) that

$$\sum_{i=1}^N f(\phi(w_i))\phi'(w_i)(\tilde{p}_i - \tilde{p}_{i-1}) + \epsilon C \geq \sum_{i=1}^N f(u_i)(p_i - p_{i-1}) \geq \sum_{i=1}^N m_i(p_i - p_{i-1}) = \mathcal{L}_P(f) .$$

Since this inequality is true for all $w_i \in [\tilde{p}_{i-1}, \tilde{p}_i]$, we get

$$\mathcal{L}_{\tilde{P}}((f \circ \phi)\phi') + \epsilon C = \sum_{i=1}^N \tilde{m}_i(\tilde{p}_i - \tilde{p}_{i-1}) + \epsilon C \geq \mathcal{L}_P(f) . \quad (2.3.8)$$

Finally, we get from (2.3.7) and (2.3.8) that

$$|\mathcal{L}_P(f) - \mathcal{L}_{\tilde{P}}((f \circ \phi)\phi')| \leq \epsilon C . \quad (2.3.9)$$

ii) Choose a partition Q_1 of $[a, b]$ and a partition Q_2 of $[\alpha, \beta]$ such that $|\mathcal{U}_{Q_1}(f) - \mathcal{U}(f)| < \epsilon$ and $|\mathcal{U}_{Q_2}((f \circ \phi)\phi') - \mathcal{U}((f \circ \phi)\phi')| < \epsilon$. Choose a partition $\tilde{P}$ such that (2.3.1) is satisfied, $Q_2 \subset \tilde{P}$ and $Q_1 \subset P = \phi(\tilde{P})$. We have from Proposition 2.1.4 that

$$|\mathcal{U}_P(f) - \mathcal{U}(f)| < \epsilon$$

and

$$|\mathcal{U}_{\tilde{P}}((f \circ \phi)\phi') - \mathcal{U}((f \circ \phi)\phi')| < \epsilon$$

because P is a refinement of Q_1 and $\tilde{P}$ is a refinement of Q_2 . It follows from (2.3.6) that

$$\begin{aligned} |\mathcal{U}(f) - \mathcal{U}((f \circ \phi)\phi')| &\leq |\mathcal{U}(f) - \mathcal{U}_P(f)| + |\mathcal{U}_P(f) - \mathcal{U}_{\tilde{P}}((f \circ \phi)\phi')| \\ &\quad + |\mathcal{U}_{\tilde{P}}((f \circ \phi)\phi') - \mathcal{U}((f \circ \phi)\phi')| \leq \epsilon(C + 2) . \end{aligned}$$

and from (2.3.9) that

$$\begin{aligned} |\mathcal{L}(f) - \mathcal{L}((f \circ \phi)\phi')| &\leq |\mathcal{L}(f) - \mathcal{L}_P(f)| + |\mathcal{L}_P(f) - \mathcal{L}_{\tilde{P}}((f \circ \phi)\phi')| \\ &\quad + |\mathcal{L}_{\tilde{P}}((f \circ \phi)\phi') - \mathcal{L}((f \circ \phi)\phi')| \leq \epsilon(C + 2) . \end{aligned}$$

Since ϵ is arbitrary, we finally get $\mathcal{U}(f) = \mathcal{U}((f \circ \phi)\phi')$ and $\mathcal{L}(f) = \mathcal{L}((f \circ \phi)\phi')$.

Thus, f is integrable (i.e. $\mathcal{U}(f) = \mathcal{L}(f)$) if and only if $(f \circ \phi)\phi'$ is integrable (i.e. $\mathcal{U}((f \circ \phi)\phi') = \mathcal{L}((f \circ \phi)\phi')$), and if they are integrable, then we have

$$\int_a^b f(x) dx = \mathcal{L}(f) = \mathcal{U}(f) = \mathcal{U}((f \circ \phi)\phi') = \mathcal{L}((f \circ \phi)\phi') = \int_{\alpha}^{\beta} f(\phi(y))\phi'(y) dy . \quad \blacksquare$$

Remark 2.3.2

Theorem 2.3.1 can be used with some modifications if $\phi : [\alpha, \beta] \rightarrow \mathbb{R}$ is a differentiable function such that $\phi'(x) < 0$ for all $x \in [\alpha, \beta]$ and such that ϕ' is integrable on $[\alpha, \beta]$.

Consider $\tilde{f}(x) = f(-x)$ for $x \in [-\phi(\alpha), -\phi(\beta)]$ and $\tilde{\phi}(x) = -\phi(x)$ for $x \in [\alpha, \beta]$. We have $\tilde{\phi}'(x) = -\phi'(x) > 0$ for all $x \in [\alpha, \beta]$. We may use Theorem 2.3.1 to get

$$\int_{\tilde{\phi}(\alpha)}^{\tilde{\phi}(\beta)} \tilde{f} = \int_{\alpha}^{\beta} (\tilde{f} \circ \tilde{\phi})\tilde{\phi}' = - \int_{\alpha}^{\beta} (f \circ \phi)\phi' .$$

We leave it to the reader to verify that $\int_{\tilde{\phi}(\alpha)}^{\tilde{\phi}(\beta)} \tilde{f} = \int_{-\phi(\alpha)}^{-\phi(\beta)} \tilde{f} = \int_{\phi(\beta)}^{\phi(\alpha)} f$ using the definition of the Riemann integral. We finally get

$$\int_{\phi(\beta)}^{\phi(\alpha)} f = - \int_{\alpha}^{\beta} (f \circ \phi)\phi' .$$



2.4 Sequence of Functions

Theorem 2.4.1

Suppose that $\{f_k\}_{k=0}^{\infty}$ is a sequence of integrable real valued functions defined on the interval $[a, b]$. Moreover, suppose that $\{f_k\}_{k=0}^{\infty}$ converges uniformly on $[a, b]$ to the function $f : [a, b] \rightarrow \mathbb{R}$. Then f is integrable on $[a, b]$ and

$$\lim_{k \rightarrow \infty} \int_a^b f_k(x) \, dx = \int_a^b f(x) \, dx .$$

Proof.

Choose $\epsilon > 0$. Since $\{f_k\}_{k=0}^{\infty}$ converges uniformly on $[a, b]$ to the function $f : [a, b] \rightarrow \mathbb{R}$, there exists $K > 0$ (K depends on ϵ) such that $|f_k(x) - f(x)| < \frac{\epsilon}{4(b-a)}$ for all $x \in [a, b]$ and $k > K$. Thus

$$f_k(x) - \frac{\epsilon}{4(b-a)} < f(x) < f_k(x) + \frac{\epsilon}{4(b-a)} \quad (2.4.1)$$

for all $x \in [a, b]$ and $k > K$.

i) Choose $m > K$. Since f_m is integrable on $[a, b]$, there exists a partition P_m of $[a, b]$ such that

$$\mathcal{U}_{P_m}(f_m) - \mathcal{L}_{P_m}(f_m) < \epsilon/2 . \quad (2.4.2)$$

Let $g_m(x) = f_m(x) - \frac{\epsilon}{4(b-a)}$ and $h_m(x) = f_m(x) + \frac{\epsilon}{4(b-a)}$ for $x \in [a, b]$. It follows from (2.4.1) and Proposition 2.1.10 that

$$\mathcal{L}_{P_m}(f_m) - \frac{\epsilon}{4} = \mathcal{L}_{P_m}(g_m) \leq \mathcal{L}_{P_m}(f) \leq \mathcal{U}_{P_m}(f) \leq \mathcal{U}_{P_m}(h_m) = \mathcal{U}_{P_m}(f_m) + \frac{\epsilon}{4} .$$

Combining this relation with (2.4.2) yields

$$\mathcal{L}_{P_m}(f_m) - \frac{\epsilon}{4} \leq \mathcal{L}_{P_m}(f) \leq \mathcal{U}_{P_m}(f) < \mathcal{L}_{P_m}(f_m) + \frac{3\epsilon}{4} .$$

It follows that $|\mathcal{U}_{P_m}(f) - \mathcal{L}_{P_m}(f)| < \epsilon$. Since $\epsilon > 0$ is arbitrary, we get from Theorem 2.1.8 that f is integrable on $[a, b]$.

ii) Now that we know that f is integrable on $[a, b]$, it follows from (2.4.1) that

$$\begin{aligned} \int_a^b f_k(x) \, dx - \frac{\epsilon}{4} &= \int_a^b \left(f_k(x) - \frac{\epsilon}{4(b-a)} \right) \, dx \leq \int_a^b f(x) \, dx \\ &\leq \int_a^b \left(f_k(x) + \frac{\epsilon}{4(b-a)} \right) \, dx = \int_a^b f_k(x) \, dx + \frac{\epsilon}{4} . \end{aligned}$$

Thus

$$\left| \int_a^b f(x) \, dx - \int_a^b f_k(x) \, dx \right| \leq \frac{\epsilon}{4} \quad (2.4.3)$$

for $k > K$. Since, for any $\epsilon > 0$, we can find K satisfying (2.4.3), we get that

$$\lim_{k \rightarrow \infty} \int_a^b f_k(x) \, dx = \int_a^b f(x) \, dx$$

by definition of limit of sequences. ■

In the previous theorem, the uniform convergence of the sequence $\{f_n\}_{n=0}^\infty$ of integrable real valued functions defined on the interval $[a, b]$ can be replaced by pointwise convergence if the sequence $\{f_n\}_{n=0}^\infty$ is uniformly bounded and converges pointwise toward an integrable function.

Theorem 2.4.2 (Bounded Convergence Theorem)

Suppose that $\{f_k\}_{k=0}^\infty$ is a sequence of integrable real valued functions defined on an interval $[a, b]$, $\{f_k\}_{k=0}^\infty$ converges pointwise to an integrable function $f : [a, b] \rightarrow \mathbb{R}$, and there exists $C > 0$ such that $|f_k(x)| \leq C$ for all $x \in [a, b]$ and $k \in \mathbb{N}$. Then

$$\lim_{k \rightarrow \infty} \int_a^b f_k(x) \, dx = \int_a^b f(x) \, dx .$$

This theorem is a corollary of the Lebesgue dominated convergence theorem (see [23]). The proof of the Lebesgue dominated convergence theorem requires the Lebesgue's theory of integration. Therefore, it will not be given here. Instead, we will present an elementary proof of Theorem 2.4.2 that was given in [16]. We first need to prove a few results.

Proposition 2.4.3 (Dini's Uniform Convergence Theorem)

Suppose that $D \subset \mathbb{R}$ is a compact set, $f : D \rightarrow \mathbb{R}$ is a continuous function, $\{f_k\}_{k=0}^\infty$ is a sequence of continuous real valued functions on D such that $f_{k+1}(x) \leq f_k(x)$ for all $x \in D$ and all $k \geq 0$ (or $f_{k+1}(x) \geq f_k(x)$ for all $x \in D$ and all $k \geq 0$), and $\lim_{k \rightarrow \infty} f_k(x) = f(x)$ for all $x \in D$. Then $f_k \rightarrow f$ uniformly on D as $k \rightarrow \infty$.

Proof.

i) We first consider the case where $\{f_k\}_{k=0}^\infty$ is a sequence of continuous real valued functions on D such that $f_{k+1}(x) \leq f_k(x)$ for all $x \in D$ and all $k \geq 0$, and $\lim_{k \rightarrow \infty} f_k(x) = 0$ for all $x \in D$.

Choose $\epsilon > 0$. For each $x \in D$, there exists $K_x \in \mathbb{N}$ such that $f_k(x) < \epsilon/2$ for $k \geq K_x$.

Since $f_k : D \rightarrow \mathbb{R}$ is continuous, there exists for each $x \in D$ a $\delta_{k,x} > 0$ such that $|f_k(x) - f_k(y)| < \epsilon/2$ for all $y \in]x - \delta_{k,x}, x + \delta_{k,x}[$.

The collection of open intervals $\{]x - \delta_{K_x,x}, x + \delta_{K_x,x}[\}_{x \in D}$ forms an open cover of the compact set D . Therefore, there is a finite subcover $\{]x_j - \delta_{K_{x_j},x_j}, x_j + \delta_{K_{x_j},x_j}[\}_{0 \leq j \leq J}$ of D .

Let $K = \max_{1 \leq j \leq J} K_{x_j}$. Given $x \in D$, we have $x \in]x_j - \delta_{K_{x_j}, x_j}, x_j + \delta_{K_{x_j}, x_j}[$ for some $0 \leq j \leq J$. Hence $|f_{K_{x_j}}(x) - f_{K_{x_j}}(x_j)| < \epsilon/2$. Moreover $0 \leq f_{K_{x_j}}(x_j) < \epsilon/2$. Hence

$$0 \leq f_k(x) \leq f_K(x) \leq f_{K_{x_j}}(x) = |f_{K_{x_j}}(x_j)| + |f_{K_{x_j}}(x) - f_{K_{x_j}}(x_j)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

for all $k \geq K$. Since $x \in D$ is arbitrary, we have $0 \leq f_k(x) < \epsilon$ for all $x \in D$ and $k \geq K$. This proves that $f_k \rightarrow 0$ uniformly on D as $k \rightarrow \infty$.

ii) To proof the proposition, it suffices to apply (i) to $\{g_k\}_{k=1}^\infty$ where $g_k = f_k - f$ for all $k \geq 0$ (or $g_k = f - f_k$ for all $k \geq 0$). ■

Lemma 2.4.4

Suppose that $f : [a, b] \rightarrow [0, \infty[$ is a bounded function. For each $\epsilon > 0$, there exists a continuous function $g : [a, b] \rightarrow [0, \infty[$ such that $g(x) \leq f(x)$ for all $x \in [a, b]$ and

$$\mathcal{L}(f) - \int_a^b g < \epsilon.$$

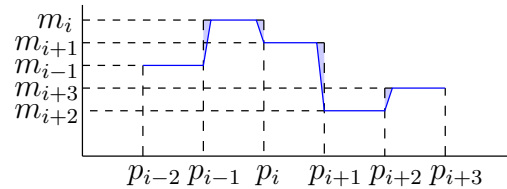
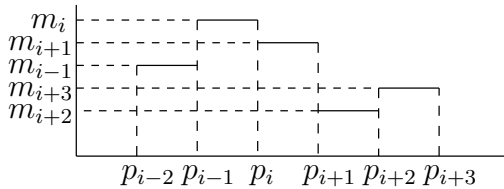
Proof.

By definition of the lower Riemann integral, there exists a partition $P = \{p_i : 0 \leq i \leq N\}$ of $[a, b]$ such that $\mathcal{L}(f) - \mathcal{L}_P(f) < \epsilon/2$. If $m_i = \inf\{f(x) : x \in [p_{i-1}, p_i]\}$ for $1 \leq i \leq N$, then we have

$$\mathcal{L}_P(f) = \sum_{i=1}^N m_i(p_i - p_{i-1}) = \int_a^b h,$$

where $h = \sum_{i=1}^N m_i \chi_{[p_{i-1}, p_i]}$. Hence $\mathcal{L}(f) - \int_a^b h < \frac{\epsilon}{2}$. We can construct a piecewise linear

function $g : [a, b] \rightarrow \mathbb{R}$ such that $g(x) \leq h(x)$ for all $x \in [a, b]$ and $\int_a^b h - \int_a^b g < \frac{\epsilon}{2}$.



The graph of g is in blue. The area in blue is smaller than $\epsilon/2$. Thus $\mathcal{L}(f) - \int_a^b g < \epsilon$. ■

Lemma 2.4.5

Suppose that $\{f_k\}_{k=0}^\infty$ is a sequence of bounded real valued functions on $[a, b]$ such that $f_{k+1}(x) \leq f_k(x)$ for all $x \in [a, b]$ and $k \geq 0$, and $\lim_{k \rightarrow \infty} f_k(x) = 0$ for all $x \in [a, b]$. Then

$$\lim_{k \rightarrow \infty} \mathcal{L}(f_k) = 0.$$

Proof.

Choose $\epsilon > 0$. It follows from the previous lemma that, for each function f_k , there exists a continuous function $g_k : [a, b] \rightarrow [0, \infty[$ such that $g_k(x) \leq f_k(x)$ for all $x \in [a, b]$ and

$$\mathcal{L}(f_k) - \int_a^b g_k < \frac{\epsilon}{2^{k+1}}.$$

Let $h_k = \min_{0 \leq j \leq k} g_j$ for $k \geq 0$. We have that $h_k : [a, b] \rightarrow \mathbb{R}$ is a continuous function because the g_j are continuous functions, and $h_{k+1}(x) \leq h_k(x)$ for all $x \in [a, b]$ and $k \geq 0$. Moreover $0 \leq h_k(x) \leq g_k(x) \leq f_k(x)$ for all $x \in [a, b]$ and $k \geq 0$. Thus $\lim_{k \rightarrow \infty} h_k(x) = 0$ for all $x \in [a, b]$ by the famous sandwich theorem. It then follows from Proposition 2.4.3 that $h_k \rightarrow 0$ uniformly on $[a, b]$ as $k \rightarrow \infty$. Thus $\lim_{k \rightarrow \infty} \int_a^b h_k = \int_0^b 0 = 0$ according to Theorem 2.4.1.

We now prove that

$$0 \leq \mathcal{L}(f_k) \leq \epsilon \left(1 - \frac{1}{2^k}\right) + \int_a^b h_k \quad (2.4.4)$$

for all k . Hence, if we let k converges to ∞ , then we get $0 \leq \lim_{k \rightarrow \infty} \mathcal{L}(f_k) \leq \epsilon$. Since ϵ is arbitrary, this will prove that $\lim_{k \rightarrow \infty} \mathcal{L}(f_k) = 0$.

Let $H_{i,k} = \max_{i \leq j \leq k} g_j$ for $0 \leq i \leq k$. For $0 \leq i \leq k$, we have that $H_{i,k} : [a, b] \rightarrow \mathbb{R}$ is a continuous function because the g_j are continuous functions. Moreover, for $0 \leq j < k$, we have

$$\begin{aligned} 0 \leq g_k(x) &= g_j(x) + (g_k(x) - g_j(x)) \leq g_j(x) + (H_{j,k}(x) - g_j(x)) \\ &\leq g_j(x) + \sum_{i=0}^{k-1} (H_{i,k}(x) - g_i(x)) \end{aligned}$$

for all $x \in [a, b]$. The previous inequality is also true for $j = k$ because $H_{i,k}(x) - g_i(x) \geq 0$ for all $x \in [a, b]$ and $0 \leq i \leq k$. Hence

$$0 \leq g_k(x) \leq h_k(x) + \sum_{i=0}^{k-1} (H_{i,k}(x) - g_i(x)) \quad (2.4.5)$$

for all $x \in [a, b]$ and $k \geq 0$. Since $H_{i,k}(x) \leq \max_{i \leq j \leq k} f_j(x) \leq f_i(x)$ for all $x \in [a, b]$ and $0 \leq i < k$, we have $\mathcal{L}_P(H_{i,k}) \leq \mathcal{L}_P(f_i)$ for all partitions P of $[a, b]$ and $0 \leq i < k$. Thus

$$\mathcal{L}(f_i) \geq \mathcal{L}(H_{i,k}) = \int_a^b H_{i,k} = \int_a^b (H_{i,k} - g_i) + \int_a^b g_i$$

for $0 \leq i < k$. We get

$$\int_a^b (H_{i,k} - g_i) \leq \mathcal{L}(f_i) - \int_a^b g_i < \frac{\epsilon}{2^{i+1}}$$

for $0 \leq i < k$. It follows from (2.4.5) that

$$0 \leq \int_a^b g_k \leq \int_a^b h_k + \sum_{i=1}^{k-1} \int_a^b (H_{i,k} + g_i) \leq \int_a^b h_k + \sum_{i=0}^{k-1} \frac{\epsilon}{2^{i+1}}.$$

Finally, we get (2.4.4) from

$$\mathcal{L}(f_k) < \int_a^b g_k + \frac{\epsilon}{2^{k+1}} \leq \int_a^b h_k + \sum_{i=0}^{k-1} \frac{\epsilon}{2^{i+1}} + \frac{\epsilon}{2^k} = \int_a^b h_k + \sum_{i=0}^k \frac{\epsilon}{2^{i+1}}. \quad \blacksquare$$

Proof (of Theorem 2.4.2).

Let $g_k = |f_k - f|$ for $k \geq 1$. We have $0 \leq g_k \leq 2C$ for all $k \geq 0$ and $\lim_{k \rightarrow \infty} g_k(x) = 0$ for all $x \in [a, b]$.

Let $h_k : [a, b] \rightarrow \mathbb{R}$ be the function defined by $h_k(x) = \sup_{i \geq k} g_i(x)$ for all $x \in [a, b]$ and $k \geq 0$. Note that $0 \leq h_k(x) \leq 2C$ for all $x \in [a, b]$ and $k \geq 0$. Moreover $h_{k+1}(x) \leq h_k(x)$ for all $x \in [a, b]$ and $k \geq 0$ and $\lim_{k \rightarrow \infty} h_k(x) = \limsup_{k \rightarrow \infty} g_k(x) = \lim_{k \rightarrow \infty} g_k(x) = 0$ for all $x \in [a, b]$ because $\lim_{k \rightarrow \infty} f_k(x) = f(x)$ for all $x \in [a, b]$. It follows from Lemma 2.4.5 that $\lim_{k \rightarrow \infty} \mathcal{L}(h_k) = 0$.

Since $g_k \leq h_k$ for all $k \geq 0$, we have $0 \leq \int_a^b g_k = \mathcal{L}(g_k) \leq \mathcal{L}(h_k) \rightarrow 0$ as $k \rightarrow \infty$.

Since f and the f_k are integrable, $f_k - f$ is integrable and we get from Proposition 2.1.12 that

$$\left| \int_a^b f_k - \int_a^b f \right| = \left| \int_a^b (f_k - f) \right| \leq \int_a^b |f_k - f| = \int_a^b g_k \rightarrow 0$$

as $k \rightarrow \infty$. \blacksquare

2.5 Exercises

Though we have not proved the Fundamental Theorem of Calculus and integration by parts, they may be useful to answer some of the problems below. The reader may consult [7, 23] for proofs of these results. We have limited the content of this chapter to the results that best fit the goal of these lecture notes.

Question 2.1

Let

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

Prove that f is not Riemann integrable on any interval.

Solution

Given any interval $[a, b]$ and any partition $P = \{p_i : 0 \leq i \leq N\}$ of $[a, b]$, we have $m_i = \inf\{f(x) : p_{i-1} \leq x \leq p_i\} = 0$ and $M_i = \sup\{f(x) : p_{i-1} \leq x \leq p_i\} = 1$ for all i because

$[p_{i-1}, p_i]$ contains rational and irrational numbers. Thus $\mathcal{L}_P(f) = \sum_{i=1}^N m_i(p_i - p_{i-1}) = 0$ and

$$\mathcal{U}_P(f) = \sum_{i=1}^N M_i(p_i - p_{i-1}) = b - a > 0.$$

Hence $\mathcal{U}(f) = \inf\{\mathcal{U}_P(f) : P \text{ is a partition of } [a, b]\} = b - a$ is not equal to $\mathcal{L}(f) = \sup\{\mathcal{L}_P(f) : P \text{ is a partition of } [a, b]\} = 0$. $\heartsuit$

Question 2.2

Suppose that $f : [a, b] \rightarrow \mathbb{R}$ is an integrable function and $c \in \mathbb{R}$. Using the definition of the integral, prove that $cf : [a, b] \rightarrow \mathbb{R}$ is integrable and $\int_a^b cf(x) dx = c \int_a^b f(x) dx$.

Solution

The result is trivial for $c = 0$. We may therefore assume that $c \neq 0$.

i) We prove that $cf : [a, b] \rightarrow \mathbb{R}$ is integrable.

Given $\epsilon > 0$, since f is integrable on $[a, b]$, there exists a partition $P = \{p_i : 0 \leq i \leq N\}$ of the interval $[a, b]$ such that $\mathcal{U}_P(f) - \mathcal{L}_P(f) < \epsilon/|c|$. Let $m_i = \inf\{f(x) : p_{i-1} \leq x \leq p_i\}$, $M_i = \sup\{f(x) : p_{i-1} \leq x \leq p_i\}$, $\tilde{m}_i = \inf\{cf(x) : p_{i-1} \leq x \leq p_i\}$ and $\tilde{M}_i = \sup\{cf(x) : p_{i-1} \leq x \leq p_i\}$.

If $c > 0$, then $\tilde{m}_i = cm_i$ and $\tilde{M}_i = cM_i$ for all i . Thus

$$\mathcal{L}_P(cf) = \sum_{i=1}^N \tilde{m}_i(p_i - p_{i-1}) = c \sum_{i=1}^N m_i(p_i - p_{i-1}) = c\mathcal{L}_P(f)$$

and

$$\mathcal{U}_P(cf) = \sum_{i=1}^N \tilde{M}_i(p_i - p_{i-1}) = c \sum_{i=1}^N M_i(p_i - p_{i-1}) = c\mathcal{U}_P(f).$$

Hence

$$\mathcal{U}_P(cf) - \mathcal{L}_P(cf) = c(\mathcal{U}_P(f) - \mathcal{L}_P(f)) < c(\epsilon/c) = \epsilon. \quad (2.5.1)$$

If $c < 0$, then $\tilde{m}_i = cM_i$ and $\tilde{M}_i = cm_i$ for all i . Thus

$$\mathcal{L}_P(cf) = \sum_{i=1}^N \tilde{m}_i(p_i - p_{i-1}) = c \sum_{i=1}^N M_i(p_i - p_{i-1}) = c\mathcal{U}_P(f)$$

and

$$\mathcal{U}_P(cf) = \sum_{i=1}^N \tilde{M}_i(p_i - p_{i-1}) = c \sum_{i=1}^N m_i(p_i - p_{i-1}) = c\mathcal{L}_P(f).$$

Hence

$$\mathcal{U}_P(cf) - \mathcal{L}_P(cf) = c(\mathcal{L}_P(f) - \mathcal{U}_P(f)) = |c|(\mathcal{U}_P(f) - \mathcal{L}_P(f)) < |c|(\epsilon/|c|) = \epsilon \quad (2.5.2)$$

We get from (2.5.1) and (2.5.2) that $\mathcal{U}_P(cf) - \mathcal{L}_P(cf) < \epsilon$ independently of the sign of c . It follows from Theorem 2.1.8 that $cf : [a, b] \rightarrow \mathbb{R}$ is Riemann integrable.

ii) We prove that

$$\int_a^b cf(x) dx = c \int_a^b f(x) dx . \quad (2.5.3)$$

Let $\epsilon > 0$ and P be a partition of $[a, b]$ such that $\mathcal{U}_P(cf) - \mathcal{L}_P(cf) < \epsilon$. From

$$\mathcal{L}_P(f) \leq \int_a^b f(x) dx \leq \mathcal{U}_P(f) ,$$

we get

$$\mathcal{L}_P(cf) = c\mathcal{L}_P(f) \leq c \int_a^b f(x) dx \leq c\mathcal{U}_P(f) = \mathcal{U}_P(cf)$$

if $c > 0$ and

$$\mathcal{L}_P(cf) = c\mathcal{U}_P(f) \leq c \int_a^b f(x) dx \leq c\mathcal{L}_P(f) = \mathcal{U}_P(cf)$$

if $c < 0$. So $\int_a^b cf(x) dx$ and $c \int_a^b f(x) dx$ are in the interval $[\mathcal{L}_P(cf), \mathcal{U}_P(cf)]$ of length at most ϵ . Since ϵ is arbitrary, we get (2.5.3). $\heartsuit$

Question 2.3

Suppose that $f : [a, b] \rightarrow \mathbb{R}$ is an integrable function on $[a, b]$ and $c \in \mathbb{R}$. Using the definition of the integral (i.e. without using substitution), prove that the function $g : [a - c, b - c] \rightarrow \mathbb{R}$ defined by $g(x) = f(x + c)$ for $x \in [a - c, b - c]$ is integrable, and that $\int_a^b f(x) dx = \int_{a-c}^{b-c} f(x + c) dx$.

Solution

i) We prove that $g : [a - c, b - c] \rightarrow \mathbb{R}$ is integrable.

Given $\epsilon > 0$, there exists a partition $P = \{p_i : 0 \leq i \leq N\}$ of the interval $[a, b]$ such that $\mathcal{U}_P(f) - \mathcal{L}_P(f) < \epsilon$ since $f : [a, b] \rightarrow \mathbb{R}$ is integrable. Let $m_i = \inf\{f(x) : p_{i-1} \leq x \leq p_i\}$ and $M_i = \sup\{f(x) : p_{i-1} \leq x \leq p_i\}$. We have

$$\mathcal{L}_P(f) = \sum_{i=1}^N m_i(p_i - p_{i-1}) \quad \text{and} \quad \mathcal{U}_P(f) = \sum_{i=1}^N M_i(p_i - p_{i-1}) .$$

Consider the function $g : [a - c, b - c] \rightarrow \mathbb{R}$ defined by $g(x) = f(x + c)$ for $x \in [a - c, b - c]$. Let $Q = \{q_i : q_i = p_i - c \text{ for } 0 \leq i \leq N\}$. We have that Q is a partition of $[a - c, b - c]$. Moreover

$$\begin{aligned} \tilde{m}_i &= \inf\{g(x) : q_{i-1} \leq x \leq q_i\} = \inf\{f(x + c) : p_{i-1} - c \leq x \leq p_i - c\} \\ &= \inf\{f(x + c) : p_{i-1} \leq x + c \leq p_i\} = \inf\{f(x) : p_{i-1} \leq x \leq p_i\} = m_i \end{aligned}$$

and

$$\tilde{M}_i = \sup\{g(x) : q_{i-1} \leq x \leq q_i\} = \sup\{f(x + c) : p_{i-1} - c \leq x \leq p_i - c\}$$

$$= \sup\{f(x+c) : p_{i-1} \leq x+c \leq p_i\} = \sup\{f(x) : p_{i-1} \leq x \leq p_i\} = M_i$$

for all i . Thus

$$\mathcal{L}_Q(g) = \sum_{i=1}^N \tilde{m}_i(q_i - q_{i-1}) = \sum_{i=1}^N m_i(p_i - p_{i-1}) = \mathcal{L}_P(f)$$

and

$$\mathcal{U}_Q(g) = \sum_{i=1}^N \tilde{M}_i(q_i - q_{i-1}) = \sum_{i=1}^N M_i(p_i - p_{i-1}) = \mathcal{U}_P(f) .$$

Hence $\mathcal{U}_Q(g) - \mathcal{L}_Q(g) = \mathcal{U}_P(f) - \mathcal{L}_P(f) < \epsilon$. Since ϵ is arbitrary, it follows from Theorem 2.1.8 that g is integrable on $[a-c, b-c]$.

ii) We prove that

$$\int_a^b f(x) dx = \int_{a-c}^{b-c} f(x+c) dx \quad (2.5.4)$$

Given $\epsilon > 0$, let $P = \{p_i : 0 \leq i \leq N\}$ be a partition of the interval $[a, b]$ such that $\mathcal{U}_P(f) - \mathcal{L}_P(f) < \epsilon$ and $Q = \{q_i : q_i = p_i - c \text{ for } 0 \leq i \leq N\}$ be the partition of $[a-c, b-c]$ associated to P as defined in (i). We have

$$\mathcal{L}_P(f) \leq \int_a^b f(x) dx \leq \mathcal{U}_P(f)$$

and

$$\mathcal{L}_P(f) = \mathcal{L}_Q(g) \leq \int_{a-c}^{b-c} g(x) dx = \int_{a-c}^{b-c} f(x+c) dx \leq \mathcal{U}_Q(g) = \mathcal{U}_P(f) .$$

Thus, $\int_a^b f(x) dx$ and $\int_{a-c}^{b-c} f(x+c) dx$ are both in the interval $[\mathcal{L}_P(f), \mathcal{U}_P(f)]$ of length at most ϵ . Since ϵ is arbitrary, we get (2.5.4). $\heartsuit$

Question 2.4

Suppose that $f : [a, b] \rightarrow \mathbb{R}$ is an integrable function on $[a, b]$ and $c > 0$. Using the definition of the integral (i.e. without using substitution), prove that the function $g : [a/c, b/c] \rightarrow \mathbb{R}$ defined by $g(x) = f(cx)$ for $x \in [a/c, b/c]$ is integrable, and that $\frac{1}{c} \int_a^b f(x) dx = \int_{a/c}^{b/c} f(cx) dx$.

Solution

i) We prove that $g : [a/c, b/c] \rightarrow \mathbb{R}$ is integrable.

Given $\epsilon > 0$, there exists a partition $P = \{p_i : 0 \leq i \leq N\}$ of the interval $[a, b]$ such that $\mathcal{U}_P(f) - \mathcal{L}_P(f) < \epsilon c$ since $f : [a, b] \rightarrow \mathbb{R}$ is integrable. Let $Q = \{q_i : q_i = p_i/c \text{ for } 0 \leq i \leq N\}$. Thus Q is a partition of $[a/c, b/c]$. Moreover

$$\begin{aligned} \tilde{m}_i &= \inf\{g(x) : q_{i-1} \leq x \leq q_i\} = \inf\{f(cx) : p_{i-1}/c \leq x \leq p_i/c\} \\ &= \inf\{f(cx) : p_{i-1} \leq cx \leq p_i\} = \inf\{f(x) : p_{i-1} \leq x \leq p_i\} = m_i \end{aligned}$$

and

$$\begin{aligned}\tilde{M}_i &= \sup\{g(x) : q_{i-1} \leq x \leq q_i\} = \sup\{f(cx) : p_{i-1}/c \leq x \leq p_i/c\} \\ &= \sup\{f(cx) : p_{i-1} \leq cx \leq p_i\} = \sup\{f(x) : p_{i-1} \leq x \leq p_i\} = M_i.\end{aligned}$$

Thus

$$\begin{aligned}\mathcal{U}_Q(g) - \mathcal{L}_Q(g) &= \sum_{i=1}^N \tilde{M}_i(q_i - q_{i-1}) - \sum_{i=1}^N \tilde{m}_i(q_i - q_{i-1}) \\ &= \frac{1}{c} \left(\sum_{i=1}^N M_i(p_i - p_{i-1}) - \sum_{i=1}^N m_i(p_i - p_{i-1}) \right) = \frac{1}{c} (\mathcal{U}_P(f) - \mathcal{L}_P(f)) < \epsilon.\end{aligned}$$

Since ϵ is arbitrary, it follows from Theorem 2.1.8 that $g : [a/c, b/c] \rightarrow \mathbb{R}$ is integrable.

ii) We prove that

$$\frac{1}{c} \int_a^b f(x) dx = \int_{a/c}^{b/c} f(cx) dx \quad (2.5.5)$$

Given $\epsilon > 0$, let $P = \{p_i : 0 \leq i \leq N\}$ be a partition of the interval $[a, b]$ such that $\mathcal{U}_P(f) - \mathcal{L}_P(f) < c\epsilon$ as in (i). As before, let $Q = \{q_i : q_i = p_i/c \text{ for } 0 \leq i \leq N\}$ be the partition of $[a/c, b/c]$ associated with P as in (i).

We have

$$\mathcal{L}_Q(g) \leq \int_{a/c}^{b/c} g(x) dx = \int_{a/c}^{b/c} f(cx) dx \leq \mathcal{U}_Q(g).$$

Moreover

$$\mathcal{L}_Q(g) = \frac{1}{c} \mathcal{L}_P(f) \leq \frac{1}{c} \int_a^b f(x) dx \leq \frac{1}{c} \mathcal{U}_P(f) = \mathcal{U}_Q(g)$$

because

$$\mathcal{L}_P(f) \leq \int_a^b f(x) dx \leq \mathcal{U}_P(f).$$

Thus

$$\left| \frac{1}{c} \int_a^b f(x) dx - \int_{a/c}^{b/c} f(cx) dx \right| \leq \mathcal{U}_Q(g) - \mathcal{L}_Q(g) < \epsilon.$$

Since ϵ is arbitrary, we get (2.5.5). ♡

Question 2.5

Suppose that $f : [a, b] \rightarrow \mathbb{R}$ is an integrable function on $[a, b]$. Using the definition of the integral (i.e. without using substitution), prove that the function $g : [-b, -a] \rightarrow \mathbb{R}$ defined by $g(x) = f(-x)$ for $x \in [-b, -a]$ is integrable, and that $\int_a^b f(x) dx = \int_{-b}^{-a} f(-x) dx$.

Question 2.6

Suppose that $f, g : [a, b] \rightarrow \mathbb{R}$ are two integrable functions. Prove that $h_m : [a, b] \rightarrow \mathbb{R}$ and $h_M : [a, b] \rightarrow \mathbb{R}$ defined by $h_m(x) = \min\{f(x), g(x)\}$ and $h_M(x) = \max\{f(x), g(x)\}$ for all

$x \in [a, b]$ are also integrable functions.

Hint: Theorem 2.1.17.

Question 2.7

Suppose that $[c, d] \subset [a, b]$ and $f[a, b] \rightarrow \mathbb{R}$ is integrable. Prove that f is integrable on $[c, d]$.

Solution

Given $\epsilon > 0$, choose a partition P of $[a, b]$ such that $\mathcal{U}_P(f) - \mathcal{L}_P(f) < \epsilon$.

Let $R = P \cup \{c, d\}$. Then $R = \{r_i : 0 \leq i \leq N\}$ is a refinement of P containing c and d . In particular, $\mathcal{U}_R(f) - \mathcal{L}_R(f) < \epsilon$.

Let $Q = R \cap [c, d]$. We have that Q is a partition of $[c, d]$ and $Q = \{r_i : J_1 \leq j \leq J_2\}$ for $0 \leq J_1 < J_2 \leq N$.

Let $M_i = \sup\{f(x) : r_{i-1} \leq x \leq r_i\}$ and $m_i = \inf\{f(x) : r_{i-1} \leq x \leq r_i\}$ for $1 \leq i \leq N$. We then have

$$\begin{aligned} \mathcal{U}_Q(f) - \mathcal{L}_Q(f) &= \sum_{j=J_1+1}^{J_2} M_j (r_j - r_{j-1}) - \sum_{j=J_1+1}^{J_2} m_j (r_j - r_{j-1}) \\ &= \sum_{j=J_1+1}^{J_2} (M_j - m_j) (r_j - r_{j-1}) \leq \sum_{j=1}^N (M_j - m_j) (r_j - r_{j-1}) \\ &\leq \sum_{j=1}^N M_j (r_j - r_{j-1}) - \sum_{j=1}^N m_j (r_j - r_{j-1}) = \mathcal{U}_R(f) - \mathcal{L}_R(f) < \epsilon. \end{aligned}$$

Since ϵ is arbitrary, it follows from Theorem 2.1.8 that f is integrable on $[c, d]$. ♡

Question 2.8

Give a second proof of Proposition 2.1.11 if $g : [m, M] \rightarrow \mathbb{R}$ is continuously differentiable.

Solution

For every x and y in $[m, M]$, the Mean Value Theorem says that there exists $\xi(x, y)$ between x and y such that $g(x) - g(y) = g'(\xi(x, y))(x - y)$. Since g' is continuous on the compact set $[m, M]$, we have that g' is bounded on $[m, M]$. Choose $K > 0$ such that $|g'(x)| < K$ for all $x \in [m, M]$. We get

$$|g(x) - g(y)| = |g'(\xi(x, y))| |x - y| < K|x - y|$$

for all x and y in $[m, M]$.

Since g is continuous on the compact set $[m, M]$, we have that g is bounded on $[m, M]$. Hence, $g \circ f : [a, b] \rightarrow \mathbb{R}$ is bounded because $f(x) \in [m, M]$ for all $x \in [a, b]$.

Given $\epsilon > 0$, choose a partition $P = \{p_i : 0 \leq i \leq N\}$ of $[a, b]$ such that $\mathcal{U}_P(f) - \mathcal{L}_P(f) < \epsilon/K$. As usual, let $m_i = \inf\{f(x) : p_{i-1} \leq x \leq p_i\}$ and $M_i = \sup\{f(x) : p_{i-1} \leq x \leq p_i\}$. Since

$$|(g \circ f)(x) - (g \circ f)(y)| = |g(f(x)) - g(f(y))| < K|f(x) - f(y)| \leq K(M_i - m_i)$$

for all x and y in $[p_{i-1}, p_i]$, we get $\tilde{M}_i - \tilde{m}_i < K(M_i - m_i)$ where $\tilde{m}_i = \inf\{(g \circ f)(x) : p_{i-1} \leq x \leq p_i\}$ and $\tilde{M}_i = \sup\{(g \circ f)(x) : p_{i-1} \leq x \leq p_i\}$. Hence

$$\begin{aligned}\mathcal{U}_P(g \circ f) - \mathcal{L}_P(g \circ f) &= \sum_{i=1}^N (\tilde{M}_i - \tilde{m}_i)(p_i - p_{i-1}) \leq K \sum_{i=1}^N (M_i - m_i)(p_i - p_{i-1}) \\ &= K(\mathcal{U}_P(f) - \mathcal{L}_P(f)) < \epsilon.\end{aligned}$$

♡

Question 2.9

Let $\{x_n\}_{n=0}^{\infty}$ be a convergent sequence in $\mathbb{R}$. Show that the set $\{x_n : n \geq 0\}$ has content zero.

Solution

Let $x = \lim_{n \rightarrow \infty} x_n$. Given $\epsilon > 0$, choose N such that $|x_n - x| < \frac{\epsilon}{4}$ for $n > N$.

Let $R_n = [x_n - \frac{\epsilon}{5^{n+1}}, x_n + \frac{\epsilon}{5^{n+1}}]$ for $0 \leq n \leq N$. Moreover, let $R_{N+1} = [x - \frac{\epsilon}{4}, x + \frac{\epsilon}{4}]$.

We have $\{x_n : n \geq 0\} \subset \bigcup_{n=0}^{N+1} R_n$ and

$$\begin{aligned}\sum_{n=0}^{N+1} |R_n| &= |R_{N+1}| + \sum_{n=0}^N |R_n| = \frac{\epsilon}{2} + \sum_{n=0}^N \frac{2\epsilon}{5^{n+1}} = \frac{\epsilon}{2} + \frac{2\epsilon}{5} \sum_{n=0}^N \left(\frac{1}{5}\right)^n \\ &= \frac{\epsilon}{2} + \frac{2\epsilon}{5} \left(\frac{1 - (1/5)^{N+1}}{1 - (1/5)}\right) < \frac{\epsilon}{2} + \frac{2\epsilon}{5} \left(\frac{1}{1 - (1/5)}\right) = \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.\end{aligned}$$

Since ϵ is arbitrary, this proves that $\{x_n : n \geq 0\}$ has content zero.

♡

Question 2.10

Suppose that $f : [a, b] \rightarrow \mathbb{R}$ is integrable and $f(x) \geq 0$ for all $x \in [a, b]$. If there exists $x_0 \in [a, b]$ where f is continuous at x_0 and $f(x_0) > 0$, then $\int_a^b f(x) dx > 0$.

Solution

i) First, we assume that $a < x_0 < b$. Let $\epsilon = f(x_0)/2$. Since f is continuous at x_0 , there exists $\delta > 0$ small enough such that $|f(x) - f(x_0)| < \epsilon = f(x_0)/2$ for all $x \in]x_0 - \delta, x_0 + \delta[\subset]a, b[$. Thus

$$\begin{aligned}\int_a^b f(x) dx &= \underbrace{\int_a^{x_0-\delta} f(x) dx}_{\geq 0} + \int_{x_0-\delta}^{x_0+\delta} f(x) dx + \underbrace{\int_{x_0+\delta}^b f(x) dx}_{\geq 0} \\ &\geq \int_{x_0-\delta}^{x_0+\delta} f(x) dx \geq \int_{x_0-\delta}^{x_0+\delta} \frac{f(x_0)}{2} dx = \delta f(x_0) > 0\end{aligned}$$

because $f(x) \geq f(x_0)/2$ for $x \in]x_0 - \delta, x_0 + \delta[$.

ii) We now assume that $x_0 = b$. Let $\epsilon = f(b)/2$. Since f is continuous at b , there exists $\delta > 0$ small enough such that $|f(x) - f(b)| < \epsilon = f(b)/2$ for all $x \in]b - \delta, b[$. Thus

$$\int_a^b f(x) dx = \underbrace{\int_a^{b-\delta} f(x) dx}_{\geq 0} + \int_{b-\delta}^b f(x) dx \geq \int_{b-\delta}^b f(x) dx \geq \int_{b-\delta}^b \frac{f(b)}{2} dx = \frac{\delta f(b)}{2} > 0$$

because $f(x) \geq f(b)/2$ for $x \in]b - \delta, b[$.

iii) Finally, a proof similar to the previous one in (ii) shows that $\int_a^b f(x) dx > 0$ if $x_0 = a$. ♥

Question 2.11

Suppose that $f : [a, b] \rightarrow \mathbb{R}$ is integrable and that $F : [a, b] \rightarrow \mathbb{R}$ is a differentiable function such that $F'(x) = f(x)$ for all $x \in [a, b]$. Prove from the definition of the integral (i.e. without the Fundamental Theorem of Calculus) that

$$F(b) - F(a) = \int_a^b f(x) dx .$$

Solution

Given $\epsilon > 0$, choose a partition $P = \{p_i : 0 \leq i \leq N\}$ of the interval $[a, b]$ such that $\mathcal{U}_P(f) - \mathcal{L}_P(f) < \epsilon$.

Since F' exists for all x , we get that F is continuous on $[a, b]$. Since F is continuous on $[p_{i-1}, p_i]$ and differentiable on $]p_{i-1}, p_i[$, we may use the Mean Value Theorem to conclude that there exists $y_i \in]p_{i-1}, p_i[$ such that

$$F(p_i) - F(p_{i-1}) = F'(y_i)(p_i - p_{i-1}) = f(y_i)(p_i - p_{i-1})$$

for $1 \leq i \leq N$. Hence

$$F(b) - F(a) = F(p_N) - F(p_0) = \sum_{i=1}^N (F(p_i) - F(p_{i-1})) = \sum_{i=1}^N f(y_i)(p_i - p_{i-1}) .$$

Since

$$m_i = \inf\{f(x) : x \in [p_{i-1}, p_i]\} \leq f(y_i) \leq M_i = \sup\{f(x) : x \in [p_{i-1}, p_i]\}$$

for $1 \leq i \leq N$, we get

$$\begin{aligned} \mathcal{L}_P(f) &= \sum_{i=1}^N m_i(p_i - p_{i-1}) \leq F(b) - F(a) = \sum_{i=1}^N f(y_i)(p_i - p_{i-1}) \\ &\leq \sum_{i=1}^N M_i(p_i - p_{i-1}) = \mathcal{U}_P(f) . \end{aligned}$$

We also have $\mathcal{L}_P(f) \leq \int_a^b f(x) dx \leq \mathcal{U}_P(f)$. Thus, both $\int_a^b f(x) dx$ and $F(b) - F(a)$ are in the interval $[\mathcal{L}_P(f), \mathcal{U}_P(f)]$ of length smaller than ϵ . Therefore

$$\left| \int_a^b f(x) dx - (F(b) - F(a)) \right| < \epsilon.$$

Since this is true for all $\epsilon > 0$, we get $\int_a^b f(x) dx = F(b) - F(a)$. ♡

Question 2.12

Suppose that $f : [a, b] \rightarrow \mathbb{R}$ is a Riemann integrable function and that $f(x) > c$ for all $x \in [a, b]$ where c is a constant. Show that $1/f$ is integrable on $[a, b]$. What can we say if the hypothesis $f(x) > c$ for all $x \in [a, b]$ is replaced by $f(x) > 0$ for all $x \in [a, b]$?

Question 2.13

If $f : [a, b] \rightarrow \mathbb{R}$ and $g : [a, b] \rightarrow \mathbb{R}$ are integrable, prove that $fg : [a, b] \rightarrow \mathbb{R}$ is integrable.

Hint: Prove first that $h : [a, b] \rightarrow \mathbb{R}$ integrable on $[a, b]$ implies that h^2 is integrable on $[a, b]$.

Solution

i) We first prove that h^2 is integrable on $[a, b]$ if $h : [a, b] \rightarrow \mathbb{R}$ is integrable on $[a, b]$. If $h : [a, b] \rightarrow \mathbb{R}$ is integrable, then h is bounded on $[a, b]$ by definition of the Riemann integral. Let $C > 0$ be a constant such that $|h(x)| < C$ for all $x \in [a, b]$.

Given $\epsilon > 0$, choose a partition $P = \{p_i : 0 \leq i \leq N\}$ of $[a, b]$ such that $\mathcal{U}_P(h) - \mathcal{L}_P(h) < \epsilon/(2C)$.

As usual, let $m_i = \inf\{h(x) : p_{i-1} \leq x \leq p_i\}$ and $M_i = \sup\{h(x) : p_{i-1} \leq x \leq p_i\}$ for $1 \leq i \leq N$. Since

$$|h^2(x) - h^2(y)| = |h(x) - h(y)| |h(x) + h(y)| \leq 2C|h(x) - h(y)| \leq 2C(M_i - m_i)$$

for all $x, y \in [p_{i-1}, p_i]$, we get

$$\tilde{M}_i - \tilde{m}_i \leq 2C(M_i - m_i)$$

where $\tilde{m}_i = \inf\{h^2(x) : p_{i-1} \leq x \leq p_i\}$ and $\tilde{M}_i = \sup\{h^2(x) : p_{i-1} \leq x \leq p_i\}$. Hence

$$\begin{aligned} \mathcal{U}_P(h^2) - \mathcal{L}_P(h^2) &= \sum_{i=1}^N (\tilde{M}_i - \tilde{m}_i)(p_i - p_{i-1}) \leq 2C \sum_{i=1}^N (M_i - m_i)(p_i - p_{i-1}) \\ &= 2C(\mathcal{U}_P(h) - \mathcal{L}_P(h)) < \epsilon. \end{aligned}$$

Since ϵ is arbitrary, it follows from Theorem 2.1.8 that h^2 is integrable on $[a, b]$.

ii) To prove that fg is integrable, we write $fg = \frac{1}{2}((f+g)^2 - f^2 - g^2)$. Since $(f+g)^2$, f^2 and g^2 are integrable, then fg is integrable. ♡

Question 2.14

Let $f(x) = \lfloor x \rfloor$ for $x \in [0, 9]$, where $\lfloor x \rfloor$ is defined as the largest integer smaller or equal to x . In other words, it is the integer part of x . Show that f is integrable on $[0, 9]$.

Question 2.15

Give a non-integrable function $f : [a, b] \rightarrow \mathbb{R}$ such that $|f|$ and f^2 are integrable on $[a, b]$.

Solution

Let

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q} \\ -1 & \text{if } x \in \mathbb{Q} \end{cases}$$

The function f is not integrable on the interval $[0, 1]$ because $\mathcal{U}(f) = 1$ and $\mathcal{L}(f) = -1$. However, $|f|$ and f^2 are integrable on the interval $[0, 1]$ because $|f(x)| = 1$ and $f^2(x) = 1$ for all $x \in [0, 1]$. $\heartsuit$

Question 2.16

Suppose that $f : [a, b] \rightarrow \mathbb{R}$ is an integrable function such that $\int_0^1 f(x)g(x) dx = 0$ for all continuous functions $g : [a, b] \rightarrow \mathbb{R}$. Prove that $f(x) = 0$ for all $x \in [a, b]$ where f is continuous. In particular, show that $f = 0$ on $[a, b]$ if $f : [a, b] \rightarrow \mathbb{R}$ is continuous.

Solution

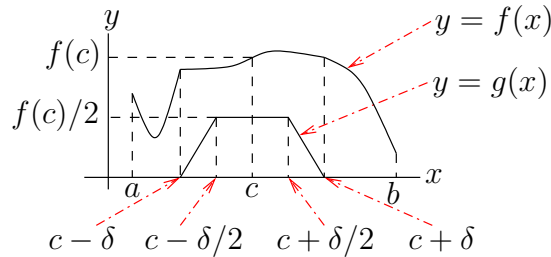
Suppose that f is continuous at $c \in]a, b[$ and $f(c) > 0$.

Since f is continuous at c , there exists $\delta > 0$ such that $|f(x) - f(c)| < f(c)/2$ for $|x - c| < \delta$ and $x \in [a, b]$. Hence $f(x) > f(c)/2$ for $|x - c| < \delta$ and $x \in [a, b]$. To simplify the discussion, we may assume that δ is small enough to have $]c - \delta, c + \delta[\subset [a, b]$.

We define a continuous function $g : [a, b] \rightarrow \mathbb{R}$ as follows

$$g(x) = \begin{cases} 0 & \text{if } |x - c| \geq \delta \\ f(c)/2 & \text{if } |x - c| \leq \delta/2 \\ (f(c)/\delta)(x - (c - \delta/2)) + f(c)/2 & \text{if } c - \delta < x < c - \delta/2 \\ -(f(c)/\delta)(x - (c + \delta/2)) + f(c)/2 & \text{if } c + \delta/2 < x < c + \delta \end{cases}$$

This function is represented in the following figure.



Hence, since g is continuous, we have

$$0 = \int_a^b f(x)g(x) dx = \int_{c-\delta}^{c+\delta} f(x)g(x) dx \geq \int_{c-\delta/2}^{c+\delta/2} \frac{f^2(c)}{4} dx = \frac{f^2(c)\delta}{4}.$$

This is a contradiction that $f(c) > 0$. If $f(c) < 0$, then the previous reasoning applied to $-f(x)$ shows that $-f(c) > 0$ is not possible. Thus $f(c) = 0$.

The cases $c = a$ or $c = b$ are handled in a similar fashion.

If f is continuous on $[a, b]$, then the previous discussion can be applied to every $x \in [a, b]$ to show that $f(x) = 0$. ♡

Question 2.17

Suppose that $f : [a, b] \rightarrow \mathbb{R}$ is continuous and $g : [a, b] \rightarrow \mathbb{R}$ is an increasing continuous differentiable function. Prove that there exists $c \in [a, b]$ such that

$$\int_a^b f(x)g(x) \, dx = g(a) \int_a^c f(x) \, dx + g(b) \int_c^b f(x) \, dx. \quad (2.5.6)$$

The result is also true if g increasing is replaced by g decreasing.

Solution

i) Suppose that $g(b) = 0$. Let $F(x) = \int_a^x f(t) \, dt$. Since f is continuous on $[a, b]$, the Fundamental Theorem of Calculus tells us that F is differentiable and that $F'(x) = f(x)$ for all x . Using integration by parts, we get

$$\int_a^b f(x)g(x) \, dx = \int_a^b F'(x)g(x) \, dx = F(x)g(x) \Big|_a^b - \int_a^b F(x)g'(x) \, dx = - \int_a^b F(x)g'(x) \, dx$$

because $F(a) = 0$ and $g(b) = 0$.

Since g is increasing on $[a, b]$, we have $g'(x) \geq 0$ for all $x \in [a, b]$. We may therefore use the Mean Value Theorem for integrals³ to conclude that

$$\int_a^b F(x)g'(x) \, dx = F(c) \int_a^b g'(x) \, dx = F(c)(g(b) - g(a)) = -F(c)g(a) = -g(a) \int_a^c f(t) \, dt$$

for some $c \in [a, b]$. Thus

$$\int_a^b f(x)g(x) \, dx = g(a) \int_a^c f(t) \, dt$$

for some $c \in [a, b]$. This is (2.5.6) for the case $g(b) = 0$.

ii) In the general case (including $g(b) = 0$), let $h(x) = g(x) - g(b)$. We have that h is an increasing continuous differentiable function because g is such a function. Moreover, $h(b) = 0$. We may apply part (i) to conclude that there exist $c \in [a, b]$ such that

$$\begin{aligned} \int_a^b f(x)h(x) \, dx &= h(a) \int_a^c f(x) \, dx \\ &\Rightarrow \int_a^b f(x)(g(x) - g(b)) \, dx = (g(a) - g(b)) \int_a^c f(x) \, dx \\ &\Rightarrow \int_a^b f(x)g(x) \, dx - g(b) \int_a^b f(x) \, dx = g(a) \int_a^c f(x) \, dx - g(b) \int_a^c f(x) \, dx \end{aligned}$$

³The reader will be asked to prove it for functions of several variables in Question 3.6 of the next chapter.

$$\begin{aligned}
\Rightarrow \int_a^b f(x)g(x) dx &= g(a) \int_a^c f(x) dx + g(b) \left(\int_a^b f(x) dx - \int_a^c f(x) dx \right) \\
&= g(a) \int_a^c f(x) dx + g(b) \int_c^b f(x) dx .
\end{aligned}$$

♡

Question 2.18

Suppose that $h : [a, b] \rightarrow \mathbb{R}$ is a non-negative, decreasing function and $f : [a, b] \rightarrow \mathbb{R}$ is continuous. Prove that there exists $\xi \in [a, b]$ such that $\int_a^b f(x)h(x) dx = h(a) \int_a^\xi f(x) dx$.

Hint: If we assume that h is also differentiable, then this is a consequence of Question 2.17 with $g(x) = h(x) - h(b)$. For the more general case, one may use the Riemann-Stieltjes integral that we have not introduced. Therefore, this question is more a little project.

Question 2.19

Consider a continuous function $p : [a, b] \rightarrow \mathbb{R}$ such that $p(x) \geq 0$ for all $x \in [a, b]$. Suppose that there exists a constant $c > 0$ such that $p(x) \leq c \int_a^x p(t) dt$ for all $x \in [a, b]$. Prove that $p(x) = 0$ for all $x \in [a, b]$.

Solution

Suppose that $p(x) \neq 0$ for some $x \in [a, b]$. Take $x_0 \in [a, b]$ as the infimum of the non-empty set $\{x \in [a, b] : p(x) > 0\}$. Since p is continuous, $x_0 < b$ and there exists $\delta > 0$ such that $x_0 + \delta \leq b$ and $p(x) > 0$ for $x \in]x_0, x_0 + \delta]$.

Let $P(x) = c \int_a^x p(t) dt$ for $x \in [a, b]$. We have $p(x) \leq P(x)$ for all x by hypothesis. If $x_0 > a$, then we have $P(x) = 0$ for $x \in [a, x_0[$ because $p(x) = 0$ for $x \in [a, x_0[$ by definition of x_0 . We have $P(x) > 0$ for $x \in]x_0, b]$ because $p(x) > 0$ for $x_0 < x < x_0 + \delta$ and $p(x) \geq 0$ for all $x \in [a, b]$. Moreover $P(x_0) = 0$ by continuity of P if $x_0 > a$ and $P(x_0) = 0$ by definition of P if $x_0 = a$.

We have $P'(x) = cp(x)$ for all x . Therefore

$$\frac{P'(x)}{P(x)} = \frac{cp(x)}{P(x)} \leq \frac{cP(x)}{P(x)} = c .$$

for $x > x_0$. It is here that we need $P(x) > 0$ for all $x > x_0$. Given $x > x_0$, take x_1 between x and x_0 . Hence

$$\begin{aligned}
\int_{x_1}^x \frac{P'(t)}{P(t)} dt &\leq c \int_{x_1}^x dt \Rightarrow \ln(P(x)) - \ln(P(x_1)) \leq c(x - x_1) \\
&\Rightarrow \frac{P(x)}{P(x_1)} \leq e^{c(x-x_1)} \Rightarrow P(x) \leq P(x_1)e^{c(x-x_1)} .
\end{aligned}$$

Thus

$$0 \leq p(x) \leq P(x) \leq \lim_{x_1 \rightarrow x_0^+} P(x_1)e^{c(x-x_1)} = P(x_0)e^{c(x-x_0)} = 0 .$$

We have $p(x) = 0$ for all $x > x_0$. This contradicts our hypothesis that $p(x) \neq 0$ for some $x \in [a, b]$.

Note: This result is a particular case of Gronwall's Lemma that we have in the theory of ordinary differential equations. ♡

Question 2.20

Suppose that $f : [a, b] \rightarrow \mathbb{R}$ is a continuous function such that $f(x) \geq 0$ for all $x \in [a, b]$ and $M = \sup\{f(x) : x \in [a, b]\}$. Prove that $M = \lim_{n \rightarrow \infty} \left(\int_a^b f^n(x) dx \right)^{1/n}$.

Solution

If $M = 0$, then there is nothing to prove. We assume that $M > 0$. Since

$$\int_a^b f^n(x) dx \leq \int_a^b M^n dx = M^n(b-a),$$

we get

$$\left(\int_a^b f^n(x) dx \right)^{1/n} \leq M(b-a)^{1/n}$$

for all $n > 0$. Since $(b-a)^{1/n} \rightarrow 1$ as $n \rightarrow \infty$, we get

$$\lim_{n \rightarrow \infty} \left(\int_a^b f^n(x) dx \right)^{1/n} \leq M.$$

Given $\epsilon > 0$ such that $M - \epsilon > 0$, let $J = \{x \in [a, b] : f(x) > M - \epsilon\}$. Since f is continuous, $J = f^{-1}(]M - \epsilon, \infty[)$ is an open subset of $[a, b]$. Thus J contains an open interval I . The interval I is obviously measurable. Let $\delta > 0$ be the length of I . We have

$$\int_a^b f^n(x) dx \geq \int_I f^n(x) dx \geq \int_I (M - \epsilon)^n dx = (M - \epsilon)^n \delta.$$

Hence

$$\left(\int_a^b f^n(x) dx \right)^{1/n} \geq (M - \epsilon) \delta^{1/n}.$$

Since $\delta^n \rightarrow 1$ as $n \rightarrow \infty$, we get

$$\lim_{n \rightarrow \infty} \left(\int_a^b f^n(x) dx \right)^{1/n} \geq (M - \epsilon).$$

Since ϵ is arbitrary small, we conclude that

$$\lim_{n \rightarrow \infty} \left(\int_a^b f^n(x) dx \right)^{1/n} \geq M.$$

♡

Question 2.21

Suppose that $f : [a, b] \rightarrow \mathbb{R}$ and $g : [a, b] \rightarrow \mathbb{R}$ are continuous. Prove that

$$\left(\int_a^b f(x)g(x) \, dx \right)^2 \leq \left(\int_a^b f^2(x) \, dx \right) \left(\int_a^b g^2(x) \, dx \right).$$

Question 2.22

Consider the functions $f_n(x) = \frac{nx}{1+nx}$ for $n \geq 1$. Let

$$f(x) = \begin{cases} 0 & \text{if } x = 0 \\ 1 & \text{if } 0 < x \leq 1 \end{cases}$$

Prove that $f_n \rightarrow f$ pointwise on $[0, 1]$ as $n \rightarrow \infty$ and, without computing any integral, that $\int_0^1 f_n(x) \, dx \rightarrow \int_0^1 f(x) \, dx$ as $n \rightarrow \infty$.

Question 2.23

Prove that $\lim_{n \rightarrow \infty} \int_a^\pi \frac{\sin(nx)}{nx} \, dx = 0$ for $0 < a < \pi$. What can be said if $a = 0$?

Solution

The result is the same for $a = 0$ or $a > 0$. We leave the proof of the case $a > 0$ to the reader. It suffices to note that $\left| \frac{\sin(nx)}{nx} \right| \leq \frac{1}{na}$ for $a \leq x < \pi$. The bounded convergence theorem can then be used to conclude that $\lim_{n \rightarrow \infty} \int_a^\pi \frac{\sin(nx)}{nx} \, dx = 0$.

We prove the result for $a = 0$. Let

$$f_n(x) = \begin{cases} \frac{\sin(nx)}{nx} & 0 < x \leq \pi \\ 1 & x = 0 \end{cases}$$

for $n \in \mathbb{N}^+$. Recall that $-1 < \frac{\sin(y)}{y} < 1$ for $y > 0$ and $\lim_{y \rightarrow 0} \frac{\sin(y)}{y} = 1$. Therefore, the functions $f_n : [0, \pi] \rightarrow \mathbb{R}$ are continuous and $|f_n(x)| \leq 1$ for all $x \in [0, \pi]$ and $n \in \mathbb{N}^+$. Moreover, $|f_n(x)| \leq \frac{1}{nx} \rightarrow 0$ as $n \rightarrow \infty$ for all $x \in]0, \pi]$ and $\lim_{n \rightarrow \infty} f_n(x) = 1$ for $x = 0$. Hence, the sequence of functions $\{f_n\}_{n=1}^\infty$ converge pointwise to the integrable function $f : [0, \pi] \rightarrow \mathbb{R}$ defined by $f(x) = 0$ for all $x \in]0, \pi]$ and $f(0) = 1$.

We may use the bounded convergence theorem to conclude that

$$\lim_{n \rightarrow \infty} \int_0^\pi \frac{\sin(nx)}{nx} \, dx = \lim_{n \rightarrow \infty} \int_0^\pi f_n(x) \, dx = \int_0^\pi f(x) \, dx = 0.$$

♡

Question 2.24

Compute $\lim_{n \rightarrow \infty} \int_0^1 f_n(x) \, dx$ for the following sequences $\{f_n\}_{n=1}^\infty$ of functions.

a) $f_n(x) = \left(1 + \frac{x}{n}\right)^n$ for $x \in [0, 1]$.

b) $f_n(x) = \sum_{k=0}^n (-x)^k$ for $x \in [0, 1]$.

Use the result in (b) to deduce a series expansion for $\ln(2)$.

Question 2.25

Give an example of a sequence of integrable functions $\{f_n\}_{n=0}^\infty$ on $[0, 1]$ that converges pointwise to a non integrable function $f : [0, 1] \rightarrow \mathbb{R}$.

Solution

First example: Let $\mathbb{Q} \cap [0, 1] = \{r_0, r_1, r_2, \dots\}$ be an enumeration of the rational numbers in the interval $[0, 1]$. Let

$$f_n(x) = \begin{cases} 1 & \text{if } x \in \{r_1, r_2, \dots, r_n\} \\ 0 & \text{if } x \in [0, 1] \setminus \{r_1, r_2, \dots, r_n\} \end{cases}$$

The functions $f_n : [0, 1] \rightarrow \mathbb{R}$ are integrable because they are continuous on $[0, 1]$ except at a finite number of points. However, we have

$$f(x) = \lim_{n \rightarrow \infty} f_n(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \cap [0, 1] \\ 0 & \text{if } x \in [0, 1] \setminus \mathbb{Q} \end{cases}$$

which is not integrable on $[0, 1]$.

Second example: Let $f_n(x) = \frac{nx}{1 + nx^2}$ for $0 \leq x \leq 1$. For each $n \geq 0$, the function f_n is integrable on $[0, 1]$ because it is continuous on $[0, 1]$. We have

$$\lim_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} \frac{nx}{1 + nx^2} = \lim_{n \rightarrow \infty} \frac{x}{(1/n) + x^2} = \frac{1}{x}$$

for $x > 0$ and $\lim_{n \rightarrow \infty} f_n(x) = 0$ for $x = 0$ because $f_n(0) = 0$ for all n . Hence

$$f(x) = \lim_{n \rightarrow \infty} f_n(x) = \begin{cases} 1/x & x \in]0, 1] \\ 0 & x = 0 \end{cases}$$

which is not integrable on $[0, 1]$.

♡

Question 2.26

Find all continuous function $f : [0, 1] \rightarrow \mathbb{R}$ such that

$$(f(x))^2 = \int_0^x f(t) dt \tag{2.5.7}$$

for $x \in [0, 1]$.

Solution

There is obviously the trivial solution $f(x) = 0$ for all $x \in [0, 1]$. Suppose that there exists $x \in [0, 1]$ such that $f(x) \neq 0$.

Let $c = \inf\{x \in [0, 1] : f(x) \neq 0\}$. Since f is continuous and $f(x) = 0$ for $x < c$, we get $f(c) = \lim_{x \rightarrow c^-} f(x) = 0$. Therefore $0 \leq c < 1$. Do not forget that $f(0) = 0$ because of (2.5.7).

i) We first show that $f(x) > 0$ for all $x \in]c, 1]$. Suppose that $f(x) = 0$ for some $x \in]c, 1]$. Since f is continuous and f is not trivially null on $]c, 1]$, there exists an interval $[a, b] \subset]c, 1]$ such that $f(a) = f(b) = 0$ and either $f(x) > 0$ for all $a < x < b$ or $f(x) < 0$ for all $a < x < b$. We then have

$$0 = (f(b))^2 = \underbrace{\int_0^a f(t) dt}_{=(f(a))^2=0} + \underbrace{\int_a^b f(t) dt}_{\neq 0},$$

where we have used the result of Question 2.10 to justify that the integral of f between a and b is non-null. This is a contradiction. Thus $f(x) \neq 0$ for all $x \in]c, 1]$; namely, either $f(x) > 0$ for all $x \in]c, 1]$ or $f(x) < 0$ for all $x \in]c, 1]$. Since $\int_0^x f(t) dt = (f(x))^2 \geq 0$, we must therefore have $f(x) > 0$ for all $x \in]c, 1]$.

ii) Since f is continuous, it follows from the Fundamental Theorem of Calculus that $F : [0, 1] \rightarrow \mathbb{R}$ defined by $F(x) = \int_0^x f(t) dt$ is differentiable. Since (i) implies that $f(x) = \left(\int_0^x f(t) dt\right)^{1/2} = (F(x))^{1/2}$ for all $x \in [0, 1]$, we get that f is differentiable.

If we differentiate with respect to x both sides of the equality (2.5.7), then we get $2f(x)f'(x) = f(x)$. Thus $f'(x) = 1/2$ for all $x > c$ because $f(x) \neq 0$ for all $x > c$. We find that $f(x) = x/2 + C$ for $x > c$ where C is a constant. If $c > 0$, then f is not differentiable at $x = c$ because $f(x) = 0$ for $x \leq c$. Therefore we must have $c = 0$.

If we substitute $f(x) = x/2 + C$ in (2.5.7) and integrate, then we get

$$\left(\frac{x}{2} + C\right)^2 = \frac{x^2}{4} + Cx.$$

If we simplify this equation, then we get $C^2 = 0$. Thus $f(x) = x/2$ for $0 \leq x \leq 1$. ♡

Question 2.27

Answer the following questions without using $\ln$. This problem introduces one of the possible definitions of $\ln$. Let $L(x) = \int_1^x \frac{1}{t} dt$. Prove that $L :]0, \infty[\rightarrow \mathbb{R}$ is continuously differentiable and that $L(xy) = L(x) + L(y)$ for all $x, y \in]0, \infty[$.

Hint: let $G(x) = L(xy)$ for $y > 0$ fixed and compute $G'(x)$.

Solution

Since $f :]0, \infty[\rightarrow \mathbb{R}$ defined by $f(t) = 1/t$ is continuous, it follows from the Fundamental Theorem of Calculus that $L'(x) = 1/x$ for $x > 0$. Hence L is of class C^1 on $]0, \infty[$.

Let $G(x) = L(xy) = \int_1^{xy} \frac{1}{t} dt$ for $y > 0$ arbitrary but fixed. Using the chain rule, we get

$$G'(x) = \left(\frac{d}{du} \left(\int_1^u \frac{1}{t} dt\right)\right) \Big|_{u=xy} \frac{d}{dx}(xy) = \left(L'(u)\right) \Big|_{u=xy} y = \left(\frac{1}{u}\right) \Big|_{u=xy} y = \frac{1}{x}$$

for $x > 0$. Thus

$$G(x) - G(1) = \int_1^x G'(t) dt = \int_1^x \frac{1}{t} dt = L(x)$$

for $x > 0$. Since $G(1) = L(y)$ and $G(x) = L(xy)$, we get $L(xy) = L(x) + L(y)$ for $x, y > 0$. $\heartsuit$

Question 2.28

Describe the continuous functions $f : [0, 1] \rightarrow \mathbb{R}$ that satisfy $\int_0^x f(t) dt = \int_x^1 f(t) dt$ for all $x \in [0, 1]$.

Solution

Since f is a continuous function, $F(x) = \int_0^x f(t) dt$ and $G(x) = \int_x^1 f(t) dt = -\int_1^x f(t) dt$ are two differentiable functions on $]0, 1[$ according to the Fundamental Theorem of Calculus. We have by hypothesis that $F(x) = G(x)$ for all $x \in [0, 1]$. Therefore

$$f(x) = F'(x) = G'(x) = -f(x)$$

for all $x \in]0, 1[$. Hence $f(x) = 0$ for all $x \in]0, 1[$. Since f is continuous on $[0, 1]$, we get $f(x) = 0$ for $x \in [0, 1]$. $\heartsuit$

Question 2.29

Consider the function

$$f(x) = \begin{cases} 1/q & \text{if } x = p/q \text{ where } p \geq 0 \text{ and } q > 0 \text{ are relatively prime} \\ 0 & \text{otherwise} \end{cases}$$

Is f Riemann integrable on $[0, 1]$? If it is, then what is the value of $\int_0^1 f(x) dx$?

Hint: If N is a positive integer, then there are only a finite number of rational numbers in the interval $[0, 1]$ of the form p/q with p and q relatively prime and $0 < q \leq N$.

Solution

We prove that f is integrable and that $\int_0^1 f(x) dx = 0$. Given $\epsilon > 0$, choose $q_0 \in \mathbb{N}^+$ large enough such that $1/q_0 < \epsilon/2$, and a partition $R = \{r_i : 0 \leq i \leq N\}$ of $[0, 1]$ such that

$$\frac{q_0(q_0 + 1)}{2} \max_{1 \leq i \leq N} (r_i - r_{i-1}) < \frac{\epsilon}{2}.$$

Note that there are less than $q_0(q_0 + 1)/2$ values of x of the form p/q with p and q relatively prime and $0 \leq q \leq q_0$ ⁴. Let $\{a_k : 1 \leq k \leq K\}$ with $K < q_0(q_0 + 1)/2$ be these values.

⁴The total number of pairs (p, q) with $0 < p \leq q \leq q_0$ is $\sum_{q=1}^{q_0} q = \frac{q_0(q_0 + 1)}{2}$. However, for a large number of pairs (p, q) , we note that p and q are not relatively prime.

Let $m_i = \inf\{f(x) : r_{i-1} \leq x \leq r_i\}$ and $M_i = \sup\{f(x) : r_{i-1} \leq x \leq r_i\}$ for $1 \leq i \leq N$. Then $M_i - m_i < 1/q_0$ if $a_k \notin [r_{i-1}, r_i]$ for all k . Hence

$$\mathcal{U}_R(f) - \mathcal{L}_R(f) = \sum_{i=1}^N (M_i - m_i)(r_i - r_{i-1}) < \sum' \frac{1}{q_0}(r_i - r_{i-1}) + \sum''(r_i - r_{i-1}),$$

where the first sum is over all sub-intervals $[r_{i-1}, r_i]$ such that $a_k \notin [r_{i-1}, r_i]$ for all k , and the second sum is over all sub-intervals $[r_{i-1}, r_i]$ such that $a_k \in [r_{i-1}, r_i]$ for some k .

We have $\sum' \frac{1}{q_0}(r_i - r_{i-1}) < \frac{\epsilon}{2}$ because $1/q_0 < \epsilon/2$ and $\sum_{i=0}^N (r_i - r_{i-1}) = 1$; the length of the interval $[0, 1]$.

Moreover $\sum''(r_i - r_{i-1}) < \frac{\epsilon}{2}$ because there are at most $q_0(q_0 + 1)/2$ intervals $[r_{i-1}, r_i]$ containing some a_k . Thus

$$\sum''(r_i - r_{i-1}) \leq \frac{q_0(q_0 + 1)}{2} \max_{1 \leq i \leq N} (r_i - r_{i-1}) < \frac{\epsilon}{2}.$$

Therefore $\mathcal{U}_R(f) - \mathcal{L}_R(f) < \epsilon$. Since ϵ is arbitrary, it follows from Theorem 2.1.8 that f is integrable on $[0, 1]$. Since $\mathcal{L}_R(f) = 0$ for all partition R of $[0, 1]$, we get $\int_0^1 f = \mathcal{L}(f) = 0$. ♡

Question 2.30

Consider

$$F(x) = \begin{cases} x^2 \sin(1/x^2) & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$$

Prove that $F'(x)$ exists for all x but that $F' : [0, 1] \rightarrow \mathbb{R}$ is not integrable. Therefore, we do not have $\int_0^1 F'(x) dx = F(1) - F(0)$.

Solution

It is clear that F is differentiable on $]0, \infty[$ and on $] - \infty, 0[$. The only interesting case is at the origin. To show that F is differentiable at the origin, we use the definition of the derivative.

$$\lim_{h \rightarrow 0^-} \frac{F(h) - F(0)}{h} = \lim_{h \rightarrow 0^-} \frac{0 - 0}{h} = 0$$

and

$$\lim_{h \rightarrow 0^+} \frac{F(h) - F(0)}{h} = \lim_{h \rightarrow 0^+} \frac{h^2 \sin(1/h^2) - 0}{h} = \lim_{h \rightarrow 0^+} h \sin\left(\frac{1}{h^2}\right) = 0$$

because $\left| h \sin\left(\frac{1}{h^2}\right) \right| \leq h \rightarrow 0$ as $h \rightarrow 0^+$. Thus $F'(0) = 0$. Hence

$$F'(x) = \begin{cases} 2x \sin\left(\frac{1}{x^2}\right) - \frac{2}{x} \cos\left(\frac{1}{x^2}\right) & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$$

Note that F' is not bounded in any neighbourhood of the origin. For instance, with $x_n = \frac{1}{\sqrt{2n\pi}}$, we have $x_n \rightarrow 0^+$ and $F'(x_n) = -2\sqrt{2n\pi} \rightarrow -\infty$ as $n \rightarrow \infty$. Therefore, F' is not integrable on $[0, 1]$. $\heartsuit$

Question 2.31

Consider a continuous function $f : [0, 1] \rightarrow \mathbb{R}$. We define inductively a sequence of real valued functions on $[0, 1]$ by $f_0(x) = f(x)$ and $f_{k+1}(x) = \int_0^x f_k(t) dt$ for $x \in [0, 1]$ and $k \in \mathbb{N}$. Prove that $|f_k(x)| \leq Mx^k/k!$ for all $k \geq 0$ and all $x \in [0, 1]$ where $M = \sup\{|f(x)| : x \in [0, 1]\}$. Moreover, prove that $\{f_k\}_{k=0}^\infty$ converges uniformly to 0 on $[0, 1]$.

Solution

i) The proof that

$$|f_k(x)| \leq \frac{M}{k!} x^k \quad (2.5.8)$$

for $k \geq 0$ is by induction on k . For $k = 0$, we have $|f_0(x)| = |f(x)| \leq M$ for all $x \in [0, 1]$ by definition of M . Hence (2.5.8) is true for $k = 0$. Suppose that (2.5.8) is true for $k \geq 0$. Then

$$|f_{k+1}(x)| = \left| \int_0^x f_k(t) dt \right| \leq \int_0^x |f_k(t)| dt \leq \int_0^x \frac{M}{k!} t^k dt = \frac{M}{k!} \left(\frac{t^{k+1}}{k+1} \right) \Big|_0^x = \frac{M}{(k+1)!} x^{k+1}$$

for all $x \in [0, 1]$, where we have used the hypothesis of induction for the second inequality. Thus (2.5.8) is true if we replace k by $k+1$. This complete the proof by induction.

ii) It follows from (i) that

$$|f_k(x)| \leq \frac{M}{k!} x^k \leq \frac{M}{k!}$$

for all $x \in [0, 1]$. Given $\epsilon > 0$, choose $K \in \mathbb{N}^+$ such that $\frac{M}{K!} < \epsilon$. Then

$$|f_k(x) - 0| \leq \frac{M}{k!} \leq \frac{M}{K!} < \epsilon$$

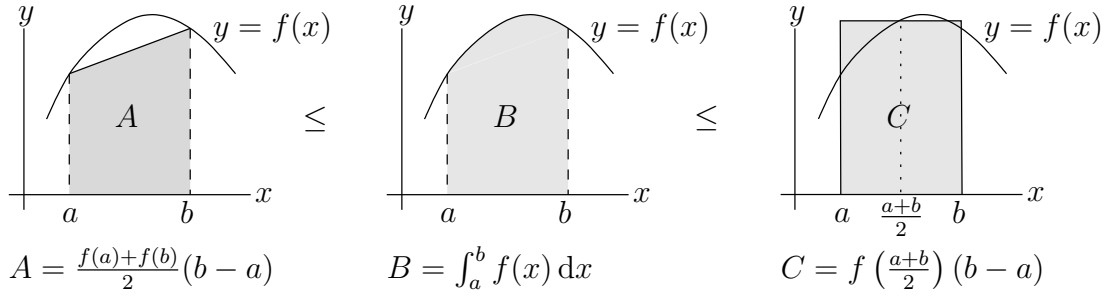
for all $x \in [0, 1]$ and $k \geq K$. Since ϵ is arbitrary, this proves that $f_k \rightarrow 0$ uniformly on $[0, 1]$ as $k \rightarrow \infty$. $\heartsuit$

Question 2.32

Suppose that $f : [a, b] \rightarrow \mathbb{R}$ is twice continuously differentiable and $f''(x) \leq 0$ for all $x \in [a, b]$. Prove that

$$\frac{f(a) + f(b)}{2}(b-a) \leq \int_a^b f(x) dx \leq (b-a)f\left(\frac{a+b}{2}\right). \quad (2.5.9)$$

In terms more common to numerical analysis, this inequality says that, for a twice differentiable and concave down function, the trapezoidal method underestimates the value of the integral while the midpoint method overestimates the value of the integral.

**Solution**

i) The Taylor's polynomial of f of order 2 centred at the point $x \in]a, b[$ is

$$f(y) = f(x) + f'(x)(y-x) + \frac{f''(\xi)}{2}(y-x)^2,$$

where $\xi = \xi(x)$ is between x and y . If $y = a$, then we get

$$f(a) = f(x) + f'(x)(a-x) + \frac{f''(\xi)}{2}(a-x)^2 \leq f(x) + f'(x)(a-x)$$

because $f''(\xi) \leq 0$ for $\xi \in [a, b]$. If we integrate $f(a) \leq f(x) + f'(x)(a-x)$ from a to b on both sides, then we get

$$\int_a^b f(a) dx \leq \int_a^b f(x) dx + \int_a^b f'(x)(a-x) dx. \quad (2.5.10)$$

Thus

$$f(a)(b-a) \leq \int_a^b f(x) dx + f(x)(a-x) \Big|_a^b + \int_a^b f(x) dx = 2 \int_a^b f(x) dx - f(b)(b-a),$$

where we have used integration by parts to compute the second integral on the right hand side in (2.5.10). If we isolate $\int_a^b f(x) dx$, then we get

$$\int_a^b f(x) dx \geq \frac{f(a) + f(b)}{2}(b-a).$$

This proves the left side of the inequality in (2.5.9).

ii) This time, we start with the Taylor's polynomial of f of order 2 centred at $\frac{a+b}{2}$.

$$f(x) = f\left(\frac{a+b}{2}\right) + f'\left(\frac{a+b}{2}\right)\left(x - \frac{a+b}{2}\right) + \frac{f''(\xi)}{2}\left(x - \frac{a+b}{2}\right)^2$$

where $\xi = \xi((a+b)/2)$ is between x and $(a+b)/2$. Since $f''(\xi) \leq 0$ for $\xi \in [a, b]$, we have

$$f(x) \leq f\left(\frac{a+b}{2}\right) + f'\left(\frac{a+b}{2}\right)\left(x - \frac{a+b}{2}\right)$$

for $x \in [a, b]$. If we integrate from a to b on both sides of this inequality, then we get

$$\int_a^b f(x) \, dx \leq f\left(\frac{a+b}{2}\right) \int_a^b dx + \underbrace{f'\left(\frac{a+b}{2}\right) \int_a^b \left(x - \frac{a+b}{2}\right) \, dx}_{=0 \text{ by symmetry}} = f\left(\frac{a+b}{2}\right) (b-a) .$$

This proves the right side of the inequality in (2.5.9). ♥

2.5.1 Supplementary Topics

There are a couple of topics related to integration that are not strictly related to the main subject of these lectures notes but are useful to know. In particular, they will be useful tools for the examples in the future chapters. These topics have been studied in Calculus. We therefore provide a few problems to give a brief overview of these topics.

Integrals Depending on a Parameter

The first topic is about integrals that depend on a parameter. The main results to remember are summarized below. The reader may consult [7, 23] for references.

Suppose that $f : [a, b] \times [c, d] \rightarrow \mathbb{R}$ is a function that satisfies the following two conditions for every $y \in [c, d]$:

1. $f_y : [a, b] \rightarrow \mathbb{R}$ defined by $f_y(x) = f(x, y)$ for $x \in [a, b]$ is Riemann integrable and,
2. given $\epsilon > 0$, there exists $\delta > 0$ (that may also depend on y) such that $\left| \frac{\partial f}{\partial y}(x, z) - \frac{\partial f}{\partial y}(x, y) \right| < \epsilon$ for all $z \in]y - \delta, y + \delta[$ and all $x \in [a, b]$.

Then $F(y) = \int_a^b f(x, y) \, dx$ for $y \in [c, d]$ is differentiable and

$$F'(y) = \int_a^b \frac{\partial f}{\partial y}(x, y) \, dx . \quad (2.5.11)$$

This formula can be generalized to functions defined by $F(y) = \int_a^{h(y)} f(x, y) \, dx$ for $y \in [c, d]$ where $h : [c, d] \rightarrow [a, b]$ is differentiable, and f and $\frac{\partial f}{\partial y}$ are continuous on $[a, b] \times [c, d]$. These last conditions ensure that (1) and (2) above are satisfied. If we set $G(u, v) = \int_a^u f(x, v) \, dx$ for $(u, v) \in [a, b] \times [c, d]$, then $F(y) = G(h(y), y)$ for $y \in [c, d]$. Using the chain rule, we find that

$$F'(y) = \frac{\partial G}{\partial u} \Big|_{u=h(y), v=y} h'(y) + \frac{\partial G}{\partial v} \Big|_{u=h(y), v=y} \frac{d}{dy}(y)$$

$$\begin{aligned}
&= \left(\frac{\partial}{\partial u} \int_a^u f(x, v) \, dx \right) \Big|_{u=h(y), v=y} h'(y) + \left(\frac{\partial}{\partial v} \int_a^u f(x, v) \, dx \right) \Big|_{u=h(y), v=y} \\
&= f(u, v) \Big|_{u=h(y), v=y} h'(y) + \int_a^u \frac{\partial f}{\partial v}(x, v) \, dx \Big|_{u=h(y), v=y} \\
&= f(h(y), y) h'(y) + \int_a^{h(y)} \frac{\partial f}{\partial y}(x, y) \, dx,
\end{aligned}$$

where we have used the Fundamental Theorem of Calculus and (2.5.11) to obtain the third equality. The reader can proceed in a similar fashion to obtain other derivative formulae.

Question 2.33

Let

$$f(x, y) = \begin{cases} \frac{x^3}{y^2} e^{-x^2/y} & \text{if } y > 0 \\ 0 & \text{if } y \leq 0 \end{cases}$$

and $f_x(y) = f_y(x) = f(x, y)$ for all x and y .

- a) Prove that $f_y \in C^1(\mathbb{R})$ for all y fixed and $f_x \in C^1(\mathbb{R})$ for all x fixed.
- b) Prove that f is not bounded in any neighbourhood of the origin.
- c) Prove that

$$\frac{\partial}{\partial x} \int_0^1 f(x, y) \, dy \Big|_{x=0} = \int_0^1 \frac{\partial f}{\partial x}(x, y) \, dy \Big|_{x=0}. \quad (2.5.12)$$

but

$$\int_0^1 \frac{\partial f}{\partial x}(x, y) \, dy \Big|_{x=0} \neq \int_0^1 \frac{\partial f}{\partial x}(0, y) \, dy.$$

Does any of this contradict the theory?

Solution

- a) We have that $f_y \in C^1(\mathbb{R})$ for all y fixed because $f_y(x) = 0$ for all $x \in \mathbb{R}$ when $y \leq 0$ and $f_y(x) = x^3 e^{-x^2/y}/y^2$ for all x when $y > 0$.

We have that f_x is continuously differentiable on $\mathbb{R} \setminus \{0\}$ for all x fixed because $f_x(y) = 0$ for all $y < 0$ and $f_x(y) = x^3 e^{-x^2/y}/y^2$ for all $y > 0$. The only issue is at the origin. We first note that f_x is continuous at the origin because $\lim_{y \rightarrow 0^-} f_x(y) = 0$ and $\lim_{y \rightarrow 0^+} f_x(y) =$

$\lim_{y \rightarrow 0^+} \frac{x^3}{y^2} e^{-x^2/y} = 0$ ⁵. We also have $\frac{df_x}{dy}(0) = 0$ because $\lim_{y \rightarrow 0^-} \frac{f_x(y) - f_x(0)}{y} = 0$ and $\lim_{y \rightarrow 0^+} \frac{f_x(y) - f_x(0)}{y} = \lim_{y \rightarrow 0^+} \frac{x^3}{y^3} e^{-x^2/y} = 0$ ⁶. Thus

$$\frac{df_x}{dy}(y) = \begin{cases} \left(\frac{-2x^3}{y^3} + \frac{x^5}{y^4} \right) e^{-x^2/y} & \text{if } y > 0 \\ 0 & \text{if } y \leq 0 \end{cases}$$

⁵One may use l'Hospital Rule to compute this limit.

⁶Again, one may use l'Hospital Rule to compute this limit.

Again, it is easy to show that $\lim_{y \rightarrow 0} \frac{df_x}{dy}(y) = \frac{df_x}{dy}(0)$. Thus $f_x \in C^1(\mathbb{R})$.

b) We have $f(\sqrt{y}, y) = \frac{1}{\sqrt{y}e} \rightarrow \infty$ as $y \rightarrow 0^+$.

c) As we have done in (a), we can show that

$$g_x(y) = \begin{cases} xe^{-x^2/y} & \text{if } y > 0 \\ 0 & \text{if } y \leq 0 \end{cases}$$

is a function of class C^1 on $\mathbb{R}$ and $\frac{dg_x}{dy} = f_x$ for all x fixed. Hence, we have from the Fundamental Theorem of Calculus that

$$\int_0^1 f(x, y) dy = \int_0^1 f_x(y) dy = g_x(y) \Big|_{y=0}^{y=1} = xe^{-x^2}.$$

Thus $\frac{\partial}{\partial x} \int_0^1 f(x, y) dy = e^{-x^2} - 2x^2 e^{-x^2}$ and

$$\frac{\partial}{\partial x} \int_0^1 f(x, y) dy \Big|_{x=0} = 1. \quad (2.5.13)$$

We have

$$\frac{\partial f}{\partial x}(x, y) = \begin{cases} \left(\frac{3x^2}{y^2} - \frac{2x^4}{y^3} \right) e^{-x^2/y} & \text{if } y > 0 \\ 0 & \text{if } y \leq 0 \end{cases}$$

Let $h_x(y) = \frac{\partial f}{\partial x}(x, y)$ for all $y \in \mathbb{R}$ and x fixed. Again, as we have done in (a), we can prove that

$$k_x(y) = \begin{cases} \left(1 - \frac{2x^2}{y} \right) e^{-x^2/y} & \text{if } y > 0 \\ 0 & \text{if } y \leq 0 \end{cases}$$

is a function of class C^1 on $\mathbb{R}$ and $\frac{dk_x}{dy} = h_x$ for all x fixed. Hence, we have from the Fundamental Theorem of Calculus that

$$\int_0^1 \frac{\partial f}{\partial x}(x, y) dy = \int_0^1 h_x(y) dy = k_x(y) \Big|_{y=0}^{y=1} = (1 - 2x^2) e^{-x^2}.$$

Hence

$$\int_0^1 \frac{\partial f}{\partial x}(x, y) dy \Big|_{x=0} = 1. \quad (2.5.14)$$

It follows from (2.5.13) and (2.5.14) that (2.5.11) is true.

However, since $\frac{\partial f}{\partial x}(0, y) = 0$ for all y , we get $\int_0^1 \frac{\partial f}{\partial x}(0, y) dy = 0$.

The fact that (2.5.11) is true does not come from the theory because $\frac{\partial f}{\partial x}$ are not bounded in any neighbourhood of the origin. Thus (2) cannot be satisfied. ♡

Question 2.34

Let $F(x) = \int_{g(x)}^{h(x)} f(x, y) dy$ where all the functions are sufficiently differentiable. Compute $F'(x)$.

Solution

Since

$$F(x) = \int_a^{h(x)} f(x, y) dy + \int_{g(x)}^a f(x, y) dy = \int_a^{h(x)} f(x, y) dy - \int_a^{g(x)} f(x, y) dy ,$$

we get

$$\begin{aligned} F'(x) &= f(x, h(x))h'(x) + \int_a^{h(x)} \frac{\partial f}{\partial x}(x, y) dy - f(x, g(x))g'(x) - \int_a^{g(x)} \frac{\partial f}{\partial x}(x, y) dy \\ &= f(x, h(x))h'(x) - f(x, g(x))g'(x) + \int_{g(x)}^{h(x)} \frac{\partial f}{\partial x}(x, y) dy . \end{aligned}$$

♡

Question 2.35

Compute the derivatives of

$$\text{a) } F(x) = \int_x^{x^2} \frac{e^{xy}}{y} dy \quad \text{b) } G(x) = \int_{2x}^{3x} \cos(x^2 y) dy$$

Solution

We use the formula found in Question 2.34.

a)

$$\begin{aligned} F'(x) &= \left(\frac{e^{xy}}{y} \right) \Big|_{y=x^2} \frac{d}{dx}(x^2) - \left(\frac{e^{xy}}{y} \right) \Big|_{y=x} \frac{d}{dx}(x) + \int_x^{x^2} \frac{\partial}{\partial x} \left(\frac{e^{xy}}{y} \right) dy \\ &= \frac{2e^{x^3}}{x} - \frac{e^{x^2}}{x} + \int_x^{x^2} e^{xy} dy = \frac{2e^{x^3}}{x} - \frac{e^{x^2}}{x} + \frac{e^{xy}}{x} \Big|_{y=x}^{y=x^2} \\ &= \frac{2e^{x^3}}{x} - \frac{e^{x^2}}{x} + \frac{1}{x} (e^{x^3} - e^{x^2}) = \frac{3e^{x^3}}{x} - \frac{2e^{x^2}}{x} . \end{aligned}$$

b)

$$\begin{aligned} F'(x) &= (\cos(x^2 y)) \Big|_{y=3x} \frac{d}{dx}(3x) - (\cos(x^2 y)) \Big|_{y=2x} \frac{d}{dx}(2x) + \int_{2x}^{3x} \frac{\partial}{\partial x} (\cos(x^2 y)) dy \\ &= 3 \cos(3x^3) - 2 \cos(2x^3) - \int_{2x}^{3x} 2xy \sin(x^2 y) dy \end{aligned}$$

$$\begin{aligned}
&= 3 \cos(3x^3) - 2 \cos(2x^3) - \frac{2}{x^3} \int_{2x^3}^{3x^3} u \sin(u) \, du \\
&= 3 \cos(3x^3) - 2 \cos(2x^3) - \frac{2}{x^3} \left(\sin(u) - u \cos(u) \right) \Big|_{u=2x^3}^{u=3x^3} \\
&= 9 \cos(3x^3) - 6 \cos(2x^3) - \frac{2}{x^3} (\cos(3x^3) - \cos(2x^3)) \quad ,
\end{aligned}$$

where we have use the substitution $u = x^2 y$ for the third equality and integration by parts for the fourth equality. ♥

Question 2.36

Given a continuous function $f : \mathbb{R} \rightarrow \mathbb{R}$, let $f_0(x) = f(x)$ and

$$f_n(x) = \frac{1}{(n-1)!} \int_0^x (x-y)^{n-1} f(y) \, dy$$

for $n \in \mathbb{N}^+$ and $x \in \mathbb{R}$. Prove that $\frac{df_n}{dx}(x) = f_{n-1}(x)$ and use this result to prove that f_n is the n^{th} order primitive of f ; namely, $\frac{d^n f_n}{dx^n} = f$.

Improper Integrals

The second topic that will be useful later is improper integrals. Our definition of Riemann integral is in some way restrictive. For instance

$$f(x) = \begin{cases} 1/x^2 & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$$

is not Riemann integrable on $[0, 1]$. Since f is not bounded on $[0, 1]$, we cannot define Riemann sum. For this reason, the concept of improper integral is introduced in calculus courses. For instance, we define the integral of f on $[0, 1]$ as

$$\int_0^1 f(x) \, dx = \lim_{a \rightarrow 0^+} \int_a^1 f(x) \, dx \quad .$$

We have that f is Riemann integrable on $[a, 1]$ for all $a > 0$.

We review below some of the results about improper integrals that we assume are known to the readers. They will be useful for the questions below.

1. $\int_1^\infty \frac{1}{x^p} \, dx$ converges for $p > 1$ and diverges for $p \leq 1$.
2. $\int_0^1 \frac{1}{x^p} \, dx$ converges for $p < 1$ and diverges for $p \geq 1$.

3. Suppose that $f, g :]a, b[\rightarrow \mathbb{R}$ are two continuous functions such that $0 \leq f(x) \leq g(x)$ for $a < x < b$. Then $\int_a^b f(x) dx$ converges if $\int_a^b g(x) dx$ converges. Moreover, $\int_a^b g(x) dx$ diverges if $\int_a^b f(x) dx$ diverges. These results are known as the **comparison test**.

4. Suppose that $f :]a, b[\rightarrow \mathbb{R}$ is a continuous function. If $\int_a^b |f(x)| dx$ converges, then $\int_a^b f(x) dx$ converges.

Question 2.37

Determine if the following improper integrals converge or diverge.

a) $\int_0^1 \frac{1}{x^{1/3}(x^2 + 2x)^{1/2}} dx$	b) $\int_0^1 \frac{x^2}{\sqrt{1-x^4}} dx$	c) $\int_0^\infty \sqrt{x}e^{-x} dx$
d) $\int_1^\infty \frac{1}{x\sqrt{x+3}} dx$	e) $\int_1^\infty \frac{x^2 - 2x - 3}{x(x^2 + 1)} dx$	f) $\int_0^\infty x^2 e^{-x^2} dx$
g) $\int_3^\infty \frac{\sin(4x)}{x^2 - x - 2} dx$	h) $\int_1^\infty \tan\left(\frac{1}{x}\right) dx$	i) $\int_0^\infty \frac{e^x}{1 + e^{2x}} dx$
j) $\int_0^\infty \frac{1 - \cos(x)}{x^2} dx$	k) $\int_0^1 \frac{1}{\sqrt{x^3 + x^2}} dx$	l) $\int_1^\infty \frac{1}{\sqrt{x^3 + x^2}} dx$
m) $\int_0^\infty e^x \sin(e^x) dx$	n) $\int_0^\infty x \sin(e^x) dx$	o) $\int_0^1 \frac{\sin(1/x)}{x} dx$
p) $\int_0^1 x \ln(x) dx$		

Solution

a) We have $x^{1/3}(x^2 + 2x)^{1/2} = x^{5/6}(x+2)^{1/2} \geq x^{5/6}\sqrt{2}$ for $0 \leq x \leq 1$. Thus $\frac{1}{x^{1/3}(x^2 + 2x)^{1/2}} \leq \frac{1}{x^{5/6}\sqrt{2}}$ for $0 < x \leq 1$. Since $\int_0^1 x^{-p} dx$ converges for $p < 1$, we have by the comparison test that $\int_0^1 \frac{1}{x^{1/3}(x^2 + 2x)^{1/2}} dx$ converges.

b) We have

$$\lim_{q \rightarrow 1^-} \int_0^q \frac{x^2}{\sqrt{1-x^4}} dx = \lim_{q \rightarrow 1^-} \int_0^q \frac{x^2}{\sqrt{(1-x^2)(1+x^2)}} dx .$$

With the substitution $u = 1 - x^2$ for $0 \leq x < 1$, we get

$$\lim_{q \rightarrow 1^-} \int_0^q \frac{x^2}{\sqrt{1-x^4}} dx = \lim_{q \rightarrow 1^-} \frac{-1}{2} \int_1^{\sqrt{1-q}} \frac{\sqrt{1-u}}{\sqrt{u(2-u)}} du = \lim_{q \rightarrow 1^-} \frac{1}{2} \int_{\sqrt{1-q}}^1 \frac{\sqrt{1-u}}{\sqrt{u(2-u)}} du .$$

Since $0 \leq \frac{\sqrt{1-u}}{\sqrt{u(2-u)}} \leq \frac{1}{\sqrt{u}}$ for $0 < u \leq 1$ and $\int_0^1 u^{-1/2} du$ converges, we get from the comparison test that $\int_0^1 \frac{\sqrt{1-u}}{\sqrt{u(2-u)}} du$ converges. Hence $\int_0^1 \frac{x^2}{\sqrt{1-x^4}} dx$ converges.

c) Since $\sqrt{x}e^{-x}$ is continuous on $[0, 1]$, the integral $\int_0^1 \sqrt{x}e^{-x} dx$ exists. We only have to consider $\int_1^\infty \sqrt{x}e^{-x} dx$.

We first prove that $h(x) = \sqrt{x}e^{-x/2}$ is bounded for $x \geq 1$. Since

$$h'(x) = \frac{1}{2}x^{-1/2}e^{-x/2} - \frac{1}{2}x^{1/2}e^{-x/2} = \frac{1}{2}x^{-1/2}(1-x)e^{-x/2} < 0$$

for $x > 1$, the function h is decreasing on $]1, \infty[$. Thus $h(x) \leq h(1) = e^{-1/2}$ for $x > 1$. Since $0 \leq \sqrt{x}e^{-x} = h(x)e^{-x/2} \leq e^{-1/2}e^{-x/2}$ for $x \geq 1$ and $\int_1^\infty e^{-x/2} dx$ converges, we get from the comparison test that $\int_1^\infty \sqrt{x}e^{-x} dx$ converges. Thus

$$\int_0^\infty \sqrt{x}e^{-x} dx = \int_0^1 \sqrt{x}e^{-x} dx + \int_1^\infty \sqrt{x}e^{-x} dx$$

converges.

d) Since $\frac{1}{x\sqrt{x+3}} \leq \frac{1}{x^{3/2}}$ and $\int_1^\infty x^{-3/2} dx$ converges because $\int_1^\infty x^{-p} dx$ converges for $p > 1$, we get from the comparison test that $\int_1^\infty \frac{1}{x\sqrt{x+3}} dx$ converges.

e) Since $\lim_{x \rightarrow \infty} \frac{x^2 - 2x - 3}{x^2 + 1} = 1$, there exists $C > 0$ such that $\frac{x^2 - 2x - 3}{x^2 + 1} > \frac{1}{2}$ for $x \geq C$. Hence $\frac{x^2 - 2x - 3}{x(x^2 + 1)} > \frac{1}{2x}$ for $x \geq C$. Since $\int_C^\infty \frac{1}{2x} du$ diverges, we get from the comparison test that $\int_C^\infty \frac{x^2 - 2x - 3}{x(x^2 + 1)} dx$ diverges. It follows from

$$\int_1^q \frac{x^2 - 2x - 3}{x(x^2 + 1)} dx = \int_1^C \frac{x^2 - 2x - 3}{x(x^2 + 1)} dx + \int_C^q \frac{x^2 - 2x - 3}{x(x^2 + 1)} dx$$

that $\int_1^\infty \frac{x^2 - 2x - 3}{x(x^2 + 1)} dx$ diverges. Note that $\int_1^C \frac{x^2 - 2x - 3}{x(x^2 + 1)} dx$ is a real number because $\frac{x^2 - 2x - 3}{x(x^2 + 1)}$ is continuous for $1 \leq x \leq C$.

f) Since $x^2e^{-x^2} \leq x^3e^{-x^2}$ for $x \geq 1$ and $\int_1^\infty x^3e^{-x^2} dx$ converges, we get from the comparison

test that $\int_1^\infty x^2 e^{-x^2} dx$ converges. Note that

$$\begin{aligned} \lim_{q \rightarrow \infty} \int_1^q x^3 e^{-x^2} dx &= \lim_{q \rightarrow \infty} \frac{1}{2} \int_1^{q^2} u e^{-u} du = \lim_{q \rightarrow \infty} -\frac{1}{2} (e^{-u} + u e^{-u}) \Big|_1^{q^2} \\ &= \lim_{q \rightarrow \infty} -\frac{1}{2} (e^{-q^2} + q^2 e^{-q^2} - 2e^{-1}) = \frac{1}{e}, \end{aligned}$$

where we have used the substitution $u = x^2$ for the first equality, integration by parts for the second equality, and l'Hospital Rule to compute $\lim_{q \rightarrow \infty} q^2 e^{-q^2} = 0$. Hence

$$\int_0^q x^2 e^{-x^2} dx = \underbrace{\int_0^1 x^2 e^{-x^2} dx}_{\in \mathbb{R}} + \int_1^q x^2 e^{-x^2} dx$$

converges as $q \rightarrow \infty$.

g) We have

$$\left| \frac{\sin(4x)}{x^2 - x - 2} \right| \leq \frac{1}{x^2 - x - 2} = \frac{1}{(x-2)(x+1)}$$

for $x \geq 3$. Moreover, $x-2 \geq \frac{x}{3} \Rightarrow \frac{1}{x-2} \leq \frac{3}{x}$ and $x+1 > x \Rightarrow \frac{1}{x+1} < \frac{1}{x}$ for $x \geq 3$. Thus $\left| \frac{\sin(4x)}{x^2 - x - 2} \right| < \frac{3}{x^2}$ for $x \geq 3$. Since $\int_3^\infty \frac{1}{x^2} dx$ converges because $\int_3^\infty x^{-p} dx$ converges for $p > 1$, we get from the comparison theorem that $\int_3^\infty \frac{\sin(4x)}{x^2 - x - 2} dx$ converges absolutely, and so converges.

h) Since $\tan(u) > u$ for $0 < u < 1$, we get $\tan\left(\frac{1}{x}\right) > \frac{1}{x}$ for $1 \leq x < \infty$. Since $\int_1^\infty \frac{1}{x} dx$ diverges, we get from the comparison test that $\int_1^\infty \tan\left(\frac{1}{x}\right) dx$ diverges. $\heartsuit$

Question 2.38

Consider the function

$$f(x) = \begin{cases} 1 & \text{if } x \in \left[n, n + \frac{1}{2^n}\right] \text{ for some } n \in \mathbb{N}^+ \\ 0 & \text{otherwise} \end{cases}$$

a) Prove that $f(x) \not\rightarrow 0$ as $x \rightarrow \infty$ but f is bounded.

b) Prove that $\int_0^\infty f(x) dx$ converges and is equal to 1.

c) Construct an unbounded function $f : \mathbb{R} \rightarrow \mathbb{R}$ such that $\int_0^\infty f(x) dx$ converges.

Solution

a) Since $f(n) = 1$ for all positive integers, we do not have that $f(x) \rightarrow 0$ as $x \rightarrow \infty$.

b) Let $F(q) = \int_0^q f(x) dx$. Since $f(x) \geq 0$ for all $x \geq 0$, the function F is monotonically increasing. Moreover

$$F(q) \leq \sum_{n=1}^{\lfloor q \rfloor} \frac{1}{2^n} = -1 + \sum_{n=0}^{\lfloor q \rfloor} \frac{1}{2^n} \leq -1 + \sum_{n=0}^{\infty} \frac{1}{2^n} = -1 + \frac{1}{1 - (1/2)} = 1$$

for all q , where $\lfloor q \rfloor$ is the integer part of q . Since F is monotonically increasing and bounded, we have that $\lim_{q \rightarrow \infty} F(q) = \sup_{q \geq 0} F(q)$ exists.

Given $\epsilon > 0$, choose M large enough such that $\left| \sum_{n=1}^m \frac{1}{2^n} - 1 \right| < \epsilon$ for $m \geq M$. This is possible because the geometric series $\sum_{n=1}^{\infty} \frac{1}{2^n}$ converges to 1. Hence, for $q > M + 1/2^M$, we have

$$1 \geq F(q) \geq F(M + 1/2^M) = \sum_{n=1}^M \frac{1}{2^n} > 1 - \epsilon.$$

Since ϵ is arbitrary, we get $\lim_{q \rightarrow \infty} F(q) = \sup_{q \geq 0} F(q) = 1$.

c) Consider the function

$$g(x) = \begin{cases} n & \text{if } x \in \left[n, n + \frac{1}{n2^n} \right] \text{ for some } n \in \mathbb{N}^+ \\ 0 & \text{otherwise} \end{cases}$$

It is clear that g is not bounded because $g(n) = n$ for all positive integers n . Let $G(q) = \int_0^q g(x) dx$. Since $g(x) \geq 0$ for all $x \geq 0$, the function G is monotonically increasing. Moreover

$$G(q) \leq \sum_{n=1}^{\lfloor q \rfloor} \frac{n}{n2^n} = -1 + \sum_{n=0}^{\lfloor q \rfloor} \frac{1}{2^n} \leq -1 + \sum_{n=0}^{\infty} \frac{1}{2^n} = -1 + \frac{1}{1 - (1/2)} = 1$$

for all q , where $\lfloor q \rfloor$ is the integer part of q . Since G is monotonically increasing and bounded, we have that $\lim_{q \rightarrow \infty} G(q) = \sup_{q \geq 0} G(q)$ exists. We also have $G(q) = F(q)$ for all positive integer q . Thus $\lim_{\substack{q \rightarrow \infty \\ q \in \mathbb{N}^+}} G(q) = \lim_{\substack{q \rightarrow \infty \\ q \in \mathbb{N}^+}} F(q) = 1$. Since $\lim_{q \rightarrow \infty} G(q)$ exists, we have $\lim_{q \rightarrow \infty} G(q) = 1$. $\heartsuit$

Question 2.39

Prove that $\int_1^{\infty} \frac{|\sin(x)|}{x} dx$ diverges but $\int_1^{\infty} \frac{\sin(x)}{x} dx$ converges

Solution

We first prove that $\int_1^{\infty} \frac{|\sin(x)|}{x} dx$ diverges. We have

$$\int_{n\pi}^{(n+1)\pi} \frac{|\sin(x)|}{x} dx \geq \int_{n\pi}^{(n+1)\pi} \frac{|\sin(x)|}{(n+1)\pi} dx = \frac{1}{(n+1)\pi} \int_{n\pi}^{(n+1)\pi} |\sin(x)| dx$$

$$= \frac{1}{(n+1)\pi} \int_0^\pi \sin(x) \, dx = \frac{2}{(n+1)\pi}.$$

and

$$\int_{n\pi}^{(n+1)\pi} \frac{1}{x} \, dx \leq \frac{1}{n\pi} \int_{n\pi}^{(n+1)\pi} dx = \frac{1}{n} \leq \frac{2}{n+1}$$

for $n \geq 1$. Thus

$$\int_{n\pi}^{(n+1)\pi} \frac{|\sin(x)|}{x} \, dx \geq \frac{2}{(n+1)\pi} \geq \frac{1}{\pi} \int_{n\pi}^{(n+1)\pi} \frac{1}{x} \, dx$$

for $n \in \mathbb{N}^+$.

Since

$$\int_\pi^{q\pi} \frac{|\sin(x)|}{x} \, dx = \sum_{n=1}^{q-1} \int_{n\pi}^{(n+1)\pi} \frac{|\sin(x)|}{x} \, dx \geq \frac{1}{\pi} \sum_{n=1}^{q-1} \int_{n\pi}^{(n+1)\pi} \frac{1}{x} \, dx = \frac{1}{\pi} \int_\pi^{q\pi} \frac{1}{x} \, dx$$

for $q \in \mathbb{N}^+$, and $\int_\pi^\infty \frac{1}{x} \, dx$ diverges, we get that $\int_\pi^\infty \frac{|\sin(x)|}{x} \, dx$ diverges and therefore $\int_1^\infty \frac{|\sin(x)|}{x} \, dx$ diverges.

To prove that $\int_1^\infty \frac{\sin(x)}{x} \, dx$ converges, we use integration by parts to get

$$\int_1^q \frac{\sin(x)}{x} \, dx = -\frac{\cos(x)}{x} \Big|_1^q - \int_1^q \frac{\cos(x)}{x^2} \, dx = \cos(1) - \frac{\cos(q)}{q} - \int_1^q \frac{\cos(x)}{x^2} \, dx.$$

Since $\frac{|\cos(x)|}{x^2} \leq \frac{1}{x^2}$ and $\int_1^\infty \frac{1}{x^2} \, dx$ converges, we get that

$$\lim_{q \rightarrow \infty} \int_1^q \frac{\sin(x)}{x} \, dx = \cos(1) - \lim_{q \rightarrow \infty} \int_1^q \frac{\cos(x)}{x^2} \, dx$$

exists. ♡

Question 2.40

Let $f : [0, \infty[\rightarrow \mathbb{R}$ be a continuous function and $g : [0, \infty[\rightarrow \mathbb{R}$ be a continuously differentiable function. Suppose that $F(x) = \int_0^x f(y) \, dy$ is bounded on $[0, \infty[$, $g'(x) \leq 0$ for $x \geq 0$, and

$\lim_{x \rightarrow \infty} g(x) = 0$. Prove that $\int_0^\infty f(x)g(x) \, dx$ converges.

Solution

Since $f : [0, \infty[\rightarrow \mathbb{R}$ is continuous, we have from the Fundamental Theorem of Calculus that F is differentiable and $F'(x) = f(x)$ for all x .

Using integration by parts, we find that

$$\int_0^q f(x)g(x) \, dx = \int_0^q F'(x)g(x) \, dx = F(x)g(x) \Big|_0^q - \int_0^q F(x)g'(x) \, dx. \quad (2.5.15)$$

From the first hypothesis, there exists M such that $F(x) \leq M$ for all $x \geq 0$. Hence $|F(q)g(q)| \leq M|g(q)| \rightarrow 0$ as $q \rightarrow \infty$ from the third hypothesis. Thus

$$\lim_{q \rightarrow \infty} F(x)g(x) \Big|_0^q = \lim_{q \rightarrow \infty} (F(q)g(q) - F(0)g(0)) = -F(0)g(0) = 0 \quad (2.5.16)$$

because $F(0) = 0$. Moreover

$$\begin{aligned} \int_0^q |F(x)g'(x)| dx &\leq M \int_a^q |g'(x)| dx = -M \int_a^q g'(x) dx \\ &= -Mg(x) \Big|_0^q = Mg(0) - Mg(q) \rightarrow Mg(0) \end{aligned}$$

as $q \rightarrow \infty$, where we have used the second hypothesis for the first equality and the third hypothesis to compute the limit. Thus $\int_0^\infty F(x)g'(x) dx$ converges absolutely and therefore converges. It follows from (2.5.15) and (2.5.16) that $\int_0^\infty f(x)g(x) dx$ converges. $\heartsuit$

Question 2.41

Suppose that $f : [0, 1] \rightarrow \mathbb{R}$ is Riemann integrable, prove that $\lim_{q \rightarrow 0^+} \int_q^1 f(x) dx = \int_0^1 f(x) dx$. This shows that the definition of the improper integral agrees with the definition of the Riemann integral when the function is Riemann integrable.

Question 2.42

Give a function $f : [a, b] \rightarrow \mathbb{R}$ such that $\lim_{q \rightarrow a^+} \int_q^b f(x) dx = \int_a^b f(x) dx$ but $\lim_{q \rightarrow a^+} \int_q^b |f(x)| dx$ does not exist. The property that f integrable implies $|f|$ integrable does not hold for improper integrals.

Question 2.43

Suppose that $f : [1, \infty[\rightarrow \mathbb{R}$ is a monotonically decreasing function such that $f(x) \geq 0$ for all $x \geq 1$. Prove that $\int_1^\infty f(x) dx$ converges if and only if the series $\sum_{k=1}^\infty f(k)$ converges. This is called the **integral test**.

Solution

$\Rightarrow$) Suppose that $\int_1^\infty f(x) dx$ converges. Let $M = \int_1^\infty f(x) dx < \infty$. Since $f(x) \geq 0$ for all $x \geq 1$, we have

$$M = \int_1^\infty f(x) dx \geq \int_1^N f(x) dx$$

for all $N > 0$. Since f is monotonically decreasing, we have

$$M \geq \int_1^N f(x) dx = \sum_{n=1}^{N-1} \int_n^{n+1} f(x) dx \geq \sum_{n=1}^{N-1} \int_n^{n+1} f(n+1) dx = \sum_{n=1}^{N-1} f(n+1) = \sum_{n=2}^N f(n)$$

for all $N > 0$. Thus $S_N = \sum_{n=1}^N f(n)$ is a non-decreasing and bounded sequence; more precisely, bounded by $M + f(1)$. It follows that $\{S_N\}_{N=1}^\infty$ converges; namely, $\sum_{n=1}^\infty f(n)$ converges.

$\Leftarrow$) Suppose that $\sum_{n=1}^\infty f(n)$ converges. Let $M = \sum_{n=1}^\infty f(n) < \infty$. We have that $F(q) = \int_1^q f(x) dx$ for $q \geq 1$ is a non-decreasing, non-negative function because $f(x) \geq 0$ for all $x \geq 1$. Given $q \geq 1$, choose $N \in \mathbb{N}^+$ such that $N \geq q$. Since $f(x) \geq 0$ for all $x \geq 1$, we have

$$F(q) = \int_1^q f(x) dx \leq \int_1^N f(x) dx .$$

Moreover, since f is monotonically decreasing, we have

$$\begin{aligned} F(q) &\leq \int_1^N f(x) dx = \sum_{n=1}^{N-1} \int_n^{n+1} f(x) dx \\ &\leq \sum_{n=1}^{N-1} \int_n^{n+1} f(n) dx = \sum_{n=1}^{N-1} f(n) \leq \sum_{n=1}^\infty f(n) = M \end{aligned}$$

Since $q \geq 1$ is arbitrary, F is also bounded by M . Since $F : [1, \infty[\rightarrow \mathbb{R}$ is a bounded and non-decreasing function, we have that $\lim_{q \rightarrow \infty} F(q)$ exists; namely, $\int_1^\infty f(x) dx$ converges. $\heartsuit$

Question 2.44

Suppose that $f : [a, \infty[\rightarrow \mathbb{R}$ is integrable on all bounded intervals of the form $[a, q]$. Prove that $\int_a^\infty f(x) dx$ converges if and only if for every $\epsilon > 0$ there exists $K > a$ such that $\left| \int_c^d f(x) dx \right| < \epsilon$ for $K \leq c \leq d$.

Solution

Let $F(q) = \int_a^q f(x) dx$ for $q \geq a$. The question can be reformulated as follows. Show that $\lim_{q \rightarrow \infty} F(q)$ exists if and only if there exists for every $\epsilon > 0$ a constant $K > a$ such that $|F(d) - F(c)| < \epsilon$ for $K \leq c \leq d$. But this is just the Cauchy Criterion of convergence for the function $F : [a, \infty[\rightarrow \mathbb{R}$. $\heartsuit$

Question 2.45

Suppose that $\{f_k\}_{k=0}^\infty$ is a sequence of functions defined on $]0, \infty[$ such that (1) f_k is integrable on $[a, b]$ for every interval $[a, b] \subset]0, \infty[$, (2) there exists a function $f :]0, \infty[\rightarrow \mathbb{R}$ such that $f_k \rightarrow f$ uniformly on compact subsets of $]0, \infty[$ as $k \rightarrow \infty$, and (3) there exists a function $g :]0, \infty[\rightarrow \mathbb{R}$ such that $|f_k| \leq g$ on $]0, \infty[$ and $\int_0^\infty g(x) dx$ converges. Prove that

$$\lim_{k \rightarrow \infty} \int_0^\infty f_k(x) dx = \int_0^\infty f(x) dx$$

This is a weak version of the Lebesgue's dominated convergence theorem.

Chapter 3

The Riemann Integral of Real Valued Functions of Several Variables

3.1 Definition and Properties

As usual, we start with some basic definitions.

Definition 3.1.1

A **rectangle** $R \subset \mathbb{R}^n$ is a set of the form $R = \prod_{i=1}^n I_i$ where the I_i are bounded intervals.

The **volume** of a rectangle $R = \prod_{i=1}^n I_i$ is the value $V(R) = \prod_{i=1}^n |I_i|$ where $|I_i|$ is the length of the interval I_i . We may use the notation $A(R)$ instead of $V(R)$ when $n = 2$.

The definition of sets of zero content, sets of measure zero and (Jordan) measurable sets in $\mathbb{R}$ (Definitions 2.1.14 and 2.1.16) can be extended to $\mathbb{R}^n$.

Definition 3.1.2

A set $S \in \mathbb{R}^n$ has **zero content** if for every $\epsilon > 0$ there exists a finite collection of closed rectangles $\{R_i\}_{0 \leq i \leq M}$ such that $S \subset \bigcup_{i=0}^M R_i$ and $\sum_{i=0}^M V(R_i) < \epsilon$.

A set $S \in \mathbb{R}^n$ has **measure zero** if for every $\epsilon > 0$ there exists a countable collection of closed rectangles $\{R_i\}_{i \geq 0}$ such that $S \subset \bigcup_{i=0}^{\infty} R_i$ and $\sum_{i=0}^{\infty} V(R_i) < \epsilon$.

Definition 3.1.3

A set $S \subset \mathbb{R}^n$ is **(Jordan) measurable** if S is bounded and ∂S has measure zero.

Remark 3.1.4

In the definition of sets of measure zero or zero content, we may use open rectangles instead of closed rectangles. Suppose that $S \subset \mathbb{R}^n$ has measure zero according to the definition using closed rectangles. Given $\epsilon > 0$, there exists a countable collection $\{R_i\}_{i \geq 0}$ of closed rectangles such that $S \subset \bigcup_{i=0}^{\infty} R_i$ and $\sum_{i=0}^{\infty} V(R_i) < \epsilon/2$. For each i , choose an open rectangle V_i such that $R_i \subset V_i$ and $V(V_i) \leq V(R_i) + \epsilon/2^{i+2}$. Then $S \subset \bigcup_{i=0}^{\infty} R_i \subset \bigcup_{i=0}^{\infty} V_i$ and

$$\sum_{i=0}^{\infty} V(V_i) \leq \sum_{i=0}^{\infty} V(R_i) + \frac{\epsilon}{4} \sum_{i=0}^{\infty} \frac{1}{2^i} = \sum_{i=0}^{\infty} V(R_i) + \frac{\epsilon}{2} < \epsilon .$$

Since ϵ is arbitrary, S has measure zero according to the definition using open rectangles.

Conversely, suppose that $S \subset \mathbb{R}^n$ has measure zero according to the definition using open rectangles. Given $\epsilon > 0$, there exists a countable collection $\{V_i\}_{i \geq 0}$ of open rectangles such that $S \subset \bigcup_{i=0}^{\infty} V_i$ and $\sum_{i=0}^{\infty} V(V_i) < \epsilon$. For each i , let $R_i = \overline{V_i}$. Then $V(R_i) = V(V_i)$ for all i . Hence $S \subset \bigcup_{i=0}^{\infty} V_i \subset \bigcup_{i=0}^{\infty} R_i$ and $\sum_{i=0}^{\infty} V(R_i) = \sum_{i=0}^{\infty} V(V_i) < \epsilon$. Since ϵ is arbitrary, S has measure zero according to the definition using closed rectangles. ♠

Given a bounded set $V \subset \mathbb{R}^n$, we define the diameter of V as the number

$$\text{diam } V = \sup\{\|\mathbf{x} - \mathbf{y}\| : \mathbf{x}, \mathbf{y} \in V\} .$$

Proposition 3.1.5

If, for every $\epsilon > 0$, there exists a countable collection of compact sets $\{K_i\}_{i \geq 0}$ such that $S \subset \bigcup_{i=0}^{\infty} K_i$ and $\sum_{i=0}^{\infty} \text{diam } K_i < \epsilon$, then the set $S \in \mathbb{R}^n$ has measure zero.

Proof.

Suppose that S is a set satisfying the hypotheses of the proposition. Given $\epsilon > 0$, choose a countable collection of compact sets $\{K_i\}_{i \geq 0}$ such that $S \subset \bigcup_{i=0}^{\infty} K_i$ and $\sum_{i=0}^{\infty} \text{diam } K_i < \epsilon^{1/n}$.

Each K_i is a subset of a rectangle R_i with edges of length $d_i = \text{diam } K_i$. Thus $V(R_i) = d_i^n$. Hence $S \subset \bigcup_{i=0}^{\infty} K_i \subset \bigcup_{i=0}^{\infty} R_i$ and

$$\sum_{i=0}^{\infty} V(R_i) = \sum_{i=0}^{\infty} d_i^n \leq \left(\sum_{i=0}^{\infty} d_i \right)^n < (\epsilon^{1/n})^n = \epsilon .$$

Since ϵ is arbitrary, the definition of sets of measure zero is satisfied. ■

Proposition 3.1.6

Compact subsets of a set of measure zero have zero content.

Proof.

Suppose that Z is a set of measure zero and that $K \subset Z$ is a compact set. We prove that K has zero content. Given $\epsilon > 0$, since Z has measure zero, there exists a countable collection $\{V_i\}_{i \geq 0}$ of open rectangles such that $K \subset Z \subset \bigcup_{i=0}^{\infty} V_i$ and $\sum_{i=0}^{\infty} V(V_i) < \epsilon$. Since K is compact, there exists a finite subcover $\{V_{i_j}\}_{0 \leq j \leq N}$ such that $K \subset \bigcup_{j=0}^N V_{i_j}$. We obviously have $\sum_{j=0}^N V(V_{i_j}) \leq \sum_{i=0}^{\infty} V(V_i) < \epsilon$. Since ϵ is arbitrary, this shows that K has zero content. ■

Definition 3.1.7

A **partition** P of a rectangle $R = \prod_{i=1}^n I_i \subset \mathbb{R}^n$, where each I_i is an interval, is a collection $P = \{P_i : 1 \leq i \leq n\}$ where $P_i = \{p_{k,i} : 0 \leq k \leq N_i\}$ is a partition of the interval I_i for $1 \leq i \leq n$.

Rectangles of the form $R_{k_1, k_2, \dots, k_n} = \prod_{i=1}^n [p_{k_i-1, i}, p_{k_i, i}]$, where $1 \leq k_i \leq N_i$, are called **subrectangles of the partition** P .

We define the integral of a bounded function on a rectangle in basically the same way that we defined the integral of a bounded function on an interval.

Definition 3.1.8

Let $R \subset \mathbb{R}^n$ be a rectangle as defined above, $f : R \rightarrow \mathbb{R}$ be a bounded function, and P be a partition of $[a, b]$. Moreover, let $m_{k_1, k_2, \dots, k_n} = \inf\{f(\mathbf{x}) : \mathbf{x} \in R_{k_1, k_2, \dots, k_n}\}$ and $M_{k_1, k_2, \dots, k_n} = \sup\{f(\mathbf{x}) : \mathbf{x} \in R_{k_1, k_2, \dots, k_n}\}$ for $1 \leq k_i \leq N_i$ and $1 \leq i \leq n$. The **lower Riemann sum** of f with respect to the partition P is

$$\mathcal{L}_P(f) = \sum_{\substack{1 \leq k_i \leq N_i \\ 1 \leq i \leq n}} m_{k_1, k_2, \dots, k_n} V(R_{k_1, k_2, \dots, k_n}) .$$

The **upper Riemann sum** of f with respect to the partition P is

$$\mathcal{U}_P(f) = \sum_{\substack{1 \leq k_i \leq N_i \\ 1 \leq i \leq n}} M_{k_1, k_2, \dots, k_n} V(R_{k_1, k_2, \dots, k_n}) .$$

Definition 3.1.9

A **refinement** for a partition $P = \{P_i : 1 \leq i \leq n\}$ of a rectangle R is another partition $Q = \{Q_i : 1 \leq i \leq n\}$ of R such that Q_i is a refinement of P_i for $1 \leq i \leq n$.

The lower and upper Riemann sums for functions of several variables have the same properties than the lower and upper Riemann sums for functions of one variable.

Definition 3.1.10

Let $R \subset \mathbb{R}^n$ be a rectangle as defined above and $f : R \rightarrow \mathbb{R}$ be a bounded function. The **lower Riemann integral** of f on R is defined as

$$\mathcal{L}(f) = \sup\{\mathcal{L}_P(f) : P \text{ is a partition of } R\} .$$

The **upper Riemann integral** of f on R is defined as

$$\mathcal{U}(f) = \inf\{\mathcal{U}_P(f) : P \text{ is partition of } R\} .$$

Definition 3.1.11

Let $R \subset \mathbb{R}^n$ be a rectangle as defined above and $f : R \rightarrow \mathbb{R}$ be a bounded function. If $\mathcal{U}(f) = \mathcal{L}(f)$, then we say that **f is (Riemann) integrable** on the rectangle R . The **Riemann integral of f** on R is defined as

$$\int_R f = \int_R f(\mathbf{x}) \, d\mathbf{x} = \mathcal{L}(f) = \mathcal{U}(f) .$$

Most of the results that we have stated about the Riemann integral of a bounded function on an interval remain true for the Riemann integral of a bounded function on a rectangle. The obvious exception is Theorem 2.3.1 that we will reformulate shortly.

Suppose that R is a rectangle in $\mathbb{R}^n$ and $f : R \rightarrow \mathbb{R}$ is a bounded function. Given $\mathbf{x} \in R$, let

$$\begin{aligned} M(\mathbf{x}, f, \delta) &= \sup\{f(\mathbf{y}) : \mathbf{y} \in R \text{ and } \|\mathbf{y} - \mathbf{x}\| < \delta\} , \\ m(\mathbf{x}, f, \delta) &= \inf\{f(\mathbf{y}) : \mathbf{y} \in R \text{ and } \|\mathbf{y} - \mathbf{x}\| < \delta\} \end{aligned}$$

and

$$o(\mathbf{x}, f) = \lim_{\delta \rightarrow 0^+} (M(\mathbf{x}, f, \delta) - m(\mathbf{x}, f, \delta)) .$$

The limit always exists because $M(\mathbf{x}, f, \delta) - m(\mathbf{x}, f, \delta) > 0$ is decreasing as δ decreases to 0. It is also easy to show that f is continuous at $\mathbf{x}$ if and only if $o(\mathbf{x}, f) = 0$.

The following result will be useful shortly.

Lemma 3.1.12

Suppose that $R \subset \mathbb{R}^n$ is a closed set. The set $Z_\epsilon = \{\mathbf{x} \in R : o(\mathbf{x}, f) \geq \epsilon\}$ is a closed subset of R .

Proof.

If $z_\epsilon = \emptyset$, then Z_ϵ is closed.

Suppose that $Z_\epsilon \neq \emptyset$. We prove that the complement of Z_ϵ in R is open relative to the topology induced on R from $\mathbb{R}^n$. Given $\mathbf{x} \in R \setminus Z_\epsilon$, there exists $\delta > 0$ such that $M(\mathbf{x}, f, \delta) - m(\mathbf{x}, f, \delta) < \epsilon$ since $o(\mathbf{x}, f) < \epsilon$. It suffices to show that $B_\delta(\mathbf{x}) \cap R \subset R \setminus Z_\epsilon$ to prove that $R \setminus Z_\epsilon$ is open.

Given $\mathbf{y} \in B_\delta(\mathbf{x}) \cap R$, choose $\delta_y < \delta - \|\mathbf{x} - \mathbf{y}\|$. We have $B_{\delta_y}(\mathbf{y}) \subset B_\delta(\mathbf{x})$. Therefore,

$$M(\mathbf{y}, f, \delta_y) - m(\mathbf{y}, f, \delta_y) < M(\mathbf{x}, f, \delta) - m(\mathbf{x}, f, \delta) < \epsilon.$$

Since $M(\mathbf{y}, f, \delta_y) - m(\mathbf{y}, f, \delta_y) \geq 0$ decreases as δ_y decreases, we have $o(\mathbf{y}, f) < \epsilon$. Thus $\mathbf{y} \in R \setminus Z_\epsilon$. ■

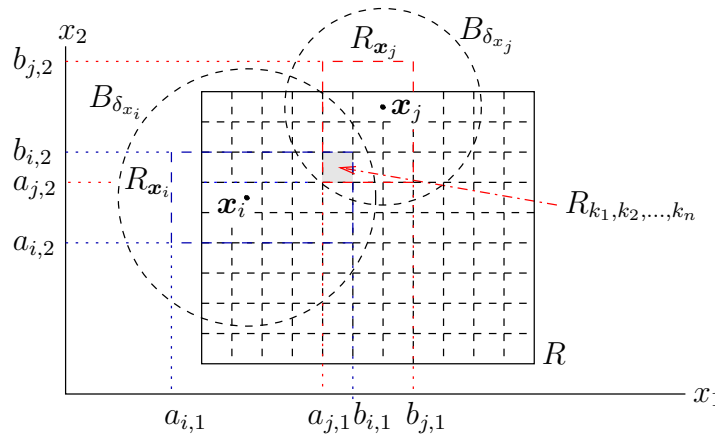
Lemma 3.1.13

Let R be a closed rectangle in $\mathbb{R}^n$ and $f : R \rightarrow \mathbb{R}$ be a bounded function such that $o(\mathbf{x}, f) < \epsilon$ for all $\mathbf{x} \in R$. Then there exists a partition P of R such that $\mathcal{U}_P(f) - \mathcal{L}_P(f) < \epsilon V(R)$.

Proof.

Since $o(\mathbf{x}, f) < \epsilon$ for all $\mathbf{x} \in R$, there exists for every $\mathbf{x} \in R$ a radius δ_x such that $M(\mathbf{x}, f, \delta_x) - m(\mathbf{x}, f, \delta_x) < \epsilon$. For each $\mathbf{x} \in R$, choose an open rectangle R_x such that $\mathbf{x} \in R_x \subset B_{\delta_x}(\mathbf{x})$. The collection $\{R_x : \mathbf{x} \in R\}$ is an open cover of the closed and bounded set R . Since R is compact, there exists a finite subcover $\{R_{\mathbf{x}_i}\}_{1 \leq i \leq N}$ of R .

Suppose that $R_{\mathbf{x}_i} = \prod_{j=1}^n]a_{i,j}, b_{i,j}[$ for $1 \leq i \leq N$. Choose a partition P of R such that $a_{i,j}, b_{i,j} \in P_i$ for all $1 \leq j \leq N_i$ and $1 \leq i \leq N$ such that $(a_{i,j}, b_{i,j}) \in R$. For each n -tuple $(k_1, k_2, \dots, k_n)$, there exists at least one $\mathbf{x}_i$ such that $R_{k_1, k_2, \dots, k_n} \subset R_{\mathbf{x}_i}$.



Therefore

$$M_{k_1, k_2, \dots, k_n} - m_{k_1, k_2, \dots, k_n} < M(\mathbf{x}_i, f, \delta_{x_i}) - m(\mathbf{x}_i, f, \delta_{x_i}) < \epsilon$$

for all n -tuples $(k_1, k_2, \dots, k_n)$. It follows that

$$\begin{aligned} \mathcal{U}_P(f) - \mathcal{L}_P(f) &= \sum_{\substack{1 \leq k_i \leq N_i \\ 1 \leq i \leq n}} (M_{k_1, k_2, \dots, k_n} - m_{k_1, k_2, \dots, k_n}) V(R_{k_1, k_2, \dots, k_n}) \\ &< \epsilon \sum_{\substack{1 \leq k_i \leq N_i \\ 1 \leq i \leq n}} V(R_{k_1, k_2, \dots, k_n}) = \epsilon V(R) . \end{aligned} \quad \blacksquare$$

Theorem 3.1.14

Let R be a rectangle in $\mathbb{R}^n$ and $f : R \rightarrow \mathbb{R}$ be a bounded function. The set of all points in R where f is discontinuous has measure zero if and only if f is Riemann integrable on R .

Proof.

Suppose that $Z \subset R$ is the set of points where f is discontinuous. Let $Z_\nu = \{\mathbf{x} \in R : o(\mathbf{x}, f) \geq \nu\}$ for $\nu > 0$.

Since f is bounded, there exists $M > 0$ such that $|f(\mathbf{x})| < M$ for all $\mathbf{x} \in R$.

$\Rightarrow$) If $Z = \emptyset$, then f is continuous on R and therefore Riemann integrable on R . Suppose that $Z \neq \emptyset$.

Given $\epsilon > 0$, we have $Z_{\epsilon/(2V(R))} \subset Z$. We may assume that ϵ is small enough to have $Z_{\epsilon/(2V(R))} \neq \emptyset$ because $Z \neq \emptyset$. Since $Z_{\epsilon/(2V(R))}$ is closed according to Lemma 3.1.12 and bounded, it is compact. Since $Z_{\epsilon/(2V(R))}$ is a compact subset of Z which has measure zero, it follows from Proposition 3.1.6 that $Z_{\epsilon/(2V(R))}$ has zero content. Thus, there exists a finite

collection $\{R_i\}_{0 \leq i \leq N}$ of closed rectangles in R such that $Z_{\epsilon/(2V(R))} \subset \bigcup_{i=0}^N R_i$ and

$$\sum_{i=0}^N V(R_i) < \frac{\epsilon}{4M} . \quad (3.1.1)$$

As in the proof of Lemma 3.1.13, we choose a partition P of R such that each subrectangle $R_{k_1, k_2, \dots, k_n}$ of the partition P with $R_{k_1, k_2, \dots, k_n} \cap Z_{\epsilon/(2V(R))} \neq \emptyset$ satisfies $R_{k_1, k_2, \dots, k_n} \subset R_i$ for some i . We group the subrectangles of the partition P into two collections: S_1 is the collection of all subrectangles $R_{k_1, k_2, \dots, k_n}$ of the partition P such that $R_{k_1, k_2, \dots, k_n} \cap Z_{\epsilon/(2V(R))} \neq \emptyset$, and S_2 is the collection of all the other subrectangles.

For each subrectangle $R_{k_1, k_2, \dots, k_n}$ of the partition P , we define as before $M_{k_1, k_2, \dots, k_n} = \sup\{f(\mathbf{x}) : \mathbf{x} \in R_{k_1, k_2, \dots, k_n}\}$ and $m_{k_1, k_2, \dots, k_n} = \inf\{f(\mathbf{x}) : \mathbf{x} \in R_{k_1, k_2, \dots, k_n}\}$. We have $M_{k_1, k_2, \dots, k_n} - m_{k_1, k_2, \dots, k_n} < 2M$ for every subrectangle $R_{k_1, k_2, \dots, k_n}$ of the partition P . Hence

$$\begin{aligned} &\sum_{R_{k_1, k_2, \dots, k_n} \in S_1} (M_{k_1, k_2, \dots, k_n} - m_{k_1, k_2, \dots, k_n}) V(R_{k_1, k_2, \dots, k_n}) \\ &\leq 2M \sum_{R_{k_1, k_2, \dots, k_n} \in S_1} V(R_{k_1, k_2, \dots, k_n}) \leq 2M \sum_{i=1}^N V(R_i) < \frac{\epsilon}{2} \end{aligned} \quad (3.1.2)$$

because of (3.1.1).

Given a subrectangle $R_{k_1, k_2, \dots, k_n} \in S_2$, we have $o(\mathbf{x}, f) < \epsilon/(2V(R))$ for all $\mathbf{x} \in R_{k_1, k_2, \dots, k_n}$ because $R_{k_1, k_2, \dots, k_n} \cap Z_{\epsilon/(2V(R))} = \emptyset$. Hence, it follows from Lemma 3.1.13 that there exists a refinement $\tilde{P}$ of the partition P such that

$$\sum_{\tilde{R}_{j_1, j_2, \dots, j_n} \subset R_{k_1, k_2, \dots, k_n}} \left(\tilde{M}_{j_1, j_2, \dots, j_n} - \tilde{m}_{j_1, j_2, \dots, j_n} \right) V(\tilde{R}_{j_1, j_2, \dots, j_n}) < \frac{\epsilon V(R_{k_1, k_2, \dots, k_n})}{2V(R)} \quad (3.1.3)$$

where the $\tilde{R}_{j_1, j_2, \dots, j_n}$ are subrectangles of the partition $\tilde{P}$ and, as expected, $\tilde{M}_{j_1, j_2, \dots, j_n} = \sup\{f(\mathbf{x}) : \mathbf{x} \in \tilde{R}_{j_1, j_2, \dots, j_n}\}$ and $\tilde{m}_{j_1, j_2, \dots, j_n} = \inf\{f(\mathbf{x}) : \mathbf{x} \in \tilde{R}_{j_1, j_2, \dots, j_n}\}$.

Repeating inductively this procedure of refinement of the partition for each $R_{k_1, k_2, \dots, k_n} \in S_2$, we end up with a refinement of P that we still call $\tilde{P}$ such that (3.1.3) is true for all $R_{k_1, k_2, \dots, k_n} \in S_2$. Note that a version of Proposition 2.1.4 in $\mathbb{R}^n$ is used to justify the last statement.

The relation (3.1.2) reminds true for the refinement $\tilde{P}$ because

$$\begin{aligned} & \sum_{\substack{\tilde{R}_{j_1, j_2, \dots, j_n} \subset R_{k_1, k_2, \dots, k_n} \\ R_{k_1, k_2, \dots, k_n} \in \tilde{S}_1}} \left(\tilde{M}_{j_1, j_2, \dots, j_n} - \tilde{m}_{j_1, j_2, \dots, j_n} \right) V(\tilde{R}_{j_1, j_2, \dots, j_n}) \\ & \leq 2M \sum_{\substack{\tilde{R}_{j_1, j_2, \dots, j_n} \subset R_{k_1, k_2, \dots, k_n} \\ R_{k_1, k_2, \dots, k_n} \in \tilde{S}_1}} V(\tilde{R}_{j_1, j_2, \dots, j_n}) = 2M \sum_{R_{k_1, k_2, \dots, k_n} \in \tilde{S}_1} V(R_{k_1, k_2, \dots, k_n}) \\ & \leq 2M \sum_{i=1}^N V(R_i) < \frac{\epsilon}{2}, \end{aligned} \quad (3.1.4)$$

where (3.1.1) was used to get the last inequality and $\tilde{R}_{j_1, j_2, \dots, j_n}$ are subrectangles of the partition $\tilde{P}$.

Finally, it follows from (3.1.3) and (3.1.4) that

$$\begin{aligned} \mathcal{U}_{\tilde{P}}(f) - \mathcal{L}_{\tilde{P}}(f) &= \sum_{\substack{\tilde{R}_{j_1, j_2, \dots, j_n} \subset R_{k_1, k_2, \dots, k_n} \\ R_{k_1, k_2, \dots, k_n} \in \tilde{S}_1}} \left(\tilde{M}_{j_1, j_2, \dots, j_n} - \tilde{m}_{j_1, j_2, \dots, j_n} \right) V(\tilde{R}_{j_1, j_2, \dots, j_n}) \\ &+ \sum_{\substack{\tilde{R}_{j_1, j_2, \dots, j_n} \subset R_{k_1, k_2, \dots, k_n} \\ R_{k_1, k_2, \dots, k_n} \in \tilde{S}_2}} \left(\tilde{M}_{j_1, j_2, \dots, j_n} - \tilde{m}_{j_1, j_2, \dots, j_n} \right) V(\tilde{R}_{j_1, j_2, \dots, j_n}) \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2V(R)} \sum_{R_{k_1, k_2, \dots, k_n} \in \tilde{S}_2} V(R_{k_1, k_2, \dots, k_n}) \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

Since ϵ is arbitrary, it follows from the version of Theorem 2.1.8 in $\mathbb{R}^n$ that f is integrable.

$\Leftarrow$) We have

$$Z = \bigcup_{m=1}^{\infty} Z_{1/m}.$$

We prove that each $Z_{1/m}$ has measure zero. Hence Z will also be of measure zero since the countable union of sets of measure zero is a set of measure zero.

Choose $\epsilon > 0$. Let P be a partition of R such that $\mathcal{U}_P(f) - \mathcal{L}_P(f) < \epsilon/(2m)$. Moreover, let S_1 be the collection of all subrectangles $R_{k_1, k_2, \dots, k_n}$ of the partition P such that $R_{k_1, k_2, \dots, k_n}^\circ \cap Z_{1/m} \neq \emptyset$. For each $R_{k_1, k_2, \dots, k_n} \in S_1$, there is at least one point of $Z_{1/m}$ which is in the interior of $R_{k_1, k_2, \dots, k_n}$. Therefore, we get $M_{k_1, k_2, \dots, k_n} - m_{k_1, k_2, \dots, k_n} \geq 1/m$ for each $R_{k_1, k_2, \dots, k_n} \in S_1$ ¹. Since

$$\begin{aligned} \frac{1}{m} \sum_{R_{k_1, k_2, \dots, k_n} \in S_1} V(R_{k_1, k_2, \dots, k_n}) &\leq \sum_{R_{k_1, k_2, \dots, k_n} \in S_1} (M_{k_1, k_2, \dots, k_n} - m_{k_1, k_2, \dots, k_n}) V(R_{k_1, k_2, \dots, k_n}) \\ &\leq \mathcal{U}_P(f) - \mathcal{L}_P(f) < \frac{\epsilon}{2m}, \end{aligned}$$

we get

$$\sum_{R_{k_1, k_2, \dots, k_n} \in S_1} V(R_{k_1, k_2, \dots, k_n}) < \frac{\epsilon}{2}. \quad (3.1.5)$$

Let $B = \bigcup_{R_{k_1, k_2, \dots, k_n}} \partial R_{k_1, k_2, \dots, k_n}$. The union is over all rectangles $R_{k_1, k_2, \dots, k_n}$ generated by the partition P of R . The set B is compact and has zero content since it is composed of a finite number of line segments. Choose a finite collection S_2 of closed rectangles such that $B \subset \bigcup_{T \in S_2} T$ and

$$\sum_{T \in S_2} V(T) < \frac{\epsilon}{2} \quad (3.1.6)$$

We have by construction that

$$Z_{1/m} \subset \bigcup_{R_{k_1, k_2, \dots, k_n} \in S_1} R_{k_1, k_2, \dots, k_n} \cup \bigcup_{T \in S_2} T.$$

Moreover, we get from (3.1.5) and (3.1.6) that

$$\sum_{R_{k_1, k_2, \dots, k_n} \in S_1} V(R_{k_1, k_2, \dots, k_n}) + \sum_{T \in S_2} V(T) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Since ϵ is arbitrary, we obtain that $Z_{1/m}$ has zero content and so is of measure zero. ■

Given $W \subset \mathbb{R}^n$, let $\mathcal{X}_W : \mathbb{R}^n \rightarrow \mathbb{R}$ be the function defined by

$$\mathcal{X}(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in W \\ 0 & \text{if } \mathbf{x} \in \mathbb{R}^n \setminus W \end{cases}$$

The only points where $\mathcal{X}_W$ is discontinuous are the points on ∂W .

¹The condition $R_{k_1, k_2, \dots, k_n} \cap Z_{1/m} \neq \emptyset$ is not sufficient to ensure that $M_{k_1, k_2, \dots, k_n} - m_{k_1, k_2, \dots, k_n} \geq 1/m$. We may have that $\mathbf{x} \in R_{k_1, k_2, \dots, k_n} \cap Z_{1/m} \subset \partial R_{k_1, k_2, \dots, k_n}$ with $M_{k_1, k_2, \dots, k_n} - m_{k_1, k_2, \dots, k_n} < 1/m$ but $o(\mathbf{x}, f) \geq 1/m$ because of the values of f in $Z_{1/m} \setminus R_{k_1, k_2, \dots, k_n}$.

Definition 3.1.15

Let $W \subset \mathbb{R}^n$ be a Jordan measurable set. Given a bounded function $f : W \rightarrow \mathbb{R}$, we define the **Riemann integral of f on W** , denoted $\int_W f$ or $\int_W f(\mathbf{x}) d\mathbf{x}$, as the integral $\int_R \chi_W f$ if this integral exists ² where R is a closed rectangle containing W . We say that f is **Riemann integrable** on W .

This definition of the integral of $f : W \rightarrow \mathbb{R}$ is independent of the rectangle R containing W since $\chi_W(\mathbf{x})f(\mathbf{x}) = 0$ for $\mathbf{x} \in R \setminus W$. Moreover, we have that ∂W is of measure zero because W is a Jordan measurable set. Hence, the set of points where $\chi_W f : R \rightarrow \mathbb{R}$ is discontinuous has measure zero if and only if the set of points where $f : W \rightarrow \mathbb{R}$ is discontinuous has measure zero for the same reasons that we have used to prove this result for functions of one variable.

Proposition 3.1.16

Let R be a rectangle in $\mathbb{R}^n$ and suppose that $Z \subset R$ has zero content. If $h : Z \rightarrow \mathbb{R}$ is a bounded function, then h is Riemann integrable on Z and $\int_Z h = 0$.

Proof.

Let $m = \inf\{h(\mathbf{x}) : \mathbf{x} \in Z\}$ and $M = \sup\{h(\mathbf{x}) : \mathbf{x} \in Z\}$.

i) We first prove that h is integrable on R . We give two proofs of this property.

The first proof uses Theorem 3.1.14. The set of points where $\chi_Z h : R \rightarrow \mathbb{R}$ is discontinuous has zero content because it is a subset of $\bar{Z}$ which has zero content ³. Hence $\bar{Z}$ has measure zero. It then follows from Theorem 3.1.14 that $\chi_Z h$ is integrable on R .

The second proof does not use Theorem 3.1.14. Given $\epsilon > 0$, there exists a finite collection of open rectangles $\{R_i\}_{0 \leq i \leq N}$ such that $Z \subset \bigcup_{i=0}^N R_i$ and $\sum_{i=0}^N V(R_i) < \frac{\epsilon}{M - m}$.

Choose a partition P of R such that the vertices of all the rectangles R_i are included in P . We then have $M_{k_1, k_2, \dots, k_n} = \sup\{\chi_Z(\mathbf{x})h(\mathbf{x}) : \mathbf{x} \in R_{k_1, k_2, \dots, k_n}\} = 0$ and $m_{k_1, k_2, \dots, k_n} = \inf\{\chi_Z(\mathbf{x})h(\mathbf{x}) : \mathbf{x} \in R_{k_1, k_2, \dots, k_n}\} = 0$ for all subrectangles $R_{k_1, k_2, \dots, k_n}$ of the partition P such

²It is understood that f is extended to R . For instance, we may set $f(\mathbf{x}) = 0$ for $\mathbf{x} \in R \setminus W$.

³Given $\epsilon > 0$, if $\{R_i\}_{0 \leq i \leq N}$ is a collection of closed rectangles such that $Z \subset \bigcup_{i=0}^N R_i$ and $\sum_{i=0}^N V(R_i) < \epsilon$, then $\bar{Z} \subset \bigcup_{i=0}^N R_i$ and we still have $\sum_{i=0}^N V(R_i) < \epsilon$.

that $R_{k_1, k_2, \dots, k_n} \cap R_i = \emptyset$ for all i . Hence

$$\begin{aligned} \mathcal{U}_P(\mathcal{X}_Z h) - \mathcal{L}_P(\mathcal{X}_Z h) &= \sum_{\substack{R_{k_1, k_2, \dots, k_n} \cap R_i \neq \emptyset \\ \text{for some } 0 \leq i \leq N}} (M_{k_1, k_2, \dots, k_n} - m_{k_1, k_2, \dots, k_n}) V(R_{k_1, k_2, \dots, k_n}) \\ &\leq (M - m) \sum_{\substack{R_{k_1, k_2, \dots, k_n} \cap R_i \neq \emptyset \\ \text{for some } 0 \leq i \leq N}} V(R_{k_1, k_2, \dots, k_n}) \leq (M - m) \sum_{i=0}^N V(R_i) < \epsilon. \end{aligned}$$

Note that the last equality comes from the fact that $R_{k_1, k_2, \dots, k_n} \cap R_i \neq \emptyset$ implies that $R_{k_1, k_2, \dots, k_n}^\circ \subset R_i$ because that R_i are open. Hence

$$\bigcup_{\substack{R_{k_1, k_2, \dots, k_n} \cap R_i \neq \emptyset \\ \text{for some } 0 \leq i \leq N}} R_{k_1, k_2, \dots, k_n}^\circ = \bigcup_{i=0}^N R_i.$$

It follows from the version of Theorem 2.1.8 in $\mathbb{R}^n$ that $\mathcal{X}_Z h$ is integrable on R .

ii) To prove that $\int_Z h = 0$, we consider $h = h^+ - h^-$ where

$$h^+(\mathbf{x}) = \begin{cases} h(\mathbf{x}) & \text{if } h(\mathbf{x}) > 0 \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad h^-(\mathbf{x}) = \begin{cases} -h(\mathbf{x}) & \text{if } h(\mathbf{x}) < 0 \\ 0 & \text{otherwise} \end{cases}$$

It follows from (i) that h^+ and h^- are integrable on R and

$$\int_Z h = \int_R \mathcal{X}_Z h = \int_R \mathcal{X}_Z h^+ - \int_R \mathcal{X}_Z h^- = \int_Z h^+ - \int_Z h^-.$$

It suffices to prove that $\int_Z h^+$ and $\int_Z h^-$ are null.

To prove that $\int_Z h^+$ is null, we prove that $\mathcal{L}_P(\mathcal{X}_Z h^+) = 0$ for all partitions P of R . Let P be a partition of R and let $\delta = \min_{\substack{0 < k_i \leq N_i \\ 1 \leq i \leq n}} V(R_{k_1, k_2, \dots, k_n})$. We have that $\delta > 0$ because there is a finite number of subrectangles. Since Z has measure zero, there exists a finite collection of open rectangles $\{R_i\}_{0 \leq i \leq N}$ such that $Z \subset \bigcup_{i=0}^N R_i$ and $\sum_{i=0}^N V(R_i) < \delta$. We have

$$m_{k_1, k_2, \dots, k_n} = \inf \{ \mathcal{X}_Z(\mathbf{x}) h^+(\mathbf{x}) : \mathbf{x} \in R_{k_1, k_2, \dots, k_n} \} = 0$$

for every subrectangle $R_{k_1, k_2, \dots, k_n}$ of the partition P because $R_{k_1, k_2, \dots, k_n} \setminus Z \supset R_{k_1, k_2, \dots, k_n} \setminus \bigcup_{i=1}^{\infty} R_i \neq \emptyset$ since

$$V\left(\bigcup_{i=1}^{\infty} R_i\right) \leq \sum_{i=1}^{\infty} V(R_i) < \delta \leq V(R_{k_1, k_2, \dots, k_n}).$$

Thus $\mathcal{L}_P(\mathcal{X}_Z h^+) = 0$. The proof that $\mathcal{L}_P(\mathcal{X}_Z h^-) = 0$ for all partitions P of R is similar. ■

Remark 3.1.17

The previous proposition is also true if Z is a closed set of measure zero. We have that the set of points where $\mathcal{X}_Z h : R \rightarrow \mathbb{R}$ is discontinuous has measure zero because it is a subset of $\overline{Z} = Z$ which has measure zero. Therefore, it follows from Theorem 3.1.14 that $\mathcal{X}_Z h$ is integrable on R . The proof that $\int_Z h = 0$ does not change.

It follows that the next corollary is also true if Z is a closed set of measure zero.

In fact, it is not hard to prove that the previous proposition and the next corollary are true if Z is a subset of a closed set of measure zero. ♠

Corollary 3.1.18

Let R be a rectangle in $\mathbb{R}^n$. Suppose that $f, g : R \rightarrow \mathbb{R}$ are two Riemann integrable functions such that the set $Z = \{\mathbf{x} \in R : f(\mathbf{x}) \neq g(\mathbf{x})\}$ has zero content. Then

$$\int_R f = \int_R g.$$
Proof.

Let $h = f - g$. We have $h = \mathcal{X}_Z h$. From the previous proposition, we get

$$\int_R f - \int_R g = \int_R (f - g) = \int_R h = \int_R \mathcal{X}_Z h = 0.$$

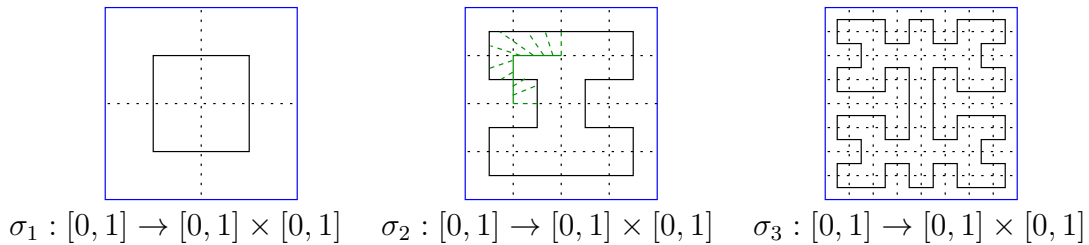
3.2 Bounded Curves in $\mathbb{R}^n$

A **curve** in $\mathbb{R}^n$ is the image of a continuous function $\sigma : [a, b] \rightarrow \mathbb{R}^n$. To simplify the discussion, we will freely refer to σ as a curve though the curve is its image. A curve σ is **closed** if $\sigma(a) = \sigma(b)$. The domain of σ could also be a semi-open or open interval.

A bounded curve in $\mathbb{R}^n$ should have zero content (measure zero), should it not? Unfortunately, the following example due to Peano shows that this is not true.

Example 3.2.1

The **Peano's curve** is the following closed curve $\sigma : [0, 1] \rightarrow [0, 1] \times [0, 1]$ that fills up the entire unit square $[0, 1] \times [0, 1]$. The curve σ is the uniform limit of the closed curves σ_i for $i \geq 1$ defined as it follows.



The curve σ_i contains the centre of all the rectangles $\left[\frac{j_1-1}{2^i}, \frac{j_1}{2^i}\right] \times \left[\frac{j_2-1}{2^i}, \frac{j_2}{2^i}\right]$ for $1 \leq j_1, j_2 \leq 2^i$ and $i \geq 1$. By choosing a proper parametric representation σ_i on $[0, 1]$ for each i , one can show that $\sup_{0 \leq x \leq 1} \|\sigma_i(x) - \sigma_{i+1}(x)\| \leq \frac{1}{2^{i+1}}$ for $i \geq 1$ as it is illustrated in green in the figure above for $i = 1$. ♣

However, if we require that the curve $\sigma : [a, b] \rightarrow \mathbb{R}^n$ be continuously differentiable, then the image of the curve σ must have zero content.

Proposition 3.2.2

Suppose that $\sigma : [a, b] \rightarrow \mathbb{R}^n$ belongs to $C^1([a, b])$. Then $\sigma([a, b])$ has zero content.

Proof.

We prove the proposition for $\sigma : [a, b] \rightarrow \mathbb{R}^2$ only. The proof is (almost) identical for $\sigma : [a, b] \rightarrow \mathbb{R}^n$.

Let $C = \sup\{\|\sigma'(t)\| : t \in [a, b]\}$. Note that $C < \infty$ because σ' is continuous on the compact set $[a, b]$. Moreover,

$$|\sigma'_i(t)| \leq \|\sigma'(t)\| = \sqrt{(\sigma'_1(t))^2 + (\sigma'_2(t))^2} \leq C$$

for all $t \in [a, b]$ and $i = 1$ or 2 .

Given $\epsilon > 0$, choose $K > 0$ large enough such that $2C^2(b-a)^2/K < \epsilon$. Let $T = \{t_i : 0 \leq i \leq K\}$ be the partition of $[a, b]$ with $t_i = a + i(b-a)/K$ for $0 \leq i \leq K$.

Given $0 < i \leq K$ and $1 \leq j \leq 2$, we have for every $s_1, s_2 \in [t_{i-1}, t_i]$ that there exists $\xi_{i,j} = \xi_{i,j}(s_1, s_2)$ between s_1 and s_2 such that

$$|\sigma_j(s_1) - \sigma_j(s_2)| = |\sigma'_j(\xi_{i,j})| |s_1 - s_2| \leq C(b-a)/K.$$

Hence

$$\|\sigma(s_1) - \sigma(s_2)\| = \sqrt{(\sigma_1(s_1) - \sigma_1(s_2))^2 + (\sigma_2(s_1) - \sigma_2(s_2))^2} \leq \sqrt{2}C(b-a)/K$$

for every s_1 and s_2 in $[t_{i-1}, t_i]$ and $0 < i \leq K$. Therefore, for each $0 < i \leq K$, we can choose a rectangle R_i with sides parallel to the axis and $A(R_i) \leq \left(\sqrt{2}C(b-a)/K\right)^2 = 2C^2(b-a)^2/K^2$

such that $\sigma([t_{i-1}, t_i]) \subset R_i$. Thus $\sigma([a, b]) \subset \bigcup_{i=1}^K R_i$ with

$$\sum_{i=1}^K A(R_i) \leq \sum_{i=1}^K \frac{2C^2(b-a)^2}{K^2} = \frac{2C^2(b-a)^2}{K} < \epsilon.$$

Since $\epsilon > 0$ is arbitrary, this proves that $\sigma([a, b])$ is of zero content. ■

3.3 Iterated Integrals

The next theorem is one of the fundamental results about the integration of functions of several variables.

Theorem 3.3.1 (Fubini)

Let I_1 and I_2 be two intervals in $\mathbb{R}$, and $f : I_1 \times I_2 \rightarrow \mathbb{R}$ be a Riemann integrable function.

1. For $x_1 \in I_1$, let $f_{x_1} : I_2 \rightarrow \mathbb{R}$ be the function defined by $f_{x_1}(x_2) = f(x_1, x_2)$ for $x_2 \in I_2$. Let $h(x_1) = \mathcal{L}(f_{x_1})$ (the lower Riemann integral of f_{x_1} on I_2) and $H(x_1) = \mathcal{U}(f_{x_1})$ (the upper Riemann integral of f_{x_1} on I_2) for $x_1 \in I_1$. Then h and H are integrable on I_1 and

$$\int_{I_1 \times I_2} f = \int_{I_1} h(x_1) dx_1 = \int_{I_1} H(x_1) dx_1 .$$

2. For $x_2 \in I_2$, let $f_{x_2} : I_1 \rightarrow \mathbb{R}$ be the function defined by $f_{x_2}(x_1) = f(x_1, x_2)$ for $x_1 \in I_1$. Let $g(x_2) = \mathcal{L}(f_{x_2})$ (the lower Riemann integral of f_{x_2} on I_1) and $G(x_2) = \mathcal{U}(f_{x_2})$ (the upper Riemann integral of f_{x_2} on I_1) for $x_2 \in I_2$. Then g and G are integrable on I_2 and

$$\int_{I_1 \times I_2} f = \int_{I_2} g(x_2) dx_2 = \int_{I_2} G(x_2) dx_2 .$$

Proof.

We prove only (1), the proof of (2) is similar.

Let $P_j = \{p_{i,j} : 0 \leq i \leq N_j\}$ be a partition of I_j for $j = 1$ and 2 . Then $P = \{P_1, P_2\}$ is a partition of $R = I_1 \times I_2$. The subrectangles $R_{i,j}$ of the partition R are of the form $R_{i,j} = [p_{i-1,1}, p_{i,1}] \times [p_{j-1,2}, p_{j,2}]$. Let $A(R_{i,j})$ be the area of the subrectangle $R_{i,j}$.

i) Let $m_{i,j} = \inf\{f(x_1, x_2) : (x_1, x_2) \in R_{i,j}\}$. We have

$$\begin{aligned} \mathcal{L}_P(f) &= \sum_{\substack{1 \leq i \leq N_1 \\ 1 \leq j \leq N_2}} m_{i,j} A(R_{i,j}) = \sum_{\substack{1 \leq i \leq N_1 \\ 1 \leq j \leq N_2}} m_{i,j} (p_{j,2} - p_{j-1,2}) (p_{i,1} - p_{i-1,1}) \\ &= \sum_{i=1}^{N_1} \left(\sum_{j=1}^{N_2} m_{i,j} (p_{j,2} - p_{j-1,2}) \right) (p_{i,1} - p_{i-1,1}) . \end{aligned} \quad (3.3.1)$$

If $m_j(x_1) = \inf\{f_{x_1}(x_2) : p_{j-1,2} \leq x_2 \leq p_{j,2}\}$ for $0 < j \leq N_2$, then $m_j(x_1) \geq m_{i,j}$ for all $x_1 \in [p_{i-1,1}, p_{i,1}]$. Hence

$$\sum_{j=1}^{N_2} m_{i,j} (p_{j,2} - p_{j-1,2}) \leq \sum_{j=1}^{N_2} m_j(x_1) (p_{j,2} - p_{j-1,2}) = \mathcal{L}_{P_2}(f_{x_1}) \leq \mathcal{L}(f_{x_1}) = h(x_1)$$

for all $x_1 \in [p_{i-1,1}, p_{i,1}]$. It follows from (3.3.1) that

$$\mathcal{L}_P(f) \leq \sum_{i=1}^{N_1} h(x_1^{[i]})(p_{i,1} - p_{i-1,1}) \quad (3.3.2)$$

for all $x_1^{[i]} \in [p_{i-1,1}, p_{i,1}]$ with $0 < i \leq N_1$. If $m_i = \inf\{h(x_1^{[i]}) : p_{i-1,1} \leq x_1^{[i]} \leq p_{i,1}\}$ for all i , then it follows from (3.3.2) that

$$\mathcal{L}_P(f) \leq \sum_{i=1}^N m_i(p_{i,1} - p_{i-1,1}) = \mathcal{L}_{P_1}(h) . \quad (3.3.3)$$

ii) Let $M_{i,j} = \sup\{f(x_1, x_2) : (x_1, x_2) \in R_{i,j}\}$. We have

$$\begin{aligned} \mathcal{U}_P(f) &= \sum_{\substack{1 \leq i \leq N_1 \\ 1 \leq j \leq N_2}} M_{i,j} A(R_{i,j}) = \sum_{\substack{1 \leq i \leq N_1 \\ 1 \leq j \leq N_2}} M_{i,j} (p_{j,2} - p_{j-1,2})(p_{i,1} - p_{i-1,1}) \\ &= \sum_{i=1}^{N_1} \left(\sum_{j=1}^{N_2} M_{i,j} (p_{j,2} - p_{j-1,2}) \right) (p_{i,1} - p_{i-1,1}) . \end{aligned} \quad (3.3.4)$$

If $M_j(x_1) = \sup\{f_{x_1}(x_2) : p_{j-1,2} \leq x_2 \leq p_{j,2}\}$ for $0 < j \leq N_2$, then $M_j(x_1) \leq M_{i,j}$ for all $x_1 \in [p_{i-1,1}, p_{i,1}]$. Hence

$$\sum_{j=1}^{N_2} M_{i,j} (p_{j,2} - p_{j-1,2}) \geq \sum_{j=1}^{N_2} M_j(x_1) (p_{j,2} - p_{j-1,2}) = \mathcal{U}_{P_2}(f_{x_1}) \geq \mathcal{U}(f_{x_1}) = H(x_1)$$

for all $x_1 \in [p_{i-1,1}, p_{i,1}]$. It follows from (3.3.4) that

$$\mathcal{U}_P(f) \geq \sum_{i=1}^{N_1} H(x_1^{[i]})(p_{i,1} - p_{i-1,1}) \quad (3.3.5)$$

for all $x_1^{[i]} \in [p_{i-1,1}, p_{i,1}]$ with $0 < i \leq N_1$. If $M_i = \sup\{H(x_1^{[i]}) : p_{i-1,1} \leq x_1^{[i]} \leq p_{i,1}\}$ for all i , then it follows from (3.3.5) that

$$\mathcal{U}_P(f) \geq \sum_{i=1}^N M_i(p_{i,1} - p_{i-1,1}) = \mathcal{U}_{P_1}(H) . \quad (3.3.6)$$

iii) We have

$$\mathcal{L}_P(f) \leq \mathcal{L}_{P_1}(h) \leq \mathcal{U}_{P_1}(h) \leq \mathcal{U}_{P_1}(H) \leq \mathcal{U}_P(f) . \quad (3.3.7)$$

The first inequality is (3.3.3), the second inequality follows from the definition of the lower and upper Riemann sums, the third inequality is true because $h(x_1) = \mathcal{L}(f_{x_1}) \leq \mathcal{U}(f_{x_1}) = H(x_1)$ for all $x_1 \in I_1$, and the last inequality follows from (3.3.6).

Choose $\epsilon > 0$. Since f is integrable on R , there exists a partition $P = (P_1, P_2)$ of R such that $\mathcal{U}_P(f) - \mathcal{L}_P(f) < \epsilon$. We get from (3.3.7) that

$$\mathcal{U}_{P_1}(h) - \mathcal{L}_{P_1}(h) \leq \mathcal{U}_P(f) - \mathcal{L}_P(f) < \epsilon .$$

Since ϵ is arbitrary, it follows from the version of Theorem 2.1.8 in $\mathbb{R}^n$ that $h : I_1 \rightarrow \mathbb{R}$ is integrable. Moreover, we get from (3.3.7) that $\mathcal{L}_P(f) \leq \mathcal{L}_{P_1}(h) \leq \mathcal{L}(h)$ for all partitions P of R . Thus $\mathcal{L}(f) \leq \mathcal{L}(h)$. Similarly, we get from (3.3.7) that $\mathcal{U}(h) \leq \mathcal{U}_{P_1}(h) \leq \mathcal{U}_P(f)$ for all partitions P of R . Thus $\mathcal{U}(h) \leq \mathcal{U}(f)$. We have shown that $\mathcal{L}(f) \leq \mathcal{L}(h) \leq \mathcal{U}(h) \leq \mathcal{U}(f)$.

Since f is integrable on R , we have $\mathcal{L}(f) = \mathcal{U}(f) = \int_R f$. Hence

$$\int_{I_1} h(x_1) dx_1 = \mathcal{L}(h) = \mathcal{U}(h) = \mathcal{L}(f) = \mathcal{U}(f) = \int_R f .$$

iv) Similarly to (iii), we have

$$\mathcal{L}_P(f) \leq \mathcal{L}_{P_1}(h) \leq \mathcal{L}_{P_1}(H) \leq \mathcal{U}_{P_1}(H) \leq \mathcal{U}_P(f) . \quad (3.3.8)$$

The first inequality is (3.3.3), the second inequality is true because $h(x_1) = \mathcal{L}(f_{x_1}) \leq \mathcal{U}(f_{x_1}) = H(x_1)$ for all $x_1 \in I_1$, the third inequality follows from the definition of the lower and upper Riemann sums, and the last inequality follows from (3.3.6).

Choose $\epsilon > 0$. Since f is integrable on R , there exists a partition $P = (P_1, P_2)$ of R such that $\mathcal{U}_P(f) - \mathcal{L}_P(f) < \epsilon$. We get from (3.3.8) that

$$\mathcal{U}_{P_1}(H) - \mathcal{L}_{P_1}(H) \leq \mathcal{U}_P(f) - \mathcal{L}_P(f) < \epsilon .$$

Since ϵ is arbitrary, it follows from the version of Theorem 2.1.8 in $\mathbb{R}^n$ that $H : I_1 \rightarrow \mathbb{R}$ is integrable. Moreover, we get from (3.3.8) that $\mathcal{U}_P(f) \geq \mathcal{U}_{P_1}(H) \geq \mathcal{U}(H)$ for all partitions P of R . Thus $\mathcal{U}(f) \geq \mathcal{U}(H)$. Similarly, we get from (3.3.8) that $\mathcal{L}(H) \geq \mathcal{L}_{P_1}(H) \geq \mathcal{L}_P(f)$ for all partitions P of R . Thus $\mathcal{L}(H) \geq \mathcal{L}(f)$. We have shown that $\mathcal{L}(f) \leq \mathcal{L}(H) \leq \mathcal{U}(H) \leq \mathcal{U}(f)$.

Since f is integrable on R , we have $\mathcal{L}(f) = \mathcal{U}(f) = \int_R f$. Hence

$$\int_{I_1} H(x_1) dx_1 = \mathcal{L}(H) = \mathcal{U}(H) = \mathcal{L}(f) = \mathcal{U}(f) = \int_R f . \quad \blacksquare$$

Remark 3.3.2

1. If $f : I_1 \times I_2 \rightarrow \mathbb{R}$ is Riemann integrable and f_{x_1} is Riemann integrable for all x_1 , then $h(x_1) = H(x_1) = \int_{I_2} f(x_1, x_2) dx_2$ and we get from Fubini's Theorem that

$$\int_{I_1 \times I_2} f = \int_{I_1} \left(\int_{I_2} f(x_1, x_2) dx_2 \right) dx_1 .$$

This is called an **iterated integral**.

2. Similarly to (1), if $f : I_1 \times I_2 \rightarrow \mathbb{R}$ is Riemann integrable and f_{x_2} is Riemann integrable for all x_2 , then $g(x_2) = G(x_2) = \int_{I_1} f(x_1, x_2) dx_1$ and we get from Fubini's Theorem that

$$\int_{I_1 \times I_2} f = \int_{I_2} \left(\int_{I_1} f(x_1, x_2) dx_1 \right) dx_2 .$$

This is another **iterated integral**.

3. (1) and (2) are satisfied if $f : I_1 \times I_2 \rightarrow \mathbb{R}$ is continuous because $f_{x_1} : I_2 \rightarrow \mathbb{R}$ and $f_{x_2} : I_1 \rightarrow \mathbb{R}$ are then also continuous for all x_1 and x_2 . It follows from Theorem 3.1.14 that f is Riemann integrable on $I_1 \times I_2$, and from Theorem 2.1.17 that f_{x_1} and f_{x_2} are Riemann integrable on I_2 and I_1 respectively for all x_1 and x_2 .
4. There are cases where (1) and (2) above cannot be used but Fubini's Theorem can be used. The following example is from [26].

Consider the function $f : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ defined by

$$f(x_1, x_2) = \begin{cases} 1 - \frac{1}{q} & \text{if } x_2 \in [0, 1] \cap \mathbb{Q} \text{ and } x_1 = \frac{p}{q} \text{ where} \\ & 0 < p \leq q \text{ are relatively prime.} \\ 1 & \text{otherwise} \end{cases}$$

- i) We first prove that f is Riemann integrable on $S = [0, 1] \times [0, 1]$ and $\int_S f = 1$. Given $\epsilon > 0$, we choose $q_0 \in \mathbb{N}$ large enough such that $1/q_0 < \epsilon/2$. We also choose a partition $P = \{P_1, P_2\}$ of S where $P_j = \{p_{i,j} : 0 \leq i \leq N_j\}$ for $1 \leq j \leq 2$ and

$$\frac{q_0(q_0 + 1)}{2} \max_{1 \leq i \leq N_1} (p_{i,1} - p_{i-1,1}) < \frac{\epsilon}{2} .$$

Note that there are less than $q_0(q_0 + 1)/2$ values of x of the form p/q with $0 < p \leq q \leq q_0$ relatively prime (See footnote page 43). Let $\{a_k : 1 \leq k \leq K\}$ be these values.

As usual, let $R_{i,j} = [p_{i-1,1}, p_{i,1}] \times [p_{j-1,2}, p_{j,2}]$ for $1 \leq i \leq N_1$ and $1 \leq j \leq N_2$, $m_{i,j} = \inf\{f(x_1, x_2) : (x_1, x_2) \in R_{i,j}\}$ and $M_{i,j} = \sup\{f(x_1, x_2) : (x_1, x_2) \in R_{i,j}\}$. We have $M_{i,j} - m_{i,j} \leq 1$ for all i and j . Moreover, $M_{i,j} - m_{i,j} \leq 1/q \leq 1/q_0$ if $a_k \notin [p_{i-1,1}, p_{i,1}]$ for all k . Hence

$$\begin{aligned} \mathcal{U}_P(f) - \mathcal{L}_P(f) &= \sum_{i,j} (M_{i,j} - m_{i,j})(p_{i,1} - p_{i-1,1})(p_{j,2} - p_{j-1,2}) \\ &\leq \sum' \frac{1}{q_0} (p_{i,1} - p_{i-1,1})(p_{j,1} - p_{j-1,1}) + \sum'' (p_{i,1} - p_{i-1,1})(p_{j,2} - p_{j-1,2}) , \end{aligned} \quad (3.3.9)$$

where the first sum is over all subrectangles $R_{i,j}$ such that $a_k \notin [x_{i-1,1}, x_{i,1}]$ for all k , and the second sum is over all subrectangles $R_{i,j}$ such that $a_k \in [x_{i-1,1}, x_{i,1}]$ for some k .

We have $\sum' \frac{1}{q_0} (p_{i,1} - p_{i-1,1})(p_{j,2} - p_{j-1,2}) < \frac{\epsilon}{2}$ because $\frac{1}{q_0} < \frac{\epsilon}{2}$ and

$$\sum' (p_{i,1} - p_{i-1,1})(p_{j,2} - p_{j-1,2}) \leq \sum_{\substack{0 < i \leq N_1 \\ 0 < j \leq N_2}} (p_{i,1} - p_{i-1,1})(p_{j,2} - p_{j-1,2}) = V(S) = 1 ,$$

the area of the rectangle S . Moreover, since there are at most $q_0(q_0 + 1)/2$ intervals $[p_{i-1,1}, p_{i,1}]$ containing some a_k , we also have

$$\begin{aligned} \sum'' (p_{i,1} - p_{i-1,1})(p_{j,2} - p_{j-1,2}) &\leq \frac{q_0(q_0 + 1)}{2} \max_{1 \leq i \leq N_1} (p_{i,1} - p_{i-1,1}) \underbrace{\sum_{j=1}^{N_2} (p_{j,2} - p_{j-1,2})}_{=1} \\ &= \frac{q_0(q_0 + 1)}{2} \max_{1 \leq i \leq N_1} (p_{i,1} - p_{i-1,1}) < \frac{\epsilon}{2} . \end{aligned}$$

We get from (3.3.9) that $\mathcal{U}_P(f) - \mathcal{L}_P(f) < \epsilon$. Since ϵ is arbitrary, it follows that f is integrable on S according to the version of Theorem 2.1.8 in $\mathbb{R}^n$. Since $\mathcal{U}_P(f) = 1$ for all partition P of S , we have $\int_S f = \mathcal{U}(f) = 1$.

ii) However, $\int_0^1 f_{x_1}(x_2) dx_2 = 1$ for $x_1 \in \{0\} \cup [0, 1] \setminus \mathbb{Q}$ but f_{x_1} is not Riemann integrable for $x_1 \in \mathbb{Q} \cap]0, 1]$. To prove that f_{x_1} is not Riemann integrable for $x_1 \in \mathbb{Q} \cap]0, 1]$, suppose that $x_1 = p/q$ where $0 < p \leq q$ are relatively prime. Then

$$f_{x_1}(x_2) = \begin{cases} 1 & \text{if } x_2 \in [0, 1] \setminus \mathbb{Q} \\ 1 - \frac{1}{q} & \text{if } x_2 \in \mathbb{Q} \cap [0, 1] \end{cases}$$

This function is not Riemann integrable according to Theorem 2.1.17 because f_{x_1} is discontinuous at all points of $[0, 1]$ which is not a set of measure zero. Since $\mathbb{Q} \cap]0, 1]$ does not have measure zero, $x_1 \mapsto \int_0^1 f(x_1, x_2) dx_2$ cannot define a Riemann integrable function on $[0, 1]$.

iii) We cannot use (1) above to define an iterated integral for $\int_S f$. However, Fubini's Theorem can be used. We have

$$h(x_1) = \begin{cases} 1 - \frac{1}{q} & \text{if } x_1 = \frac{p}{q} \text{ where } 0 < p \leq q \text{ are relatively prime.} \\ 1 & \text{if } x_1 \in \{0\} \cup [0, 1] \setminus \mathbb{Q} \end{cases}$$

and $\int_0^1 h(x_1) dx_1 = 1$. The proof that this last integral is equal to 1 is similar to (in fact, is a stripped version of) the proof that $\int_S f = 1$ above.

5. It is possible for both iterated integrals in (1) and (2) to exist, and even be equal, but that the function $f : I_1 \times I_2 \rightarrow \mathbb{R}$ not be Riemann integrable.

To illustrate this situation, we first construct a set $Q \subset S = I_1 \times I_2$ where $I_1 = I_2 = I = [0, 1]$ such that there is at most one point of Q on each horizontal and vertical lines in S , and such that $\partial Q = S$. Let

$$I_{i,j,k} = \left[\frac{j-1}{2^i}, \frac{j}{2^i} \right] \times \left[\frac{k-1}{2^i}, \frac{k}{2^i} \right]$$

for $0 < j, k \leq 2^i$. For each $i > 0$, we have $S = \bigcup_{0 < j, k \leq 2^i} I_{i,j,k}$. For $i = 1$, the set S is the union of four squares; namely,

$$S = \bigcup_{0 < j, k \leq 2} I_{1,j,k} = \left[0, \frac{1}{2} \right] \times \left[0, \frac{1}{2} \right] \cup \dots \cup \left[\frac{1}{2}, 1 \right] \times \left[\frac{1}{2}, 1 \right].$$

Let Q_1 be the set containing one point from each square such that each of these points is the unique point of Q_1 on the horizontal and vertical lines on which it belongs. For $i = 2$, the set S is the union of sixteen squares. Let Q_2 be the set containing one point from each square such that each of these points is the unique point of $Q_1 \cup Q_2$ on the horizontal and vertical lines on which it belongs. We have by construction that $Q_1 \cap Q_2 = \emptyset$. Suppose that we have constructed $Q_1, Q_2, \dots, Q_{s-1}$. For $i = s$, the set S is the union of 2^{2s} squares. Let Q_s be the set containing one point from each square such that each of these points is the unique point of $\bigcup_{i=1}^s Q_i$ on the horizontal and

vertical lines on which it belongs. We have by construction that $Q_s \cap \left(\bigcup_{i=1}^{s-1} Q_i \right) = \emptyset$.

We have an inductive method to define all the sets Q_s .

The set Q is the set $Q = \bigcup_{s=1}^{\infty} Q_s$. We have that $\overline{Q} = S$ because every open subset of S contains squares of the form $I_{i,j,k}$ for i large enough and each of these squares contains a point of Q . We also have that $Q^\circ = \emptyset$ because there is at most one point of Q on each horizontal and vertical lines. Therefore, every open subset contains points that are not in Q . Thus $\partial Q = S$.

The function χ_Q is not Riemann integrable on S because the set of points where χ_Q is discontinuous is $\partial Q = S$ which is not of measure zero. However $x_2 \mapsto \chi_Q(x_1, x_2)$ is Riemann integrable on I for all $x_1 \in I$ because $\chi_Q(x_1, x_2) = 0$ for all $x_2 \in I$ but at most one point. Since $\int_I \chi_Q(x_1, x_2) dx_2 = 0$ for all $x_1 \in I$, we get $\int_I \left(\int_I \chi_Q(x_1, x_2) dx_2 \right) dx_1 = 0$. Similarly, we get $\int_I \left(\int_I \chi_Q(x_1, x_2) dx_1 \right) dx_2 = 0$.



3.4 Partition of Unity

Theorem 3.4.1

Let $\{U_\alpha\}_{\alpha \in A}$ be an open cover of $W \subset \mathbb{R}^n$; namely, U_α is an open set for every α and $W \subset \bigcup_{\alpha \in A} U_\alpha$. There exist an open set $V \supset W$ and a collection $\{\phi_j\}_{j \in \mathbb{N}^+}$ of functions in $C^\infty(V)$ such that:

1. $0 \leq \phi_j(\mathbf{x}) \leq 1$ for all $\mathbf{x} \in V$ and all $j \in \mathbb{N}^+$.
2. For each $\mathbf{x} \in W$, there exists an open neighbourhood $W_{\mathbf{x}} \subset V$ of $\mathbf{x}$ such that $\phi_j|_{W_{\mathbf{x}}} = 0$ for all but a finite number of ϕ_j .
3. For each $\mathbf{x} \in W$, we have $\sum_{j \in \mathbb{N}^+} \phi_j(\mathbf{x}) = 1$. This sum is finite according to (2).
4. For each $j \in \mathbb{N}^+$, we have that $\text{supp } \phi_j$ is a compact set and $\text{supp } \phi_j \subset U_\alpha$ for some $\alpha \in A$.

Proof.

i) If W is compact, then there exists a finite subcover $\{U_{\alpha_i} : 1 \leq i \leq N\}$ of W . For each $\mathbf{x} \in W$, there exist $0 < r_{\mathbf{x},1} < r_{\mathbf{x},2}$ such that

$$B_{r_{\mathbf{x},1}}(\mathbf{x}) \subset B_{r_{\mathbf{x},2}}(\mathbf{x}) \subset \overline{B_{r_{\mathbf{x},2}}(\mathbf{x})} \subset U_{\alpha_i}$$

for $\alpha_i \in A$. Since $\{B_{r_{\mathbf{x},1}}(\mathbf{x})\}_{\mathbf{x} \in W}$ is an open cover of the compact set W , there exists a finite subcover $\{B_{r_{\mathbf{x}_k,1}}(\mathbf{x}_k) : 1 \leq k \leq M\}$ of W . For each $1 \leq k \leq M$, let

$$\psi_k(\mathbf{x}) = \begin{cases} 1 & \text{if } \|\mathbf{x} - \mathbf{x}_k\| \leq r_{\mathbf{x}_k,1} \\ e^{1/(r_{\mathbf{x}_k,2}^2 - r_{\mathbf{x}_k,1}^2)} e^{-1/(r_{\mathbf{x}_k,2}^2 - \|\mathbf{x} - \mathbf{x}_k\|^2)} \left(1 - e^{-1/(\|\mathbf{x} - \mathbf{x}_k\|^2 - r_{\mathbf{x}_k,1}^2)}\right) & \text{if } r_{\mathbf{x}_k,1} < \|\mathbf{x} - \mathbf{x}_k\| < r_{\mathbf{x}_k,2} \\ 0 & \text{if } \|\mathbf{x} - \mathbf{x}_k\| \geq r_{\mathbf{x}_k,2} \end{cases}$$

for $\mathbf{x} \in \mathbb{R}^n$. We have $\psi_k \in C^\infty(\mathbb{R}^n)$, $0 \leq \psi_k(\mathbf{x}) \leq 1$ for all $\mathbf{x} \in \mathbb{R}^n$, $\psi_k(\mathbf{x}) = 1$ for all $\mathbf{x} \in B_{r_{\mathbf{x}_k,1}}(\mathbf{x}_k)$, and $\text{supp } \psi_k \subset \overline{B_{r_{\mathbf{x}_k,2}}(\mathbf{x}_k)} \subset U_{\alpha_i}$ for some $\alpha_i \in A$.

Let $\phi_1 = \psi_1$ and $\phi_{j+1} = \psi_{j+1} \prod_{k=1}^j (1 - \psi_k)$ for $1 < j < M$.

1. We have $0 \leq \phi_j(\mathbf{x}) \leq 1$ for all $\mathbf{x} \in \mathbb{R}^n$ and $1 \leq j \leq M$ because ϕ_j is the product of functions whose image is in $[0, 1]$.
2. For each $\mathbf{x} \in W$, it is obvious that there exists an open neighbourhood $W_{\mathbf{x}} \subset \mathbb{R}^n$ of $\mathbf{x}$ such that $\phi_j|_{W_{\mathbf{x}}} = 0$ for all but a finite number of ϕ_j because there is only a finite number of ϕ_j .

3. To prove that $\sum_{j=1}^M \phi_j(\mathbf{x}) = 1$ for all $\mathbf{x} \in W$, we first prove by induction that

$$\sum_{j=1}^m \phi_j = 1 - \prod_{j=1}^m (1 - \psi_j) \quad (3.4.1)$$

for $1 \leq m \leq M$. We obviously have that (3.4.1) is true for $m = 1$. Assume that (3.4.1) is true for $m < M$. Using the hypothesis of induction, we get

$$\begin{aligned} \sum_{j=1}^{m+1} \phi_j &= \sum_{j=1}^m \phi_j + \phi_{m+1} = 1 - \prod_{j=1}^m (1 - \psi_j) + \psi_{m+1} \prod_{j=1}^m (1 - \psi_j) \\ &= 1 - \prod_{j=1}^m (1 - \psi_j)(1 - \psi_{m+1}) = 1 - \prod_{j=1}^{m+1} (1 - \psi_j). \end{aligned}$$

Thus (3.4.1) is true with m replaced by $m + 1$. This completes the proof by induction. It follows from (3.4.1) that

$$\sum_{j=1}^M \phi_j(\mathbf{x}) = 1 - \prod_{j=1}^M (1 - \psi_j(\mathbf{x})) = 1$$

for all $\mathbf{x} \in W$ because every $\mathbf{x} \in W$ is element of $B_{r_{\mathbf{x}_k,1}}(\mathbf{x}_k)$ for at least one k and $\psi_k(\mathbf{y}) = 1$ for all $\mathbf{y} \in B_{r_{\mathbf{x}_k,1}}(\mathbf{x}_k)$. Therefore, $\sum_{j=1}^M \phi_j(\mathbf{x}) = 1$ for all $\mathbf{x} \in W$.

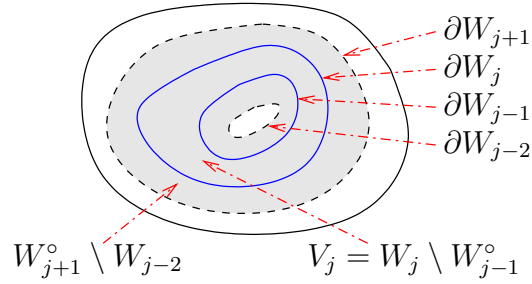
4. Since $\text{supp } \phi_j \subset \text{supp } \psi_j \subset \overline{B_{r_{\mathbf{x}_j,2}}(\mathbf{x}_j)}$, we have that $\text{supp } \phi_j$ is a compact set because it is closed and bounded. We also have $\text{supp } \phi_j \subset U_{\alpha_j}$.

All the conclusions of the theorem are satisfied by the collection $\{\phi_j\}_{1 \leq j \leq M}$.

ii) Suppose that $W = \bigcup_{j=0}^{\infty} W_j$ where each W_j is compact and $W_j \subset W_{j+1}^\circ$ for all j . We also assume that $W_0 = \emptyset$ to simplify the discussion. Let $V_1 = W_1$ and $V_j = W_j \setminus W_{j-1}^\circ$ for $j \geq 2$.

Let $U_{\alpha,1} = U_\alpha \cap W_2^\circ$ for all $\alpha \in A$. Since V_1 is a compact set and $\{U_{\alpha,1}\}_{\alpha \in A}$ is an open cover of V_1 , we may use (i) to find $\{\psi_{k,1} : 1 \leq k \leq N_1\}$ such that $0 \leq \psi_{k,1}(\mathbf{x}) \leq 1$ for all $\mathbf{x} \in \mathbb{R}^n$ and $1 \leq k \leq N_1$, $\sum_{k=1}^{N_1} \psi_{k,1}(\mathbf{x}) = 1$ for all $\mathbf{x} \in V_1$, and $\text{supp } \psi_{k,1}$ is a compact set with $\text{supp } \psi_{k,1} \subset U_{\alpha_{k,1}}$ for some $\alpha_{k,1} \in A$.

For $j \geq 2$, let $U_{\alpha,j} = U_\alpha \cap (W_{j+1}^\circ \setminus W_{j-2})$ for all $\alpha \in A$. Since V_j is a compact set and $\{U_{\alpha,j}\}_{\alpha \in A}$ is an open cover of V_j , we may use (i) to find $\{\psi_{k,j} : 1 \leq k \leq N_j\}$ such that $0 \leq \psi_{k,j}(\mathbf{x}) \leq 1$ for all $\mathbf{x} \in \mathbb{R}^n$ and $1 \leq k \leq N_j$, $\sum_{k=1}^{N_j} \psi_{k,j}(\mathbf{x}) = 1$ for all $\mathbf{x} \in V_j$, and $\text{supp } \psi_{k,j}$ is a compact set with $\text{supp } \psi_{k,j} \subset U_{\alpha_{k,j}}$ for some $\alpha_{k,j} \in A$.



Let $\psi(\mathbf{x}) = \sum_{j=1}^{\infty} \sum_{k=1}^{N_j} \psi_{k,j}(\mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^n$. We first note that the sum is finite for each $\mathbf{x}$ because $\psi_{k,j}(\mathbf{x}) = 0$ if $\mathbf{x} \notin W_{j+1}^\circ \setminus W_{j-2}$ by construction. Moreover, if $\mathbf{x} \in W$, then $\mathbf{x} \in V_q$ for some $q \geq 1$ because $\mathbf{x} \in W = \bigcup_{j=1}^{\infty} V_j$. Hence

$$\psi(\mathbf{x}) = \sum_{\substack{j \text{ such that} \\ \mathbf{x} \in W_{j+1}^\circ \setminus W_{j-2}}} \sum_{k=1}^{N_j} \psi_{k,j}(\mathbf{x}) \geq \sum_{k=1}^{N_q} \psi_{k,q}(\mathbf{x}) = 1.$$

Choose an open set $V \supset W$ such that $\psi(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in V$ ⁴. Let

$$\rho_{k,j}(\mathbf{x}) = \frac{\psi_{k,j}(\mathbf{x})}{\psi(\mathbf{x})}$$

for $\mathbf{x} \in V$, $j \geq 1$ and $1 \leq k \leq N_j$. We have that $\rho_{k,j} \in C^\infty(V)$ because $\psi_{k,j}$ and ψ are in $C^\infty(V)$, and $\psi(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in V$.

1. We have by construction that $0 \leq \rho_{k,j}(\mathbf{x}) \leq 1$ for all $\mathbf{x} \in V$, $j \geq 1$ and $1 \leq k \leq N_j$.
2. There exists an open neighbourhood $W_{\mathbf{x}} \subset V$ of $\mathbf{x} \in W$ such that $\rho_{k,j}|_{W_{\mathbf{x}}} = 0$ for all but a finite number of the $\rho_{k,j}$. More precisely, given $\mathbf{x} \in W$, we have $\mathbf{x} \in V_j$ for some j . Since $\rho_{k,q}(\mathbf{y}) = 0$ for all $\mathbf{y} \in W_{j+1}^\circ \setminus W_{j-2}$ if $q < j-2$ and $q > j+1$, it suffices to choose an open neighbourhood $W_{\mathbf{x}} \subset W_{j+1}^\circ \setminus W_{j-2}$ of $\mathbf{x}$.
3. We also have by construction that $\sum_{j=1}^{\infty} \sum_{k=1}^{N_j} \rho_{k,j}(\mathbf{x}) = 1$ for all $\mathbf{x} \in W$.
4. Since $\text{supp } \rho_{k,j} = \text{supp } \psi_{k,j}$, we have that $\rho_{k,j}$ has a compact support because $\psi_{k,j}$ has a compact support. Moreover, $\text{supp } \rho_{k,j} = \text{supp } \psi_{k,j} \subset U_{\alpha_{k,j}}$ for some $\alpha_{k,j} \in A$ for each $j \geq 1$ and $1 \leq k \leq N_j$.

⁴Since $\psi(\mathbf{x}) = 1$ for all $\mathbf{x} \in W$ and ψ is continuous, we can find for each $\mathbf{x} \in W$ an open neighbourhood $V_{\mathbf{x}}$ of $\mathbf{x}$ such that $\psi(\mathbf{y}) > 1/2$ for all $\mathbf{y} \in V_{\mathbf{x}}$. We may take $V = \bigcup_{\mathbf{x} \in W} V_{\mathbf{x}}$.

The collection $\{\phi_j\}_{j \in \mathbb{N}^+}$ that satisfies all the conclusions of the theorem is defined as it follows. $\phi_j = \rho_{j,1}$ for $1 \leq j \leq N_1$ and $\phi_j = \rho_{k,q}$ for $j = \sum_{i=1}^{q-1} N_i + k$ and $1 \leq k \leq N_q$ with $q > 1$.

iii) If W is an open set, then $W = \bigcup_{j=1}^{\infty} W_j$ where $W_j = \{\mathbf{x} : \text{dist}(\mathbf{x}, \mathbb{R}^n \setminus W) \geq 1/j\} \cap \{\mathbf{x} : \|\mathbf{x}\| \leq j\}$. The sets W_j are compact and $W_j \subset W_{j+1}^\circ$ for all $j \geq 1$. We may therefore use (ii). If W is not open, then we use (ii) with the open set $\bigcup_{\alpha \in A} U_\alpha \supset W$ instead of W . ■

Remark 3.4.2

In the previous theorem, if $\{U_i\}_{1 \leq i \leq N}$ is a finite open cover of a compact set $W \subset \mathbb{R}^n$, then the first part of the proof proves that we can choose a finite partition of unity $\{\phi_i\}_{1 \leq i \leq N}$ subordinate to the open cover $\{U_i\}_{1 \leq i \leq N}$ of W such that $\text{supp } \phi_i \subset U_i$ for $1 \leq i \leq N$. We may have to add together the ϕ_j 's given in the proof that have their support in the same U_i to get a new ϕ_i , and to reorder the ϕ_i . ♠

Definition 3.4.3

The collection $\{\phi_j\}_{j \in \mathbb{N}^+}$ of the previous theorem is called a **partition of unity** subordinated to the open cover $\{U_\alpha\}_{\alpha \in A}$ of W .

Proposition 3.4.4

Suppose that $\{\phi_j\}_{j \in \mathbb{N}^+}$ is a partition of unity subordinated to the open cover $\{U_\alpha\}_{\alpha \in A}$ of W . If $K \subset W$ is a compact set, then there exist only a finite number of functions ϕ_j such that $\phi_j|_K \neq 0$.

Proof.

For each $\mathbf{x} \in K$, let $W_{\mathbf{x}}$ be the open set given in (2) of Theorem 3.4.1. Then $\{W_{\mathbf{x}}\}_{\mathbf{x} \in K}$ is an open cover of the compact set K . Therefore, there exists a finite subcover $\{W_{\mathbf{x}_j} : 1 \leq j \leq N\}$ of K .

A function ϕ_j such that $\phi_j|_K \neq 0$ must also be one of the functions such that $\phi_j|_{W_{\mathbf{x}_j}} \neq 0$ for some j . There are only a finite number of them according to (2) of Theorem 3.4.1. ■

Definition 3.4.5

An open cover $\{U_\alpha\}_{\alpha \in A}$ of an open set W is called **admissible** if $U_\alpha \subset W$ for all α .

3.5 General Definition of the Integral

Definition 3.5.1

Let W be a Jordan measurable subset of $\mathbb{R}^n$. In this and the next sections, we denote the Riemann integral of a Riemann integrable function $h : W \rightarrow \mathbb{R}$ by $\int_W h$.

The reason for this new notation is that we will shortly generalize the definition of integrable functions. One of our goal is to generalize to $\mathbb{R}^n$ the concept of improper integral for functions of one variable that we have seen in the previous chapter.

Lemma 3.5.2

Let W be an open subset of $\mathbb{R}^n$ and $f : W \rightarrow \mathbb{R}$ be a function which is bounded on bounded open subsets of W and continuous on W except possibly on a set of measure zero. If $\psi : W \rightarrow \mathbb{R}$ is a continuous function with compact support in W , then $\psi f : W \rightarrow \mathbb{R}$ is Riemann integrable on W .

Proof.

Since $\text{supp } \psi f \subset \text{supp } \psi$ and $\text{supp } \psi$ is a compact subset of W , we get that $\text{supp } \psi f \subset W$ is compact. Thus $\text{supp } \psi f \subset W$ is bounded and so it is contained in a bounded open subset of W . Moreover, there exists a bounded rectangle $R \subset \mathbb{R}^n$ such that $\text{supp } \psi f \subset R$. We may assume that $\psi(\mathbf{x}) = f(\mathbf{x})$ for $\mathbf{x} \in R \setminus W$.

The set of points where ψf is discontinuous is a subset of the set of points where f is discontinuous. Thus the set of points where ψf is discontinuous is of measure zero. We also have that ψf is bounded on R because f is bounded on bounded open subsets of W by hypothesis and ψ is bounded because it is a continuous function with compact support.

Therefore, we have from Theorem 3.1.14 that ψf is Riemann integrable on R . Moreover,
$$\int_W \psi f = \int_R \psi f. \quad \blacksquare$$

Proposition 3.5.3

Let W be an open subset of $\mathbb{R}^n$ and $f : W \rightarrow \mathbb{R}$ be a function which is bounded on bounded open subsets of W and continuous on W except possibly on a set of measure zero. Suppose that $\{\phi_j\}_{j \in \mathbb{N}^+}$ and $\{\tilde{\phi}_k\}_{k \in \mathbb{N}^+}$ are two partitions of unity subordinated respectively to the admissible open covers $\{U_\alpha\}_{\alpha \in A}$ and $\{\tilde{U}_{\tilde{\alpha}}\}_{\tilde{\alpha} \in \tilde{A}}$ of the open set W . Then $\sum_{j \in \mathbb{N}^+} \int_W \phi_j |f|$ converges if and only if $\sum_{k \in \mathbb{N}^+} \int_W \tilde{\phi}_k |f|$ converges. Moreover, the two sums are equal when they converges.

Proof.

According to the previous lemma, all the functions $\phi_j |f|$, $\tilde{\phi}_k |f|$ and $\phi_j \tilde{\phi}_k |f|$ are Riemann

integrable on W .

Let $K_j = \text{supp } \phi_j|f|$ for $j \in \mathbb{N}^+$. The set K_j is compact according to (4) of Theorem 3.4.1.

We then have from Proposition 3.4.4 that there are only a finite number of functions $\tilde{\phi}_k$ such that $\tilde{\phi}_k|_{K_j} \neq 0$. It follows that the sum $\sum_{k \in \mathbb{N}^+} \tilde{\phi}_k \phi_j|f|$ is a finite sum for each j . A similar argument proves that $\sum_{j \in \mathbb{N}^+} \phi_j \tilde{\phi}_k|f|$ is a finite sum for each k . Hence,

$$\begin{aligned} \sum_{j \in \mathbb{N}^+} \int_W \phi_j|f| &= \sum_{j \in \mathbb{N}^+} \int_W \left(\sum_{k \in \mathbb{N}^+} \tilde{\phi}_k \right) \phi_j|f| = \sum_{j \in \mathbb{N}^+} \sum_{k \in \mathbb{N}^+} \int_W \tilde{\phi}_k \phi_j|f| \\ &= \sum_{k \in \mathbb{N}^+} \sum_{j \in \mathbb{N}^+} \int_W \tilde{\phi}_k \phi_j|f| = \sum_{k \in \mathbb{N}^+} \int_W \left(\sum_{j \in \mathbb{N}^+} \phi_j \right) \tilde{\phi}_k|f| \\ &= \sum_{k \in \mathbb{N}^+} \int_W \tilde{\phi}_k|f|. \end{aligned} \quad (3.5.1)$$

The first equality comes from $\sum_{k \in \mathbb{N}^+} \tilde{\phi}_k(\mathbf{x}) = 1$ for all $\mathbf{x} \in W$. Since the summation with respect to k is finite on $\text{supp } \phi_j$, we may pull the summation with respect to k outside the integral to get the second equality. The third equality is true because we may change the order of summation of positive series. Since the summation with respect to j is finite on $\text{supp } \tilde{\phi}_k$, we may bring the summation with respect to j inside the integral to get the fourth equality. The last equality comes from $\sum_{j \in \mathbb{N}^+} \phi_j(\mathbf{x}) = 1$ for all $\mathbf{x} \in W$. ■

The previous proposition is used to justify the following definition.

Definition 3.5.4

Suppose that $\{\phi_j\}_{j \in \mathbb{N}^+}$ is a partition of unity subordinated to an admissible open cover $\{U_\alpha\}_{\alpha \in A}$ of the open set W . Let $f : W \rightarrow \mathbb{R}$ be a function which is bounded on bounded open subsets of W and continuous on W except possibly on a set of measure zero. We say that f is **integrable** on W if $\sum_{j \in \mathbb{N}^+} \int_W \phi_j|f|$ converges.

As a consequence of the previous proposition, this definition is independent of the admissible open cover and partition of unity.

In this section and the next section, when we refer to an integrable function, we will mean an integrable function in the new sense given in Definition 3.5.4. To refer to a function integrable in the sense of Riemann, we will say that the function is Riemann integrable.

Proposition 3.5.5

Suppose that $\{\phi_j\}_{j \in \mathbb{N}^+}$ and $\{\tilde{\phi}_j\}_{j \in \mathbb{N}^+}$ are two partitions of unity subordinated respectively to the admissible open cover $\{U_\alpha\}_{\alpha \in A}$ and $\{\tilde{U}_{\tilde{\alpha}}\}_{\tilde{\alpha} \in \tilde{A}}$ of the open set W . Let $f : W \rightarrow \mathbb{R}$ be a function which is bounded on bounded open subsets of W and continuous on W except possibly on a set of measure zero. If $f : W \rightarrow \mathbb{R}$ is an integrable function, then $\sum_{j \in \mathbb{N}^+} \int_W \phi_j f = \sum_{j \in \mathbb{N}^+} \int_W \tilde{\phi}_j f$.

Proof.

Since f is integrable, $\sum_{j \in \mathbb{N}^+} \int_W \phi_j |f|$ and $\sum_{k \in \mathbb{N}^+} \int_W \tilde{\phi}_k |f|$ converge (and are equal). Thus $\sum_{j \in \mathbb{N}^+} \int_W \phi_j f$ and $\sum_{k \in \mathbb{N}^+} \int_W \tilde{\phi}_k f$ converge absolutely because $\left| \int_W \phi_j f \right| \leq \int_W \phi_j |f|$ and for all $j \in \mathbb{N}^+$ and $\left| \int_W \tilde{\phi}_k f \right| \leq \int_W \tilde{\phi}_k |f|$ for all $k \in \mathbb{N}^+$.

As in the proof of Proposition 3.5.3, we have that $\sum_{k \in \mathbb{N}^+} \tilde{\phi}_k \phi_j f$ is a finite sum for each j and $\sum_{j \in \mathbb{N}^+} \phi_j \tilde{\phi}_k f$ is a finite sum for each k . It also follows from (3.5.1) in the proof of Proposition 3.5.3 that $\sum_{j \in \mathbb{N}^+} \sum_{k \in \mathbb{N}^+} \int_W \tilde{\phi}_k \phi_j f$ converges absolutely. Hence,

$$\begin{aligned} \sum_{j \in \mathbb{N}^+} \int_W \phi_j f &= \sum_{j \in \mathbb{N}^+} \int_W \left(\sum_{k \in \mathbb{N}^+} \tilde{\phi}_k \right) \phi_j f = \sum_{j \in \mathbb{N}^+} \sum_{k \in \mathbb{N}^+} \int_W \tilde{\phi}_k \phi_j f \\ &= \sum_{k \in \mathbb{N}^+} \sum_{j \in \mathbb{N}^+} \int_W \tilde{\phi}_k \phi_j f = \sum_{k \in \mathbb{N}^+} \int_W \left(\sum_{j \in \mathbb{N}^+} \phi_j \right) \tilde{\phi}_k f = \sum_{k \in \mathbb{N}^+} \int_W \tilde{\phi}_k f \end{aligned}$$

The first equality comes from $\sum_{k \in \mathbb{N}^+} \tilde{\phi}_k(\mathbf{x}) = 1$ for all $\mathbf{x}$. Since the summation with respect to k is finite, we may pull the summation with respect to k outside of the integral to get the second equality. The third equality is true because we may change the order of summation of absolutely convergent series. Since the summation with respect to j is finite, we may bring the summation with respect to j inside the integral to get the fourth equality. The last equality comes from $\sum_{j \in \mathbb{N}^+} \phi_j(\mathbf{x}) = 1$ for all $\mathbf{x}$. ■

The previous proposition proves that the following definition is independent of the admissible open cover and partition of unity.

Definition 3.5.6

Suppose that $\{\phi_j\}_{j \in \mathbb{N}^+}$ is a partition of unity subordinated to an admissible open cover $\{U_\alpha\}_{\alpha \in A}$ of the open set W . Let $f : W \rightarrow \mathbb{R}$ be a function which is bounded on bounded open subsets of W and continuous on W except possibly on a set of measure zero. If $f : W \rightarrow \mathbb{R}$ is integrable, then the **integral** of f on W is defined as

$$\int_W f = \sum_{j \in \mathbb{N}^+} \int_W \phi_j f .$$

Remark 3.5.7

We stated Definitions 3.5.4 and 3.5.6 using the C^∞ functions ϕ_j provided by the partition of unity because they came for free from Theorem 3.4.1. However, we may only assumed that the ϕ_j are continuous. This will not change the class of integrable functions as none of the previous propositions required the smoothness of the ϕ_j . ♠

Theorem 3.5.8

Let W be an open and Jordan measurable subset of $\mathbb{R}^n$ and $f : W \rightarrow \mathbb{R}$ be a Riemann integrable function on W . Then f is integrable on W in the new sense.

Proof.

Suppose that $\{U_\alpha\}_{\alpha \in A}$ is an admissible open cover of W and that $\{\phi_j\}_{j \in \mathbb{N}^+}$ is a partition of unity subordinate to $\{U_\alpha\}_{\alpha \in A}$.

Since W is bounded, W is contained in a closed and bounded rectangle K . Since f is Riemann integrable, f is bounded and thus $M = \sup\{f(\mathbf{x}) : \mathbf{x} \in W\} < \infty$. For every finite subset $Q \subset \mathbb{N}^+$, we have

$$\sum_{j \in Q} \int_W \phi_j |f| \leq M \sum_{j \in Q} \int_W \phi_j = M \int_W \sum_{j \in Q} \phi_j \leq M \int_W 1 \leq MV(K) < \infty ,$$

where the second inequality comes from $\sum_{j \in Q} \phi_j(\mathbf{x}) \leq \sum_{j \in \mathbb{N}^+} \phi_j(\mathbf{x}) = 1$ for all $\mathbf{x} \in W$ and the

third inequality comes from $\int_W \chi_W = V(W)$ and $W \subset K$. Since $\sum_{j \in \mathbb{N}^+} \int_W \phi_j |f|$ is a series of positive terms, we get

$$\sum_{j \in \mathbb{N}^+} \int_W \phi_j |f| = \sup_{\substack{Q \subset \mathbb{N} \\ Q \text{ finite}}} \sum_{j \in Q} \int_W \phi_j |f| \leq MV(K) .$$

Therefore f is integrable. ■

The next proposition shows that if $f : W \rightarrow \mathbb{R}$ is Riemann integrable, then the Riemann integral of f is equal to the integral of f as we have defined in Definition 3.5.6.

Theorem 3.5.9

Suppose that $W \subset \mathbb{R}^n$ is an open and Jordan measurable set. Moreover, suppose that $f : W \rightarrow \mathbb{R}$ is Riemann integrable (and so integrable in the new sense according to Theorem 3.5.8). Then the Riemann integral of f on W is equal to the integral of f as defined in Definition 3.5.6.

Proof.

Since f is Riemann integrable, it is bounded on W . Choose $M > 0$ such that $|f(\mathbf{x})| < M$ for all $\mathbf{x} \in W$.

i) Given $\epsilon > 0$, there exists a compact subset $K \subset W$ such that $\int_{W \setminus K} 1 < \epsilon/M$. More precisely, to find K , we use a partition P of a rectangle R containing W such that $\mathcal{L}(\mathcal{X}_W) - \mathcal{L}_P(\mathcal{X}_W) < \epsilon/M$. Note that $\mathcal{X}_W$ is integrable because W is a Jordan measurable set and so the set of points ∂W where $\mathcal{X}_W$ is discontinuous has measure zero. We have

$$\mathcal{L}_P(\mathcal{X}_W) = \sum_{R_{k_1, k_2, \dots, k_n} \subset W} V(R_{k_1, k_2, \dots, k_n})$$

where as usual $R_{k_1, k_2, \dots, k_n}$ is a subrectangle of R generated by the partition P and $V(R_{k_1, k_2, \dots, k_n})$ is the volume of $R_{k_1, k_2, \dots, k_n}$. Note that $m_{k_1, k_2, \dots, k_n} = \inf\{\mathcal{X}_W(\mathbf{x}) : \mathbf{x} \in R_{k_1, k_2, \dots, k_n}\} = 0$ if $R_{k_1, k_2, \dots, k_n} \cap (R \setminus W) \neq \emptyset$.

Let $K = \bigcup_{R_{k_1, k_2, \dots, k_n} \subset W} R_{k_1, k_2, \dots, k_n}$. Then

$$\int_K 1 = \int_R \mathcal{X}_K = \sum_{R_{k_1, k_2, \dots, k_n} \subset W} V(R_{k_1, k_2, \dots, k_n}) = \mathcal{L}_P(\mathcal{X}_W).$$

Moreover

$$\int_W 1 = \int_R \mathcal{X}_W = \mathcal{L}(\mathcal{X}_W)$$

because $\mathcal{X}_W$ is integrable. It follows that

$$\int_{W \setminus K} 1 = \int_W 1 - \int_K 1 = \mathcal{L}(\mathcal{X}_W) - \mathcal{L}_P(\mathcal{X}_W) < \epsilon/M.$$

ii) Suppose that $\{U_\alpha\}_{\alpha \in A}$ is an admissible open cover of W and that $\{\phi_j\}_{j \in \mathbb{N}^+}$ is a partition of unity subordinate to $\{U_\alpha\}_{\alpha \in A}$.

According to Proposition 3.4.4, there are only finitely many ϕ_j such that $\phi_j|_K \neq 0$. Let $Q \subset \mathbb{N}^+$ be the set of indices $j \in \mathbb{N}^+$ such that $\phi_j|_K \neq 0$. Then, for any finite set $G \subset \mathbb{N}^+$ such that $Q \subset G$, we have

$$\left| \int_W f - \sum_{j \in G} \int_W \phi_j f \right| = \left| \int_W f \left(1 - \sum_{j \in G} \phi_j \right) \right| \leq \int_W |f| \left(1 - \sum_{j \in G} \phi_j \right)$$

$$\leq M \int_W \left(1 - \sum_{j \in G} \phi_j\right) \leq M \int_W \left(1 - \sum_{j \in Q} \phi_j\right) \leq M \int_{W \setminus K} 1 < \epsilon .$$

To get the last two inequalities, we have used the relation

$$\sum_{j \in Q} \phi_j(\mathbf{x}) \leq \sum_{j \in G} \phi_j(\mathbf{x}) \leq \sum_{j \in \mathbb{N}^+} \phi_j(\mathbf{x}) = 1$$

for all $\mathbf{x} \in W$ and $\sum_{j \in Q} \phi_j(\mathbf{x}) = 1$ for all $\mathbf{x} \in K$ because the other ϕ_j are null on K . In

particular, this last inequality implies that $1 - \sum_{j \in Q} \phi_j(\mathbf{x})$ may be non zero only if $\mathbf{x} \in W \setminus K$.

Since $G \subset \mathbb{N}^+$ with $Q \subset G$ is arbitrary, we get

$$\left| \int_W f - \sum_{j \in \mathbb{N}^+} \int_W \phi_j f \right| \leq \epsilon .$$

Since ϵ is arbitrary, we get

$$\int_W f = \sum_{j \in \mathbb{N}^+} \int_W \phi_j f = \int_W f .$$

■

3.6 Change of Variables

The main result of this section is the following theorem.

Theorem 3.6.1

Let $W \subset \mathbb{R}^n$ be an open set, and $g : W \rightarrow \mathbb{R}^n$ be a one-to-one and continuously differentiable function such that $\det Dg(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in W$. If $f : g(W) \rightarrow \mathbb{R}$ is integrable, then

$$\int_{g(W)} f = \int_W (f \circ g) |\det Dg| . \quad (3.6.1)$$

Remark 3.6.2

The notation used in the statement of (3.6.1) is not the notation generally used in standard calculus textbooks. If $\mathbf{x} = g(\mathbf{w})$, they use the **Jacobian** defined by

$$\frac{\partial(x_1, x_2, \dots, x_n)}{\partial(w_1, w_2, \dots, w_n)} = \det Dg ,$$

So (3.6.1) is stated as

$$\int_{g(W)} f = \int_W f(g(\mathbf{w})) \left| \frac{\partial(x_1, x_2, \dots, x_n)}{\partial(w_1, w_2, \dots, w_n)}(\mathbf{w}) \right| d\mathbf{w} .$$

♠

The following proposition is at the heart of the proof of Theorem 3.6.1.

Proposition 3.6.3

Let $W \subset \mathbb{R}^n$ be an open set, and $g : W \rightarrow \mathbb{R}^n$ be a one-to-one and continuously differentiable function such that $\det Dg(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in W$. If $f : g(W) \rightarrow \mathbb{R}$ is Riemann integrable with compact support in $g(W)$, then $(f \circ g) |\det Dg| : W \rightarrow \mathbb{R}$ is Riemann integrable with compact support in W and

$$\int_{g(W)} f = \int_W (f \circ g) |\det Dg| .$$

To prove this proposition, we first give several lemmas which will reduce the proof to a simple situation.

Lemma 3.6.4

Let $W \subset \mathbb{R}^n$ be an open set and $f : W \rightarrow \mathbb{R}^n$ be continuously differentiable function. If $A \subset W$ has measure zero, then $f(A)$ has measure zero.

Proof.

We can write W as the countable union of compact sets Q_j for $j \in \mathbb{N}^+$ such that $Q_j \subset Q_{j+1}^\circ \subset Q_{j+1} \subset W$ for all j . For instance, we may write $W = \bigcup_{j \in \mathbb{N}^+} Q_j$ where Q_j is the compact set $Q_j = \{\mathbf{x} \in W : \|\mathbf{x}\| \leq j \text{ and } \text{dist}(\partial W, \mathbf{x}) \leq 1/j\}$. Since

$$f(A) = f(A \cap W) = f\left(A \cap \bigcup_{j \in \mathbb{N}^+} Q_j\right) = f\left(\bigcup_{j \in \mathbb{N}^+} (A \cap Q_j)\right) \subset \bigcup_{j \in \mathbb{N}^+} f(A \cap Q_j) ,$$

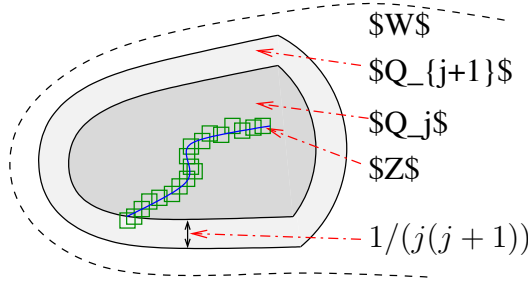
it suffices to prove that $f(A \cap Q_j)$ is of measure zero. Recall that a subset of a set of measure zero has measure zero and the countable union of sets of measure zero has measure zero.

We consider an arbitrary but fixed $j \in \mathbb{N}^+$. Let M be the maximum of $\|Df\|$ on Q_{j+1} . We have that $M < \infty$ because Df is continuous on the compact set Q_{j+1} , so $\|Df\|$ reaches its maximum at a point of Q_{j+1} .

Given $\epsilon > 0$, since $Z = A \cap Q_j$ has measure zero, there exists a collection of rectangles $\{R_i\}_{i \geq 0}$ such that $Z \subset \bigcup_{i=0}^{\infty} R_i$ and $\sum_{i=0}^{\infty} V(R_i) < \frac{\epsilon}{n^{n/2} M^n}$. By subdividing the rectangles R_i

if necessary we may assume that the R_i are squares with edges of length $d_i < \frac{1}{j(j+1)\sqrt{n}}$.

Hence $\text{diam } R_i < \frac{1}{j(j+1)}$. Thus $R_i \subset Q_{j+1}^\circ$ for all i .



Using the Mean Value Theorem for functions from $\mathbb{R}^n$ to $\mathbb{R}^n$ (Theorem 9.19 in [23]), we find that $\|f(\mathbf{x}) - f(\mathbf{y})\| \leq M\|\mathbf{x} - \mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in R_i$ and all i . It follows that

$$\text{diam } f(R_i) \leq M \text{diam } R_i = \sqrt{n} M d_i$$

for all i . Thus, each $f(R_i)$ is a subset of a square $S_i \subset \mathbb{R}^n$ with edges of length $\sqrt{n} M d_i$. The volume of S_i is $V(S_i) = n^{n/2} M^n d_i^n$. We have

$$f(Z) \subset f\left(\bigcup_{i=1}^{\infty} R_i\right) = \bigcup_{i=1}^{\infty} f(R_i) \subset \bigcup_{i=1}^{\infty} S_i$$

with

$$\sum_{i=1}^{\infty} V(S_i) = \sum_{i=1}^{\infty} n^{n/2} M^n d_i^n = n^{n/2} M^n \sum_{i=1}^{\infty} d_i^n < \epsilon$$

because $V(R_i) = d_i^n$. Since ϵ is arbitrary, this proves that $f(Z)$ has measure zero. ■

Lemma 3.6.5

Let $W \subset \mathbb{R}^n$ be an open set, and $g : W \rightarrow \mathbb{R}^n$ be a one-to-one and continuously differentiable function such that $\det Dg(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in W$. If $f : g(W) \rightarrow \mathbb{R}$ is Riemann integrable with compact support, then $(f \circ g) |\det Dg| : W \rightarrow \mathbb{R}$ is Riemann integrable with compact support.

Proof.

Let $K = \text{supp } f \subset g(W)$. The set K is compact and $f(\mathbf{x}) = 0$ for $\mathbf{x} \in W \setminus K$.

i) The support of $(f \circ g) |\det Dg| : W \rightarrow \mathbb{R}$ is a compact subset of W .

We have that $g : W \rightarrow \mathbb{R}^n$ is a homeomorphism of W onto $g(W)$ because g one-to-one, g continuously differentiable, and $\det Dg(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in W$ implies that g is an open mapping⁵ thanks to the Inverse Function Theorem. Thus $g^{-1}(K) \subset W$ is a compact set because continuous functions map compact sets onto compact sets.

Since $(f \circ g)(\mathbf{x}) = 0$ for $\mathbf{x} \in W \setminus g^{-1}(K)$, the support of $(f \circ g) |\det Dg| : W \rightarrow \mathbb{R}$ is a subset of the compact set $g^{-1}(K)$. Therefore the support is compact because a closed subset of a compact set is compact.

⁵Namely, $g(V)$ is an open subset of $\mathbb{R}^n$ for all open sets $V \subset W$. Note that open subsets of W with the induced topology from $\mathbb{R}^n$ are open subsets of $\mathbb{R}^n$ because W is open.

ii) $(f \circ g) |\det Dg| : W \rightarrow \mathbb{R}$ is bounded.

Let $M = \sup\{|f(\mathbf{x})| : \mathbf{x} \in g(W)\}$. We have that $M < \infty$ because f is bounded on $g(W)$.

Since $|\det Dg| : W \rightarrow \mathbb{R}$ is a continuous function because g is continuously differentiable, the function $|\det Dg|$ restricted to the compact set $g^{-1}(K)$ reaches its maximum at a point in $g^{-1}(K)$. Therefore $L = \max\{|\det Dg(\mathbf{x})| : \mathbf{x} \in g^{-1}(K)\} < \infty$. It follows that

$$|((f \circ g) |\det Dg|)(\mathbf{x})| = |f(g(\mathbf{x}))| |\det Dg(\mathbf{x})| \leq ML$$

for all $\mathbf{x} \in W$ because $f(g(\mathbf{x})) = 0$ for all $\mathbf{x} \in W \setminus g^{-1}(K)$.

iii) $(f \circ g) |\det Dg| : W \rightarrow \mathbb{R}$ is continuous except possibly on a set of measure zero. In fact, if $Z \subset K$ is the set of measure zero where f is discontinuous, then we show that $g^{-1}(Z) \subset g^{-1}(K) \subset W$ is the set of measure zero where $(f \circ g) |\det Dg| : W \rightarrow \mathbb{R}$ is discontinuous.

As we say in (i), since $g : W \rightarrow g(W)$ is one-to-one and onto $g(W)$, g is continuously differentiable and $\det Dg(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in W$, we get from the Inverse Function Theorem that g is an open mapping. Moreover, we also get that $Dg^{-1}(\mathbf{y})$ exists for all $\mathbf{y} \in g(W)$. In particular, $Dg^{-1}(\mathbf{y}) = (Dg)^{-1}(g^{-1}(\mathbf{y}))$ for all $\mathbf{y} \in g(W)$ according to the Inverse Function Theorem. Thus $g : W \rightarrow g(W)$ is a diffeomorphism.

It follows from Lemma 3.6.4 that $g^{-1}(Z)$ has measure zero.

iv) We get from (ii) and (iii) that $(f \circ g) |\det Dg| : W \rightarrow \mathbb{R}$ is Riemann integrable according to Theorem 3.1.14. ■

Lemma 3.6.6

Let $W \subset \mathbb{R}^n$ be an open set, and $g : W \rightarrow \mathbb{R}^n$ be a one-to-one and continuously differentiable function such that $\det Dg(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in W$. If Proposition 3.6.3 is true for Riemann integrable functions $f : g(W) \rightarrow [0, \infty[$ with compact support, then it is true for Riemann integrable functions $f : g(W) \rightarrow \mathbb{R}$ with compact support.

Proof.

Suppose that Proposition 3.6.3 is true for Riemann integrable functions $f : g(W) \rightarrow [0, \infty[$ with compact support.

Given a Riemann integrable function $h : g(W) \rightarrow \mathbb{R}$ with compact support, let $h^+ = (h + |h|)/2$ and $h^- = (|h| - h)/2$. We have $h = h^+ - h^-$ with $h^+(\mathbf{x}) \geq 0$ and $h^-(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in g(W)$. Moreover, since $h : g(W) \rightarrow \mathbb{R}$ and $|h| : g(W) \rightarrow \mathbb{R}$ are Riemann integrable functions with compact support according to the version of Proposition 2.1.12 for functions of several variables, we have that h^+ and h^- are Riemann integrable functions on $g(W)$ with compact support. Hence

$$\begin{aligned} \int_{g(W)} h &= \int_{g(W)} h^+ - \int_{g(W)} h^- = \int_W (h^+ \circ g) |\det Dg| - \int_W (h^- \circ g) |\det Dg| \\ &= \int_W (h \circ g) |\det Dg| \end{aligned}$$

by linearity of the Riemann integral, where the second equality comes from our assumption that Proposition 3.6.3 is true for Riemann integrable functions $f : g(W) \rightarrow [0, \infty[$ with compact support. ■

Lemma 3.6.7

Let $W \subset \mathbb{R}^n$ be an open set, and $g : W \rightarrow \mathbb{R}^n$ be a one-to-one and continuously differentiable function such that $\det Dg(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in W$. If Proposition 3.6.3 is true for $f = \chi_Q$ where $Q \subset g(W)$ is a closed rectangle, then Proposition 3.6.3 is true for all Riemann integrable function $f : g(W) \rightarrow \mathbb{R}$ with compact support.

Proof.

We assume that

$$\int_{g(W)} \chi_Q = \int_W (\chi_Q \circ g) |\det Dg| = \int_W \chi_{g^{-1}(Q)} |\det Dg|$$

for all closed rectangles $Q \subset g(W)$. By linearity of the Riemann integral, we also have

$$\int_{g(W)} \alpha \chi_Q = \alpha \int_{g(W)} \chi_Q = \alpha \int_W (\chi_Q \circ g) |\det Dg| = \int_W \alpha (\chi_Q \circ g) |\det Dg|. \quad (3.6.2)$$

for all closed rectangles $Q \subset g(W)$ and $\alpha \in \mathbb{R}$.

According to Lemma 3.6.6, we only have to consider Riemann integrable functions $f : g(W) \rightarrow [0, \infty[$ with compact support. From Lemma 3.6.5, we have that $(f \circ g) |\det Dg| : W \rightarrow [0, \infty[$ is Riemann integrable and has compact support. In the following discussion, we assume without loss of generality that $f : \mathbb{R}^n \rightarrow [0, \infty[$ by setting $f(\mathbf{x}) = 0$ for $\mathbf{x} \in \mathbb{R}^n \setminus g(W)$.

i) Let R be a rectangle containing the support of f and P be a partition of R . As usual, the subrectangles generated by the partition are denoted $R_{k_1, k_2, \dots, k_n}$ and $m_{k_1, k_2, \dots, k_n} = \inf\{f(\mathbf{y}) : \mathbf{y} \in R_{k_1, k_2, \dots, k_n}\}$. Note that $m_{k_1, k_2, \dots, k_n} = 0$ if $R_{k_1, k_2, \dots, k_n} \cap (R \setminus \text{supp } f) \neq \emptyset$.

We have

$$\begin{aligned} \mathcal{L}_P(f) &= \sum_{R_{k_1, k_2, \dots, k_n} \subset g(W)} m_{k_1, k_2, \dots, k_n} V(R_{k_1, k_2, \dots, k_n}) \\ &= \sum_{R_{k_1, k_2, \dots, k_n} \subset g(W)} \int_{g(W)} m_{k_1, k_2, \dots, k_n} \chi_{R_{k_1, k_2, \dots, k_n}} \\ &= \sum_{R_{k_1, k_2, \dots, k_n} \subset g(W)} \int_W (m_{k_1, k_2, \dots, k_n} \chi_{R_{k_1, k_2, \dots, k_n}} \circ g) |\det Dg| \\ &= \int_W \left(\left(\sum_{R_{k_1, k_2, \dots, k_n} \subset g(W)} m_{k_1, k_2, \dots, k_n} \chi_{R_{k_1, k_2, \dots, k_n}} \right) \circ g \right) |\det Dg| \\ &\leq \int_W (f \circ g) |\det Dg|, \end{aligned} \quad (3.6.3)$$

where the third equality comes from (3.6.2) with Q replaced by $R_{k_1, k_2, \dots, k_n}$, and the fourth equality is a consequence of the linearity of the Riemann integral. Note that the sum is finite. The inequality is true because $0 \leq \sum_{R_{k_1, k_2, \dots, k_n} \subset g(W)} m_{k_1, k_2, \dots, k_n} \mathcal{X}_{R_{k_1, k_2, \dots, k_n}} \leq f$.

Since (3.6.3) is true for all partition P of R , we get

$$\int_{g(W)} f = \mathcal{L}(f) \leq \int_W (f \circ g) |\det Dg|, \quad (3.6.4)$$

ii) We now prove that

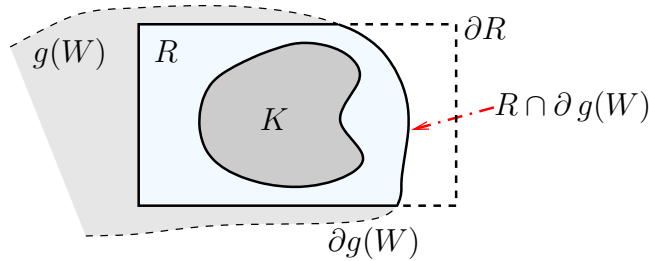
$$\int_{g(W)} f \geq \int_W (f \circ g) |\det Dg|, \quad (3.6.5)$$

following the approach used in (i).

Let R be a rectangle containing $K = \text{supp } f$ and P be a partition of R . As before, the subrectangles generated by the partition are denoted $R_{k_1, k_2, \dots, k_n}$ and $M_{k_1, k_2, \dots, k_n} = \sup\{f(\mathbf{y}) : \mathbf{y} \in R_{k_1, k_2, \dots, k_n}\}$. Since K is a compact subset of the open set $g(W)$, we may select the partition P fine enough such that the $R_{k_1, k_2, \dots, k_n} \subset g(W)$ if $R_{k_1, k_2, \dots, k_n} \cap \text{supp } f \neq \emptyset$. For instance, we may assume that the closed rectangle R is such that $K \subset R$ with $K \cap \partial R = \emptyset$ and let

$$\delta = \min\{\|\mathbf{x} - \mathbf{y}\| : \mathbf{x} \in K \text{ and } \mathbf{y} \in \partial(R \cap g(W))\}.$$

We have that $\delta > 0$ because $\partial(R \cap g(W))$ is a compact set that does not intersect K . Therefore, there is a pair $(\mathbf{x}, \mathbf{y}) \in K \times \partial(R \cap g(W))$ where the minimum δ is reached. Note that $\partial g(W) \cap K = \emptyset$ because $g(W)$ is open.



It suffices to take the partition P such that $\text{diam } R_{k_1, k_2, \dots, k_n} < \delta/2$ for all indices $k_1, k_2, \dots, k_n$. Therefore $R_{k_1, k_2, \dots, k_n} \subset g(W)$ for all indices such that $R_{k_1, k_2, \dots, k_n} \cap K \neq \emptyset$.

We have

$$\begin{aligned} \mathcal{U}_P(f) &= \sum_{R_{k_1, k_2, \dots, k_n} \cap \text{supp } f \neq \emptyset} M_{k_1, k_2, \dots, k_n} V(R_{k_1, k_2, \dots, k_n}) \\ &= \sum_{R_{k_1, k_2, \dots, k_n} \cap \text{supp } f \neq \emptyset} \int_{g(W)} M_{k_1, k_2, \dots, k_n} \mathcal{X}_{R_{k_1, k_2, \dots, k_n}} \\ &= \sum_{R_{k_1, k_2, \dots, k_n} \cap \text{supp } f \neq \emptyset} \int_W (M_{k_1, k_2, \dots, k_n} \mathcal{X}_{R_{k_1, k_2, \dots, k_n}} \circ g) |\det Dg| \end{aligned}$$

$$\begin{aligned}
&= \int_W \left(\left(\sum_{R_{k_1, k_2, \dots, k_n} \cap \text{supp } f \neq \emptyset} M_{k_1, k_2, \dots, k_n} \chi_{R_{k_1, k_2, \dots, k_n}} \right) \circ g \right) |\det Dg| \\
&\geq \int_W (f \circ g) |\det Dg|, \tag{3.6.6}
\end{aligned}$$

where the third equality comes from (3.6.2) with Q replaced by $R_{k_1, k_2, \dots, k_n}$, and the fourth equality is a consequence of the linearity of the Riemann integral. The inequality is true because $\text{supp } f \subset \bigcup_{R_{k_1, k_2, \dots, k_n} \cap \text{supp } f \neq \emptyset} R_{k_1, k_2, \dots, k_n}$ and $f \leq \sum_{R_{k_1, k_2, \dots, k_n} \cap \text{supp } f \neq \emptyset} M_{k_1, k_2, \dots, k_n} \chi_{R_{k_1, k_2, \dots, k_n}}$.

Since (3.6.6) is true for all fine enough partition P of R , we get

$$\int_{g(W)} f = \mathcal{U}(f) \geq \int_W (f \circ g) |\det Dg|.$$

This proves (3.6.5).

It follows from (3.6.4) and (3.6.5) that

$$\int_{g(W)} f = \int_W (f \circ g) |\det Dg|. \quad \blacksquare$$

Lemma 3.6.8

Let $W \subset \mathbb{R}^n$ be an open set. Proposition 3.6.3 is true if $g : W \rightarrow \mathbb{R}^n$ is an invertible linear map.

Proof.

Note that g is a continuous differentiable function and $\det Dg(\mathbf{x}) = \det g \neq 0$ for all $\mathbf{x} \in W$ because g is an invertible linear map.

According to Lemma 3.6.7, we only have to prove that

$$\int_{g(W)} \chi_Q = \int_W \chi_{g^{-1}(Q)} |\det Dg|$$

for all closed rectangles $Q \subset g(W)$. We have used the fact that g is one-to-one to get $\chi_Q \circ g = \chi_{g^{-1}(Q)}$. This equation can be restated as

$$\int_Q = \int_{g^{-1}(Q)} |\det g|. \tag{3.6.7}$$

Note that $Dg(\mathbf{x}) = g$ for all $\mathbf{x} \in W$ because g is a linear map. The equation (3.6.7) simply states that the volume of Q is equal to $|\det g|$ times the volume of $g^{-1}(Q)$ because $|\det g|$ is constant.

Since $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a linear map such that $\det g \neq 0$, we may express g as $g = G_1 G_2 \dots G_s$ where the G_i are linear maps $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ of the following forms.

1. For some $1 \leq k \leq n$ and $a \neq 1$, the components $g_{i,j}$ of the matrix G are

$$g_{i,j} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \neq k \\ a & \text{if } i = j = k \end{cases}$$

$G\mathbf{x}$ represents a stretching of $\mathbf{x}$ in the direction of $\mathbf{e}_k$ by a factor of $|a|$ with a possible reflection through the space orthogonal to $\mathbf{e}_k$ if $a < 0$.

2. For some $1 \leq k_1 < k_2 \leq n$, the components $g_{i,j}$ of the matrix G are

$$g_{i,j} = \begin{cases} 0 & \text{if } i \neq j, (i,j) \neq (k_1, k_2) \text{ and } (i,j) \neq (k_2, k_1) \\ 1 & \text{if } (i,j) = (k_1, k_2) \text{ or } (i,j) = (k_2, k_1) \\ 0 & \text{if } i = j = k_1 \text{ or } i = j = k_2 \\ 1 & \text{if } i = j \neq k_1 \text{ and } i = j \neq k_2 \end{cases}$$

$G\mathbf{x}$ is a permutation of the k_1 and k_2 coordinates of $\mathbf{x}$.

3. For some $1 \leq k_1, k_2 \leq n$ with $k_1 \neq k_2$, the components $g_{i,j}$ of the matrix G are

$$g_{i,j} = \begin{cases} 0 & \text{if } i \neq j \text{ and } (i,j) \neq (k_1, k_2) \\ 1 & \text{if } (i,j) = (k_1, k_2) \\ 1 & \text{if } i = j \end{cases}$$

$G\mathbf{x}$ adds the k_2 -component of $\mathbf{x}$ to the k_1 -component of $\mathbf{x}$ if $k_1 < k_2$ or the k_1 -component of $\mathbf{x}$ to the k_2 -component of $\mathbf{x}$ if $k_2 < k_1$.

One learns in Linear algebra that transformations of the form (2) and (3) preserve volume and that transformations of the form (1) change the volume by a factor of $|a|$. We have that $|\det G| = |a|$ for transformations G of the form (1) and $|\det G| = 1$ for transformation G of the form (2) and (3). Hence, for any set $U \subset \mathbb{R}^n$, the volume of $G(U)$ is the volume of U multiplied by $|\det G|$ for all the linear matrices of the form (1), (2) and (3).

The volume of Q is therefore the product of the volume of $g^{-1}(Q)$ by $|\det G_1| |\det G_2| \dots |\det G_s| = |\det g|$ as claimed by (3.6.7). ■

Lemma 3.6.9

Let $W \subset \mathbb{R}^n$ be an open set. If Proposition 3.6.3 is true for the one-to-one and continuously differentiable functions $g : W \rightarrow \mathbb{R}^n$ such that $\det Dg(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in W$ and $Dg(\mathbf{a}) = \text{Id}$ for some $\mathbf{a} \in W$, then Proposition 3.6.3 is true in general.

Proof.

Suppose that Proposition 3.6.3 is true for one-to-one and continuously differentiable functions $g : W \rightarrow \mathbb{R}^n$ such that $\det Dg(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in W$ and $Dg(\mathbf{a}) = \text{Id}$ for some $\mathbf{a} \in W$.

Suppose that $h : W \rightarrow \mathbb{R}^n$ is a one-to-one and continuously differentiable function such that $\det D h(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in W$. Given $\mathbf{a} \in W$, let $T = D h(\mathbf{a})$. We have that $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a linear invertible map because $\det D h(\mathbf{a}) \neq 0$.

Suppose that $f : h(W) \rightarrow \mathbb{R}$ is a Riemann integrable function with compact support in $h(W)$. Since $T : T^{-1}(h(W)) \rightarrow \mathbb{R}^n$ is a one-to-one and continuously differentiable function such that $\det D T(\mathbf{x}) = \det T \neq 0$ for all $\mathbf{x} \in W$, we get from Lemma 3.6.5 that $(f \circ T)|\det T| : T^{-1}(h(W)) \rightarrow \mathbb{R}$ is a Riemann integrable function with compact support. Note that $T^{-1}(h(W)) \subset \mathbb{R}^n$ is an open set and recall that $D T(\mathbf{x}) = T$ for all $\mathbf{x}$ because T is a linear map. Furthermore, since T is an invertible linear map, we get from Lemma 3.6.8 that

$$\int_{h(W)} f = \int_{T^{-1}(h(W))} (f \circ T)|\det T| . \quad (3.6.8)$$

Since $(f \circ T)|\det T| : T^{-1}(h(W)) \rightarrow \mathbb{R}$ is a Riemann integrable function with compact support and $g = T^{-1} \circ h : W \rightarrow \mathbb{R}^n$ is a one-to-one and continuously differentiable function such that

$$\det(D g(\mathbf{x})) = \det(D(T^{-1} \circ h)(\mathbf{x})) = \det(T^{-1} D h(\mathbf{x})) = \det(T^{-1}) \det(D h(\mathbf{x})) \neq 0$$

for all $\mathbf{x} \in W$, we get from Lemma 3.6.5 that

$$\begin{aligned} (((f \circ T)|\det T|) \circ g) |\det D g| &= ((f \circ T \circ (T^{-1} \circ h)) |\det T|) |\det D(T^{-1} \circ h)| \\ &= ((f \circ T \circ (T^{-1} \circ h)) |\det T|) |\det T^{-1}| |\det D h| \\ &= (f \circ h) |\det D h| \end{aligned}$$

is a Riemann integrable function with compact support. Moreover

$$D g(\mathbf{a}) = D(T^{-1} \circ h)(\mathbf{a}) = T^{-1} D h(\mathbf{a}) = T^{-1} \circ T = \text{Id} .$$

Therefore, we get from our assumption at the beginning of the proof that

$$\begin{aligned} \int_{T^{-1}(h(W))} (f \circ T)|\det T| &= \int_{g(W)} (f \circ T)|\det T| \\ &= \int_W (((f \circ T)|\det T|) \circ g) |\det D g| = \int_W (f \circ h) |\det D h| . \end{aligned} \quad (3.6.9)$$

It follows from (3.6.8) and (3.6.9) that

$$\int_{h(W)} f = \int_W (f \circ h) |\det D h| . \quad \blacksquare$$

Proof (of Proposition 3.6.3).

We proceed by induction on the dimension n .

i) We prove that Proposition 3.6.3 is true for $n = 1$.

We may assume that W is connected. If W is not connected and so is the union of non-intersecting open intervals, then we may use the linearity of the integral to reduce the integral on W to a sum of integrals on intervals.

We have from Lemma 3.6.5 that $(f \circ g)g'$ is Riemann integrable with compact support in W . Since $g^{-1}(\text{supp } f) \subset W$ is a compact set because $\text{supp } f$ is compact and $g^{-1} : g(W) \rightarrow \mathbb{R}$ is continuous as we have shown in the proof of Lemma 3.6.5, and since W is connected, there exists a closed interval $[\alpha, \beta] \subset W$ such that $g^{-1}(\text{supp } f) \subset [\alpha, \beta]$. Therefore $\text{supp } f \subset g([\alpha, \beta])$.

The condition that $g'(x) \neq 0$ for all $x \in W$ implies that g is strictly increasing or decreasing on W . If $g'(x) > 0$ for all x , then $g(b) > g(a)$ and it follows from Theorem 2.3.1 that

$$\int_{g(\alpha)}^{g(\beta)} f(x) \, dx = \int_{\alpha}^{\beta} (f \circ g)(x) g'(x) \, dx . \quad (3.6.10)$$

We may rewrite (3.6.10) as

$$\int_{g([\alpha, \beta])} f = \int_{[\alpha, \beta]} (f \circ g) |g'| .$$

If $g'(x) < 0$ for all x , then $g(b) < g(a)$ and it follows from Remark 2.3.2 that

$$\int_{g(\beta)}^{g(\alpha)} f(x) \, dx = - \int_{\alpha}^{\beta} (f \circ g)(x) g'(x) \, dx = \int_{\alpha}^{\beta} (f \circ g)(x) |g'(x)| \, dx \quad (3.6.11)$$

because $-g'(x) = |g'(x)|$ for all x . Again, we can rewrite (3.6.11) as

$$\int_{g([\alpha, \beta])} f = \int_{[\alpha, \beta]} (f \circ g) |g'| .$$

In either cases, we get Proposition 3.6.3 for $n = 1$.

ii) We assume that Proposition 3.6.3 is true for $n < k$.

iii) We prove that Proposition 3.6.3 is true for $n = k$.

We first prove that for every $\mathbf{a} \in W$, there exists an open set $U_{\mathbf{a}} \subset W$ containing $\mathbf{a}$ such that

$$\int_{g(U_{\mathbf{a}})} f = \int_{U_{\mathbf{a}}} (f \circ g) |\det Dg| . \quad (3.6.12)$$

for all Riemann integrable functions $f : g(U_{\mathbf{a}}) \rightarrow \mathbb{R}$ with compact support in $g(U_{\mathbf{a}})$.

Given $\mathbf{a} \in W$, let $T = Dg(\mathbf{a})$ and $\tilde{g} = T^{-1} \circ g$. The map $\tilde{g} : W \rightarrow \mathbb{R}^k$ is a one-to-one and continuously differentiable map such that $\det D\tilde{g}(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in W$ and $D\tilde{g}(\mathbf{a}) = \text{Id}$.

Consider

$$Q_1 : W \rightarrow \mathbb{R}^k$$

$$\mathbf{x} \mapsto (\tilde{g}_1(\mathbf{x}) \quad \tilde{g}_2(\mathbf{x}) \quad \dots \quad \tilde{g}_{k-1}(\mathbf{x}) \quad x_k)^\top$$

We have

$$DQ_1(\mathbf{x}) = \begin{pmatrix} \frac{\partial \tilde{g}_1}{\partial x_1}(\mathbf{x}) & \frac{\partial \tilde{g}_1}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial \tilde{g}_1}{\partial x_k}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \tilde{g}_{k-1}}{\partial x_1}(\mathbf{x}) & \frac{\partial \tilde{g}_{k-1}}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial \tilde{g}_{k-1}}{\partial x_k}(\mathbf{x}) \\ 0 & 0 & \cdots & 1 \end{pmatrix}. \quad (3.6.13)$$

In particular, we have

$$DQ_1(\mathbf{a}) = \begin{pmatrix} \text{Id}_{k-1} & 0 \\ 0 & 1 \end{pmatrix} = \text{Id}. \quad (3.6.14)$$

Since $DQ_1(\mathbf{a})$ is invertible and $DQ_1 : W \rightarrow \mathbb{R}^k$ is continuous, we get that $DQ_1(\mathbf{x})$ is invertible for all $\mathbf{x}$ in a neighbourhood of $\mathbf{a}$. Hence, we may use the Inverse Function Theorem to find an open neighbourhood U_1 of $\mathbf{a}$ such that $Q_1 : U_1 \rightarrow \mathbb{R}^k$ is a diffeomorphism of the open set U_1 onto the open set $Q_1(U_1)$.

Consider

$$Q_2 : Q_1(U_1) \rightarrow \mathbb{R}^k \\ \mathbf{x} \mapsto (x_1 \ x_2 \ \dots \ x_{k-1} \ \tilde{g}_k(Q_1^{-1}(\mathbf{x})))^\top$$

We have

$$DQ_2(\mathbf{x}) = \begin{pmatrix} 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \\ \frac{\partial}{\partial x_1}(\tilde{g}_k \circ Q_1^{-1})(\mathbf{x}) & \cdots & \frac{\partial}{\partial x_{k-1}}(\tilde{g}_k \circ Q_1^{-1})(\mathbf{x}) & \frac{\partial}{\partial x_k}(\tilde{g}_k \circ Q_1^{-1})(\mathbf{x}) \end{pmatrix}. \quad (3.6.15)$$

Using the Inverse Function Theorem for Q_1 , we get

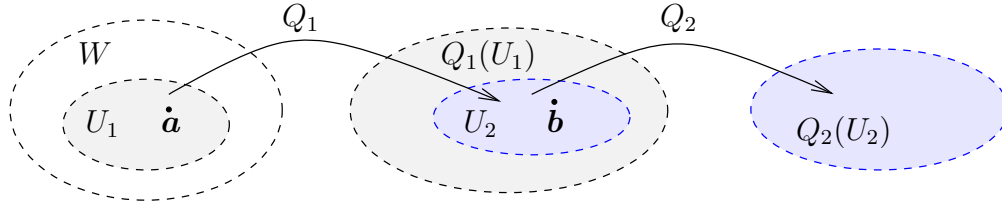
$$D(\tilde{g}_k \circ Q_1^{-1})(Q_1(\mathbf{a})) = D\tilde{g}_k(\mathbf{a}) DQ_1^{-1}(Q_1(\mathbf{a})) = D\tilde{g}_k(\mathbf{a})(DQ_1(\mathbf{a}))^{-1} = D\tilde{g}_k(\mathbf{a})$$

because $DQ_1(\mathbf{a}) = \text{Id}$. It follows that

$$\frac{\partial}{\partial x_j}(\tilde{g}_k \circ Q_1^{-1})(Q_1(\mathbf{a})) = \frac{\partial \tilde{g}_k}{\partial x_j}(\mathbf{a}) = \delta_{j,k}$$

for $1 \leq j \leq k$ because $D\tilde{g}(\mathbf{a}) = \text{Id}_k$, where $\delta_{j,k}$ is the **Kronecker delta function** defined by $\delta_{j,k} = 1$ when $j = k$ and 0 otherwise. Hence $DQ_2(Q_1(\mathbf{a})) = \text{Id}_k$.

Since $DQ_2(Q_1(\mathbf{a}))$ is invertible and $DQ_2 : Q_1(U_1) \rightarrow \mathbb{R}^k$ is continuous, we get that $DQ_2(\mathbf{x})$ is invertible for all $\mathbf{x}$ in a neighbourhood of $\mathbf{b} = Q_1(\mathbf{a})$. Hence, we may use the Inverse Function Theorem to find an open neighbourhood $U_2 \subset Q_1(U_1)$ of $\mathbf{b}$ such that $Q_2 : U_2 \rightarrow \mathbb{R}^k$ is a diffeomorphism of the open set U_2 onto the open set $Q_2(U_2)$.



iii.a) We first prove that

$$\int_{Q_1(U_1)} f = \int_{U_1} (f \circ Q_1) |\det D Q_1| . \quad (3.6.16)$$

for all Riemann integrable functions $f : Q_1(U_1) \rightarrow \mathbb{R}$ with compact support in $Q_1(U_1)$. According to Lemma 3.6.7, it suffices to prove (3.6.16) for $f = \chi_R$ where R is a closed rectangle in $Q_1(U_1)$; namely, it suffices to prove that

$$\int_R = \int_{Q_1(U_1)} \chi_R = \int_{U_1} (\chi_R \circ Q_1) |\det D Q_1| = \int_{Q_1^{-1}(R)} |\det D Q_1| .$$

Suppose that $R = R_1 \times [a, b]$ where $R_1 \subset \mathbb{R}^{k-1}$ is a closed rectangle such that $R_1 \times [a, b] \subset Q_1(U_1)$.

From Fubini's Theorem, we get

$$\int_R = \int_{[a, b]} \left(\int_{R_1} d\mathbf{y} \right) dx_k .$$

Let

$$Q_{1,x_k}(\mathbf{y}) = (\tilde{g}_1(\mathbf{y}, x_k) \quad \tilde{g}_2(\mathbf{y}, x_k) \quad \dots \quad \tilde{g}_{k-1}(\mathbf{y}, x_k))^T .$$

for all $(\mathbf{y}, x_k) \in U_1$. The important observation is that

$$Q_1^{-1}(R) = \bigcup_{x_k \in [a, b]} (Q_{1,x_k}^{-1}(R_1) \times \{x_k\}) .$$

Note that Q_{1,x_k} is one-to-one because $Q_1 : U_1 \rightarrow \mathbb{R}^k$ is one-to-one. Moreover

$$\det D_{\mathbf{y}} Q_{1,x_k}(\mathbf{y}) = \det D Q_1(\mathbf{y}, x_k) \neq 0$$

for all $(\mathbf{y}, x_k) \in U_1$ follows from (3.6.13) and $\det D Q_1(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in U_1$. Therefore Q_{1,x_k} satisfies Proposition 3.6.3 in dimension $k-1$ for all $x_k \in [a, b]$. Hence, by our hypothesis of induction, we have

$$\int_{R_1} d\mathbf{y} = \int_{Q_{1,x_k}^{-1}(R_1)} |\det D_{\mathbf{y}} Q_{1,x_k}(\mathbf{y})| d\mathbf{y} = \int_{Q_{1,x_k}^{-1}(R_1)} |\det D Q_1(\mathbf{y}, x_k)| d\mathbf{y}$$

for all $x_k \in [a, b]$. Thus

$$\int_R = \int_{[a, b]} \left(\int_{Q_{1,x_k}^{-1}(R_1)} |\det D Q_1(\mathbf{y}, x_k)| d\mathbf{y} \right) dx_k = \int_{Q_1^{-1}(R)} |\det D Q_1| ,$$

where the last equality comes from Fubini's Theorem since $\det D Q_1$ is continuous on U_1 . This last statement needs some clarifications. Since $Q_1 : U_1 \rightarrow \mathbb{R}^k$ is a one-to-one and continuously differentiable function with $\det D Q_1(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in U_1$, we have that $Q_1 : U_1 \rightarrow Q(U_1)$ is a diffeomorphism. It follows from Lemma 3.6.4 that $Q_1^{-1}(R)$ is Jordan measurable because $\partial Q_1^{-1}(R) = Q_1^{-1}(\partial R)$ is of measure zero since ∂R is of measure zero. We also have that $Q_1^{-1}(R)$ is bounded because R is a compact set and thus $Q_1^{-1}(R)$ is a compact set in $\mathbb{R}^k$. The Fubini's Theorem is applied to the function $\mathcal{X}_{Q_1^{-1}(R)} |\det D Q_1|$ on a rectangle containing $Q_1^{-1}(R)$.

iii.b) We prove that

$$\int_{Q_2(U_2)} f = \int_{U_2} (f \circ Q_2) |\det D Q_2| . \quad (3.6.17)$$

for all Riemann integrable functions $f : Q_2(U_2) \rightarrow \mathbb{R}$ with compact support in $Q_2(U_2)$. According to Lemma 3.6.7, it suffices to prove (3.6.17) for $f = \mathcal{X}_R$ where R is a closed rectangle in $Q_2(U_2)$; namely, it suffices to prove that

$$\int_R = \int_{Q_2(U_2)} \mathcal{X}_R = \int_{U_2} (\mathcal{X}_R \circ Q_2) |\det D Q_2| = \int_{Q_2^{-1}(R)} |\det D Q_2| .$$

Suppose that $R = R_1 \times [a, b]$ where $R_1 \subset \mathbb{R}^{k-1}$ is a closed rectangle such that $R_1 \times [a, b] \subset Q_2(U_2)$.

From Fubini's Theorem, we have

$$\int_R = \int_{R_1} \left(\int_{[a, b]} dx_k \right) d\mathbf{y} .$$

Let $Q_{2,\mathbf{y}}(x_k) = \tilde{g}_k(Q_1^{-1}(\mathbf{y}, x_k))$ for all $(\mathbf{y}, x_k) \in U_2$. The important observation is that

$$Q_2^{-1}(R) = \bigcup_{\mathbf{y} \in R_1} (\{\mathbf{y}\} \times Q_{2,\mathbf{y}}^{-1}([a, b])) .$$

Note that $Q_{2,\mathbf{y}}$ is one-to-one because $Q_2 : U_2 \rightarrow \mathbb{R}^k$ is one-to-one. Moreover

$$D_x Q_{2,\mathbf{y}}(x_k) = \frac{d}{dx} \tilde{g}_k(Q_1^{-1}(\mathbf{y}, x)) \Big|_{x=x_k} \neq 0$$

for all $x_k \in [a, b]$ follows from (3.6.15) and $\det D Q_2(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in U_2$. Therefore $Q_{2,\mathbf{y}}$ satisfies Proposition 3.6.3 in dimension 1 for all $\mathbf{y} \in R_1$. Hence, by our hypothesis of induction, we have

$$\int_{[a, b]} dx_k = \int_{Q_{2,\mathbf{y}}^{-1}([a, b])} |\det D_{x_k} Q_{2,\mathbf{y}}(x_k)| dx_k = \int_{Q_{2,\mathbf{y}}^{-1}([a, b])} |\det D Q_2(\mathbf{y}, x_k)| dx_k$$

for all $\mathbf{y} \in R_1$ because $\det D Q_2(\mathbf{y}, x_k) = D_{x_k} Q_{2,\mathbf{y}}(x_k)$ for all $(\mathbf{y}, x_k) \in U_2$ according to (3.6.15). Hence

$$\int_R h = \int_{R_1} \left(\int_{Q_{2,\mathbf{y}}^{-1}([a, b])} |\det D Q_2(\mathbf{y}, x_k)| dx_k \right) d\mathbf{y} = \int_{Q_2^{-1}(R)} |\det D Q_2| ,$$

where the last equality comes from Fubini's Theorem since $\det D Q_2$ is continuous on U_2 . As in (iii.a), this last statement needs some clarifications. Since $Q_2 : U_2 \rightarrow \mathbb{R}^k$ is a one-to-one and continuously differentiable function with $\det D Q_2(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in U_2$, we have that $Q_2 : U_2 \rightarrow Q(U_2)$ is a diffeomorphism. It follows as in (iii.a) that $Q_2^{-1}(R)$ is a bounded and Jordan measurable set. The Fubini's Theorem is applied to the function $\mathcal{X}_{Q_2^{-1}(R)} |\det D Q_2|$ on a rectangle containing $Q_2^{-1}(R)$.

iii.c) We now prove (3.6.12) with $U_{\mathbf{a}} = U_1$. We consider the Riemann integrable function $h : \tilde{g}(U_1) \rightarrow \mathbb{R}$ with compact support in $\tilde{g}(U_1)$ defined by $h = f \circ T$ where $f : g(U_1) = (T \circ \tilde{g})(U_1) \rightarrow \mathbb{R}$ is a Riemann integrable functions with compact support in $g(U_1)$. Recall that $\tilde{g} = Q_2 \circ Q_1 = T^{-1} \circ g$ on U_1 .

We get from Lemma 3.6.5 that $(h \circ \tilde{g}) |\det D \tilde{g}| : U_1 \rightarrow \mathbb{R}$ is a Riemann integrable function with compact support.

Let $\tilde{h} : U_2 = Q_1(U_1) \rightarrow \mathbb{R}$ be the function defined by $\tilde{h}(\mathbf{x}) = (h \circ Q_2)(\mathbf{x}) |\det D Q_2(\mathbf{x})|$ for $\mathbf{x} \in U_2$. As usual, according to Lemma 3.6.5, we have that $\tilde{h}$ is a Riemann integrable function with compact support in U_2 because $h : \tilde{g}(U_1) = Q_2(U_2) \rightarrow \mathbb{R}$ is a Riemann integrable function with compact support.

We may apply (iii.a) to $\tilde{h}$ to get

$$\begin{aligned} \int_{U_1} (h \circ \tilde{g}) |\det \tilde{g}| &= \int_{U_1} h(Q_2(Q_1(\mathbf{x}))) |\det(D Q_2)(Q_1(\mathbf{x}))| |\det D Q_1(\mathbf{x})| d\mathbf{x} \\ &= \int_{U_1} (\tilde{h} \circ Q_1) |\det D Q_1| = \int_{Q_1(U_1)} \tilde{h} = \int_{U_2} (h \circ Q_2) |\det D Q_2|. \end{aligned}$$

We may also apply (iii.b) to h to get

$$\int_{U_2} (h \circ Q_2) |\det D Q_2| = \int_{Q_2(U_2)} h = \int_{Q_2(Q_1(U_1))} h = \int_{\tilde{g}(U_1)} h.$$

Combining the last two equations, we get

$$\int_{U_1} (h \circ \tilde{g}) |\det \tilde{g}| = \int_{\tilde{g}(U_1)} h. \quad (3.6.18)$$

Since T is an invertible linear map, we get from Lemma 3.6.8 and (3.6.18) that

$$\begin{aligned} \int_{g(U_1)} f &= \int_{T^{-1}(\tilde{g}(U_1))} (f \circ T) |\det T| = \int_{\tilde{g}(U_1)} h |\det T| = \int_{U_1} (h \circ \tilde{g}) |\det T| |\det \tilde{g}| \\ &= \int_{U_1} (f \circ g) |\det g|. \end{aligned}$$

Therefore, we may take U_1 as the set $U_{\mathbf{a}}$ that we were looking for at the beginning of (iii).

iii.d) Since f in the statement of Proposition 3.6.3 has a compact support, there exists a finite collection $\{U_{\mathbf{a}_i}\}_{1 \leq i \leq N}$ of open subsets of W such that $\text{supp } f \subset \bigcup_{i=1}^N g(U_{\mathbf{a}_i})$ and (3.6.12)

is true for all Riemann integrable functions with compact support in $g(U_{\mathbf{a}_i})$. Do not forget that $g : W \rightarrow \mathbb{R}^n$ is an open mapping and so $g(U_{\mathbf{a}_i})$ is an open set for all i .

Let $\{\phi_i\}_{1 \leq i \leq N}$ be a partition of unity subordinate to the open cover $\{g(U_{\mathbf{a}_i})\}_{1 \leq i \leq N}$ of $\text{supp } f$ as described in Remark 3.4.2. In particular, $\text{supp } \phi_i \subset g(U_{\mathbf{a}_i})$ for all i . We have

$$f = \sum_{i=1}^N \phi_i f \text{ on } g(W) \text{ and}$$

$$\begin{aligned} \int_{g(W)} f &= \int_{g(W)} \sum_{i=1}^N \phi_i f = \sum_{i=1}^N \int_{g(W)} \phi_i f = \sum_{i=1}^N \int_{g(U_{\mathbf{a}_i})} \phi_i f \\ &= \sum_{i=1}^N \int_{U_{\mathbf{a}_i}} ((\phi_i f) \circ g) |\det Dg| = \sum_{i=1}^N \int_W ((\phi_i f) \circ g) |\det Dg| \\ &= \int_W \left(\left(\sum_{i=1}^N \phi_i f \right) \circ g \right) |\det Dg| = \int_W (f \circ g) |\det Dg|, \end{aligned}$$

where the fourth equality comes from (3.6.12) since, for each i , $\phi_i f$ is a Riemann integrable function with compact support in $g(U_{\mathbf{a}_i})$. We have also used several times the linearity of the Riemann integral.

This completes the proof by induction and the proof of Proposition 3.6.3. ■

Proof (of Theorem 3.6.1).

Since $f : g(W) \rightarrow \mathbb{R}$ is integrable, there exist an admissible open cover $\{\tilde{U}_\alpha\}_{\alpha \in A}$ of $g(W)$ and a partition of unity $\{\phi_j\}_{j \in \mathbb{N}^+}$ subordinate to $\{\tilde{U}_\alpha\}_{\alpha \in A}$ such that

$$\int_{g(W)} f = \sum_{j \in \mathbb{N}^+} \int_{g(W)} \phi_j f.$$

In particular, the sum converges absolutely. For each $j \in \mathbb{N}^+$, we have that $\text{supp } \phi_j \subset \tilde{U}_{\alpha_j}$ for some $\alpha_j \in A$.

Since $g : W \rightarrow g(W)$ is a diffeomorphism as we have shown in (iii) of the proof of Lemma 3.6.5, the collection $\{U_\alpha\}_{\alpha \in A}$ where $U_\alpha = g^{-1}(\tilde{U}_\alpha)$ for all α is an admissible open cover of W and $\{\phi_j \circ g\}_{j \in \mathbb{N}^+}$ is a partition of unity subordinate to $\{U_\alpha\}_{\alpha \in A}$ of class C^1 (see Remark 3.5.7). We have that $\sum_{j \in \mathbb{N}^+} (\phi_j \circ g)(\mathbf{x}) = 1$ for all $\mathbf{x} \in W$ and $\text{supp}(\phi_j \circ g) \subset U_{\alpha_j} = g^{-1}(\tilde{U}_{\alpha_j})$ for all $j \in \mathbb{N}^+$.

Since $\phi_j f : \tilde{U}_{\alpha_j} \rightarrow \mathbb{R}$ is Riemann integrable according to Lemma 3.5.2 and has a compact support in $g(W)$ because ϕ_j has a compact support in $\tilde{U}_{\alpha_j}$, we may use Proposition 3.6.3 to conclude that

$$\int_{g(U_{\alpha_j})} \phi_j f = \int_{U_{\alpha_j}} ((\phi_j f) \circ g) |\det Dg|$$

for $h = |f|$ and $h = f$, and all $j \in \mathbb{N}^+$. Thus

$$\begin{aligned} \int_{g(W)} h &= \sum_{j \in \mathbb{N}^+} \int_{g(W)} \phi_j h = \sum_{j \in \mathbb{N}^+} \int_{g(U_{\alpha_j})} \phi_j h = \sum_{j \in \mathbb{N}^+} \int_{U_{\alpha_j}} ((\phi_j h) \circ g) |\det Dg| \\ &= \sum_{j \in \mathbb{N}^+} \int_W (\phi_j \circ g)(h \circ g) |\det Dg| = \int_W (h \circ g) |\det Dg| \end{aligned}$$

for $h = |f|$ and $h = f$. Thus, $(f \circ g) |\det Dg|$ is integrable on W and (3.6.1) is satisfied. ■

3.7 Sard's Theorem

If $f : U \rightarrow \mathbb{R}$ is an integrable function on an open set U , then it follows from Remark 3.1.17 that the value of the integral of f on U is equal to the value of the integral of f on $U \setminus A$ where A is a closed set of measure zero (just consider $g(\mathbf{x}) = f(\mathbf{x})$ for $\mathbf{x} \in U \setminus A$ and $g(\mathbf{x}) = 0$ for $\mathbf{x} \in A$). It follows from the next theorem that Theorem 3.6.1 is also true if the set $A = \{\mathbf{x} \in W : \det Dg(\mathbf{x}) = 0\}$ is a closed set of measure zero because the set $g(A)$ is always of measure zero.

Definition 3.7.1

Let U be an open subset of $\mathbb{R}^n$ and $f : U \rightarrow \mathbb{R}^n$ be a continuously differentiable function.

A point $\mathbf{x} \in U$ is a **critical point** of f if $\det(Df(\mathbf{x})) = 0$ or, equivalently, $Df(\mathbf{x})$ has rank less than n or $Df(\mathbf{x})$ is not invertible.

A point $\mathbf{y} \in \mathbb{R}^n$ is a **regular value** of f if none of the points $\mathbf{x} \in f^{-1}(\{\mathbf{y}\})$ is a critical point.

Before stating the main result of this section, we state a result about regular points that will be useful later.

Proposition 3.7.2

Let U be an open subset of $\mathbb{R}^n$ and $f : U \rightarrow \mathbb{R}^n$ be a continuously differentiable function. If $\mathbf{y} \in \mathbb{R}^n$ is a regular value of f , then $f^{-1}(\{\mathbf{y}\})$ is the union of isolated points.

Proof.

Suppose that $\mathbf{x} \in f^{-1}(\{\mathbf{y}\})$. Then $\det(Df(\mathbf{x})) \neq 0$ because $\mathbf{y}$ is regular. We get from the Inverse Function Theorem that f is an isomorphism from an open neighbourhood $U_{\mathbf{x}} \subset U$ of $\mathbf{x}$ to an open neighbourhood $V_{\mathbf{y}}$ of $\mathbf{y}$. In particular, $U_{\mathbf{x}} \cap f^{-1}(\{\mathbf{y}\}) = \{\mathbf{x}\}$. ■

The next theorem says that almost all points in the image of a continuously differentiable function are regular values. The proof that we give closely follows the proof given in [26].

Another proof is given in [10] which, according to the authors, is “virtually verbatim” from [19].

Theorem 3.7.3 (Sard's Theorem)

Let U be an open subset of $\mathbb{R}^n$ and $f : U \rightarrow \mathbb{R}^n$ be a continuously differentiable function. Moreover, let $A = \{\mathbf{x} \in U : \det(Df(\mathbf{x})) = 0\}$, Then $f(A)$ is a set of measure zero.

Proof.

i) Suppose that $R \subset U$ is a closed rectangle with sides of length L . Since Df is continuous on the compact set R , we have that $\frac{\partial f_i}{\partial x_j}$ is uniformly continuous on R for $1 \leq i, j \leq n$.

Given $\epsilon > 0$, there exists $\delta > 0$ such that $\left| \frac{\partial f_i}{\partial x_j}(\mathbf{y}) - \frac{\partial f_i}{\partial x_j}(\mathbf{x}) \right| < \frac{\epsilon}{n}$ for $\|\mathbf{y} - \mathbf{x}\| < \delta$ and $1 \leq i, j \leq n$. Choose N large enough such that $n(L/N)^2 < \delta^2$ and partition the rectangle R into N^n subrectangles with sides of length L/N . It follows from our choice of N that the diameter $(L\sqrt{n})/N$ of the subrectangles is less than δ . Hence, if $R_{k_1, k_2, \dots, k_n}$ is one of the subrectangles, then $\left| \frac{\partial f_i}{\partial x_j}(\mathbf{y}) - \frac{\partial f_i}{\partial x_j}(\mathbf{x}) \right| < \frac{\epsilon}{n}$ for all $\mathbf{x}, \mathbf{y} \in R_{k_1, k_2, \dots, k_n}$ and $1 \leq i, j \leq n$ because $\|\mathbf{y} - \mathbf{x}\| < \delta$ for all $\mathbf{x}, \mathbf{y} \in R_{k_1, k_2, \dots, k_n}$.

ii) Suppose that $R_{k_1, k_2, \dots, k_n}$ is one of the subrectangles of R and choose $\mathbf{x} \in R_{k_1, k_2, \dots, k_n}$. Let $g : R_{k_1, k_2, \dots, k_n} \rightarrow \mathbb{R}^n$ be the function defined by $g(\mathbf{y}) = (Df(\mathbf{x}))\mathbf{y} - f(\mathbf{y})$ for $\mathbf{y} \in R_{k_1, k_2, \dots, k_n}$. Since $\left| \frac{\partial g_i}{\partial x_j}(\mathbf{y}) \right| = \left| \frac{\partial f_i}{\partial x_j}(\mathbf{x}) - \frac{\partial f_i}{\partial x_j}(\mathbf{y}) \right| < \frac{\epsilon}{n}$ for $\mathbf{y} \in R_{k_1, k_2, \dots, k_n}$, we get from the Mean Value Theorem and Schwarz inequality that

$$\begin{aligned} |g_i(\mathbf{y}) - g_i(\mathbf{x})| &= \left| \sum_{j=1}^n (g_i(y_1, y_2, \dots, y_j, x_{j+1}, \dots, x_n) - g_i(y_1, y_2, \dots, y_{j-1}, x_j, \dots, x_n)) \right| \\ &= \left| \sum_{j=1}^n \frac{\partial g_i}{\partial x_j}(\xi_j)(y_j - x_j) \right| \leq \left\| \left(\frac{\partial g_i}{\partial x_1}(\xi_j), \frac{\partial g_i}{\partial x_2}(\xi_j), \dots, \frac{\partial g_i}{\partial x_n}(\xi_j) \right) \right\| \|\mathbf{y} - \mathbf{x}\| \\ &\leq \sqrt{n} \left(\frac{\epsilon}{n} \right) \|\mathbf{y} - \mathbf{x}\| \end{aligned} \quad (3.7.1)$$

for $\mathbf{y} \in R_{k_1, k_2, \dots, k_n}$, where $\xi_j \in R_{k_1, k_2, \dots, k_n}$ on the segment joining $(y_1, y_2, \dots, y_j, x_{j+1}, \dots, x_n)$ and $(y_1, y_2, \dots, y_{j-1}, x_j, \dots, x_n)$ is given by the Mean Value Theorem for $1 \leq j \leq n$. Hence

$$\|g(\mathbf{y}) - g(\mathbf{x})\| = \left(\sum_{i=1}^n |g_i(\mathbf{y}) - g_i(\mathbf{x})|^2 \right)^{1/2} \leq \epsilon \|\mathbf{y} - \mathbf{x}\| \quad (3.7.2)$$

for all $\mathbf{y} \in R_{k_1, k_2, \dots, k_n}$. We get

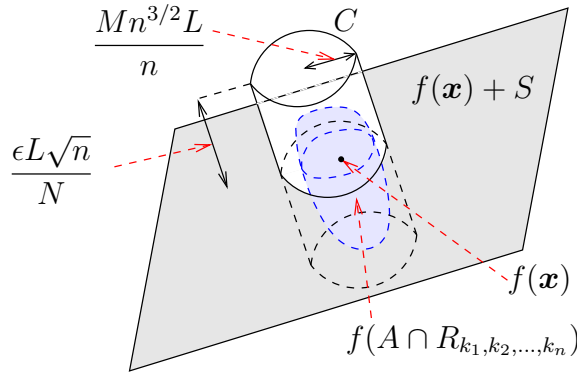
$$\|(Df(\mathbf{x}))(\mathbf{y} - \mathbf{x}) - (f(\mathbf{y}) - f(\mathbf{x}))\| = \|g(\mathbf{y}) - g(\mathbf{x})\| \leq \epsilon \|\mathbf{y} - \mathbf{x}\| < \frac{\epsilon L \sqrt{n}}{N} \quad (3.7.3)$$

for $\mathbf{x}, \mathbf{y} \in R_{k_1, k_2, \dots, k_n}$.

iii) Suppose that $A \cap R_{k_1, k_2, \dots, k_n} \neq \emptyset$ for some subrectangle $R_{k_1, k_2, \dots, k_n}$. Choose $\mathbf{x} \in A \cap R_{k_1, k_2, \dots, k_n}$. Since $Df(\mathbf{x})$ is not onto $\mathbb{R}^n$, its image is a subspace $S \subset \mathbb{R}^n$ of dimension at most $n-1$. In particular, $(Df(\mathbf{x}))(\mathbf{y} - \mathbf{x}) \in S$ for all $\mathbf{y} \in R_{k_1, k_2, \dots, k_n}$. It follows from (3.7.3) that $\text{dist}(S + f(\mathbf{x}), f(\mathbf{y})) = \text{dist}(S, f(\mathbf{y}) - f(\mathbf{x})) \leq (\epsilon L \sqrt{n})/N$ for all $\mathbf{y} \in R_{k_1, k_2, \dots, k_n}$.

Moreover, since $\frac{\partial f_i}{\partial x_j}$ is continuous on the compact set R for $1 \leq i, j \leq n$, they all reach their maximum on R . Thus, there exists $M > 0$ such that $\left| \frac{\partial f_i}{\partial x_j}(\mathbf{z}) \right| \leq M$ for all $\mathbf{z} \in R$ and $1 \leq i, j \leq n$. Proceeding as we did for (3.7.1) and (3.7.2) with f instead of g , we get $|f_i(\mathbf{x}) - f_i(\mathbf{y})| \leq M\sqrt{n} \|\mathbf{x} - \mathbf{y}\|$ and $\|f(\mathbf{x}) - f(\mathbf{y})\| \leq Mn \|\mathbf{x} - \mathbf{y}\| \leq (Mn^{3/2}L)/N$ for all $\mathbf{y} \in R$.

Therefore, when $A \cap R_{k_1, k_2, \dots, k_n} \neq \emptyset$, we have that $f(A \cap R_{k_1, k_2, \dots, k_n})$ is contained in a cylinder C of radius $(Mn^{3/2}L)/N$ and height $2(\epsilon L \sqrt{n})/N$ with an axis perpendicular to the space S .



The volume of C is

$$\begin{aligned} & \left(\frac{2\pi^{(n-1)/2}}{(n-1)\Gamma((n-1)/2)} \right) \left(\frac{Mn^{3/2}L}{N} \right)^{n-1} \left(\frac{2\epsilon L \sqrt{n}}{N} \right) \\ &= \underbrace{\epsilon \left(\frac{4\pi^{(n-1)/2} n^{(3n-1)/2} M^{n-1}}{(n-1)\Gamma((n-1)/2)} \right)}_{\equiv B} \left(\frac{L}{N} \right)^n = \epsilon B \left(\frac{L}{N} \right)^n, \end{aligned}$$

where we have used the fact that the “area” of the cross section perpendicular to the axis of the cylinder is the volume of a sphere of radius $(Mn^{3/2}L)/N$ in the subspace S of dimension $n-1$. The volume of a sphere of radius r in $\mathbb{R}^m$ is $(2\pi^{m/2}r^m)/(m\Gamma(m/2))$ ⁶

Hence $f(A \cap R)$ is contained in the union of N^n cylinders whose total volume is at most $\epsilon B(L/N)^n N^n = \epsilon B L^n$. Since ϵ is arbitrary, we have that $f(A \cap R)$ is of zero content.

⁶Integrating $e^{-\pi\|\mathbf{x}\|_2^2}$ over $\mathbb{R}^m$ using spherical coordinates (i.e. coordinates of the form $\mathbf{x} = r\mathbf{y}$ with $\|\mathbf{y}\| = 1$), the reader can verify that the volume of $B_1(\mathbf{0})$ is $2\pi^{m/2}/(m\Gamma(m/2))$. In particular, the volume of $B_1(\mathbf{0})$ in $\mathbb{R}^2$ and $\mathbb{R}^3$ are respectively π and $4\pi/3$.

iv) Since U is the union of closed rectangles $\{R_i\}_{i \in \mathbb{N}}$ (see remark below) and the countable union of sets of zero content is a set of measure zero, we have that $f(A) = \bigcup_{i \in \mathbb{N}} f(A \cap R_i)$ is of measure zero. ■

Remark 3.7.4

There are several methods to prove that an open set $V \subset \mathbb{R}^n$ is the countable union of closed rectangles. We present two methods. The first one is quick and simple but lacks the constructive approach which is often really useful.

1st Method For each $\mathbf{x} \in U$, choose a closed rectangle $R_{\mathbf{x}} = [a_1(\mathbf{x}), b_1(\mathbf{x})] \times [a_2(\mathbf{x}), b_2(\mathbf{x})] \times \cdots \times [a_n(\mathbf{x}), b_n(\mathbf{x})] \subset U$ such that $\mathbf{x} \in R_{\mathbf{x}}$ and $a_j(\mathbf{x}), b_j(\mathbf{x}) \in \mathbb{Q}$ for $1 \leq j \leq n$. Then $U = \bigcup_{\mathbf{x} \in V} R_{\mathbf{x}}$. Surprisingly, the number of rectangles is countable because $\{(a_1(\mathbf{x}), b_1(\mathbf{x}), a_2(\mathbf{x}), b_2(\mathbf{x}), \dots, a_n(\mathbf{x}), b_n(\mathbf{x})) : \mathbf{x} \in U\} \subset \mathbb{Q}^{2n}$ and a subset of a countable set is a countable set.

2nd Method Let $U_j = \{\mathbf{x} \in U : \|\mathbf{x}\| < j \text{ and } \text{dist}(\mathbf{x}, \partial U) > 1/j\}$ for $j > 0$ and $V_0 = \emptyset$. Note that many of the sets U_j may be empty but that does not matter. Moreover, let $K_j = \overline{U_j} \setminus U_{j-1}$ for $j > 0$. We have that the K_j are compact subsets of U and $U = \bigcup_{j>0} K_j$.

For each $\mathbf{x} \in K_j$, choose a closed rectangle $R_{\mathbf{x}} \subset U$ such that $\mathbf{x} \in R_{\mathbf{x}}$. Since $\{R_{\mathbf{x}}^\circ\}_{\mathbf{x} \in K_j}$ is an open cover of the compact set K_j , there exists a finite subcover $\{R_{j,i}^\circ\}_{1 \leq i \leq I_j}$ of K_j . The union of the $\{R_{j,i}\}_{1 \leq i \leq I_j}$ for $j > 0$ is a countable collection of closed rectangles such that $U = \bigcup_{1 \leq i \leq I_j, j>0} R_{j,i}$. ♠

3.8 Exercises

Question 3.1

Suppose that $f : [a, b] \rightarrow \mathbb{R}$ is Riemann integrable. Prove that the graph of f has zero content.

Solution

Given $\epsilon > 0$, we may choose a partition $P = \{p_i : 0 \leq i \leq N\}$ of $[a, b]$ such that $\mathcal{U}_P(f) - \mathcal{L}_P(f) < \epsilon$ because f is Riemann integrable.

As usual, let $m_i = \inf\{f(x) : p_{i-1} \leq x \leq p_i\}$ and $M_i = \sup\{f(x) : p_{i-1} \leq x \leq p_i\}$ for $0 < i \leq N$. Consider the rectangles $E_i = \{(x, y) : p_{i-1} \leq x \leq p_i \text{ and } m_i \leq y \leq M_i\}$ for $0 < i \leq N$. We have that the graph of f is a subset of $E = \bigcup_{i=1}^N E_i$ and the area of E is

$$A(E) = \sum_{i=1}^N (M_i - m_i)(p_i - p_{i-1}) = \mathcal{U}_P(f) - \mathcal{L}_P(f) < \epsilon.$$

Since ϵ is arbitrary, this proves that the graph of f has content zero. ♥

Question 3.2

Suppose that $f : [a, b] \rightarrow [0, \infty[$ is Riemann integrable. Let $S = \{(x, y) : a \leq x \leq b \text{ and } 0 \leq y \leq f(x)\}$. Prove that S is Jordan measurable and the area $A(S)$ of S is equal to $\int_a^b f(x) dx$.

Solution

i) Since f is Riemann integrable, f is bounded. Thus, there exists $C > 0$ such that $\sup\{f(x) : a \leq x \leq b\} < C$. Since $S \subset R = \{(x, y) : a \leq x \leq b \text{ and } 0 \leq y \leq C\}$, S is bounded. Moreover, ∂S is the union of four sets of zero content and so of measure zero. The graph of f (including the vertical lines joining points of discontinuity) has content zero according to Question 3.1. The sets $\{(x, y) : x = a \text{ and } 0 \leq y \leq f(a)\}$, $\{(x, y) : x = b \text{ and } 0 \leq y \leq f(b)\}$ and $\{(x, y) : y = 0 \text{ and } a \leq x \leq b\}$ have all zero content according to Proposition 3.2.2. This proves that S is Jordan measurable.

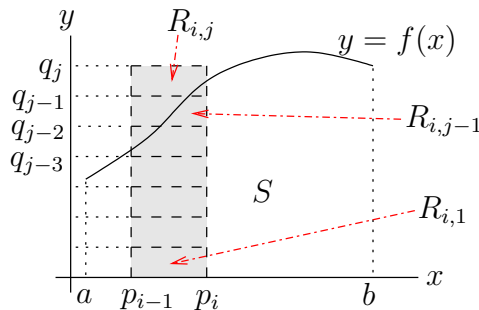
ii) Since S is Jordan measurable, we have that the characteristic function $\mathcal{X}_S$ is Riemann integrable on R defined above. Given $\epsilon > 0$, choose a partition $\{P, Q\}$ of R such that

$$\mathcal{U}_{\{P, Q\}}(\mathcal{X}_S) - \mathcal{L}_{\{P, Q\}}(\mathcal{X}_S) < \epsilon. \quad (3.8.1)$$

We assume that $P = \{p_i : 0 \leq i \leq N\}$ and $Q = \{q_j : 0 \leq j \leq M\}$. Let $m_i = \inf\{f(x) : p_{i-1} \leq x \leq p_i\}$, $M_i = \sup\{f(x) : p_{i-1} \leq x \leq p_i\}$ and $R_{i,j} = \{(x, y) : p_{i-1} \leq x \leq p_i \text{ and } q_{j-1} \leq y \leq q_j\}$. Then

$$\begin{aligned} \mathcal{U}_P(f) &= \sum_{i=1}^N M_i(p_i - p_{i-1}) \leq \sum_{i=1}^N \left(\sum_{R_{i,j} \cap S \neq \emptyset} (q_j - q_{j-1}) \right) (p_i - p_{i-1}) \\ &= \sum_{R_{i,j} \cap S \neq \emptyset} (q_j - q_{j-1})(p_i - p_{i-1}) = \mathcal{U}_{\{P, Q\}}(\mathcal{X}_S). \end{aligned}$$

The following figure shows the rectangle $R_{i,j}$ included in the previous summation.



Similarly,

$$\mathcal{L}_P(f) = \sum_{i=1}^N m_i(p_i - p_{i-1}) \geq \sum_{i=1}^N \left(\sum_{R_{i,j} \subset S} (q_j - q_{j-1}) \right) (p_i - p_{i-1})$$

$$= \sum_{R_{i,j} \subset S} (q_j - q_{j-1})(p_i - p_{i-1}) = \mathcal{L}_{\{P,Q\}}(\mathcal{X}_S) .$$

Thus

$$\mathcal{L}_{\{P,Q\}}(\mathcal{X}_S) \leq \mathcal{L}_P(f) \leq \int_a^b f(x) dx \leq \mathcal{U}_P(f) \leq \mathcal{U}_{\{P,Q\}}(\mathcal{X}_S)$$

and

$$\mathcal{L}_{\{P,Q\}}(\mathcal{X}_S) \leq A(S) = \int_R \mathcal{X}_S \leq \mathcal{U}_{\{P,Q\}}(\mathcal{X}_S) .$$

It follows from (3.8.1) that $\left| A(S) - \int_a^b f(x) dx \right| < \epsilon$. Since ϵ is arbitrary, we get $A(S) = \int_a^b f(x) dx$. ♥

Question 3.3

Let S be a bounded subset of $\mathbb{R}^2$. Prove that S and S° have the same inner area; namely, show that $\mathcal{L}(\mathcal{X}_S) = \mathcal{L}(\mathcal{X}_{S^\circ})$. Note that we do not require S to be Jordan measurable. If it was, then the question would be easy to answer because $S \setminus S^\circ \subset \partial S$ would be a subset of the closed set of measure zero ∂S and we will be able to use Remark 3.1.17.

Solution

i) We first prove that $\mathcal{L}(\mathcal{X}_{S^\circ}) \leq \mathcal{L}(\mathcal{X}_S)$. Let R be a rectangle containing S and $\{P, Q\}$ be a partition of R . We assume that $P = \{p_i : 0 \leq i \leq N\}$ and $Q = \{q_j : 0 \leq j \leq M\}$, and let $R_{i,j} = \{(x, y) : p_{i-1} \leq x \leq p_i \text{ and } q_{j-1} \leq y \leq q_j\}$ for $0 < i \leq N$ and $0 < j \leq M$. Given $V \subset \mathbb{R}^2$, let

$$m_{i,j}(V) = \inf\{\mathcal{X}_V(x, y) : (x, y) \in R_{i,j}\} = \begin{cases} 1 & \text{if } R_{i,j} \subset V \\ 0 & \text{otherwise} \end{cases}$$

for $0 < i \leq N$ and $0 < j \leq M$. We have that $m_{i,j}(S^\circ) \leq m_{i,j}(S)$ for $0 < i \leq N$ and $0 < j \leq M$. Hence

$$\begin{aligned} \mathcal{L}_{\{P,Q\}}(\mathcal{X}_{S^\circ}) &= \sum_{R_{i,j} \subset S^\circ} (p_i - p_{i-1})(q_j - q_{j-1}) \leq \sum_{R_{i,j} \subset S} (p_i - p_{i-1})(q_j - q_{j-1}) \\ &= \mathcal{L}_{\{P,Q\}}(\mathcal{X}_S) \leq \mathcal{L}(\mathcal{X}_S) . \end{aligned}$$

Since this is true for all partition $\{P, Q\}$ of R , we get

$$\mathcal{L}(\mathcal{X}_{S^\circ}) = \sup \{ \mathcal{L}_{\{P,Q\}}(\mathcal{X}_{S^\circ}) : \{P, Q\} \text{ is a partition of } R \} \leq \mathcal{L}(\mathcal{X}_S) .$$

ii) We prove that $\mathcal{L}(\mathcal{X}_S) \leq \mathcal{L}(\mathcal{X}_{S^\circ})$. Let R be a rectangle containing S and $\{P, Q\}$ be a partition of R . We assume that P, Q and $R_{i,j}$ are as defined in (i).

Given $\epsilon > 0$, choose $\delta < \frac{\epsilon}{\mathcal{L}(\mathcal{X}_S)}$. We assume that $\mathcal{L}(\mathcal{X}_S) > 0$ otherwise $0 = \mathcal{L}(\mathcal{X}_S) \leq \mathcal{L}(\mathcal{X}_{S^\circ})$ and there is nothing to prove. For each rectangle $R_{i,j} \subset S$, we construct a rectangle $\check{R}_{i,j} \subset R_{i,j}^\circ \subset S^\circ$ by shrinking the rectangle $R_{i,j}$ by a factor of $\sqrt{1 - \delta}$ in both directions. We

have $A(\check{R}_{i,j}) = (1 - \delta)A(R_{i,j})$ and $\check{m}_{i,j} = \inf\{\mathcal{X}_{S^\circ}(x, y) : (x, y) \in \check{R}_{i,j}\} = m_{i,j}(S) = 1$ for all $R_{i,j} \subset S$.

We choose a refinement $\{\tilde{P}, \tilde{Q}\}$ of the partition $\{P, Q\}$ that includes all the vertices (i.e. the corners) of the rectangles $\check{R}_{i,j}$. Let $\tilde{R}_{r,s}$ be the rectangles associated to the partition $\{\tilde{P}, \tilde{Q}\}$. Note that each rectangle $\check{R}_{i,j}$ is represented by one and only one of the $\tilde{R}_{r,s}$. Therefore,

$$\begin{aligned} \mathcal{L}_{\{\tilde{P}, \tilde{Q}\}}(\mathcal{X}_{S^\circ}) &= \sum_{\tilde{R}_{r,s} \subset S^\circ} A(\tilde{R}_{r,s}) \geq \sum_{R_{i,j} \subset S} A(\check{R}_{i,j}) = (1 - \delta) \sum_{R_{i,j} \subset S} m_{i,j} A(R_{i,j}) \\ &= (1 - \delta) \mathcal{L}_{\{P, Q\}}(\mathcal{X}_S) \geq \left(1 - \frac{\epsilon}{\mathcal{L}(\mathcal{X}_S)}\right) \mathcal{L}_{\{P, Q\}}(\mathcal{X}_S). \end{aligned}$$

Since this relation is true for all refinements $\{\tilde{P}, \tilde{Q}\}$ of the partitions $\{P, Q\}$ containing the $\check{R}_{i,j}$, we get

$$\mathcal{L}(\mathcal{X}_{S^\circ}) \geq \left(1 - \frac{\epsilon}{\mathcal{L}(\mathcal{X}_S)}\right) \mathcal{L}_{\{P, Q\}}(\mathcal{X}_S).$$

Since this relation is true for all partitions $\{P, Q\}$ of R , we get

$$\mathcal{L}(\mathcal{X}_{S^\circ}) \geq \left(1 - \frac{\epsilon}{\mathcal{L}(\mathcal{X}_S)}\right) \mathcal{L}(\mathcal{X}_S) = \mathcal{L}(\mathcal{X}_S) - \epsilon.$$

Finally, since ϵ is arbitrary, we get $\mathcal{L}(\mathcal{X}_{S^\circ}) \geq \mathcal{L}(\mathcal{X}_S)$. ♡

Question 3.4

Let S be a bounded subset of $\mathbb{R}^2$. Prove that S and $\bar{S}$ have the same outer area; namely, show that $\mathcal{U}(\mathcal{X}_S) = \mathcal{U}(\mathcal{X}_{\bar{S}})$. As for question 3.3, we do not require S to be Jordan measurable.

Solution

The proof is similar to the proof given for Question 3.3. ♡

Question 3.5

Let S be a bounded subset of $\mathbb{R}^2$. Prove that outer area of S is equal to the inner area of S plus the outer area of ∂S ; namely, prove that $\mathcal{U}(\mathcal{X}_S) = \mathcal{L}(\mathcal{X}_{\bar{S}}) + \mathcal{U}(\mathcal{X}_{\partial S})$. Again, as for question 3.3, we do not require S to be Jordan measurable.

Solution

We have that $\partial S = \bar{S} \setminus S^\circ$.

i) We first prove that $\mathcal{U}(\mathcal{X}_S) \leq \mathcal{L}(\mathcal{X}_S) + \mathcal{U}(\mathcal{X}_{\partial S})$. According to the result of Question 3.3, it is enough to prove that $\mathcal{U}(\mathcal{X}_S) \leq \mathcal{L}(\mathcal{X}_{S^\circ}) + \mathcal{U}(\mathcal{X}_{\partial S})$.

Let $\{P, Q\}$ be a partition of a rectangle $R \supset \bar{S}$. As usual, we assume that $P = \{p_i : 0 \leq i \leq N\}$ and $Q = \{q_j : 0 \leq j \leq M\}$, and let $R_{i,j} = \{(x, y) : p_{i-1} \leq x \leq p_i \text{ and } q_{j-1} \leq y \leq q_j\}$ for $0 < i \leq N$ and $0 < j \leq M$. We have

$$M_{i,j}(V) = \sup\{\mathcal{X}_V(x, y) : (x, y) \in R_{i,j}\} = \begin{cases} 1 & \text{if } R_{i,j} \cap V \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

and

$$m_{i,j}(V) = \inf\{\mathcal{X}_V(x, y) : (x, y) \in R_{i,j}\} = \begin{cases} 1 & \text{if } R_{i,j} \subset V \\ 0 & \text{otherwise} \end{cases}$$

for all $V \subset \mathbb{R}^2$. Note that $m_{i,j}(V) = M_{i,j}(V)$ for $R_{i,j} \subset V$. Hence

$$\begin{aligned} \mathcal{U}(\mathcal{X}_S) &= \mathcal{U}(\mathcal{X}_{\bar{S}}) \leq \mathcal{U}_{\{P,Q\}}(\mathcal{X}_{\bar{S}}) = \sum_{R_{i,j} \cap \bar{S} \neq \emptyset} A(R_{i,j}) = \sum_{R_{i,j} \subset S^\circ} A(R_{i,j}) + \sum_{R_{i,j} \cap \partial S \neq \emptyset} A(R_{i,j}) \\ &= \mathcal{L}_{\{P,Q\}}(\mathcal{X}_{S^\circ}) + \mathcal{U}_{\{P,Q\}}(\mathcal{X}_{\partial S}) \leq \mathcal{L}(\mathcal{X}_{S^\circ}) + \mathcal{U}_{\{P,Q\}}(\mathcal{X}_{\partial S}), \end{aligned}$$

where the first equality is a consequence of the result of Question 3.4. Since this is true for all partitions $\{P, Q\}$ of R , we get $\mathcal{U}(\mathcal{X}_S) \leq \mathcal{L}(\mathcal{X}_{S^\circ}) + \mathcal{U}(\mathcal{X}_{\partial S})$.

ii) We now prove that $\mathcal{U}(\mathcal{X}_S) \geq \mathcal{L}(\mathcal{X}_S) + \mathcal{U}(\mathcal{X}_{\partial S})$. Again, according to the result of Question 3.3, it is enough to prove that $\mathcal{U}(\mathcal{X}_S) \geq \mathcal{L}(\mathcal{X}_{S^\circ}) + \mathcal{U}(\mathcal{X}_{\partial S})$.

Let R be rectangle containing $\bar{S}$. Given $\epsilon > 0$, choose a partition $\{\tilde{P}, \tilde{Q}\}$ of R such that $\mathcal{L}_{\{\tilde{P}, \tilde{Q}\}}(\mathcal{X}_{S^\circ}) + \epsilon > \mathcal{L}(\mathcal{X}_{S^\circ})$. Let $\{P, Q\}$ be a refinement of $\{\tilde{P}, \tilde{Q}\}$. We assume that P, Q and $R_{i,j}$ are as defined in (i). Then

$$\begin{aligned} \mathcal{U}_{\{P,Q\}}(\mathcal{X}_S) &= \mathcal{U}_{\{P,Q\}}(\mathcal{X}_{\bar{S}}) = \sum_{R_{i,j} \cap \bar{S} \neq \emptyset} A(R_{i,j}) = \sum_{R_{i,j} \subset S^\circ} A(R_{i,j}) + \sum_{R_{i,j} \cap \partial S \neq \emptyset} A(R_{i,j}) \\ &= \mathcal{L}_{\{P,Q\}}(\mathcal{X}_{S^\circ}) + \mathcal{U}_{\{P,Q\}}(\mathcal{X}_{\partial S}) \geq \mathcal{L}_{\{\tilde{P}, \tilde{Q}\}}(\mathcal{X}_{S^\circ}) + \mathcal{U}(\mathcal{X}_{\partial S}) \\ &> \mathcal{L}(\mathcal{X}_{S^\circ}) - \epsilon + \mathcal{U}(\mathcal{X}_{\partial S}), \end{aligned}$$

where the first equality is again a consequence of the result of Question 3.4. Since this is true for any refinement $\{P, Q\}$ of $\{\tilde{P}, \tilde{Q}\}$, we get $\mathcal{U}(\mathcal{X}_S) \geq \mathcal{L}(\mathcal{X}_{S^\circ}) - \epsilon + \mathcal{U}(\mathcal{X}_{\partial S})$. Since this last inequality is true for all $\epsilon > 0$, we finally get $\mathcal{U}(\mathcal{X}_S) \geq \mathcal{L}(\mathcal{X}_{S^\circ}) + \mathcal{U}(\mathcal{X}_{\partial S})$. $\heartsuit$

Question 3.6

Prove the **Mean Value Theorem for integrals**. Namely, let $S \subset \mathbb{R}^n$ be a compact, convex and Jordan measurable set. If $g : S \rightarrow \mathbb{R}$ is Riemann integrable, $g(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in S$, and $f : S \rightarrow \mathbb{R}$ is continuous, then there exists $\mathbf{a} \in S$ such that $\int_S f(\mathbf{x})g(\mathbf{x}) d\mathbf{x} =$

$$f(\mathbf{a}) \int_S g(\mathbf{x}) d\mathbf{x}.$$

Solution

Since $f, g : S \rightarrow \mathbb{R}$ are Riemann integrable, we have that $fg : S \rightarrow \mathbb{R}$ is Riemann integrable (Question 2.13). Let $m = \inf\{f(\mathbf{x}) : \mathbf{x} \in S\}$ and $M = \sup\{f(\mathbf{x}) : \mathbf{x} \in S\}$. m and M are real number because f is continuous on the compact set S . In fact, $m = f(\mathbf{x}_1)$ and $M = f(\mathbf{x}_2)$ for some $\mathbf{x}_1$ and $\mathbf{x}_2$ in S .

Since $m \leq f(\mathbf{x}) \leq M$ for all $\mathbf{x} \in S$, we get

$$m \int_S g(\mathbf{x}) d\mathbf{x} \leq \int_S g(\mathbf{x})f(\mathbf{x}) d\mathbf{x} \leq M \int_S g(\mathbf{x}) d\mathbf{x}.$$

The result is trivial if $\int_S g(\mathbf{x}) \, d\mathbf{x} = 0$. We therefore assume that $\int_S g(\mathbf{x}) \, d\mathbf{x} > 0$. Thus

$$m \leq \frac{\int_S g(\mathbf{x}) f(\mathbf{x}) \, d\mathbf{x}}{\int_S g(\mathbf{x}) \, d\mathbf{x}} \leq M.$$

Since f is continuous on S and the line $\ell = \{\mathbf{x}_1 + t(\mathbf{x}_2 - \mathbf{x}_1) : 0 \leq t \leq 1\}$ between $\mathbf{x}_1$ and $\mathbf{x}_2$ is completely inside S because S is convex, we may use the Intermediate Value Theorem to find $t_1 \in [0, 1]$ such that

$$f(\mathbf{x}_1 + t_1(\mathbf{x}_2 - \mathbf{x}_1)) = \frac{\int_S g(\mathbf{x}) f(\mathbf{x}) \, d\mathbf{x}}{\int_S g(\mathbf{x}) \, d\mathbf{x}}.$$

Hence, we may take $\mathbf{a} = \mathbf{x}_1 + t_1(\mathbf{x}_2 - \mathbf{x}_1)$.

♥

Question 3.7

Let

$$f(x, y) = \begin{cases} 1 & \text{if } \mathbf{x} \in \mathbb{Q} \\ 2y & \text{if } \mathbf{x} \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

Prove that $\int_0^1 \int_0^1 f(x, y) \, dy \, dx = 1$ but $\int_0^1 \int_0^1 f(x, y) \, dx \, dy$ does not exist.

Question 3.8

Consider the function

$$f(x, y) = \begin{cases} \frac{1}{q} & \text{if } x = p/q \text{ where } 0 < p \leq q \text{ are relatively prime} \\ 0 & \text{otherwise} \end{cases}$$

Determine if f is Riemann integrable on $[0, 1] \times [0, 1]$.

Question 3.9

Consider the function

$$f(x, y) = \begin{cases} y^{-2} & \text{if } 0 < x < y < 1 \\ -x^{-2} & \text{if } 0 < y < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Let S be the square $S = [0, 1] \times [0, 1]$.

- Prove that f is not integrable on S .
- Prove that $x \rightarrow f(x, y)$ is integrable on $[0, 1]$ for all $y \in [0, 1]$ and compute the integral.
- Prove that $y \rightarrow f(x, y)$ is integrable on $[0, 1]$ for all $x \in [0, 1]$ and compute the integral.
- Prove that $\int_0^1 \int_0^1 f(x, y) \, dx \, dy$ and $\int_0^1 \int_0^1 f(x, y) \, dy \, dx$ exist but are not equal.

Solution

- The function f is not integrable on S because f is not bounded on S . For instance, $f(z/2, z) = 1/z^2 \rightarrow \infty$ as $z \rightarrow 0$ and $f(z, z/2) = -1/z^2 \rightarrow -\infty$ as $z \rightarrow 0$.

b) For $y \in]0, 1[$ fixed, we have

$$f_y(x) = f(x, y) = \begin{cases} y^{-2} & \text{if } 0 < x < y < 1 \\ -x^{-2} & \text{if } 0 < y < x < 1 \\ 0 & \text{if } x = 0, y \text{ or } 1 \end{cases}$$

The function f_y is integrable because it is a bounded function and the set of points where f_y is discontinuous is $\{0, y, 1\}$ which is a set of zero content. We have

$$\begin{aligned} \int_0^1 f(x, y) dx &= \int_0^1 f_y(x) dx = \int_0^y f_y(x) dx + \int_y^1 f_y(x) dx = \int_0^y \frac{1}{y^2} dx - \int_y^1 \frac{1}{x^2} dx \\ &= \frac{x}{y^2} \Big|_{x=0}^y + \frac{1}{x} \Big|_{x=y}^1 = \frac{1}{y} + 1 - \frac{1}{y} = 1. \end{aligned}$$

For $y = 0$ or $y = 1$, we have that $f_y(x) = f(x, y) = 0$ for $0 \leq x \leq 1$. Hence f_y is integrable and $\int_0^1 f_y(x) dx = 0$.

c) For $x \in]0, 1[$ fixed, we have

$$f_x(y) = f(x, y) = \begin{cases} -x^{-2} & \text{if } 0 < y < x < 1 \\ y^{-2} & \text{if } 0 < x < y < 1 \\ 0 & \text{if } y = 0, x \text{ or } 1 \end{cases}$$

The function f_x is integrable because it is a bounded function and the set of points where f_x is discontinuous is the set $\{0, x, 1\}$ of zero content. We have

$$\begin{aligned} \int_0^1 f(x, y) dy &= \int_0^1 f_x(y) dy = \int_0^x f_x(y) dy + \int_x^1 f_x(y) dy = - \int_0^x \frac{1}{x^2} dy + \int_x^1 \frac{1}{y^2} dy \\ &= -\frac{y}{x^2} \Big|_{y=0}^x - \frac{1}{y} \Big|_{y=x}^1 = -\frac{1}{x} - 1 + \frac{1}{x} = -1. \end{aligned}$$

For $x = 0$ or $x = 1$, we have that $f_x(y) = f(x, y) = 0$ for $0 \leq y \leq 1$. Hence f_x is integrable and $\int_0^1 f_x(y) dy = 0$.

d) Let $g(x) = \int_0^1 f_x(y) dy$ for $0 \leq x \leq 1$ and $h(y) = \int_0^1 f_y(x) dx$ for $0 \leq y \leq 1$. We have

$$g(x) = \begin{cases} -1 & \text{if } 0 < x < 1 \\ 0 & \text{if } x = 0, 1 \end{cases} \quad \text{and} \quad h(y) = \begin{cases} 1 & \text{if } 0 < y < 1 \\ 0 & \text{if } y = 0, 1 \end{cases}$$

Since they are two continuous functions on $]0, 1[$ and $\{0, 1\}$ is a set of zero content, they are both Riemann integrable on $[0, 1]$. In fact, $\int_0^1 \int_0^1 f(x, y) dy dx = \int_0^1 g(x) dx = -1$ and

$$\int_0^1 \int_0^1 f(x, y) dx dy = \int_0^1 h(y) dy = 1.$$

♡

3.8.1 Supplementary Topics

The following questions are traditional questions that can be found in advanced calculus textbooks. Nevertheless, we include them here for the benefit of the readers who may not have had a solid advanced calculus course. They are problems that any mathematics students should be able to solve.

Multiple Integrals

The most common double integrals are illustrated below. The reader should have not problem to generalize this presentation to multiple integrals in higher dimensions. If the domain of integration is $R = \{(x, y) : a \leq x \leq b \text{ and } g_1(x) \leq y \leq g_2(x)\}$ (Figure 3.1), then it follows from Fubini's theorem on a rectangle containing R that

$$\iint_R f = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x, y) \, dy \, dx .$$

Likewise, if the domain of integration is $R = \{(x, y) : c \leq y \leq d \text{ and } h_1(y) \leq x \leq h_2(y)\}$ (Figure 3.1), it follows from Fubini's theorem on a rectangle containing R that

$$\iint_R f = \int_c^d \int_{h_1(y)}^{h_2(y)} f(x, y) \, dx \, dy .$$

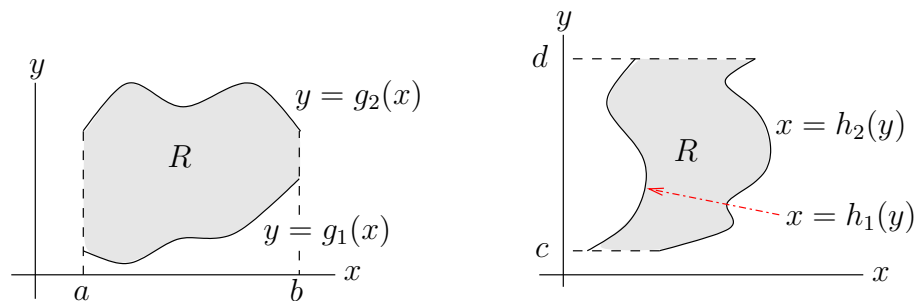


Figure 3.1: We have drawn two of the most frequent domains of integration for double integrals.

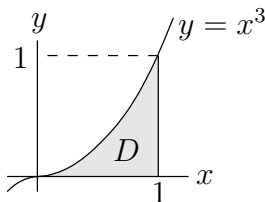
In the questions below, we will often use Theorem 3.6.1 in situations that do not perfectly meet the hypothesis of this theorem for a change of variables. If $f : U \rightarrow \mathbb{R}$ is an integrable function on an open set U and the change of variables meet all the hypothesis on U except on a closed set A of measure zero, then we may use Theorem 3.6.1 on $U \setminus A$. It follows from Remark 3.1.17 that the value of the integral of f on U is equal to the value of the integral of f on $U \setminus A$. This is particularly useful when using polar or spherical coordinates because these changes of variables are not properly defined at the origin.

Question 3.10

Consider the region D enclosed by the curve $y = x^3$, and the lines $x = 1$ and $y = 0$. Give two integrals to compute the mass of this thin region if the density is given by $\rho(x, y) = \sqrt{x^4 + 1}$. Compute the mass of this region using one of the integrals that you have given.

Solution

We have to compute the mass of the region represented in the following figure.



The mass is given by the formula $m = \iint_D \rho = \int_0^1 \int_0^{x^3} \rho(x, y) \, dy \, dx = \int_0^1 \int_{y^{1/3}}^1 \rho(x, y) \, dx \, dy$.

It is easier to compute the mass of the region D with the first double integral than the second one (the reader should try to compute the second integral.) Hence, the mass is

$$\begin{aligned} m &= \int_0^1 \int_0^{x^3} \sqrt{x^4 + 1} \, dy \, dx = \int_0^1 y \sqrt{x^4 + 1} \Big|_{y=0}^{x^3} \, dx = \int_0^1 x^3 \sqrt{x^4 + 1} \, dx \\ &= \frac{1}{6} (x^4 + 1)^{3/2} \Big|_0^1 = \frac{1}{6} (2\sqrt{2} - 1) \end{aligned}$$

♡

Question 3.11

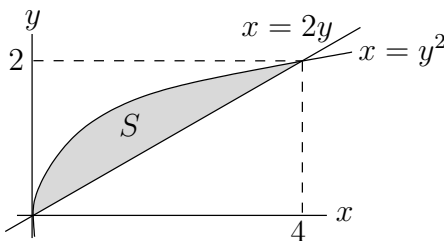
Evaluate $I = \iint_S (x^2 - \sqrt{y})$, where S is the region enclosed by the parabola $x = y^2$ and the line $x = 2y$.

Solution

The ordinates (y values) of the points of intersection of $x = y^2$ and $x = 2y$ are given by

$$y^2 = 2y \Rightarrow y(y - 2) = 0 \Rightarrow y = 0 \text{ or } y = 2.$$

We find the two intersection points $(0, 0)$ and $(4, 2)$. We have drawn the region below.



We have $I = \int_0^4 \int_{x/2}^{\sqrt{x}} (x^2 - \sqrt{y}) \, dy \, dx = \int_0^2 \int_{y^2}^{2y} (x^2 - \sqrt{y}) \, dx \, dy$. We evaluate the second integral.

$$\begin{aligned} I &= \int_0^2 \left(\frac{x^3}{3} - x\sqrt{y} \right) \Big|_{y^2}^{2y} \, dx = \int_0^2 \left(\frac{8y^3}{3} - 2y^{3/2} - \frac{y^6}{3} + y^{5/2} \right) \, dy \\ &= \left(\frac{2y^4}{3} - \frac{4y^{5/2}}{5} - \frac{y^7}{21} + \frac{2y^{7/2}}{7} \right) \Big|_0^2 = \frac{32}{7} - \frac{32\sqrt{2}}{35} . \end{aligned}$$

♡

Question 3.12

Express the following double integrals $\iint_S f$ in terms of iterated integrals in two different ways.

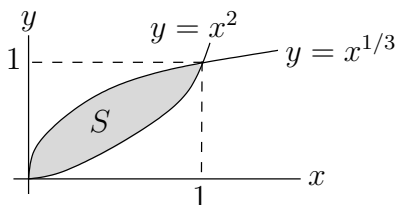
- a) S is the region enclosed by $y = x^{1/3}$ and $y = x^2$ for $y \geq 0$.
- b) S is the triangle with vertices at $(0, 0)$, $(-1, 1)$ and $(2, 1)$.
- c) S is the region enclosed by the x axis, the line $x = 2$ and the curve $y = \ln(x)$.

Solution

- a) The abscissas (x values) of the points of intersection of $y = x^2$ with $y = x^{1/3}$ are given by

$$x^2 = x^{1/3} \Rightarrow x(x^5 - 1) = 0 \Rightarrow x = 0 \text{ or } x = 1 .$$

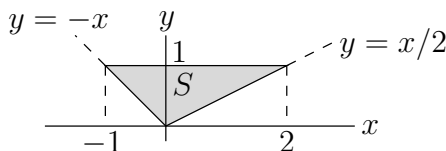
The points of intersection are $(0, 0)$ and $(1, 1)$. The domain of integration is shown in the following figure.



Hence

$$\iint_S f = \int_0^1 \int_{x^2}^{x^{1/3}} f(x, y) \, dy \, dx = \int_0^1 \int_{y^3}^{\sqrt{y}} f(x, y) \, dx \, dy .$$

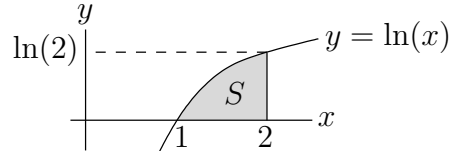
- b) The domain of integration is shown in the following figure.



Hence

$$\iint_S f = \int_0^1 \int_{-y}^{2y} f(x, y) \, dx \, dy = \int_{-1}^0 \int_{-x}^1 f(x, y) \, dy \, dx + \int_0^2 \int_{x/2}^1 f(x, y) \, dy \, dx .$$

c) The domain of integration is shown in the following figure.



Hence

$$\iint_S f = \int_1^2 \int_0^{\ln(x)} f(x, y) \, dy \, dx = \int_0^{\ln(2)} \int_{e^y}^2 f(x, y) \, dx \, dy .$$

♡

Question 3.13

Express the following double integrals $\iint_S f$ in terms of iterated integrals in two different ways.

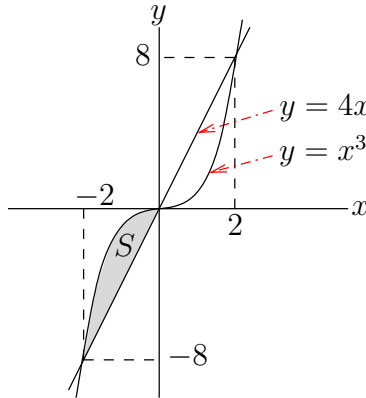
- a) S is the region enclosed by $y = x^3$ and $y = 4x$ for $y \leq 0$.
- b) S is the triangle with vertices at $(0, 0)$, $(2, 2)$ and $(3, 1)$.
- c) S is the region enclosed by the parabolas $y = x^2$ and $y = 6 - 4x - x^2$.

Solution

a) The abscissas (x values) of the points of intersection of $y = x^3$ with $y = 4x$ are given by

$$x^3 = 4x \Rightarrow x(x^2 - 4) = 0 \Rightarrow x = 0, x = -2 \text{ or } x = 2 .$$

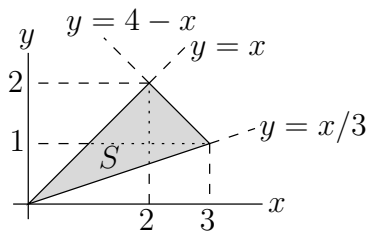
The points of intersection are $(-2, -8)$, $(0, 0)$ and $(2, 8)$. The domain of integration is shown in the following figure.



Hence

$$\iint_S f = \int_{-2}^0 \int_{4x}^{x^3} f(x, y) \, dy \, dx = \int_{-8}^0 \int_{y^{1/3}}^{y/4} f(x, y) \, dx \, dy .$$

b) The domain of integration is shown in the following figure.



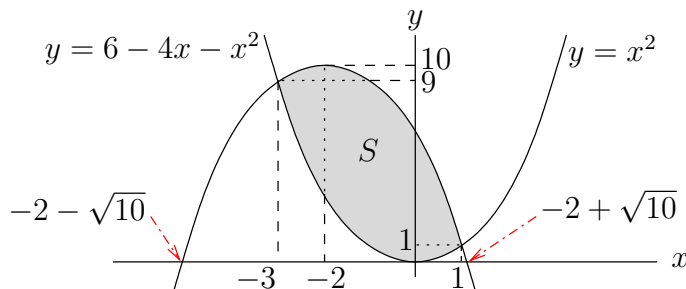
Hence

$$\begin{aligned} \iint_S f &= \int_0^2 \int_{x/3}^x f(x, y) \, dy \, dx + \int_2^3 \int_{x/3}^{4-x} f(x, y) \, dy \, dx \\ &= \int_0^1 \int_y^{3y} f(x, y) \, dx \, dy + \int_1^2 \int_y^{4-y} f(x, y) \, dx \, dy . \end{aligned}$$

c) The abscissas of the points of intersection of the two parabolas $y = x^2$ and $y = 6 - 4x - x^2$ are given by

$$x^2 = 6 - 4x - x^2 \Rightarrow 2x^2 + 4x - 6 = 0 \Rightarrow 2(x - 1)(x + 3) = 0 \Rightarrow x = -3 \text{ or } x = 1 .$$

The points of intersection are $(-3, 9)$ and $(1, 1)$. The domain of integration is shown in the following figure.



Hence

$$\begin{aligned} \iint_S f &= \int_{-3}^1 \int_{x^2}^{6-4x-x^2} f(x, y) \, dy \, dx = \int_0^1 \int_{-\sqrt{y}}^{\sqrt{y}} f(x, y) \, dx \, dy \\ &\quad + \int_1^9 \int_{-2+\sqrt{10-y}}^{-2-\sqrt{10-y}} f(x, y) \, dx \, dy + \int_9^{10} \int_{-2-\sqrt{10-y}}^{-2+\sqrt{10-y}} f(x, y) \, dx \, dy . \end{aligned}$$

To get the limits of integration, we have use the relation

$$y = 6 - 4x - x^2 \Leftrightarrow x^2 + 4x + (y - 6) = 0 \Leftrightarrow x = \frac{-4 \pm \sqrt{4^2 - 4(y - 6)}}{2} = -2 \pm \sqrt{10 - y}$$

for $y \leq 10$.



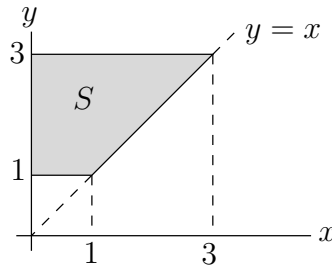
Question 3.14

Evaluate the integral $\iint_S f$ in each of the following cases.

- a) $f(x, y) = ye^{2x}$ and S is the region enclosed by $y = x$, $y = 1$, $y = 3$ and the y axis.
- b) $f(x, y) = \cos(y^3 + 1)$ and S is the region enclosed by $y = \sqrt{x}$, $y = 1$ and the y axis.
- c) $f(x, y) = ye^{xy}$ and S is the region enclosed by $y = 1/x$, $x = 2$ and $y = 1$.

Solution

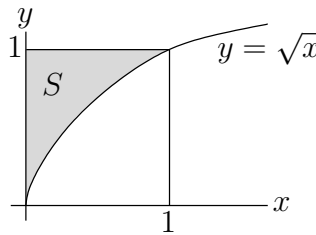
- a) The domain of integration is shown in the following figure.



To avoid having to split the domain of integration in two, it is preferable to integrate with respect to x first than with respect to y . Thus

$$\begin{aligned} \iint_S ye^{2x} &= \int_1^3 \int_0^y ye^{2x} dx dy = \int_1^3 \left(\frac{ye^{2x}}{2} \Big|_{x=0}^y \right) dy \\ &= \frac{1}{2} \int_1^3 (ye^{2y} - y) dy = \frac{1}{2} \left(\frac{ye^{2y}}{2} - \frac{e^{2y}}{4} - \frac{y^2}{2} \right) \Big|_1^3 = \frac{5e^6}{8} - \frac{e^2}{8} - 2. \end{aligned}$$

- b) The domain of integration is shown in the following figure.



Hence

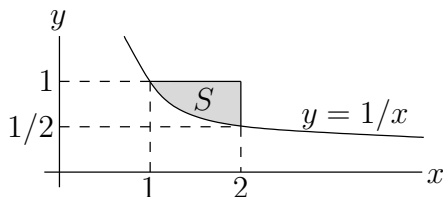
$$\iint_S \cos(y^3 + 1) = \int_0^1 \int_{\sqrt{x}}^1 \cos(y^3 + 1) dy dx = \int_0^1 \int_0^{y^2} \cos(y^3 + 1) dx dy.$$

It is clear that the last integral is the simplest to compute (The reader should try to compute the middle integral.) Thus

$$\iint_S \cos(y^3 + 1) = \int_0^1 \int_0^{y^2} \cos(y^3 + 1) dx dy = \int_0^1 (x \cos(y^3 + 1)) \Big|_{x=0}^{y^2} dy$$

$$= \int_0^1 y^2 \cos(y^3 + 1) dy = \frac{1}{3} \sin(y^3 + 1) \Big|_{y=0}^1 = \frac{\sin(2) - \sin(1)}{3} .$$

c) The domain of integration is shown in the following figure.



Hence

$$\iint_S ye^{xy} = \int_1^2 \int_{1/x}^1 ye^{xy} dy dx = \int_{1/2}^1 \int_{1/y}^2 ye^{xy} dx dy .$$

The last integral is simpler to compute because only a simple substitution is required. The middle integral requires integration by parts. Thus

$$\begin{aligned} \iint_S ye^{xy} &= \int_{1/2}^1 \int_{1/y}^2 ye^{xy} dx dy = \int_{1/2}^1 \left(e^{xy} \Big|_{x=1/y}^2 \right) dy \\ &= \int_{1/2}^1 (e^{2y} - e) dy = \left(\frac{e^{2y}}{2} - ey \right) \Big|_{y=1/2}^1 = \frac{e^2}{2} - e . \end{aligned}$$

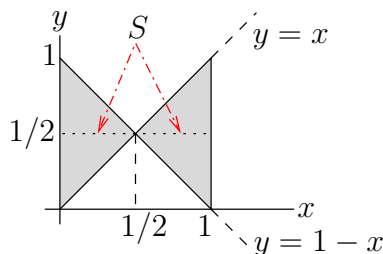
♡

Question 3.15

Let $f : [0, 1] \rightarrow \mathbb{R}$ be a continuous function. Is the integral $\int_0^1 \int_x^{1-x} f(y) dy dx$ equal to the integral of f over the region delimited by the limits of integration? Write the integral of f over the region delimited by the boundary of integration as the sum of two integrals of one variable.

Solution

The region delimited by the limits of integration is shown below.



The integral over this domain is

$$\iint_S f = \int_0^{1/2} \int_x^{1-x} f(y) dy dx + \int_{1/2}^1 \int_{1-x}^x f(y) dy dx .$$

This is not equal to $\int_0^1 \int_x^{1-x} f(y) \, dy \, dx$ because

$$\begin{aligned} \int_0^1 \int_x^{1-x} f(y) \, dy \, dx &= \int_0^{1/2} \int_x^{1-x} f(y) \, dy \, dx + \int_{1/2}^1 \int_x^{1-x} f(y) \, dy \, dx \\ &= \int_0^{1/2} \int_x^{1-x} f(y) \, dy \, dx - \int_{1/2}^1 \int_{1-x}^x f(y) \, dy \, dx . \end{aligned}$$

To express $\iint_S f$ as the sum of two integrals of one variable, we change the order of integration.

$$\begin{aligned} \iint_S f &= \int_0^{1/2} \int_0^y f(y) \, dx \, dy + \int_{1/2}^1 \int_0^{1-y} f(y) \, dx \, dy \\ &\quad + \int_0^{1/2} \int_{1-y}^1 f(y) \, dx \, dy + \int_{1/2}^1 \int_y^1 f(y) \, dx \, dy \\ &= \int_0^{1/2} y f(y) \, dy + \int_{1/2}^1 (1-y) f(y) \, dy + \int_0^{1/2} y f(y) \, dy + \int_{1/2}^1 (1-y) f(y) \, dy \\ &= 2 \int_0^{1/2} y f(y) \, dy + 2 \int_{1/2}^1 (1-y) f(y) \, dy . \end{aligned}$$

The fact that we get 2 times each integral should not be surprising to the reader because f is independent of x . Thus

$$(b-a)f(y) = \int_a^b f(y) \, dx = \int_c^d f(y) \, dx = (d-c)f(y)$$

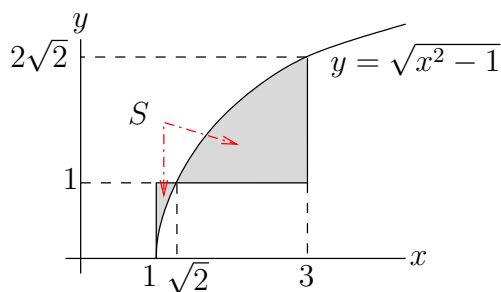
as long as $b-a = d-c$. In the present case, we have that $a=0$, $b=y$, $c=1-y$ and $d=1$. ♡

Question 3.16

Reverse the order of integration of the integral $\int_1^3 \int_1^{\sqrt{x^2-1}} f(x,y) \, dy \, dx$. Is this integral equal to the integral of f over the region enclosed by the limits of integration? Justify your answer.

Solution

The region delimited by the limits of integration is shown below.



We have

$$\begin{aligned}
 \int_1^3 \int_1^{\sqrt{x^2-1}} f(x, y) \, dy \, dx &= \int_0^{\sqrt{2}} \int_1^{\sqrt{x^2+1}} f(x, y) \, dy \, dx + \int_{\sqrt{2}}^3 \int_1^{\sqrt{x^2-1}} f(x, y) \, dy \, dx \\
 &= - \int_0^{\sqrt{2}} \int_{\sqrt{x^2+1}}^1 f(x, y) \, dy \, dx + \int_{\sqrt{2}}^3 \int_1^{\sqrt{x^2-1}} f(x, y) \, dy \, dx \\
 &= - \int_0^1 \int_1^{\sqrt{y^2+1}} f(x, y) \, dx \, dy + \int_1^{2\sqrt{2}} \int_{\sqrt{y^2+1}}^3 f(x, y) \, dx \, dy .
 \end{aligned}$$

This integral is not equal to the integral of f over the region S enclosed by the limits of integration. The integral of f over S is given by

$$\begin{aligned}
 \iint_S f &= \int_0^1 \int_1^{\sqrt{y^2+1}} f(x, y) \, dx \, dy + \int_1^{2\sqrt{2}} \int_{\sqrt{y^2+1}}^3 f(x, y) \, dx \, dy \\
 &= \int_1^{\sqrt{2}} \int_{\sqrt{x^2-1}}^1 f(x, y) \, dy \, dx + \int_{\sqrt{2}}^3 \int_1^{\sqrt{x^2-1}} f(x, y) \, dy \, dx .
 \end{aligned}$$

♡

Question 3.17

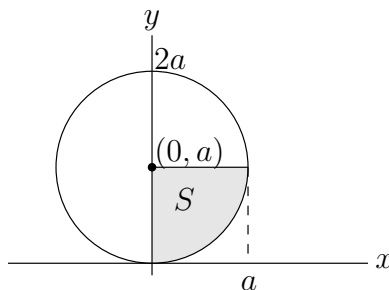
Reverse the order of integration of the integral $\int_{-6}^8 \int_{x^{1/3}}^{(x+6)/7} xy \, dy \, dx$.

Question 3.18

Reverse the order of integration in the integral $\int_0^a \int_0^{\sqrt{2ay-y^2}} f(x, y) \, dx \, dy$.

Solution

From $x = \sqrt{2ay - y^2}$, we get that $x^2 = 2ay - y^2$. By completing the square with respect to the variable y , we get that $x^2 - a^2 = -(y - a)^2$. Thus, the boundary of the domain of integration S is an arc of the circle $x^2 + (y - a)^2 = a^2$. The following figure shows the domain of integration.



Solving $x^2 + (y - a)^2 = a^2$ for y , we get that $y = a - \sqrt{a^2 - x^2}$ on the lower part of the circle. Hence

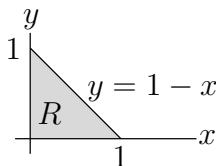
$$\int_0^a \int_0^{\sqrt{2ay-y^2}} f(x, y) \, dx \, dy = \int_0^a \int_{a-\sqrt{a^2-x^2}}^a f(x, y) \, dy \, dx .$$

**Question 3.19**

Find the volume of the solid whose base is the triangle with vertices $(0, 0)$, $(1, 0)$ and $(0, 1)$, and which is bounded above by the surface $z = 6xy(1 - x - y)$.

Solution

The domain of integration R is shown in the figure below.



Since $6xy(1 - x - y) \geq 0$ for all $(x, y) \in R$, the volume is

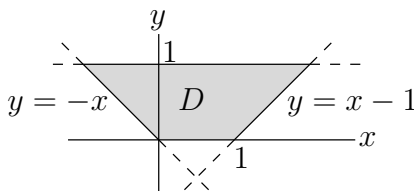
$$\begin{aligned} V &= \int_0^1 \int_0^{1-x} 6xy(1 - x - y) \, dy \, dx = 6 \int_0^1 \int_0^{1-x} (xy - x^2y - xy^2) \, dy \, dx \\ &= 6 \int_0^1 \left(\frac{xy^2}{2} - \frac{x^2y^2}{2} - \frac{xy^3}{3} \right) \Big|_{y=0}^{1-x} dx = \int_0^1 (x - 3x^2 + 3x^3 - x^4) \, dx \\ &= \left(\frac{x^2}{2} - x^3 + \frac{3x^4}{4} - \frac{x^5}{5} \right) \Big|_{x=0}^1 = \frac{1}{20}. \end{aligned}$$

**Question 3.20**

Find the value of the integral of $f(x, y) = 2xy$ over the region D enclosed by the lines $y = x - 1$, $y = -x$, $y = 0$ and $y = 1$.

Solution

The region D is shown in the figure below.



It is advantageous to integrate with respect to x first and then with respect to y to avoid having to split the domain of integration.

$$\iint_D f = \int_0^1 \int_{-y}^{1+y} 2xy \, dx \, dy = \int_0^1 x^2y \Big|_{x=-y}^{1+y} dy = \int_0^1 (y + 2y^2) \, dy = \left(\frac{y^2}{2} + \frac{2y^3}{3} \right) \Big|_{y=0}^1 = \frac{7}{6}.$$

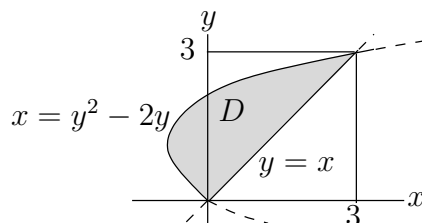


Question 3.21

Compute the integral of $f(x, y) = x^2 + y^2$ over the region D enclosed by the parabola $x = y^2 - 2y$ and the line $y = x$.

Solution

The region D is shown in the figure below.



To find the points of intersection of the parabola with the line $y = x$, we first find the ordinates of these points of intersection.

$$y = y^2 - 2y \Rightarrow y(y - 3) = 0 \Rightarrow y = 0 \text{ or } 3.$$

The points of intersection are $(0, 0)$ and $(3, 3)$. We integrate with respect to x first and then with respect to y to avoid having to split the domain of integration.

$$\begin{aligned} \iint_D f &= \int_0^3 \int_{y^2-2y}^y (x^2 + y^2) dx dy = \int_0^3 \left(\frac{x^3}{3} + xy^2 \right) \Big|_{x=y^2-2y}^y dy \\ &= \int_0^3 \left(-\frac{y^6}{3} + 2y^5 - 5y^4 + 6y^3 \right) dy = \left(-\frac{y^7}{21} + \frac{y^6}{3} - y^5 + \frac{3y^4}{2} \right) \Big|_{y=0}^3 = \frac{3^5}{14}. \end{aligned}$$

♥

Question 3.22

Evaluate $I = \iint_S (x + 3y^3)$ where S is the upper half of the disk $x^2 + y^2 \leq 4$.

Solution

Because of the shape of the domain of integration, it is advantageous to use polar coordinates to compute this integral. Let $x = r \cos(\theta)$ and $y = r \sin(\theta)$ for $0 \leq r \leq 2$ and $0 \leq \theta \leq \pi$.

Since $\left| \frac{\partial(x, y)}{\partial(r, \theta)} \right| = r$, we get

$$\begin{aligned} I &= \int_0^2 \int_0^\pi (r \cos(\theta) + 3r^3 \sin^3(\theta)) r d\theta dr \\ &= \int_0^2 \int_0^\pi r^2 \cos(\theta) d\theta dr + 3 \int_0^2 \int_0^\pi r^4 (\sin(\theta) - \cos^2(\theta) \sin(\theta)) d\theta dr \\ &= \int_0^2 r^2 \sin(\theta) \Big|_{\theta=0}^\pi dr + 3 \int_0^2 r^4 \left(-\cos(\theta) + \frac{1}{3} \cos^3(\theta) \right) \Big|_{\theta=0}^\pi dr \\ &= 0 + 3 \int_0^2 r^4 \left(\frac{4}{3} \right) dr = 4 \frac{r^5}{5} \Big|_0^2 = \frac{128}{5}, \end{aligned}$$

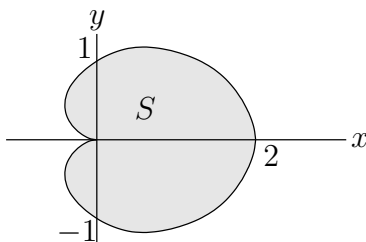
where the famous trigonometric identity $\sin^2(\theta) = 1 - \cos^2(\theta)$ has been used to get the second equality above. ♡

Question 3.23

Find the area of the region S enclosed by the cardioid $r = 1 + \cos(\theta)$.

Solution

The region S is shown in the following figure.



We need to compute $A = \iint_S 1$. Since S is given using polar coordinates, this suggests that polar coordinates should be used to compute the integral. Let $x = r \cos(\theta)$ and $y = r \sin(\theta)$ for $0 \leq r \leq 1 + \cos(\theta)$ and $0 \leq \theta \leq 2\pi$. We have that $\left| \frac{\partial(x, y)}{\partial(r, \theta)} \right| = r$. Thus

$$\begin{aligned} A &= \iint_S 1 = \int_0^{2\pi} \int_0^{1+\cos(\theta)} r \, dr \, d\theta = \int_0^{2\pi} \left(\frac{r^2}{2} \Big|_{r=0}^{1+\cos(\theta)} \right) d\theta \\ &= \int_0^{2\pi} \left(\frac{1}{2} + \cos(\theta) + \frac{\cos^2(\theta)}{2} \right) d\theta = \int_0^{2\pi} \left(\frac{3}{4} + \cos(\theta) + \frac{\cos(2\theta)}{4} \right) d\theta \\ &= \left(\frac{3\theta}{4} + \sin(\theta) + \frac{\sin(2\theta)}{8} \right) \Big|_{\theta=0}^{2\pi} = \frac{3\pi}{2}, \end{aligned}$$

where the famous trigonometric identity $\cos^2(\theta) = (1 + \cos(2\theta))/2$ has been used to get the fifth equality. ♡

Question 3.24

Evaluate $\iint_S e^{-(x^2+xy+y^2)}$ where $S = \{(x, y) : x^2 + xy + y^2 \leq 1\}$.

Solution

Before computing the integral, it may be useful to review a little bit of linear algebra. We have the quadratic form

$$x^2 + xy + y^2 = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Let $A = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix}$. Since A is symmetric, it has only real eigenvalues and an orthogonal matrix may be used to diagonalize A . More precisely, suppose that λ_1 and λ_2 are the eigenvalues of A . If $\lambda_1 \neq \lambda_2$, and $\mathbf{v}_1$ and $\mathbf{v}_2$ are eigenvectors of norm 1 associated to λ_1

and λ_2 respectively, then $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthonormal. If $\lambda_1 = \lambda_2$, then we can always find two linearly independent eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ associated to $\lambda_1 = \lambda_2$. We may select them (using Gram-Schmidt orthogonalization) such that they are orthonormal. Hence, with

$$R = (\mathbf{v}_1 \ \mathbf{v}_2), \text{ we have } R^\top A R = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

It is easy to see that $\mathbf{v}_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$ and $\mathbf{v}_2 = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$ are eigenvectors of A associated to the eigenvalues $\lambda_1 = 3/2$ and $\lambda_2 = 1/2$ of A . Thus, with $R = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$, we get $R^\top A R = \begin{pmatrix} 3/2 & 0 \\ 0 & 1/2 \end{pmatrix}$. Note that $R = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \Big|_{\theta=\pi/4}$ is a counterclockwise rotation by $\pi/4$ about the origin.

We may use this information to choose a change of variables $g : D \rightarrow S$ to compute the integral. We choose $\begin{pmatrix} x \\ y \end{pmatrix} = g(u, v) = R \begin{pmatrix} u \\ v \end{pmatrix}$. Thus $|\det D g| = \left| \det \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \right| = 1$. This was to be expected because rotations do not change the volume of objects.

To determine the domain D of g , we note that

$$\begin{aligned} x^2 + xy + y^2 &= (x \ y) A \begin{pmatrix} x \\ y \end{pmatrix} = (x \ y) R \begin{pmatrix} 3/2 & 0 \\ 0 & 1/2 \end{pmatrix} R^\top \begin{pmatrix} x \\ y \end{pmatrix} \\ &= \left(R^\top \begin{pmatrix} x \\ y \end{pmatrix} \right)^\top \begin{pmatrix} 3/2 & 0 \\ 0 & 1/2 \end{pmatrix} \left(R^\top \begin{pmatrix} x \\ y \end{pmatrix} \right) = (u \ v) \begin{pmatrix} 3/2 & 0 \\ 0 & 1/2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \frac{3}{2} u^2 + \frac{1}{2} v^2 \end{aligned}$$

Hence g maps $D = \left\{ (u, v) : \frac{3u^2}{2} + \frac{v^2}{2} \leq 1 \right\}$ onto S . We get

$$\iint_S e^{-(x^2+xy+y^2)} = \iint_D e^{-(3u^2/2)-(v^2/2)}.$$

We now use a second change of variables; namely, $\begin{pmatrix} u \\ v \end{pmatrix} = h(r, s) = \begin{pmatrix} \sqrt{2/3} r \cos(s) \\ \sqrt{2} r \sin(s) \end{pmatrix}$ for $0 \leq r \leq 1$ and $0 \leq \theta \leq 2\pi$. Thus $|\det D h| = 2r/\sqrt{3}$ and

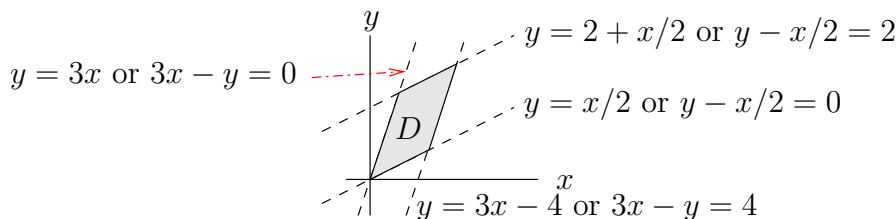
$$\iint_D e^{-(3u^2/2)-(v^2/2)} = \frac{2}{\sqrt{3}} \int_0^{2\pi} \int_0^1 e^{-r^2} r \, dr \, d\theta = -\frac{2\pi}{\sqrt{3}} e^{-r^2} \Big|_{r=0}^1 = \frac{2\pi}{\sqrt{3}} (1 - e^{-1}).$$

♡

Question 3.25

Let D be the region bounded by $y = x/2$, $y = (x/2) + 2$, $y = 3x$ and $y = 3x - 4$. Use an appropriate change of variables to compute $\iint_D xy$.

Solution



Let $u = 3x - y$ and $v = y - x/2$. We get that $x = 2(u+v)/5$ and $y = (u+6v)/5$. We consider the change of variables $\begin{pmatrix} x \\ y \end{pmatrix} = g(u, v) = \begin{pmatrix} 2(u+v)/5 \\ (u+6v)/5 \end{pmatrix}$. We have that $|\det Dg(u, v)| = \left| \det \begin{pmatrix} 2/5 & 2/5 \\ 1/5 & 6/5 \end{pmatrix} \right| = \frac{2}{5}$. Moreover, $u = 3x - y$ varies from 0 to 4, and $v = y - x/2$ varies from 0 to 2. Thus

$$\begin{aligned} \iint_D xy &= \int_0^4 \int_0^2 \left(\frac{2(u+v)}{5} \right) \left(\frac{u+6v}{5} \right) \frac{2}{5} dv du = \frac{4}{5^3} \int_0^4 \int_0^2 (u^2 + 7uv + 6v^2) dv du \\ &= \frac{4}{5^3} \int_0^4 \left(u^2 v + \frac{7uv^2}{2} + 2v^3 \right) \Big|_{v=0}^2 du = \frac{8}{5^3} \int_0^4 (u^2 + 7u + 8) du \\ &= \frac{8}{5^3} \left(\frac{u^3}{3} + \frac{7u^2}{2} + 8u \right) \Big|_{u=0}^4 = \frac{2624}{375}. \end{aligned}$$

♡

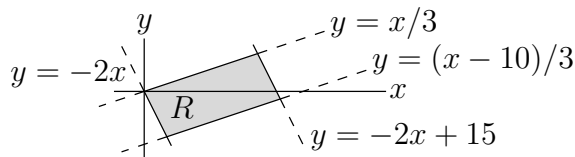
Question 3.26

Find the coordinates of the centroid (namely, the coordinates of the centre of mass) of the parallelogram R enclosed by the lines $x - 3y = 0$, $2x + y = 0$, $x - 3y = 10$ and $2x + y = 15$. The density ρ of this thin plate is constant.

Solution

Recall that the coordinates $(\bar{x}, \bar{y})$ of the **centroid** of a thin plate D of density $\rho(x, y)$ at $(x, y) \in D$ are found as it follows. Let $m = \iint_D \rho$ be the mass of the thin plate. Then $\bar{x} = \frac{1}{m} \iint_D x\rho$ and $\bar{y} = \frac{1}{m} \iint_D y\rho$.

The parallelogram R is shown in the figure below.



Let $u = x - 3y$ and $v = 2x + y$. We get that $x = (u + 3v)/7$ and $y = (v - 2u)/7$. This is the change of variables that we use with $0 \leq u \leq 10$ and $0 \leq v \leq 15$. We have that $\left| \frac{\partial(x, y)}{\partial(u, v)} \right| = \left| \det \begin{pmatrix} 1/7 & 3/7 \\ -2/7 & 1/7 \end{pmatrix} \right| = \frac{1}{7}$. Thus, the mass is

$$m = \iint_R \rho = \rho \int_0^{10} \int_0^{15} \frac{1}{7} dv du = \frac{150\rho}{7}.$$

The coordinates $(\bar{x}, \bar{y})$ of the centroid are

$$\bar{x} = \frac{1}{m} \iint_R x \rho = \frac{7}{150} \int_0^{10} \int_0^{15} \left(\frac{u+3v}{7} \right) \frac{1}{7} dv du = \frac{55}{14}$$

and

$$\bar{y} = \frac{1}{m} \iint_R y \rho = \frac{7}{150} \int_0^{10} \int_0^{15} \left(\frac{v-2u}{7} \right) \frac{1}{7} dv du = -\frac{5}{14}.$$

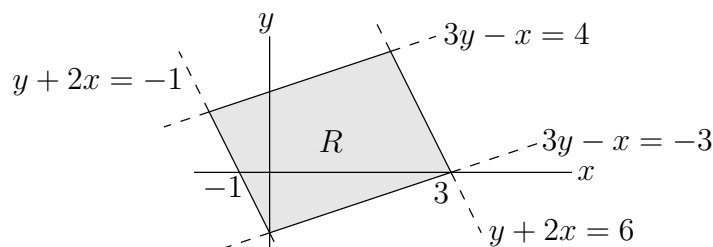
♡

Question 3.27

Compute the integral $\iint_R (2x+y)^2(3y-x)^{-2}$ where R is the parallelogram given by $-1 \leq 2x+y \leq 6$ and $-3 \leq 3y-x \leq 4$.

Solution

The parallelogram R is shown in the figure below.



Let $u = 2x + y$ and $v = 3y - x$. We get that $x = (3u - v)/7$ and $y = (u + 2v)/7$. This is the change of variables that we use with $-1 \leq u \leq 6$ and $-3 \leq v \leq 4$. We have that $\left| \frac{\partial(x, y)}{\partial(u, v)} \right| = \left| \det \begin{pmatrix} 3/7 & -1/7 \\ 1/7 & 2/7 \end{pmatrix} \right| = \frac{1}{7}$. Thus

$$\begin{aligned} \iint_R (2x+y)^2(3y-x)^{-2} &= \int_{-1}^6 \int_{-3}^4 \frac{u^2 v^{-2}}{7} dv du = \frac{1}{7} \int_{-1}^6 \left(-\frac{u^2}{v} \right) \Big|_{v=-3}^4 du \\ &= -\frac{1}{12} \int_{-1}^6 u^2 du = -\frac{1}{12} \left(\frac{u^3}{3} \right) \Big|_{u=-1}^6 = -\frac{217}{36}. \end{aligned}$$

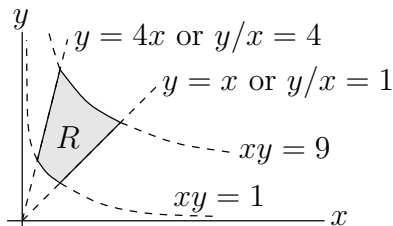
♡

Question 3.28

Let R be the region enclosed by the curves $xy = 1$ and $xy = 9$, and the lines $y = x$ and $y = 4x$ for $x, y > 0$. Find the coordinates of the centroid of this region assuming that the density ρ is constant. For those who may have forgotten the formulae used to compute the coordinates of a centroid or may not have seen them before, they are given in Question 3.26.

Solution

The region R is shown in the figure below.



Let $u = xy$ and $v = y/x$. We get that $x = \sqrt{u/v}$ and $y = \sqrt{uv}$. This is the change of variables that we use with $1 \leq u \leq 9$ and $1 \leq v \leq 4$. We have that $\left| \frac{\partial(x, y)}{\partial(u, v)} \right| =$

$$\left| \det \begin{pmatrix} \frac{1}{2}u^{-1/2}v^{-1/2} & -\frac{1}{2}u^{1/2}v^{-3/2} \\ \frac{1}{2}u^{-1/2}v^{1/2} & \frac{1}{2}u^{1/2}v^{-1/2} \end{pmatrix} \right| = \frac{1}{2v}. \text{ Thus, the mass is}$$

$$m = \iint_R \rho = \rho \int_1^9 \int_1^4 \frac{1}{2v} dv du = 9 \ln(2) \rho.$$

The coordinates $(\bar{x}, \bar{y})$ of the centroid are

$$\bar{x} = \frac{1}{m} \iint_R x \rho = \frac{1}{9 \ln(2)} \int_1^9 \int_1^4 \frac{1}{2} u^{1/2} v^{-1/2} dv du = \frac{52}{27 \ln(2)}$$

and

$$\bar{y} = \frac{1}{m} \iint_R y \rho = \frac{1}{9 \ln(2)} \int_1^9 \int_1^4 \frac{1}{2} u^{1/2} v^{1/2} dv du = \frac{364}{81 \ln(2)}.$$

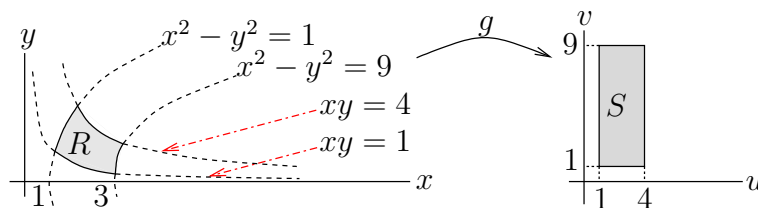
♡

Question 3.29

Let R be the region enclosed by the hyperbolas $xy = 1$, $xy = 4$, $x^2 - y^2 = 1$ and $x^2 - y^2 = 9$ for $x, y > 0$. Compute the mass of the thin plate represented by this region if the density is given by $\rho(x, y) = (x^2 + y^2)^3$.

Solution

The region R is represented in the figure below.



We want to compute the mass $m = \iint_R (x^2 + y^2)^3$. For this question, we apply the change of variables in the opposite direction. Namely, we assume that

$$\iint_R (x^2 + y^2)^3 = \iint_R f(g(x, y)) |\det Dg(x, y)| = \iint_S f(u, v)$$

for some function $f : S \rightarrow \mathbb{R}$ where $(x^2 + y^2)^3 = f(g(x, y)) |\det Dg(x, y)|$ and $g : R \rightarrow S$ is our change of variables. Let $u = xy$ and $v = x^2 - y^2$. If we try to express x and y in terms of u and v , then we get ugly formulae. Therefore, we proceed slightly differently than usual. Let $\begin{pmatrix} u \\ v \end{pmatrix} = g(x, y) = \begin{pmatrix} xy \\ x^2 - y^2 \end{pmatrix}$ with $S = \{(u, v) : 1 \leq u \leq 4 \text{ and } 1 \leq v \leq 9\}$. We have that $|\det Dg(x, y)| = \left| \frac{\partial(u, v)}{\partial(x, y)} \right| = \left| \det \begin{pmatrix} y & x \\ 2x & -2y \end{pmatrix} \right| = 2(x^2 + y^2)$ and $f(u, v) = \frac{v^2 + 4u^2}{2} = \frac{(x^2 + y^2)^2}{2}$. Hence, the mass is

$$\begin{aligned} m &= \iint_R (x^2 + y^2)^3 = \iint_R \underbrace{\frac{(x^2 + y^2)^2}{2} 2(x^2 + y^2)}_{=f(g(x,y)) |\det Dg(x,y)|} = \iint_S \frac{v^2 + 4u^2}{2} \\ &= \frac{1}{2} \int_1^4 \int_1^9 (v^2 + 4u^2) dv du = 700. \end{aligned}$$

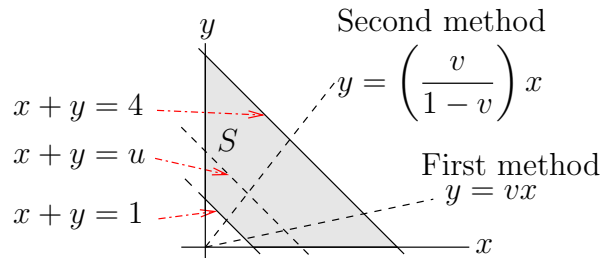
♡

Question 3.30

Evaluate the integral $\iint_S \frac{1}{x+y}$ where S is the region bounded by the x and y axes, and the lines $x+y=1$ and $x+y=4$.

Solution

The domain of integration S is shown in the following figure.



First method) The obvious change of variables comes from $u = x+y$ and $v = y/x$. We then get that $x = u/(1+v)$ and $y = uv/(1+v)$. This is the change of variables that we use with $1 \leq u \leq 4$ and $0 < v < \infty$. We have to ignore $v = 0$ to ensure that the change of variables is injective.

Unfortunately, this change of variables requires the computation of an improper integral. We have that $\left| \frac{\partial(x, y)}{\partial(u, v)} \right| = \left| \det \begin{pmatrix} 1/(1+v) & -u/(1+v)^2 \\ v/(1+v) & u/(1+v)^2 \end{pmatrix} \right| = \frac{u}{(1+v)^2}$. Thus

$$\begin{aligned} \iint_S \frac{1}{x+y} &= \lim_{R \rightarrow \infty} \int_1^4 \int_0^R \frac{1}{u} \left(\frac{u}{(1+v)^2} \right) dv du = \lim_{R \rightarrow \infty} \int_1^4 \int_0^R \frac{1}{(1+v)^2} dv du \\ &= \lim_{R \rightarrow \infty} \int_1^4 \left(\frac{-1}{1+v} \right) \Big|_{v=0}^R du = \lim_{R \rightarrow \infty} \int_1^4 \left(1 - \frac{1}{1+R} \right) du \end{aligned}$$

$$= \lim_{R \rightarrow \infty} 3 \left(1 - \frac{1}{1+R} \right) = 3.$$

Second method) If we do not want to introduce an improper integral, then we have to slightly modify the previous change of variables.

We still set $u = x + y$ for $1 \leq u \leq 4$. Instead of using $y = vx$ for $0 < v < \infty$ as we did before, we use $y = \left(\frac{v}{1-v} \right) x$ for $0 < v < 1$. This substitution yields $y = 0$ when $v = 0$ and approaches $x = 0$ when $v \rightarrow 1^-$. If we solve for v , then we get $v = \frac{y}{x+y}$. Thus, the change of variables that we use is $x = u - uv$ and $y = uv$ for $1 \leq u \leq 4$ and $0 < v < 1$. We have to ignore $v = 0$ to ensure that the change of variables is injective. As u varies from 1 to 4, we have all the lines $x + y = u$ that cover S . When v varies from 0 to 1 (excluding the end points), we have that $y = \left(\frac{v}{1-v} \right) x$ yields all the lines through the origin that cover S except for the segments $\{(x, 0) : 1 \leq x \leq 4\}$ and $\{(0, y) : 1 \leq y \leq 4\}$.

We have that $\left| \frac{\partial(x, y)}{\partial(u, v)} \right| = \left| \det \begin{pmatrix} 1-v & -u \\ v & u \end{pmatrix} \right| = u$. Thus

$$\iint_S \frac{1}{x+y} = \int_1^4 \int_0^1 \frac{1}{u} u \, dv \, du = \int_1^4 \int_0^1 dv \, du = 3.$$

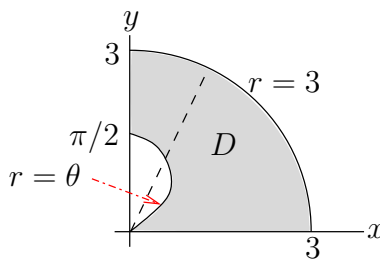
♡

Question 3.31

Compute the area of the region in the first quadrant enclosed by the curves $r = \theta$ and $r = 3$ for $0 \leq \theta \leq \pi/2$, and the x and y axes.

Solution

The region is shown in the following figure.



We use polar coordinates to compute the area of this region. Let $x = r \cos(\theta)$ and $y = r \sin(\theta)$ with $0 \leq \theta < 2\pi$ and $\theta \leq r \leq 3$. We have that $\left| \frac{\partial(x, y)}{\partial(r, \theta)} \right| = r$ and the area is

$$A = \iint_D 1 = \int_0^{\pi/2} \int_{\theta}^3 r \, dr \, d\theta = \int_0^{\pi/2} \left(\frac{r^2}{2} \Big|_{r=\theta}^3 \right) d\theta$$

$$= \int_0^{\pi/2} \left(\frac{9}{2} - \frac{\theta^2}{2} \right) d\theta = \left(\frac{9\theta}{2} - \frac{\theta^3}{6} \right) \Big|_{\theta=0}^{\pi/2} = \frac{9\pi}{4} - \frac{\pi^3}{48}.$$

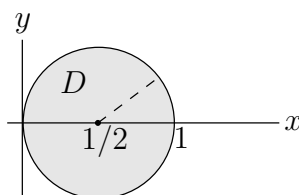
♡

Question 3.32

Compute the volume of the right cylinder whose base is the region of the x, y plane bounded by the curve $x^2 + y^2 = x$, and whose height is given by $z = 4 - y$.

Solution

One can rewrite $x^2 + y^2 = x$ as $\left(x - \frac{1}{2}\right)^2 + y^2 = \frac{1}{4}$. This is a circle of radius $1/2$ centred at $(1/2, 0)$. The bounded region is shown in the following figure.



We use polar coordinates $x = 1/2 + r \cos(\theta)$ and $y = r \sin(\theta)$ for $0 \leq \theta < 2\pi$ and $0 < r < 1/2$.

We have that $\left| \frac{\partial(x, y)}{\partial(r, \theta)} \right| = r$ and the volume of the cylinder is

$$\begin{aligned} V &= \iint_D (4 - y) = \int_0^{2\pi} \int_0^{1/2} (4 - r \sin(\theta)) r dr d\theta = \int_0^{2\pi} \left(2r^2 - \frac{r^3}{3} \sin(\theta) \right) \Big|_{r=0}^{1/2} d\theta \\ &= \int_0^{2\pi} \left(\frac{1}{2} - \frac{1}{8} \sin(\theta) \right) d\theta = \left(\frac{\theta}{2} + \frac{1}{8} \cos(\theta) \right) \Big|_{\theta=0}^{2\pi} = \pi. \end{aligned}$$

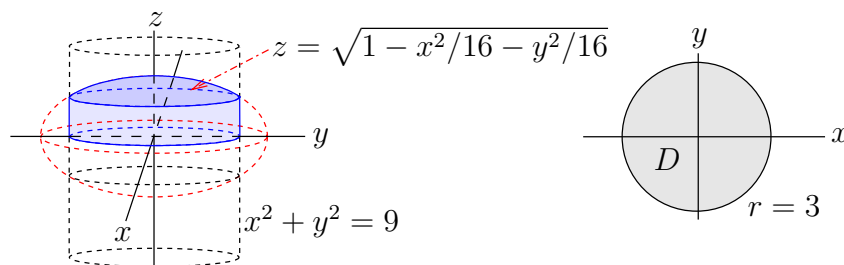
♡

Question 3.33

Compute the volume of the solid obtained if we cut out the ellipsoid given by the $\frac{x^2}{16} + \frac{y^2}{16} + z^2 \leq 1$ with a cylinder of radius 3 whose axis is the z axis.

Solution

We use polar coordinates $x = r \cos(\theta)$ and $y = r \sin(\theta)$ for $0 \leq \theta < 2\pi$ and $0 \leq r \leq 3$ to compute the volume of the region. The following figure shows half of the solid, the section above the x, y plane. Since the solid is symmetric with respect to the plane $z = 0$, the volume of the solid is twice the volume of this section.



We have that $\left| \frac{\partial(x, y)}{\partial(r, \theta)} \right| = r$ and the volume is

$$\begin{aligned} V &= 2 \iint_D \sqrt{1 - \frac{x^2}{16} - \frac{y^2}{16}} = 2 \int_0^{2\pi} \int_0^3 \sqrt{1 - \frac{r^2}{16}} r \, dr \, d\theta \\ &= 2 \int_0^{2\pi} \left(-\frac{16}{3} \left(1 - \frac{r^2}{16} \right)^{3/2} \right) \Big|_{r=0}^3 d\theta = 2 \left(-\frac{16}{3} \left(\frac{7}{16} \right)^{3/2} + \frac{16}{3} \right) \int_0^{2\pi} d\theta \\ &= \frac{\pi}{3} (64 - 7\sqrt{7}) . \end{aligned}$$

♡

Question 3.34

Compute the volume of the ellipsoid E given by $(2x + y + z)^2 + (2x - y)^2 + 4z^2 \leq 4$.

Solution

Let $u = 2x + y + z$, $v = 2x - y$ and $w = 2z$. We get that $x = \frac{u}{4} + \frac{v}{4} - \frac{w}{8}$, $y = \frac{u}{2} - \frac{v}{2} - \frac{w}{4}$ and $z = \frac{w}{2}$. This is the change of variables that we use. We have that $\left| \frac{\partial(x, y, z)}{\partial(u, v, w)} \right| = \left| \det \begin{pmatrix} 1/4 & 1/4 & -1/8 \\ 1/2 & -1/2 & -1/4 \\ 0 & 0 & 1/2 \end{pmatrix} \right| = \frac{1}{8}$. Thus, the volume of E is

$$V = \iiint_E 1 = \frac{1}{8} \iiint_{u^2+v^2+w^2 \leq 4} du \, dv \, dw .$$

We could compute this integral using spherical coordinates but it is easier to use the formula $\frac{4r^3\pi}{3}$ to compute the volume of a sphere of radius r ; a proof of this formulae is given in

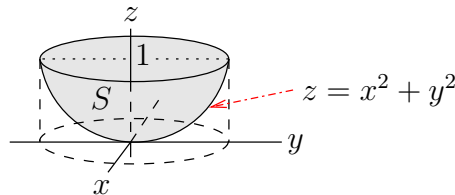
Question 3.46 for the benefit of the reader. We get that $V = \frac{4\pi}{3}$. ♡

Question 3.35

If S is the solid enclosed by the paraboloid $z = x^2 + y^2$ and the plane $z = 1$, write the iterated integrals of a function f over the region S in the orders $dz \, dy \, dx$, $dy \, dz \, dx$ and $dx \, dy \, dz$.

Solution

The solid S is represented in the following figure.



We have

$$\iiint_S f = \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \int_{x^2+y^2}^1 f(x, y, z) \, dz \, dy \, dx = \int_{-1}^1 \int_{x^2}^1 \int_{-\sqrt{z-x^2}}^{\sqrt{z-x^2}} f(x, y, z) \, dy \, dz \, dx$$

$$= \int_0^1 \int_{-\sqrt{z}}^{\sqrt{z}} \int_{-\sqrt{z-y^2}}^{\sqrt{z-y^2}} f(x, y, z) \, dx \, dy \, dz .$$



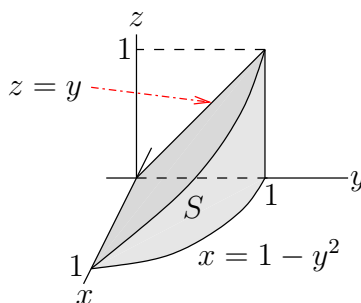
Question 3.36

For each of the following integrals, sketch its domain of integration and express the iterated integral in the requested orders.

- a) $\int_0^1 \int_0^{1-y^2} \int_0^y f(x, y, z) \, dz \, dx \, dy$ in the orders $dz \, dy \, dx$ and $dy \, dz \, dx$.
- b) $\int_0^2 \int_0^{1-x/2} \int_0^{x/2} f(x, y, z) \, dz \, dy \, dx$ in the orders $dy \, dx \, dz$ and $dx \, dy \, dz$.
- c) $\int_0^1 \int_0^z \int_0^y f(x, y, z) \, dx \, dy \, dz$ in the orders $dz \, dy \, dx$ and $dy \, dx \, dz$.
- d) $\int_0^1 \int_0^y \int_0^{x/\sqrt{3}} f(x, y, z) \, dz \, dx \, dy$ in the orders $dz \, dy \, dx$ and $dy \, dz \, dx$.
- e) $\int_1^2 \int_1^z \int_{1/y}^2 f(x, y, z) \, dx \, dy \, dz$ in the orders $dz \, dy \, dx$ and $dy \, dz \, dx$.
- f) $\int_0^4 \int_{\sqrt{x}}^2 \int_0^{2-y} f(x, y, z) \, dz \, dy \, dx$ in the orders $dx \, dz \, dy$ and $dy \, dx \, dz$.
- g) $\int_0^2 \int_1^{2-x/2} \int_x^2 f(x, y, z) \, dz \, dy \, dx$ in the orders $dy \, dx \, dz$ and $dx \, dy \, dz$.

Solution

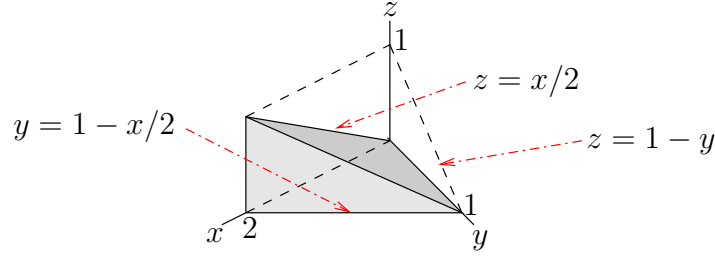
a) The solid S represented by the region enclosed by the limits of integration is shown in the following figure.



S is the solid in the quadrant $x, y, z \geq 0$ which is enclosed by the planes $x = 0$, $y = 0$ and $z = y$, and the surface $x = 1 - y^2$. We have

$$\iiint_S f = \int_0^1 \int_0^{\sqrt{1-x}} \int_0^y f(x, y, z) \, dz \, dy \, dx = \int_0^1 \int_0^{\sqrt{1-x}} \int_z^{\sqrt{1-x}} f(x, y, z) \, dy \, dz \, dx .$$

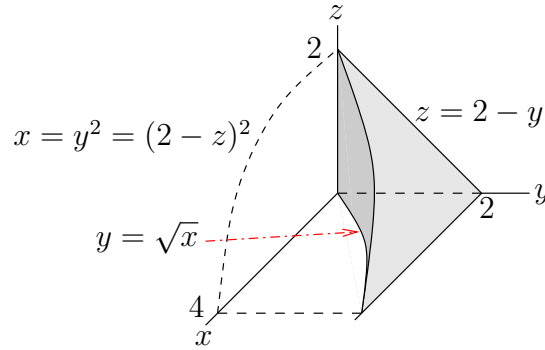
b) The domain of integration is represented in the following figure.



We have

$$\begin{aligned} \int_0^2 \int_0^{1-x/2} \int_0^{x/2} f(x, y, z) \, dz \, dy \, dx &= \int_0^1 \int_{2z}^2 \int_0^{1-x/2} f(x, y, z) \, dy \, dx \, dz \\ &= \int_0^1 \int_0^{1-z} \int_{2z}^{2-2y} f(x, y, z) \, dx \, dy \, dz . \end{aligned}$$

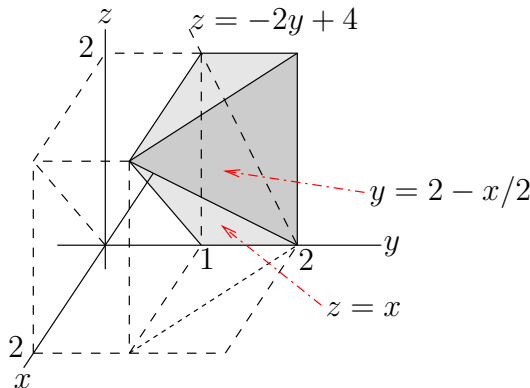
f) The domain of integration is represented in the following figure.



We have

$$\begin{aligned} \int_0^4 \int_{\sqrt{x}}^2 \int_0^{2-y} f(x, y, z) \, dz \, dy \, dx &= \int_0^2 \int_0^{2-y} \int_0^{y^2} f(x, y, z) \, dx \, dz \, dy \\ &= \int_0^2 \int_0^{(2-z)^2} \int_{\sqrt{x}}^{2-z} f(x, y, z) \, dy \, dx \, dz . \end{aligned}$$

g) The domain of integration is represented in the figure below.



We have

$$\int_0^2 \int_1^{2-x/2} \int_x^2 f(x, y, z) \, dz \, dy \, dx = \int_0^2 \int_0^z \int_1^{2-x/2} f(x, y, z) \, dy \, dx \, dz .$$

The projection on the y, z plane of the line of intersection of the planes $z = x$ and $y = 2 - x/2$ is the line given by $z = 4 - 2y$. Hence, we get

$$\begin{aligned} \int_0^2 \int_1^{2-x/2} \int_x^2 f(x, y, z) \, dz \, dy \, dx &= \int_0^2 \int_1^{2-z/2} \int_0^z f(x, y, z) \, dx \, dy \, dz \\ &\quad + \int_0^2 \int_{2-z/2}^2 \int_0^{4-2y} f(x, y, z) \, dx \, dy \, dz . \end{aligned}$$

♡

Question 3.37

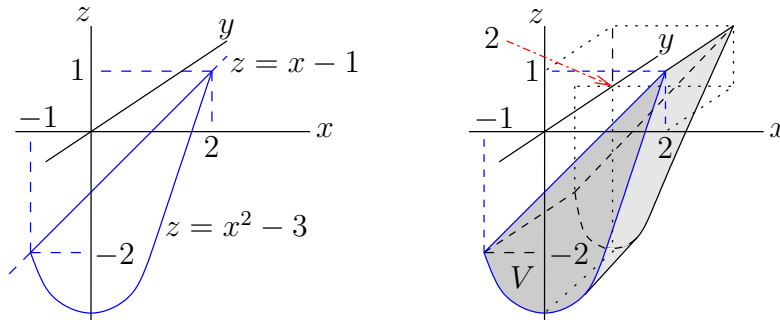
A solid V is defined by the region enclosed by the surfaces $z = x^2 - 3$, $z = x - 1$, $y = 0$ and $y = 2$. If the density of the solid is given by $\rho(x, y, z) = x^2 y$, compute the mass of the solid.

Solution

The abscissas of the points of intersection of the curve $z = x^2 - 3$ and the line $z = x - 1$ are given by

$$x^2 - 3 = x - 1 \Leftrightarrow x^2 - x - 2 = (x - 2)(x + 1) = 0 \Leftrightarrow x = -1 \text{ or } x = 2 .$$

Thus the points of intersection are $(x, z) = (-1, -2)$ and $(x, z) = (2, 1)$. The two surfaces in $\mathbb{R}^3$ given by $z = x^2 - 3$ and $z = x - 1$ intersect each along the lines $\{(-1, y, 2) : y \in \mathbb{R}\}$ and $\{(2, y, 1) : y \in \mathbb{R}\}$. The solid V is represented in the following figure.



The mass m of the solid is

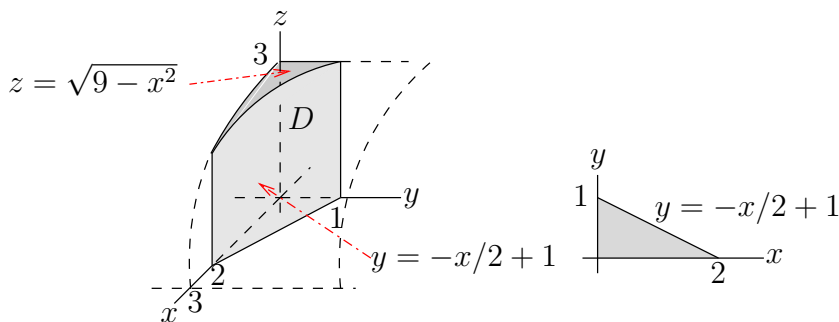
$$\begin{aligned} m &= \int_0^2 \int_{-1}^2 \int_{x^2-3}^{x-1} x^2 y \, dz \, dx \, dy = \int_0^2 \int_{-1}^2 (x^2 y z) \Big|_{z=x^2-3}^{x-1} dx \, dy \\ &= \int_0^2 \int_{-1}^2 y(-x^4 + x^3 + 2x^2) \, dx \, dy = \int_0^2 y \left(-\frac{x^5}{5} + \frac{x^4}{4} + \frac{2x^3}{3} \right) \Big|_{x=-1}^2 dy \\ &= \frac{63}{20} \int_0^2 y \, dy = \frac{63}{20} \frac{y^2}{2} \Big|_0^2 = \frac{63}{10} . \end{aligned}$$

**Question 3.38**

Compute the volume of the solid D in the region $x, y, z \geq 0$ bounded by the cylinder $x^2 + z^2 = 9$ and the planes $y = -x/2 + 1$, $x = 0$, $y = 0$ and $z = 0$.

Solution

The solid D is represented in the following figure.



The volume V is

$$\begin{aligned}
 V &= \iiint_D 1 = \int_0^2 \int_0^{-x/2+1} \int_0^{\sqrt{9-x^2}} dz \, dy \, dx = \int_0^2 \int_0^{-x/2+1} z \Big|_{y=0}^{\sqrt{9-x^2}} dy \, dx \\
 &= \int_0^2 \int_0^{-x/2+1} \sqrt{9-x^2} \, dy \, dx = \int_0^2 \left(y \sqrt{9-x^2} \right) \Big|_0^{-x/2+1} dx \\
 &= \int_0^2 \left(1 - \frac{x}{2} \right) \sqrt{9-x^2} \, dx = \left(\frac{x\sqrt{9-x^2}}{2} + \frac{9}{2} \arcsin \left(\frac{x}{3} \right) + \frac{(9-x^2)^{3/2}}{6} \right) \Big|_{x=0}^2 \\
 &= \frac{11\sqrt{5}}{6} + \frac{9}{2} \arcsin \left(\frac{2}{3} \right) - \frac{9}{2}.
 \end{aligned}$$

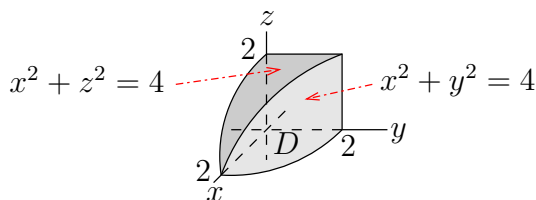
We have used the substitution $u = 9 - x^2$ to compute the integral $\int_0^2 \frac{x}{2} \sqrt{9-x^2} \, dx$, and the trigonometric substitution $x = 3 \sin(\theta)$ to compute the integral $\int_0^2 \sqrt{9-x^2} \, dx$. ♡

Question 3.39

Let R be the solid obtained from the intersection of the cylinders $x^2 + y^2 = 4$ and $x^2 + z^2 = 4$. Compute the volume of R and its mass if the density is given by $\rho(x, y, z) = x^2$.

Solution

The portion of the solid R in the region $x, y, z \geq 0$ is represented by the solid D in the figure below.



The solid D is one eighth of the solid R . The volume of R is

$$\begin{aligned} V &= \iiint_R 1 = 8 \iiint_D 1 = 8 \int_0^2 \int_0^{\sqrt{4-x^2}} \int_0^{\sqrt{4-x^2}} dz dy dx \\ &= 8 \int_0^2 \int_0^{\sqrt{4-x^2}} z \Big|_{z=0}^{\sqrt{4-x^2}} dy dx = 8 \int_0^2 \int_0^{\sqrt{4-x^2}} \sqrt{4-x^2} dy dx \\ &= 8 \int_0^2 (y\sqrt{4-x^2}) \Big|_{y=0}^{\sqrt{4-x^2}} dx = 8 \int_0^2 (4-x^2) dx = 8 \left(4x - \frac{x^3}{3} \right) \Big|_{x=0}^2 = \frac{128}{3}. \end{aligned}$$

Since the density is independent of y and z , and even with respect to x , the mass of R is

$$\begin{aligned} m &= \iiint_R \rho = 8 \iiint_D x^2 = 8 \int_0^2 \int_0^{\sqrt{4-x^2}} \int_0^{\sqrt{4-x^2}} x^2 dz dy dx \\ &= 8 \int_0^2 x^2(4-x^2) dx = 8 \left(\frac{4x^3}{3} - \frac{x^5}{5} \right) \Big|_{x=0}^2 = \frac{512}{15}. \end{aligned}$$

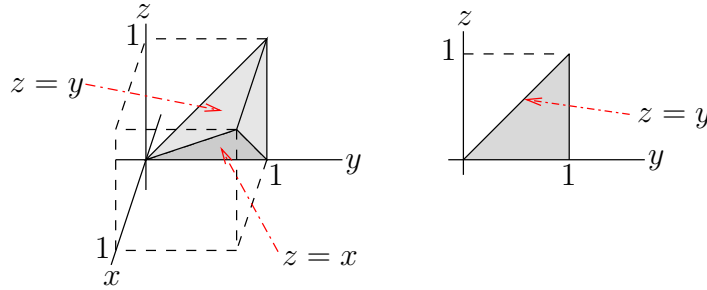
♥

Question 3.40

If D is the pyramid with vertices at $(0, 0, 0)$, $(1, 1, 1)$, $(0, 1, 0)$ and $(0, 1, 1)$, compute $\iiint_D xz$.

Solution

The pyramid D is represented in the following figure.



There are several possibilities for the order of integration, each as good as the other. We choose $dx dz dy$. Hence

$$\begin{aligned} \iiint_D xz &= \int_0^1 \int_0^y \int_0^z xz dx dz dy = \int_0^1 \int_0^y \frac{x^2 z}{2} \Big|_{x=0}^z dz dy \\ &= \int_0^1 \int_0^y \frac{z^3}{2} dz dy = \int_0^1 \frac{z^4}{8} \Big|_{z=0}^y dy = \int_0^1 \frac{y^4}{8} dy = \frac{y^5}{40} \Big|_{y=0}^1 = \frac{1}{40}. \end{aligned}$$

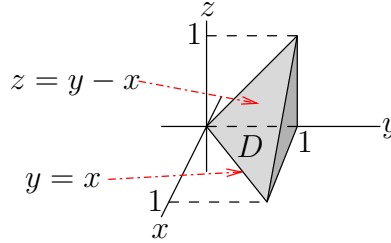
♥

Question 3.41

If D is the pyramid with vertices at $(0, 0, 0)$, $(1, 1, 0)$, $(0, 1, 0)$ and $(0, 1, 1)$, compute $\iiint_D xz$.

Solution

The pyramid D is represented in the following figure.



To find the equation of the plane through the points $(0,0,0)$, $(1,1,0)$ and $(0,1,1)$, we consider the vectors $\mathbf{v}_1 = (1, 1, 0)$ and $\mathbf{v}_2 = (0, 1, 1)$ in the plane. The vector $\mathbf{n} = \mathbf{v}_1 \times \mathbf{v}_2 = (1, -1, 1)$ is perpendicular to the plane. The equation of the plane is of the form $\langle \mathbf{n}, (x, y, z) \rangle = x - y + z = C$ where C is a constant. Since the point $(0,0,0)$ is on that plane, we must have $C = 0$. Thus, the equation of the plane is $z = y - x$. Hence

$$\begin{aligned} \iiint_D xz &= \int_0^1 \int_x^1 \int_0^{y-x} xz \, dz \, dy \, dx = \int_0^1 \int_x^1 \frac{xz^2}{2} \Big|_{z=0}^{y-x} \, dy \, dx \\ &= \int_0^1 \int_x^1 \frac{x(y-x)^2}{2} \, dy \, dx = \frac{1}{6} \int_0^1 x(y-x)^3 \Big|_{y=x}^1 \, dx \\ &= \frac{1}{6} \int_0^1 (-x^4 + 3x^3 - 3x^2 + x) \, dx = \frac{1}{6} \left(-\frac{x^5}{5} + \frac{3x^4}{4} - x^3 + \frac{x^2}{2} \right) \Big|_{x=0}^1 = \frac{1}{120}. \end{aligned}$$

♡

Question 3.42

If D is the pyramid with vertices at $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ and $(0, 0, 0)$, compute

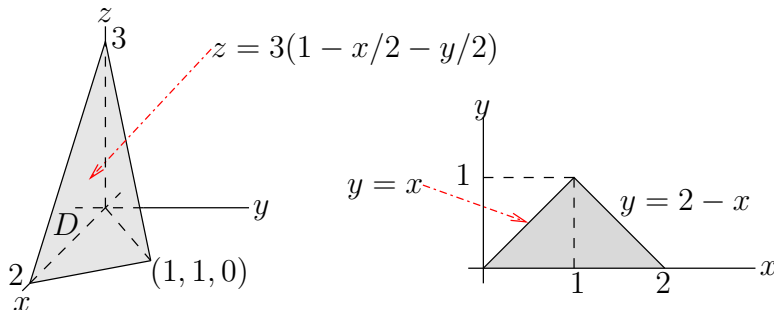
$$\iiint_D (xy + z).$$

Question 3.43

If D is the pyramid with vertices at $(0, 0, 0)$, $(2, 0, 0)$, $(1, 1, 0)$ and $(0, 0, 3)$, compute $\iiint_D xy$.

Solution

We have drawn the pyramid D in the following figure.



To find the equation of the plane through the points $(2, 0, 0)$, $(1, 1, 0)$ and $(0, 0, 3)$, we consider the vectors $\mathbf{v}_1 = (2, 0, -3)$ and $\mathbf{v}_2 = (1, 1, -3)$ that are parallel to the plane. The vector $\mathbf{n} = \mathbf{v}_1 \times \mathbf{v}_2 = (3, 3, 2)$ is perpendicular to the plane. The equation of the plane is of the form $\langle \mathbf{n}, (x, y, z) \rangle = 3x + 3y + 2z = C$ where C is a constant. Since the point $(2, 0, 0)$ is on that plane, we must have $C = 6$. Thus, the equation of the plane is $z = 3 \left(1 - \frac{x}{2} - \frac{y}{2}\right)$. Hence

$$\begin{aligned} \iiint_D xy &= \int_0^1 \int_y^{2-y} \int_0^{3(1-x/2-y/2)} xy \, dz \, dx \, dy = \int_0^1 \int_y^{2-y} xyz \Big|_{z=0}^{3(1-x/2-y/2)} dx \, dy \\ &= \int_0^1 \int_y^{2-y} \left(3xy - \frac{3x^2y}{2} - \frac{3xy^2}{2} \right) dx \, dy = \int_0^1 \left(\frac{3x^2y}{2} - \frac{x^3y}{2} - \frac{3x^2y^2}{4} \right) \Big|_{x=y}^{2-y} dy \\ &= \int_0^1 (y^4 - 3y^2 + 2y) \, dy = \left(\frac{y^5}{5} - y^3 + y^2 \right) \Big|_{y=0}^1 = \frac{1}{5}. \end{aligned}$$

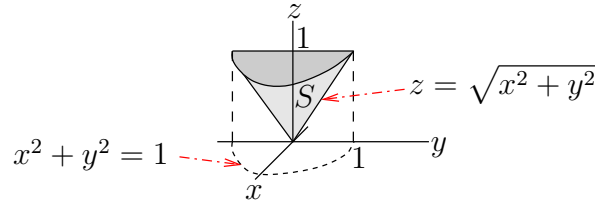
♥

Question 3.44

Find the coordinates of the centroid of the half-cone given by $\sqrt{x^2 + y^2} \leq z \leq 1$ and $x \geq 0$. We assume that the density ρ is constant to 1; namely, $\rho(x, y, z) = 1$ for all $(x, y, z) \in S$.

Solution

We have drawn the half-cone S in the figure below.



We use the cylindrical coordinates $x = r \cos(\theta)$, $y = r \sin(\theta)$ and $z = z$ for $-\pi/2 \leq \theta \leq \pi/2$, $0 \leq r \leq 1$ and $r = \sqrt{x^2 + y^2} \leq z \leq 1$. We have that $\left| \frac{\partial(x, y, z)}{\partial(r, \theta, z)} \right| = r$. The mass is

$$m = \iiint_S \rho = \int_{-\pi/2}^{\pi/2} \int_0^1 \int_r^1 r \, dz \, dr \, d\theta = \frac{\pi}{6}.$$

The coordinates $(\bar{x}, \bar{y}, \bar{z})$ of the centroid are

$$\bar{x} = \frac{1}{m} \iiint_S x \rho = \frac{6}{\pi} \int_{-\pi/2}^{\pi/2} \int_0^1 \int_r^1 r^2 \cos(\theta) \, dz \, dr \, d\theta = \frac{1}{\pi},$$

$$\bar{y} = \frac{1}{m} \iiint_S y \rho = \frac{6}{\pi} \int_{-\pi/2}^{\pi/2} \int_0^1 \int_r^1 r^2 \sin(\theta) \, dz \, dr \, d\theta = 0$$

and

$$\bar{z} = \frac{1}{m} \iiint_S z \rho = \frac{6}{\pi} \int_{-\pi/2}^{\pi/2} \int_0^1 \int_r^1 zr \, dz \, dr \, d\theta = \frac{3}{4}.$$

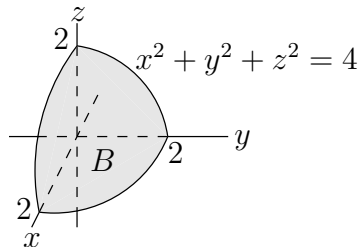
That $\bar{y} = 0$ should not be a surprise the reader because the half-cone above is symmetric with respect to the plane $y = 0$ and the density is constant. ♡

Question 3.45

Find the coordinates of the centroid of the portion B of the ball $x^2 + y^2 + z^2 \leq 4$ for $x, y, z \geq 0$. The density ρ of the ball is constant.

Solution

The portion B of the ball is represented in the figure below.



We use the spherical coordinates $x = r \cos(\theta) \sin(\phi)$, $y = r \sin(\theta) \sin(\phi)$ and $z = r \cos(\phi)$ for $0 \leq \theta \leq \pi/2$, $0 \leq \phi \leq \pi/2$ and $0 \leq r \leq 2$. We have that $\left| \frac{\partial(x, y, z)}{\partial(r, \theta, \phi)} \right| = r^2 \sin(\phi)$. The mass is

$$m = \iiint_B \rho = \rho \int_0^{\pi/2} \int_0^{\pi/2} \int_0^2 r^2 \sin(\phi) dr d\phi d\theta = \frac{4\pi\rho}{3}.$$

The coordinates $(\bar{x}, \bar{y}, \bar{z})$ of the centroid are:

$$\bar{x} = \frac{1}{m} \iiint_B x\rho = \frac{3}{4\pi} \int_0^{\pi/2} \int_0^{\pi/2} \int_0^2 r^3 \cos(\theta) \sin^2(\phi) dr d\phi d\theta = \frac{3}{4},$$

$$\bar{y} = \frac{1}{m} \iiint_B y\rho = \frac{3}{4\pi} \int_0^{\pi/2} \int_0^{\pi/2} \int_0^2 r^3 \sin(\theta) \sin^2(\phi) dr d\phi d\theta = \frac{3}{4}$$

and

$$\bar{z} = \frac{1}{m} \iiint_B z\rho = \frac{3}{4\pi} \int_0^{\pi/2} \int_0^{\pi/2} \int_0^2 r^3 \cos(\phi) \sin(\phi) dr d\phi d\theta = \frac{3}{4}.$$

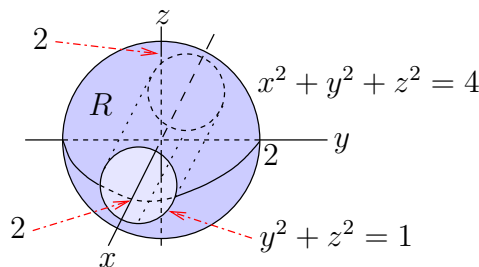
It is not surprising that all three coordinates of the centroid are equal since the density is constant and the solid B is invariant under permutations of the axes. ♡

Question 3.46

Compute the volume V of the region R enclosed by the sphere $x^2 + y^2 + z^2 = 4$ and the cylinder $y^2 + z^2 = 1$.

Solution

The easiest way to compute the volume V is to subtract the volume V_1 of the part of the cylinder inside the sphere from the volume V_2 of the sphere. The region R is represented in the following figure.



i) We compute the volume V_1 of the part of the cylinder inside the sphere. We use the cylindrical coordinates $x = r \cos(\theta)$, $y = r \sin(\theta)$ and $z = \sqrt{4 - r^2}$ for $-\sqrt{4 - r^2} \leq z \leq \sqrt{4 - r^2}$, $0 \leq r \leq 1$ and $0 \leq \theta \leq 2\pi$. We have that $\left| \frac{\partial(x, y, z)}{\partial(r, \theta, z)} \right| = r$. Hence

$$\begin{aligned} V_1 &= \int_0^{2\pi} \int_0^1 \int_{-\sqrt{4-r^2}}^{\sqrt{4-r^2}} r \, dz \, dr \, d\theta = \int_0^{2\pi} \int_0^1 (r \cdot 2\sqrt{4-r^2}) \, dr \, d\theta = \int_0^{2\pi} \int_0^1 2r\sqrt{4-r^2} \, dr \, d\theta \\ &= \int_0^{2\pi} \left(-\frac{2}{3}(4-r^2)^{3/2} \right) \Big|_{r=0}^1 d\theta = \frac{2(8-3\sqrt{3})}{3} \int_0^{2\pi} d\theta = \frac{4(8-3\sqrt{3})\pi}{3}. \end{aligned}$$

ii) Assuming that we have forgotten the formula to compute the volume of a sphere and that we do not have access to tables of formulae, we compute the volume V_2 of the sphere using integrals. This also provides a proof of the formulae to compute the volume of a sphere. We use the spherical coordinates $x = r \cos(\theta) \sin(\phi)$, $y = r \sin(\theta) \sin(\phi)$ and $z = r \cos(\phi)$ for $0 \leq r \leq 2$, $0 \leq \theta \leq 2\pi$ and $0 \leq \phi \leq \pi$. We have that $\left| \frac{\partial(x, y, z)}{\partial(r, \theta, \phi)} \right| = r^2 \sin(\phi)$. Hence

$$\begin{aligned} V_2 &= \int_0^{2\pi} \int_0^\pi \int_0^2 r^2 \sin(\phi) \, dr \, d\phi \, d\theta = \int_0^{2\pi} \int_0^\pi \left(\frac{r^3}{3} \sin(\phi) \right) \Big|_{r=0}^2 d\phi \, d\theta \\ &= \frac{8}{3} \int_0^{2\pi} \int_0^\pi \sin \phi \, d\phi \, d\theta = \frac{8}{3} \int_0^{2\pi} (-\cos(\phi)) \Big|_{\phi=0}^\pi d\theta = \frac{16}{3} \int_0^{2\pi} d\theta = \frac{32\pi}{3}. \end{aligned}$$

iii) Therefore, the volume of the region is $V = \frac{32\pi}{3} - \frac{4(8-3\sqrt{3})\pi}{3} = 4\sqrt{3}\pi$. ♡

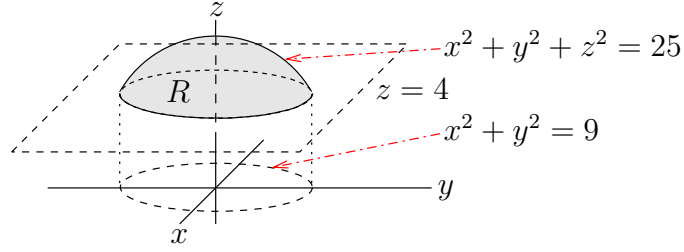
Question 3.47

Find the volume V of the solid R enclosed between the upper part of the sphere $x^2 + y^2 + z^2 = 25$ and the plane $z = 4$.

- Setup but do not compute the integral in Cartesian coordinates.
- Setup but do not compute the integral in cylindrical coordinates.
- Setup but do not compute the integral in polar coordinates.
- Compute the easiest integral to get the volume.

Solution

The intersection of the sphere $x^2 + y^2 + z^2 = 25$ with the plane $z = 4$ is the circle $x^2 + y^2 = 9$. We have drawn the solid R in the figure below.

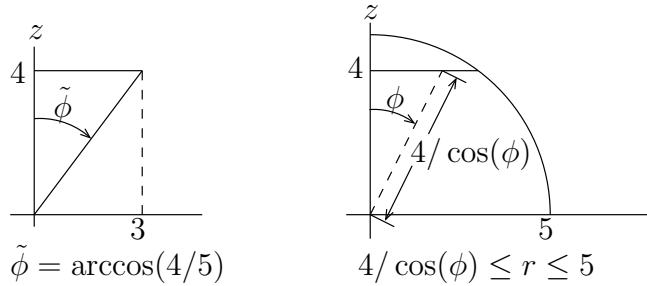


a) In Cartesian coordinates, the volume is $V = \int_{-3}^3 \int_{-\sqrt{9-x^2}}^{\sqrt{9-x^2}} \int_4^{\sqrt{25-x^2-y^2}} dz dy dx$.

b) We use the cylindrical coordinates $x = r \cos(\theta)$, $y = r \sin(\theta)$ and $z = z$ for $4 \leq z \leq \sqrt{25-r^2}$, $0 \leq r \leq 3$ and $0 \leq \theta \leq 2\pi$. We have that $\left| \frac{\partial(x, y, z)}{\partial(r, \theta, z)} \right| = r$. The volume is

$$V = \int_0^{2\pi} \int_0^3 \int_4^{\sqrt{25-r^2}} r dz dr d\theta.$$

c) We use the spherical coordinates $x = r \cos(\theta) \sin(\phi)$, $y = r \sin(\theta) \sin(\phi)$ and $z = r \cos(\phi)$ for $4/\cos(\phi) \leq r \leq 5$, $0 \leq \phi \leq \pi/3$ and $0 \leq \theta \leq 2\pi$.



We have that $\left| \frac{\partial(x, y, z)}{\partial(r, \theta, \phi)} \right| = r^2 \sin(\phi)$. The volume is

$$V = \int_0^{2\pi} \int_0^{\arccos(4/5)} \int_{4/\cos(\phi)}^5 r^2 \sin(\phi) dr d\phi d\theta.$$

d) The easiest integral to compute is in cylindrical coordinates though the integral in spherical coordinates is not too bad either. The integral in Cartesian coordinates requires a lot more work. The integral in (b) gives

$$\begin{aligned} V &= \int_0^{2\pi} \int_0^3 \int_4^{\sqrt{25-r^2}} r dz dr d\theta = \int_0^{2\pi} \int_0^3 (rz) \Big|_{z=4}^{\sqrt{25-r^2}} dr d\theta \\ &= \int_0^{2\pi} \int_0^3 \left(r\sqrt{25-r^2} - 4r \right) dr d\theta = \int_0^{2\pi} \left(-\frac{1}{3}(25-r^2)^{3/2} - 2r^2 \right) \Big|_{r=0}^3 d\theta \\ &= \left(-\frac{64}{3} - 18 + \frac{125}{3} \right) \int_0^{2\pi} d\theta = \frac{14\pi}{3}. \end{aligned}$$

The integral in (c) gives

$$\begin{aligned}
 V &= \int_0^{2\pi} \int_0^{\arccos(4/5)} \int_{4/\cos(\phi)}^5 r^2 \sin(\phi) \, dr \, d\phi \, d\theta \\
 &= \int_0^{2\pi} \int_0^{\arccos(4/5)} \left(\frac{r^3}{3} \sin(\phi) \right) \Big|_{r=4/\cos(\phi)}^5 d\phi \, d\theta \\
 &= \int_0^{2\pi} \int_0^{\arccos(4/5)} \left(\frac{125}{3} \sin(\phi) - \frac{64 \sin(\phi)}{3 \cos^3(\phi)} \right) d\phi \, d\theta \\
 &= \int_0^{2\pi} \left(-\frac{125}{3} \cos(\phi) - \frac{64}{6 \cos^2(\phi)} \right) \Big|_{\phi=0}^{\arccos(4/5)} d\theta = \frac{14\pi}{3}.
 \end{aligned}$$

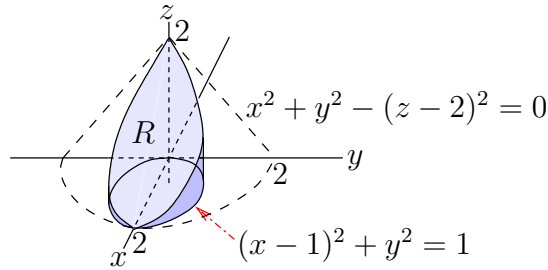
♡

Question 3.48

Compute the volume V of the bounded region R enclosed part the cylinder $(x-1)^2 + y^2 = 1$, the cone $x^2 + y^2 - (z-2)^2 = 0$ and the x, y plane.

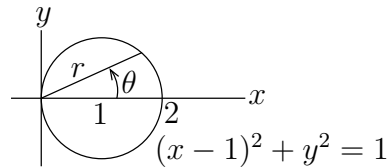
Solution

We have drawn the region R in the figure below.



Because of the cylinder, we are first inclined to use the cylindrical coordinates $x = 1 + r \cos(\theta)$, $y = r \sin(\theta)$ and $z = z$ for $0 \leq z \leq 2 - \sqrt{x^2 + y^2} = 2 - \sqrt{r^2 + 2r \cos(\theta) + 1}$, $0 \leq r \leq 1$ and $0 \leq \theta \leq 2\pi$. Unfortunately, the reader may verify that this leads to a difficult integral to compute.

We are still going to use cylindrical coordinates but in a non-usual way. We consider $x = r \cos(\theta)$, $y = r \sin(\theta)$ and $z = z$ for $0 \leq z \leq 2 - r$, $0 \leq r \leq 2 \cos(\theta)$ and $-\pi/2 \leq \theta \leq \pi/2$.



To determine the range of r , we use

$$(x-1)^2 + y^2 = 1 \Rightarrow (r \cos(\theta) - 1)^2 + r^2 \sin^2(\theta) = 1 \Rightarrow r^2 - 2r \cos(\theta) = 0$$

$$\Rightarrow r = 0 \text{ or } r = 2 \cos(\theta) .$$

To determine the range of z , we use

$$x^2 + y^2 - (z - 2)^2 = 0 \Rightarrow r^2 - (z - 2)^2 = 0 \Rightarrow z = 2 - r$$

for $z - 2, r \geq 0$. Since $\left| \frac{\partial(x, y, z)}{\partial(r, \theta, z)} \right| = r$, the volume is

$$\begin{aligned} V &= \int_{-\pi/2}^{\pi/2} \int_0^{2 \cos(\theta)} \int_0^{2-r} r \, dz \, dr \, d\theta = \int_{-\pi/2}^{\pi/2} \int_0^{2 \cos(\theta)} (zr) \Big|_{z=0}^{2-r} dr \, d\theta \\ &= \int_{-\pi/2}^{\pi/2} \int_0^{2 \cos(\theta)} (2r - r^2) \, dr \, d\theta = \int_{-\pi/2}^{\pi/2} \left(r^2 - \frac{r^3}{3} \right) \Big|_0^{2 \cos(\theta)} d\theta \\ &= \int_{-\pi/2}^{\pi/2} \left(4 \cos^2(\theta) - \frac{8 \cos^3(\theta)}{3} \right) d\theta \\ &= \int_{-\pi/2}^{\pi/2} \left(2(1 + \cos(2\theta)) - \frac{8 \cos(\theta)}{3} + \frac{8 \sin^2(\theta) \cos(\theta)}{3} \right) d\theta \\ &= \left(2\theta + \sin(2\theta) - \frac{8 \sin(\theta)}{3} + \frac{8 \sin^3(\theta)}{9} \right) \Big|_{-\pi/2}^{\pi/2} = 2\pi - \frac{32}{9} . \end{aligned}$$

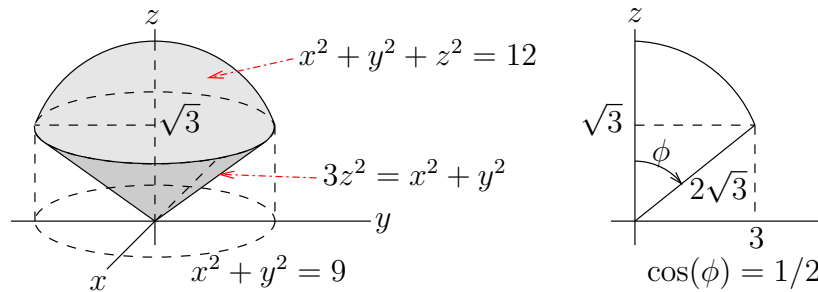
♡

Question 3.49

Set up an integral in cylindrical coordinates to compute the volume of the solid S enclosed by the sphere $x^2 + y^2 + z^2 = 12$ and the cone $3z^2 = x^2 + y^2$ for $z \geq 0$. Answer the same question in spherical coordinates.

Solution

The solid S is represented in the following figure.



On the intersection of the sphere $x^2 + y^2 + z^2 = 12$ and the cone $3z^2 = x^2 + y^2$ where $z \geq 0$, we have

$$3z^2 = x^2 + y^2 = 12 - z^2 \Rightarrow 4z^2 = 12 \Rightarrow z = \pm\sqrt{3} .$$

Thus, the intersection of the two surfaces is the circle $x^2 + y^2 = 9$ with $z = \sqrt{3}$.

In cylindrical coordinates, we have that $x = r \cos(\theta)$, $y = r \sin(\theta)$ and $z = z$ for $0 \leq \theta \leq 2\pi$, $0 \leq r \leq 3$ and $r/\sqrt{3} \leq z \leq \sqrt{12-r^2}$. Since $|\det Dg(r, \theta, z)| = r$, the volume of the solid is

$$V = \int_S 1 = \int_0^{2\pi} \int_0^3 \int_{r/\sqrt{3}}^{\sqrt{12-r^2}} r \, dz \, dr \, d\theta .$$

In spherical coordinates, we have that $x = r \cos(\theta) \sin(\phi)$, $y = r \sin(\theta) \sin(\phi)$ and $z = r \cos(\phi)$ for $0 \leq \theta \leq 2\pi$, $0 \leq \phi \leq \pi/3$ and $0 \leq r \leq 2\sqrt{3}$. Since $|\det Dg| = r^2 \sin(\phi)$, the volume of the solid is

$$V = \int_S 1 = \int_0^{2\pi} \int_0^{\pi/3} \int_0^{2\sqrt{3}} r \sin^2(\phi) \, dr \, d\phi \, d\theta .$$

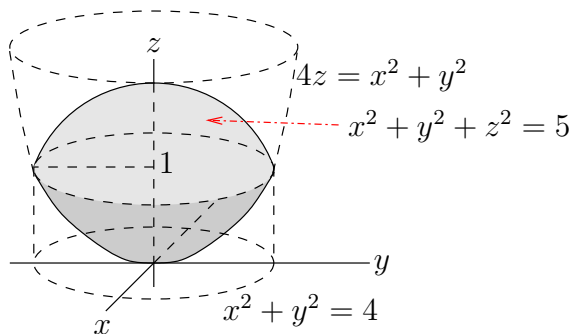
♡

Question 3.50

Set up an integral in cylindrical coordinates to compute the volume of the solid S enclosed by the sphere $x^2 + y^2 + z^2 = 5$ and the paraboloid $4z = x^2 + y^2$. Answer the same question in spherical coordinates.

Solution

The solid S is represented in the following figure.



On the intersection of the sphere $x^2 + y^2 + z^2 = 5$ and the paraboloid $4z = x^2 + y^2$ where $z \geq 0$, we have

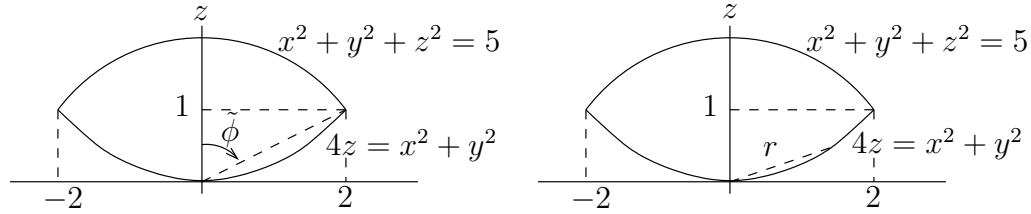
$$4z = x^2 + y^2 = 5 - z^2 \Rightarrow z^5 + 4z - 5 = (z + 5)(z - 1) = 0 \Rightarrow z = -5 \text{ or } z = 1 .$$

Thus, the intersection of the two surfaces is the circle $x^2 + y^2 = 4$ with $z = 1$.

In cylindrical coordinates, we have that $x = r \cos(\theta)$, $y = r \sin(\theta)$ and $z = z$ for $0 \leq \theta \leq 2\pi$, $0 \leq r \leq 2$ and $r^2/4 \leq z \leq \sqrt{5-r^2}$. Since $|\det Dg| = r$, the volume of the solid is

$$V = \int_0^{2\pi} \int_0^2 \int_{r^2/4}^{\sqrt{5-r^2}} r \, dz \, dr \, d\theta .$$

A cross section of the region through the z axis gives the following figure.



In spherical coordinate, we have that $x = r \cos(\theta) \sin(\phi)$, $y = r \sin(\theta) \sin(\phi)$ and $z = r \cos(\phi)$ for $0 \leq \theta \leq 2\pi$, $0 \leq \phi \leq \tilde{\phi}$ and $0 \leq r \leq \sqrt{5}$, and for $0 \leq \theta \leq 2\pi$, $\tilde{\phi} \leq \phi \leq \pi/2$ and $0 \leq r \leq 4 \cos(\phi) / \sin^2(\phi)$. We have that $\tilde{\phi} = \arctan(2)$. To determine the range of r for $\tilde{\phi} \leq \phi \leq \pi/2$, we used

$$\begin{aligned} 4z = x^2 + y^2 &\Rightarrow 4r \cos(\phi) = r^2 \cos^2(\theta) \sin^2(\phi) + r^2 \sin^2(\theta) \sin^2(\phi) = r^2 \sin^2(\phi) \\ &\Rightarrow r = 0 \text{ or } r = 4 \cos(\phi) / \sin^2(\phi) . \end{aligned}$$

Since $|\det Dg| = r^2 \sin(\phi)$, the volume of the solid is

$$\begin{aligned} V = \int_S 1 &= \int_0^{2\pi} \int_0^{\arctan(2)} \int_0^{\sqrt{5}} r^2 \sin(\phi) dr d\phi d\theta \\ &+ \int_0^{2\pi} \int_{\arctan(2)}^{\pi/2} \int_0^{4 \cos(\phi) / \sin^2(\phi)} r^2 \sin(\phi) dr d\phi d\theta . \end{aligned}$$

♡

Improper integrals

In this subsection, we provide another approach to expand the Riemann integral to unbounded domain or to domain where the function is unbounded. It is important to note that this approach is not equivalent to our general definition of integrable functions given in Definition 3.5.4. Our general definition requires that the functions be absolutely integrable on bounded open sets and that the sums obtained from partitions of unity be absolutely convergent. This is not required for the approach in this subsection. Do not forget that we now work only with functions which are continuous except on a set of measure zero. Functions integrable according to our general definition will be integrable according to the approach of this subsection but the converse is not true.

The approach in this subsection is a generalization of the concept of improper integrals for functions of one variable seen at the end of the previous chapter. The approach of this subsection introduces limits in the computations of the integrals but does not require absolute convergence of the limits. We could not ignore this approach because it is often used in engineering applications, often without proper justifications.

The improper integral is formally defined as $\int_S f = \lim_{q \rightarrow \infty} \int_{S(q)} f$ where $S(q_1) \subsetneq S(q_2) \subset S$ for $q_1 < q_2$ and $S(q) \rightarrow S$ as $q \rightarrow \infty$, or $S(q_1) \subsetneq S(q_2) \subset S$ for $q_2 < q_1$ and $S(q) \rightarrow S$ as $q \rightarrow 0^+$. We will not provide any rigorous definition of convergence for sets (though this

could be done) since we will not use improper integrals in higher dimensions in these lecture notes.

As an example of the difference between the definition of improper integrals given above and the general definition of integrals given in Definition 3.5.4, we consider the integral $\int_1^\infty \frac{\sin(x)}{x} dx$. We have seen in Question 2.39 that this improper integral converges but $\int_1^\infty \frac{|\sin(x)|}{x} dx$ diverges. Thus $\frac{\sin(x)}{x}$ is a converging improper integral on $[1, \infty[$ but it is not integrable on $[1, \infty[$ according to Definition 3.5.4.

The next questions will illustrate the general idea about improper integrals in higher dimensions.

Question 3.51

Determine if the following improper integrals converge or diverge. If they converge, then compute their value.

$$\begin{array}{ll} \text{a)} & \iiint_{\mathbb{R}^3} \frac{1}{1+x^2+y^2+z^2} \\ \text{c)} & \iiint_{x^2+y^2+z^2 < 1} \frac{z^2}{(x^2+y^2+z^2)^{3/2}} \\ \text{e)} & \iint_{x^2+y^2 < 1} \frac{x^2}{(x^2+y^2)^2} \end{array} \quad \begin{array}{ll} \text{b)} & \iint_{x,y>0} \frac{1}{(1+x^2+y^2)^3} \\ \text{d)} & \iint_{x>0} x e^{-x^2-y^2} \end{array}$$

Solution

a) We use spherical coordinates $x = r \cos(\theta) \sin(\phi)$, $y = r \sin(\theta) \sin(\phi)$ and $z = r \cos(\phi)$ for $0 \leq \theta < 2\pi$, $0 \leq \phi < \pi$ and $0 \leq r < q$ where $q \rightarrow \infty$. We have that $S(q) = B_q(\mathbf{0}) \subset \mathbb{R}^3$ in our formal definition of improper integrals. Since $\left| \frac{\partial(x, y, z)}{\partial(r, \theta, \phi)} \right| = r^2 \sin(\phi)$, we get

$$\begin{aligned} \iiint_{\mathbb{R}^3} \frac{1}{1+x^2+y^2+z^2} &= \lim_{q \rightarrow \infty} \int_0^q \int_0^\pi \int_0^{2\pi} \frac{r^2}{1+r^2} \sin(\phi) d\theta d\phi dr \\ &= \lim_{q \rightarrow \infty} 2\pi \int_0^q \int_0^\pi \frac{r^2}{1+r^2} \sin(\phi) d\phi dr = \lim_{q \rightarrow \infty} 2\pi \int_0^q \frac{r^2}{1+r^2} (-\cos(\phi)) \Big|_{\phi=0}^\pi dr \\ &= \lim_{q \rightarrow \infty} 4\pi \int_0^q \left(1 - \frac{1}{1+r^2} \right) dr = \lim_{q \rightarrow \infty} 4\pi (r - \arctan(r)) \Big|_{r=0}^q \\ &= \lim_{q \rightarrow \infty} 4\pi (q - \arctan(q)) = +\infty. \end{aligned}$$

The integral diverges.

b) We use polar coordinates $x = r \cos(\theta)$ and $y = r \sin(\theta)$ for $0 \leq \theta < \pi/2$ and $0 \leq r < q$ where $q \rightarrow \infty$. We have that $S(q) = B_q(\mathbf{0}) \cap \{(x, y) : x, y \geq 0\} \subset \mathbb{R}^2$ in our formal definition of improper integrals. Since $\left| \frac{\partial(x, y)}{\partial(r, \theta)} \right| = r$, we get

$$\iint_{x,y>0} \frac{1}{(1+x^2+y^2)^3} = \lim_{q \rightarrow \infty} \int_0^q \int_0^{\pi/2} \frac{r}{(1+r^2)^3} d\theta dr = \lim_{q \rightarrow \infty} \frac{\pi}{2} \int_0^q \frac{r}{(1+r^2)^3} dr$$

$$= \lim_{q \rightarrow \infty} -\frac{\pi}{8} \left(\frac{1}{(1+r^2)^2} \right) \Big|_{r=1}^q = \lim_{q \rightarrow \infty} -\frac{\pi}{8} \left(\frac{1}{(1+q^2)^2} - 1 \right) = \frac{\pi}{8}.$$

The integral converges.

c) We use spherical coordinates $x = r \cos(\theta) \sin(\phi)$, $y = r \sin(\theta) \sin(\phi)$ and $z = r \cos(\phi)$ for $0 \leq \theta < 2\pi$, $0 \leq \phi < \pi$ and $q \leq r \leq 1$ where $q \rightarrow 0$. We have that $S(q) = B_1(\mathbf{0}) \setminus B_q(\mathbf{0}) \subset \mathbb{R}^3$ in our formal definition of improper integrals. Since $\left| \frac{\partial(x, y, z)}{\partial(r, \theta, \phi)} \right| = r^2 \sin(\phi)$, we get

$$\begin{aligned} \iiint_{x^2+y^2+z^2 < 1} \frac{z^2}{(x^2+y^2+z^2)^{3/2}} &= \lim_{q \rightarrow 0^+} \int_q^1 \int_0^\pi \int_0^{2\pi} r \cos^2(\phi) \sin(\phi) d\theta d\phi dr \\ &= \lim_{q \rightarrow 0^+} 2\pi \int_q^1 \int_0^\pi r \cos^2(\phi) \sin(\phi) d\phi dr = \lim_{q \rightarrow 0^+} 2\pi \int_q^1 r \left(\frac{-\cos^3(\phi)}{3} \right) \Big|_{\phi=0}^\pi dr \\ &= \lim_{q \rightarrow 0^+} \frac{4\pi}{3} \int_q^1 r dr = \lim_{q \rightarrow 0^+} \frac{4\pi}{3} \left(\frac{r^2}{2} \right) \Big|_{r=q}^1 = \lim_{q \rightarrow 0^+} \frac{4\pi}{3} \left(\frac{1}{2} - \frac{q^2}{2} \right) = \frac{2\pi}{3}. \end{aligned}$$

The integral converges.

d) We use polar coordinates $x = r \cos(\theta)$ and $y = r \sin(\theta)$ for $-\pi/2 \leq \theta \leq \pi/2$ and $0 \leq r < q$ where $q \rightarrow \infty$. We have that $S(q) = B_q(\mathbf{0}) \cap \{(x, y) : x \geq 0\} \subset \mathbb{R}^2$ in our formal definition of improper integrals. Since $\left| \frac{\partial(x, y)}{\partial(r, \theta)} \right| = r$, we get

$$\begin{aligned} \iint_{x>0} x e^{-x^2-y^2} &= \lim_{q \rightarrow \infty} \int_0^q \int_{-\pi/2}^{\pi/2} r^2 \cos(\theta) e^{-r^2} d\theta dr = \lim_{q \rightarrow \infty} \int_0^q r^2 e^{-r^2} \sin(\theta) \Big|_{\theta=-\pi/2}^{\pi/2} dr \\ &= \lim_{q \rightarrow \infty} 2 \int_0^q r^2 e^{-r^2} dr = \lim_{q \rightarrow \infty} 2 \left(-\frac{r e^{-r^2}}{2} \Big|_0^q + \int_0^q \frac{e^{-r^2}}{2} dr \right) \\ &= \lim_{q \rightarrow \infty} 2 \left(-\frac{q e^{-q^2}}{2} + \int_0^q \frac{e^{-r^2}}{2} dr \right) = \int_0^\infty e^{-r^2} dr = \sqrt{\pi} \end{aligned}$$

because $\lim_{q \rightarrow \infty} q e^{-q^2} = 0$ as can be shown with l'Hospital Rule. We have use integration by parts to get the fourth equality above.

e) We use polar coordinates $x = r \cos(\theta)$ and $y = r \sin(\theta)$ for $0 \leq \theta < 2\pi$ and $q \leq r \leq 1$ where $q \rightarrow 0$. We have that $S(q) = B_1(\mathbf{0}) \setminus B_q(\mathbf{0}) \subset \mathbb{R}^2$ in our formal definition of improper integrals. Since $\left| \frac{\partial(x, y)}{\partial(r, \theta)} \right| = r$, we get

$$\begin{aligned} \iint_{x^2+y^2 < 1} \frac{x^2}{(x^2+y^2)^2} &= \lim_{q \rightarrow 0^+} \int_q^1 \int_0^{2\pi} \frac{1}{r} \cos^2(\theta) d\theta dr = \lim_{q \rightarrow 0^+} \frac{1}{2} \int_q^1 \int_0^{2\pi} \frac{1}{r} (1 + \cos(2\theta)) d\theta dr \\ &= \lim_{q \rightarrow 0^+} \frac{1}{2} \int_q^1 \frac{1}{r} \left(\theta + \frac{\sin(2\theta)}{2} \right) \Big|_0^{2\pi} dr = \lim_{q \rightarrow 0^+} \pi \int_q^1 \frac{1}{r} dr = \lim_{q \rightarrow 0^+} \pi \ln(r) \Big|_q^1 \\ &= -\lim_{q \rightarrow 0^+} \pi \ln(q) = \infty. \end{aligned}$$

The integral diverges.

♡

Chapter 4

Differential Forms

4.1 Multilinear Algebra

Before undertaking the study of differential forms and the integration of these forms, we need to introduce some concepts that are sometime covered in books on linear algebra.

Definition 4.1.1

Let H be a vector space over $\mathbb{R}$. A map $\tau : H^k \rightarrow \mathbb{R}$ is **multilinear** if

$$\tau(\mathbf{v}_1, \mathbf{v}_2, \dots, a\mathbf{v}_i + b\tilde{\mathbf{v}}_i, \dots, \mathbf{v}_k) = a\tau(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_i, \dots, \mathbf{v}_k) + b\tau(\mathbf{v}_1, \mathbf{v}_2, \dots, \tilde{\mathbf{v}}_i, \dots, \mathbf{v}_k)$$

for all $a, b \in \mathbb{R}$ and $0 \leq i \leq k$. A **k-tensor** on H is a multilinear map $\tau : H^k \rightarrow \mathbb{R}$. The set of all k -tensors on V is denoted $\mathcal{T}^k(H)$.

The space $\mathcal{T}^k(H)$ is a vector space over $\mathbb{R}$ where the sum of vectors and the product by a scalar are defined by

$$(a\tau_1 + b\tau_2)(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) = a\tau_1(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) + b\tau_2(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$$

for all $\tau_1, \tau_2 \in \mathcal{T}^k(H)$ and all $a, b \in \mathbb{R}$. It is also important to note that $\mathcal{T}^1(H) = H^*$, the dual space of H .

Definition 4.1.2

The **product of two tensors** $\tau_1 \in \mathcal{T}^{k_1}(H)$ and $\tau_2 \in \mathcal{T}^{k_2}(H)$ is the tensor $\tau_1 \otimes \tau_2 \in \mathcal{T}^{k_1+k_2}(H)$ defined by

$$(\tau_1 \otimes \tau_2)(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{k_1+k_2}) = \tau_1(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{k_1}) \tau_2(\mathbf{v}_{k_1+1}, \mathbf{v}_{k_1+2}, \dots, \mathbf{v}_{k_1+k_2}) .$$

Theorem 4.1.3

Let $\mathcal{B} = \{\mathbf{v}_i\}_{i=1}^n$ be a basis of H and let $\mathcal{B}^* = \{\phi_i\}_{i=1}^n$ be the dual basis associated to $\mathcal{B}$ ¹. Then the set $\mathcal{M} = \{\phi_{i_1} \otimes \phi_{i_2} \otimes \dots \otimes \phi_{i_k} : 1 \leq i_m \leq n, 1 \leq m \leq k\}$ is a basis of $\mathcal{T}^k(H)$.

Proof.

The proof is a standard proof in linear algebra. We note that

$$\begin{aligned} (\phi_{i_1} \otimes \phi_{i_2} \otimes \dots \otimes \phi_{i_k})(\mathbf{v}_{j_1}, \mathbf{v}_{j_2}, \dots, \mathbf{v}_{j_k}) &= \prod_{m=1}^k \phi_{i_m}(\mathbf{v}_{j_m}) = \prod_{m=1}^k \delta_{i_m, j_m} \\ &= \begin{cases} 1 & \text{if } j_m = i_m \text{ for } 1 \leq m \leq k \\ 0 & \text{otherwise} \end{cases} \end{aligned} \quad (4.1.1)$$

i) We prove that any $\tau \in \mathcal{T}^k(H)$ can be expressed as a linear combination of elements in $\mathcal{M}$.

Given $\mathbf{w}_j \in H$ for $1 \leq j \leq k$. Since $\mathcal{B}$ is a basis of H , we may express each $\mathbf{w}_j$ as $\mathbf{w}_j = \sum_{m_j=1}^n a_{m_j, j} \mathbf{v}_{m_j}$ for some $a_{m_j, j} \in \mathbb{R}$. We get from (4.1.1) that

$$\begin{aligned} &(\phi_{i_1} \otimes \phi_{i_2} \otimes \dots \otimes \phi_{i_k})(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k) \\ &= \sum_{\substack{1 \leq m_j \leq n \\ 1 \leq j \leq k}} a_{m_1, 1} a_{m_2, 2} \dots a_{m_k, k} (\phi_{i_1} \otimes \phi_{i_2} \otimes \dots \otimes \phi_{i_k})(\mathbf{v}_{m_1}, \mathbf{v}_{m_2}, \dots, \mathbf{v}_{m_k}) = a_{i_1, 1} a_{i_2, 2} \dots a_{i_k, k} . \end{aligned}$$

Hence

$$\begin{aligned} \tau(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k) &= \sum_{\substack{1 \leq i_m \leq n \\ 1 \leq m \leq k}} a_{i_1, 1} a_{i_2, 2} \dots a_{i_k, k} \tau(\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \dots, \mathbf{v}_{i_k}) \\ &= \sum_{\substack{1 \leq i_m \leq n \\ 1 \leq m \leq k}} \tau(\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \dots, \mathbf{v}_{i_k}) (\phi_{i_1} \otimes \phi_{i_2} \otimes \dots \otimes \phi_{i_k})(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k) . \end{aligned}$$

Since this is true for all $\mathbf{w}_i \in H$ with $1 \leq i \leq k$, we get

$$\tau = \sum_{\substack{1 \leq i_m \leq n \\ 1 \leq m \leq k}} \tau(\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \dots, \mathbf{v}_{i_k}) (\phi_{i_1} \otimes \phi_{i_2} \otimes \dots \otimes \phi_{i_k}) .$$

ii) We prove that the elements of $\mathcal{M}$ are linearly independent. Suppose that

$$\sum_{\substack{1 \leq i_m \leq n \\ 1 \leq m \leq k}} a_{i_1, i_2, \dots, i_k} \phi_{i_1} \otimes \phi_{i_2} \otimes \dots \otimes \phi_{i_k} = 0 .$$

¹Namely, $\phi_j(\mathbf{v}_i) = \delta_{i,j}$ for all i and j , where $\delta_{i,j}$ is the Kronecker delta function.

If we evaluate this expression at $(\mathbf{v}_{m_1}, \mathbf{v}_{m_2}, \dots, \mathbf{v}_{m_k})$ and use (4.1.1), then we get $a_{m_1, m_2, \dots, m_k} = 0$ for $1 \leq m_j \leq n$ and $1 \leq j \leq k$. ■

Definition 4.1.4

An **inner product** on H is a 2-tensor $\tau \in \mathcal{T}^2(H)$ such that τ is **symmetric** and **positive definite**².

This definition of the inner product corresponds to the standard definition of inner product on a vector space.

Theorem 4.1.5

Suppose that $\tau \in \mathcal{T}^2(H)$ is an inner product on a vector space H of dimension n over the real. There exists a basis $\mathcal{B} = \{\mathbf{v}_i\}_{i=1}^n$ of H such that $\tau(\mathbf{v}_i, \mathbf{v}_j) = \delta_{i,j}$ for all i and j .

The theorem is proved exactly as it is done in linear algebra using Gram-Schmidt orthogonalization process with the inner product $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle \equiv \tau(\mathbf{v}_1, \mathbf{v}_2)$ for $\mathbf{v}_1, \mathbf{v}_2 \in H$ and the norm $\|\mathbf{v}\| \equiv \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$ for $\mathbf{v} \in H$.

Definition 4.1.6

A basis $\mathcal{B}$ as in the statement of the Theorem 4.1.5 is called an **orthonormal basis** with respect to the inner product τ .

Let $\mathcal{B} = \{\mathbf{v}_i\}_{i=1}^n$ be a basis of H , it is possible to choose an inner product on H such that $\mathcal{B}$ is an orthonormal basis. Let $\mathcal{E} = \{\mathbf{e}_i\}_{i=1}^n$ be the canonical basis of $\mathbb{R}^n$ and $\langle, \rangle$ be the standard inner product on $\mathbb{R}^n$ defined by $\langle \mathbf{x}_1, \mathbf{x}_2 \rangle = \sum_{j=1}^n x_{j,1} x_{j,2}$ for $\mathbf{x}_i = \sum_{j=1}^n x_{j,i} \mathbf{e}_j$ and $1 \leq i \leq 2$. Moreover, let $f : H \rightarrow \mathbb{R}^n$ be the linear isomorphism defined by $f(\mathbf{v}_i) = \mathbf{e}_i$ for $1 \leq i \leq n$. We define the inner product $\tau \in \mathcal{T}^2(H)$ by

$$\begin{aligned} \tau(\mathbf{w}_1, \mathbf{w}_2) &= \langle f(\mathbf{w}_1), f(\mathbf{w}_2) \rangle = \left\langle \sum_{j=1}^n w_{j,1} f(\mathbf{v}_j), \sum_{j=1}^n w_{j,2} f(\mathbf{v}_j) \right\rangle = \left\langle \sum_{j=1}^n w_{j,1} \mathbf{e}_j, \sum_{j=1}^n w_{j,2} \mathbf{e}_j \right\rangle \\ &= \sum_{j=1}^n w_{j,1} w_{j,2} \end{aligned}$$

for $\mathbf{w}_i = \sum_{j=1}^n w_{j,i} \mathbf{v}_j$ and $1 \leq i \leq 2$. The basis $\mathcal{B}$ is an orthogonal basis with respect to the inner product τ .

²The 2-tensor τ is symmetric if $\tau(\mathbf{v}_1, \mathbf{v}_2) = \tau(\mathbf{v}_2, \mathbf{v}_1)$ for all $\mathbf{v}_1$ and $\mathbf{v}_2$ in H . The 2-tensor τ is positive definite if $\tau(\mathbf{v}, \mathbf{v}) > 0$ for all $\mathbf{v} \in H$ with $\mathbf{v} \neq \mathbf{0}$.

Definition 4.1.7

A **permutation** σ on $J_k = \{1, 2, \dots, k\}$ is a one-to-one and onto function $\sigma : J_k \rightarrow J_k$. A **transposition** on J_k is a permutation on J_k that interchanges only two elements of J_k and leaves the others fixed.

The set of all permutation on J_k forms a group with respect to the composition of functions that we denote S_k . There is a famous result in algebra that says that any permutation on J_k may be expressed as the composition of transpositions on J_k . Moreover, there exists a function sgn acting on the permutations on J_k such that $\text{sgn}(\sigma) = -1$ if σ is a transposition on J_k and $\text{sgn}(\sigma_1 \circ \sigma_2) = \text{sgn}(\sigma_1) \text{sgn}(\sigma_2)$ for all permutations σ_1 and σ_2 on J_k . In particular, for a permutation σ on J_k , $\text{sgn}(\sigma) = 1$ if and only if σ is the composition of an even number of transpositions. Therefore $\text{sgn}(\sigma) = -1$ if and only if σ is the composition of an odd number of transpositions. All these results are proved in standard books in (linear) algebra, for instance in [13].

Definition 4.1.8

A k -tensor $\tau \in \mathcal{T}^k(H)$ is called **alternating** if

$$\tau(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) = -\tau(\mathbf{v}_{\sigma(1)}, \mathbf{v}_{\sigma(2)}, \dots, \mathbf{v}_{\sigma(k)})$$

for any transposition σ on J_k . The set of all alternating k -tensor on H is denoted $\Omega^k(H)$.

It is easy to verify that $\Omega^k(H)$ is a linear subspace of $\mathcal{T}^k(H)$.

Proposition 4.1.9

Given $\tau \in \mathcal{T}^k(H)$, let

$$\text{Alt}(\tau)(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) = \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma) \tau(\mathbf{v}_{\sigma(1)}, \mathbf{v}_{\sigma(2)}, \dots, \mathbf{v}_{\sigma(k)})$$

for $\mathbf{v}_i \in H$ and $1 \leq i \leq k$, where S_k is the group of permutations on $J_k = \{1, 2, \dots, k\}$. Then $\text{Alt} : \mathcal{T}^k(H) \rightarrow \Omega^k(H)$.

Proof.

Suppose that $\tau \in \mathcal{T}^k(H)$ and $\tilde{\sigma}$ is a fixed transposition on J_k . We have that $\tilde{\sigma}^{-1} = \tilde{\sigma}$ and

$$\begin{aligned} \tilde{\sigma} : S_k &\rightarrow S_k \\ \sigma &\mapsto \sigma \circ \tilde{\sigma} \end{aligned}$$

is one-to-one and onto. Hence

$$\text{Alt}(\tau)(\mathbf{v}_{\tilde{\sigma}(1)}, \mathbf{v}_{\tilde{\sigma}(2)}, \dots, \mathbf{v}_{\tilde{\sigma}(k)}) = \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma) \tau(\mathbf{v}_{(\sigma \circ \tilde{\sigma})(1)}, \mathbf{v}_{(\sigma \circ \tilde{\sigma})(2)}, \dots, \mathbf{v}_{(\sigma \circ \tilde{\sigma})(k)})$$

$$\begin{aligned}
&= \frac{1}{k!} \sum_{\sigma \circ \tilde{\sigma} \in S_k} \operatorname{sgn}(\sigma \circ \tilde{\sigma}) \tau(\mathbf{v}_{\sigma(1)}, \mathbf{v}_{\sigma(2)}, \dots, \mathbf{v}_{\sigma(k)}) \\
&= \frac{1}{k!} \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) \operatorname{sgn}(\tilde{\sigma}) \tau(\mathbf{v}_{\sigma(1)}, \mathbf{v}_{\sigma(2)}, \dots, \mathbf{v}_{\sigma(k)}) = -\frac{1}{k!} \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) \tau(\mathbf{v}_{\sigma(1)}, \mathbf{v}_{\sigma(2)}, \dots, \mathbf{v}_{\sigma(k)}) \\
&= -\operatorname{Alt}(\tau)(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)
\end{aligned}$$

for all $\mathbf{v}_i \in H$ and $1 \leq i \leq k$. ■

Proposition 4.1.10

1. If $\tau \in \Omega^k(H)$, then $\operatorname{Alt}(\tau) = \tau$.
2. If $\tau \in \mathcal{T}^k(H)$, then $\operatorname{Alt}(\operatorname{Alt}(\tau)) = \operatorname{Alt}(\tau)$.
3. If $\tau_i \in \Omega^{k_i}(H)$ for $1 \leq i \leq 2$ and $\operatorname{Alt}(\tau_1) = 0$, then $\operatorname{Alt}(\tau_1 \otimes \tau_2) = \operatorname{Alt}(\tau_2 \otimes \tau_1) = 0$.
4. If $\tau_i \in \Omega^{k_i}(H)$ for $1 \leq i \leq 3$, then $\operatorname{Alt}(\operatorname{Alt}(\tau_1 \otimes \tau_2) \otimes \tau_3) = \operatorname{Alt}(\tau_1 \otimes \operatorname{Alt}(\tau_2 \otimes \tau_3)) = \operatorname{Alt}(\tau_1 \otimes \tau_2 \otimes \tau_3)$.

Proof.

1) Since $\sigma \in S_k$ is the product of transpositions, we have that $\tau(\mathbf{v}_{\sigma(1)}, \mathbf{v}_{\sigma(2)}, \dots, \mathbf{v}_{\sigma(k)}) = \operatorname{sgn}(\sigma) \tau(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$ for all $\mathbf{v}_i \in H$ and $1 \leq i \leq k$. Moreover, the number of different functions from J_k to J_k is $k!$. Namely, the number of elements of S_k , denoted $|S_k|$, is $k!$. Hence, for $\tau \in \Omega^k(H)$, we have

$$\begin{aligned}
\operatorname{Alt}(\tau)(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) &= \frac{1}{k!} \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) \tau(\mathbf{v}_{\sigma(1)}, \mathbf{v}_{\sigma(2)}, \dots, \mathbf{v}_{\sigma(k)}) \\
&= \left(\frac{1}{k!} \sum_{\sigma \in S_k} \underbrace{\operatorname{sgn}(\sigma) \operatorname{sgn}(\sigma)}_{=1} \right) \tau(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) = \tau(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)
\end{aligned}$$

for all $\mathbf{v}_i \in H$ and $1 \leq i \leq k$.

2) This follows from (1) because $\operatorname{Alt}(\tau) \in \Omega^k(H)$.

3) Let $\tilde{S}_{k_1} \subset S_{k_1+k_2}$ be the set of all permutations $\tilde{\sigma}$ on $J_{k_1+k_2}$ such that $\tilde{\sigma}(i) = i$ for $k_1 + 1 \leq i \leq k_2$. Namely, $\tilde{\sigma}$ permutes only the elements of the set $\{1, 2, \dots, k_1\}$. We have that $\tilde{S}_{k_1}$ is a subgroup of $S_{k_1+k_2}$. We choose a representative $\sigma_j \in S_{k_1+k_2}$ for each of the elements (i.e. equivalence classes) of the quotient group $S_{k_1+k_2}/\tilde{S}_{k_1}$. We end up with a set of permutations $S = \{\sigma_j : 1 \leq j \leq J\}$. We have that $S_{k_1+k_2} = \bigcup_{1 \leq j \leq J} \{\sigma_j \circ \tilde{\sigma} : \tilde{\sigma} \in \tilde{S}_{k_1}\}$ and each

element of $S_{k_1+k_2}$ has a unique representation of the form $\sigma_j \circ \tilde{\sigma}$ with $\tilde{\sigma} \in \tilde{S}_{k_1}$ and $\sigma_j \in S$. Hence

$$\operatorname{Alt}(\tau_1 \otimes \tau_2)(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{k_1+k_2}) = \frac{1}{(k_1+k_2)!} \sum_{\sigma \in S_{k_1+k_2}} \operatorname{sgn}(\sigma) \tau(\mathbf{v}_{\sigma(1)}, \mathbf{v}_{\sigma(2)}, \dots, \mathbf{v}_{\sigma(k_1+k_2)})$$

$$\begin{aligned}
&= \frac{1}{(k_1 + k_2)!} \sum_{j=1}^J \sum_{\tilde{\sigma} \in \tilde{S}_{k_1}} \operatorname{sgn}(\sigma_j \circ \tilde{\sigma})(\tau_1 \otimes \tau_2)(\mathbf{v}_{(\sigma_j \circ \tilde{\sigma})(1)}, \mathbf{v}_{(\sigma_j \circ \tilde{\sigma})(2)}, \dots, \mathbf{v}_{(\sigma_j \circ \tilde{\sigma})(k_1+k_2)}) \\
&= \frac{1}{(k_1 + k_2)!} \sum_{j=1}^J \sum_{\tilde{\sigma} \in \tilde{S}_{k_1}} \operatorname{sgn}(\sigma_j \circ \tilde{\sigma}) \operatorname{sgn}(\sigma_j)(\tau_1 \otimes \tau_2)(\mathbf{v}_{\tilde{\sigma}(1)}, \mathbf{v}_{\tilde{\sigma}(2)}, \dots, \mathbf{v}_{\tilde{\sigma}(k_1+k_2)}) \\
&= \frac{1}{(k_1 + k_2)!} \sum_{j=1}^J \sum_{\tilde{\sigma} \in \tilde{S}_{k_1}} \operatorname{sgn}(\tilde{\sigma})(\tau_1 \otimes \tau_2)(\mathbf{v}_{\tilde{\sigma}(1)}, \mathbf{v}_{\tilde{\sigma}(2)}, \dots, \mathbf{v}_{\tilde{\sigma}(k_1+k_2)}) \\
&= \frac{k_1!}{(k_1 + k_2)!} \sum_{j=1}^J \left(\frac{1}{k_1!} \sum_{\tilde{\sigma} \in \tilde{S}_{k_1}} \operatorname{sgn}(\tilde{\sigma}) \tau_1(\mathbf{v}_{\tilde{\sigma}(1)}, \mathbf{v}_{\tilde{\sigma}(2)}, \dots, \mathbf{v}_{\tilde{\sigma}(k_1)}) \right) \tau_2(\mathbf{v}_{k_1+1}, \mathbf{v}_{k_1+2}, \dots, \mathbf{v}_{k_1+k_2}) \\
&= \frac{J k_1!}{(k_1 + k_2)!} \operatorname{Alt}(\tau_1)(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{k_1}) \tau_2(\mathbf{v}_{k_1+1}, \mathbf{v}_{k_1+2}, \dots, \mathbf{v}_{k_1+k_2}) = 0
\end{aligned}$$

for all $\mathbf{v}_i \in H$ and $1 \leq i \leq k_1 + k_2$. The proof that $\operatorname{Alt}(\tau_2 \otimes \tau_1) = 0$ is similar. It suffices to use the subgroup $\tilde{S}_{k_1} \subset S_{k_1+k_2}$ of all permutations $\tilde{\sigma}$ on $J_{k_1+k_2}$ such that $\tilde{\sigma}(i) = i$ for $1 \leq i \leq k_2$.

4) It follows from (2) that $\operatorname{Alt}(\operatorname{Alt}(\tau_2 \otimes \tau_3) - \tau_2 \otimes \tau_3) = \operatorname{Alt}(\tau_2 \otimes \tau_3) - \operatorname{Alt}(\tau_2 \otimes \tau_3) = 0$. Hence, we get from (3) that

$$\begin{aligned}
0 &= \operatorname{Alt}(\tau_1 \otimes (\operatorname{Alt}(\tau_2 \otimes \tau_3) - \tau_2 \otimes \tau_3)) = \operatorname{Alt}(\tau_1 \otimes \operatorname{Alt}(\tau_2 \otimes \tau_3) - \tau_1 \otimes \tau_2 \otimes \tau_3) \\
&= \operatorname{Alt}(\tau_1 \otimes \operatorname{Alt}(\tau_2 \otimes \tau_3)) - \operatorname{Alt}(\tau_1 \otimes \tau_2 \otimes \tau_3) .
\end{aligned}$$

Similarly, we can prove that $\operatorname{Alt}(\operatorname{Alt}(\tau_1 \otimes \tau_2) \otimes \tau_3) - \operatorname{Alt}(\tau_1 \otimes \tau_2 \otimes \tau_3) = 0$. ■

Definition 4.1.11

The **wedge product** $\tau_1 \wedge \tau_2$ of $\tau_1 \in \Omega^{k_1}(H)$ and $\tau_2 \in \Omega^{k_2}(H)$ is the tensor $\tau_1 \wedge \tau_2 \in \Omega^{k_1+k_2}(H)$ defined by

$$\tau_1 \wedge \tau_2 = \frac{(k_1 + k_2)!}{k_1! k_2!} \operatorname{Alt}(\tau_1 \otimes \tau_2) .$$

Let H_1 and H_2 be two vector spaces over $\mathbb{R}$. Given a linear map $f : H_1 \rightarrow H_2$, we can define a map $f^* : \mathcal{T}^k(H_2) \rightarrow \mathcal{T}^k(H_1)$ by

$$(f^*(\tau))(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) = \tau(f(\mathbf{v}_1), f(\mathbf{v}_2), \dots, f(\mathbf{v}_k))$$

for all $\tau \in \mathcal{T}^k(H_2)$ and $\mathbf{v}_i \in H_1$ for $1 \leq i \leq k$. We say that $f^*(\tau)$ is the **pull-back** of τ by f .

The wedge product satisfies many relations. Some of them are listed in the next proposition.

Proposition 4.1.12

1. If $\tau_1, \tau_2 \in \Omega^{k_1}$ and $\tau_3 \in \Omega^{k_2}(H)$, then $(\tau_1 + \tau_2) \wedge \tau_3 = \tau_1 \wedge \tau_3 + \tau_2 \wedge \tau_3$.
2. If $\tau_1 \in \Omega^{k_1}(H)$ and $\tau_2, \tau_3 \in \Omega^{k_2}(H)$, then $\tau_1 \wedge (\tau_2 + \tau_3) = \tau_1 \wedge \tau_2 + \tau_1 \wedge \tau_3$.
3. If $\tau_i \in \Omega^{k_i}(H)$ for $1 \leq i \leq 2$ and $a \in \mathbb{R}$, then $(a\tau_1) \wedge \tau_2 = \tau_1 \wedge (a\tau_2) = a(\tau_1 \wedge \tau_2)$.
4. If $\tau_i \in \Omega^{k_i}(H)$ for $1 \leq i \leq 2$, then $\tau_1 \wedge \tau_2 = (-1)^{k_1 k_2} \tau_2 \wedge \tau_1$.
5. If $\tau_i \in \Omega^{k_i}(H)$ for $1 \leq i \leq 2$ and $f : H_1 \rightarrow H_2$ is a linear map from the vector space H_1 to the vector space H_2 , then $f^*(\tau_1 \wedge \tau_2) = f^*\tau_1 \wedge f^*\tau_2$.
6. If $\tau_i \in \Omega^{k_i}(H)$ for $1 \leq i \leq 3$, then

$$(\tau_1 \wedge \tau_2) \wedge \tau_3 = \tau_1 \wedge (\tau_2 \wedge \tau_3) = \frac{(k_1 + k_2 + k_3)!}{k_1! k_2! k_3!} \text{Alt}(\tau_1 \otimes \tau_2 \otimes \tau_3) .$$

Proof.

We leave the proof of (1), (2), (3) and (5) to the reader.

4) Given $\sigma \in S_{k_1+k_2}$, we have

$$\begin{aligned} & (\tau_1 \otimes \tau_2)(\mathbf{v}_{\sigma(1)}, \mathbf{v}_{\sigma(2)}, \dots, \mathbf{v}_{\sigma(k_1+k_2)}) \\ &= \tau_1(\mathbf{v}_{\sigma(1)}, \mathbf{v}_{\sigma(2)}, \dots, \mathbf{v}_{\sigma(k_1)}) \tau_2(\mathbf{v}_{\sigma(k_1+1)}, \mathbf{v}_{\sigma(k_1+2)}, \dots, \mathbf{v}_{\sigma(k_1+k_2)}) \\ &= \tau_2(\mathbf{v}_{\sigma(k_1+1)}, \mathbf{v}_{\sigma(k_1+2)}, \dots, \mathbf{v}_{\sigma(k_1+k_2)}) \tau_1(\mathbf{v}_{\sigma(1)}, \mathbf{v}_{\sigma(2)}, \dots, \mathbf{v}_{\sigma(k_1)}) \\ &= \tau_2(\mathbf{v}_{\tilde{\sigma}(1)}, \mathbf{v}_{\tilde{\sigma}(2)}, \dots, \mathbf{v}_{\tilde{\sigma}(k_2)}) \tau_1(\mathbf{v}_{\tilde{\sigma}(k_2+1)}, \mathbf{v}_{\tilde{\sigma}(k_2+2)}, \dots, \mathbf{v}_{\tilde{\sigma}(k_2+k_1)}) \\ &= (\tau_2 \otimes \tau_1)(\mathbf{v}_{\tilde{\sigma}(1)}, \mathbf{v}_{\tilde{\sigma}(2)}, \dots, \mathbf{v}_{\tilde{\sigma}(k_1+k_2)}) , \end{aligned}$$

where $\tilde{\sigma} \in S_{k_1+k_2}$ is defined by

$$\tilde{\sigma}(i) = \begin{cases} \sigma(k_1 + i) & \text{if } 1 \leq i \leq k_2 \\ \sigma(i - k_2) & \text{if } k_2 + 1 \leq i \leq k_1 + k_2 \end{cases}$$

Note that $\sigma \mapsto \tilde{\sigma}$ for $\sigma \in S_{k_1+k_2}$ defines a one-to-one and onto map from $S_{k_1+k_2}$ into itself.

We need to determine the relation between $\text{sgn}(\tilde{\sigma})$ and $\text{sgn}(\sigma)$. Let $\sigma_{i_1, i_2} \in S_{k_1+k_2}$ be the transposition such that $\sigma_{i_1, i_2}(i_1) = i_2$, $\sigma_{i_1, i_2}(i_2) = i_1$ and $\sigma_{i_1, i_2}(i) = i$ for $i \neq i_1, i_2$.

Let $\tilde{\sigma}_0 = \sigma$ and $\tilde{\sigma}_j = \tilde{\sigma}_{j-1} \circ \sigma_{k_1-j+k_2, k_1-j+k_2+1} \circ \dots \circ \sigma_{k_1-j+2, k_1-j+3} \circ \sigma_{k_1-j+1, k_1-j+2}$ for $1 \leq j \leq k_1$. We have

$$\begin{aligned} \text{sgn}(\tilde{\sigma}_j) &= \text{sgn}(\tilde{\sigma}_{j-1}) \underbrace{\text{sgn}(\sigma_{k_1-j+k_2, k_1-j+k_2+1}) \cdots \text{sgn}(\sigma_{k_1-j+2, k_1-j+3}) \text{sgn}(\sigma_{k_1-j+1, k_1-j+2})}_{k_2 \text{ times}} \\ &= (-1)^{k_2} \text{sgn}(\tilde{\sigma}_{j-1}) \end{aligned}$$

for $1 \leq j \leq k_1$. By induction, it is easy to verify that $\text{sgn}(\tilde{\sigma}_{k_1}) = (-1)^{k_1 k_2} \text{sgn}(\tilde{\sigma}_0) = (-1)^{k_1 k_2} \text{sgn}(\sigma)$.

We claim that $\tilde{\sigma}_{k_1} = \tilde{\sigma}$. We take a close look at $\tilde{\sigma}_j$. For $1 \leq i < k_1 - j + 1$ or $k_1 - j + k_2 + 1 < i \leq k_1 + k_2$, we have

$$\begin{aligned} & i \xrightarrow{\sigma_{k_1-j+1, k_1-j+2}} i \xrightarrow{\sigma_{k_1-j+2, k_1-j+3}} i \xrightarrow{\sigma_{k_1-j+3, k_1-j+4}} \dots \\ & \xrightarrow{\sigma_{k_1-j+k_2-1, k_1-j+k_2}} i \xrightarrow{\sigma_{k_1-j+k_2, k_1-j+k_2+1}} i \xrightarrow{\tilde{\sigma}_{j-1}} \tilde{\sigma}_{j-1}(i). \end{aligned}$$

For $k_1 - j + 1 < i \leq k_1 - j + k_2 + 1$, say $i = k_1 - j + m + 1$ with $1 \leq m \leq k_2$, we have

$$\begin{aligned} & i \xrightarrow{\sigma_{k_1-j+1, k_1-j+2}} i \xrightarrow{\sigma_{k_1-j+2, k_1-j+3}} i \xrightarrow{\sigma_{k_1-j+3, k_1-j+4}} \dots \xrightarrow{\sigma_{k_1-j+m-1, k_1-j+m}} i \xrightarrow{\sigma_{k_1-j+m, k_1-j+m+1}} \\ & i - 1 \xrightarrow{\sigma_{k_1-j+m+1, k_1-j+m+2}} i - 1 \xrightarrow{\sigma_{k_1-j+m+2, k_1-j+m+3}} \dots \xrightarrow{\sigma_{k_1-j+k_2-1, k_1-j+k_2}} \\ & i - 1 \xrightarrow{\sigma_{k_1-j+k_2, k_1-j+k_2+1}} i - 1 \xrightarrow{\tilde{\sigma}_{j-1}} \tilde{\sigma}_{j-1}(i - 1). \end{aligned}$$

Finally, for $i = k_1 - j + 1$, we have

$$\begin{aligned} & i \xrightarrow{\sigma_{k_1-j+1, k_1-j+2}} i + 1 \xrightarrow{\sigma_{k_1-j+2, k_1-j+3}} i + 2 \xrightarrow{\sigma_{k_1-j+3, k_1-j+4}} \dots \xrightarrow{\sigma_{k_1-j+k_2-1, k_1-j+k_2}} \\ & i + k_2 - 1 \xrightarrow{\sigma_{k_1-j+k_2, k_1-j+k_2+1}} i + k_2 \xrightarrow{\tilde{\sigma}_{j-1}} \tilde{\sigma}_{j-1}(i + k_2). \end{aligned}$$

To summarize, we have shown that

$$\tilde{\sigma}_j(i) = \begin{cases} \tilde{\sigma}_{j-1}(i) & \text{if } 1 \leq i \leq k_1 - j \\ \tilde{\sigma}_{j-1}(i + 1) & \text{if } k_1 - j + 1 \leq i < k_1 - j + k_2 + 1 \\ \tilde{\sigma}_{j-1}(i - k_2) & \text{if } k_1 - j + k_2 + 1 \leq i \leq k_1 + k_2 \end{cases}$$

With this information, a simple proof by induction shows that

$$\tilde{\sigma}_j(i) = \begin{cases} \sigma(i) & \text{if } 1 \leq i \leq k_1 - j \\ \sigma(i + j) & \text{if } k_1 - j + 1 \leq i \leq k_1 + k_2 - j \\ \sigma(i - k_2) & \text{if } k_1 + k_2 - j + 1 \leq i \leq k_1 + k_2 \end{cases}$$

for $1 \leq j \leq k_1$. In particular, $\tilde{\sigma}_{k_1} = \tilde{\sigma}$ as claimed.

We therefore have that $\text{sgn}(\sigma) = (-1)^{k_1 k_2} \text{sgn}(\tilde{\sigma})$. Hence

$$\begin{aligned} (\tau_1 \wedge \tau_2)(\mathbf{v}_1, \mathbf{v}, \dots, \mathbf{v}_{k_1+k_2}) &= \frac{1}{k_1! k_2!} \sum_{\sigma \in S_{k_1+k_2}} \text{sgn}(\sigma) (\tau_1 \otimes \tau_2)(\mathbf{v}_{\sigma(1)}, \mathbf{v}_{\sigma(2)}, \dots, \mathbf{v}_{\sigma(k_1+k_2)}) \\ &= \frac{1}{k_1! k_2!} \sum_{\tilde{\sigma} \in S_{k_1+k_2}} \text{sgn}(\sigma) (\tau_2 \otimes \tau_1)(\mathbf{v}_{\tilde{\sigma}(1)}, \mathbf{v}_{\tilde{\sigma}(2)}, \dots, \mathbf{v}_{\tilde{\sigma}(k_1+k_2)}) \\ &= (-1)^{k_1 k_2} \frac{1}{k_1! k_2!} \sum_{\tilde{\sigma} \in S_{k_1+k_2}} \text{sgn}(\tilde{\sigma}) (\tau_2 \otimes \tau_1)(\mathbf{v}_{\tilde{\sigma}(1)}, \mathbf{v}_{\tilde{\sigma}(2)}, \dots, \mathbf{v}_{\tilde{\sigma}(k_1+k_2)}) \\ &= (-1)^{k_1 k_2} (\tau_2 \wedge \tau_1)(\mathbf{v}_1, \mathbf{v}, \dots, \mathbf{v}_{k_1+k_2}) \end{aligned}$$

for all $\mathbf{v}_i \in H$ and $1 \leq i \leq k_1 + k_2$.

6) We have

$$\begin{aligned}
 \tau_1 \wedge (\tau_2 \wedge \tau_3) &= \frac{(k_1 + k_2 + k_3)!}{k_1! (k_2 + k_3)!} \text{Alt}(\tau_1 \otimes (\tau_2 \wedge \tau_3)) \\
 &= \frac{(k_1 + k_2 + k_3)!}{k_1! (k_2 + k_3)!} \text{Alt} \left(\tau_1 \otimes \left(\frac{(k_2 + k_3)!}{k_2! k_3!} \text{Alt}(\tau_2 \otimes \tau_3) \right) \right) \\
 &= \frac{(k_1 + k_2 + k_3)!}{k_1! (k_2 + k_3)!} \frac{(k_2 + k_3)!}{k_2! k_3!} \text{Alt}(\tau_1 \otimes (\text{Alt}(\tau_2 \otimes \tau_3))) \\
 &= \frac{(k_1 + k_2 + k_3)!}{k_1! k_2! k_3!} \text{Alt}(\tau_1 \otimes \tau_2 \otimes \tau_3),
 \end{aligned}$$

where we have used the definition of the wedge product to get the first two equality and (4) of Proposition 4.1.10 to get the last equality. The other relation in (6) is similarly proved. ■

From (6) of the previous theorem, we have associativity of the wedge product. Hence, we may define $\tau_1 \wedge \tau_2 \wedge \tau_3$ as $(\tau_1 \wedge \tau_2) \wedge \tau_3 = \tau_1 \wedge (\tau_2 \wedge \tau_3)$ for all $\tau_i \in \Omega^{k_i}(H)$ with $1 \leq i \leq 3$. Inductively, we may define $\tau_1 \wedge \tau_2 \wedge \dots \wedge \tau_j$ for all $\tau_i \in \Omega^{k_i}(H)$ with $1 \leq i \leq j$.

Corollary 4.1.13

$$\tau_1 \wedge \tau_2 \wedge \dots \wedge \tau_j = \frac{(k_1 + k_2 + \dots + k_j)!}{k_1! k_2! \dots k_j!} \text{Alt}(\tau_1 \otimes \tau_2 \otimes \dots \otimes \tau_j) \text{ for } \tau_i \in \Omega^{k_i}(H) \text{ and } 1 \leq i \leq j.$$

The proof is left to the reader.

Proposition 4.1.14

Given $\tau_i \in \Omega^1(H)$ for $1 \leq i \leq k$, we have

$$\tau_{\rho(1)} \wedge \tau_{\rho(2)} \wedge \dots \wedge \tau_{\rho(k)} = \text{sgn}(\rho) \tau_1 \wedge \tau_2 \wedge \dots \wedge \tau_k \quad (4.1.2)$$

for all $\rho \in S_k$, the group of permutations on $J_k = \{1, 2, \dots, k\}$. In particular, if $\tau_p = \tau_q$ for some $1 \leq p < q \leq k$, then $\tau_1 \wedge \tau_2 \wedge \dots \wedge \tau_k = 0$.

Proof.

We first note that the map from S_k to itself defined by $\sigma \mapsto \sigma \circ \rho^{-1}$ for $\sigma \in S_k$ is one-to-one and onto. Moreover, for all $1 \leq m \leq k$, we have that $\tau_{\rho(m)}(\mathbf{w}_{\sigma(m)}) = \tau_p(\mathbf{w}_{\sigma(\rho^{-1}(p))})$ for $p = \rho(m)$ and $\mathbf{w}_i \in H$ and $1 \leq i \leq k$. Hence, we get from Corollary 4.1.13 that

$$\begin{aligned}
 &(\tau_{\rho(1)} \wedge \tau_{\rho(2)} \wedge \dots \wedge \tau_{\rho(k)})(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k) \\
 &= k! \text{Alt}(\tau_{\rho(1)} \otimes \tau_{\rho(2)} \otimes \dots \otimes \tau_{\rho(k)})(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k) \\
 &= \sum_{\sigma \in S_k} \text{sgn}(\sigma) \tau_{\rho(1)}(\mathbf{w}_{\sigma(1)}) \tau_{\rho(2)}(\mathbf{w}_{\sigma(2)}) \dots \tau_{\rho(k)}(\mathbf{w}_{\sigma(k)}) \\
 &= \sum_{\sigma \in S_k} \text{sgn}(\sigma) \tau_1(\mathbf{w}_{(\sigma \circ \rho^{-1})(1)}) \tau_2(\mathbf{w}_{(\sigma \circ \rho^{-1})(2)}) \dots \tau_k(\mathbf{w}_{(\sigma \circ \rho^{-1})(k)})
 \end{aligned}$$

$$\begin{aligned}
&= \sum_{\tilde{\sigma} \in S_k} \operatorname{sgn}(\tilde{\sigma} \circ \rho) \tau_1(\mathbf{w}_{\tilde{\sigma}(1)}) \tau_2(\mathbf{w}_{\tilde{\sigma}(2)}) \cdots \tau_k(\mathbf{w}_{\tilde{\sigma}(k)}) \\
&= \operatorname{sgn}(\rho) \sum_{\tilde{\sigma} \in S_k} \operatorname{sgn}(\tilde{\sigma}) \tau_1(\mathbf{w}_{\tilde{\sigma}(1)}) \tau_2(\mathbf{w}_{\tilde{\sigma}(2)}) \cdots \tau_k(\mathbf{w}_{\tilde{\sigma}(k)}) \\
&= \operatorname{sgn}(\rho) k! \operatorname{Alt}(\tau_1 \otimes \tau_2 \otimes \cdots \otimes \tau_k)(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k) \\
&= \operatorname{sgn}(\rho) (\tau_1 \wedge \tau_2 \wedge \cdots \wedge \tau_k)(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k)
\end{aligned}$$

for all $\mathbf{w}_i \in H$ and $1 \leq i \leq k$.

Suppose the $\tau_p = \tau_q$ for some $1 \leq p < q \leq k$. As usual, let $\sigma_{p,q} \in S_k$ be the transposition of p and q . Then

$$\begin{aligned}
\tau_1 \wedge \tau_2 \wedge \cdots \wedge \tau_k &= \tau_1 \wedge \tau_2 \wedge \cdots \wedge \tau_q \wedge \cdots \wedge \tau_p \wedge \cdots \wedge \tau_k \\
&= \tau_{\sigma_{p,q}(1)} \wedge \tau_{\sigma_{p,q}(2)} \wedge \cdots \wedge \tau_{\sigma_{p,q}(k)} = \operatorname{sgn}(\sigma_{p,q}) \tau_1 \wedge \tau_2 \wedge \cdots \wedge \tau_j = -\tau_1 \wedge \tau_2 \wedge \cdots \wedge \tau_j,
\end{aligned}$$

where the first equality comes from $\tau_p = \tau_q$. Thus $\tau_1 \wedge \tau_2 \wedge \cdots \wedge \tau_j = 0$. ■

Theorem 4.1.15

As in Theorem 4.1.3, let $\mathcal{B}^* = \{\phi_i\}_{i=1}^n$ be the dual basis associated to a basis $\mathcal{B} = \{\mathbf{v}_i\}_{i=1}^n$ of H . The set $\mathcal{N} = \{\phi_{i_1} \wedge \phi_{i_2} \wedge \cdots \wedge \phi_{i_k} : 1 \leq i_1 < i_2 < \cdots < i_k \leq n\}$ is a basis of $\Omega^k(H)$. In particular, the dimension of $\Omega^k(H)$ is $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.

Proof.

i) We first show that any $\tau \in \Omega^k(H)$ can be written as a linear combination of elements in $\mathcal{N}$.

It follows from Theorem 4.1.3 that we may write τ as $\tau = \sum_{\substack{1 \leq i_j \leq n \\ 1 \leq j \leq k}} a_{i_1, i_2, \dots, i_k} \phi_{i_1} \otimes \phi_{i_2} \otimes \cdots \otimes$

ϕ_{i_k} for some $a_{i_1, i_2, \dots, i_k} \in \mathbb{R}$. Hence, using (1) of Proposition 4.1.10, Corollary 4.1.13 and Proposition 4.1.14 (with $\tau_j = \phi_{i_j}$ for $1 \leq j \leq k$), we get

$$\begin{aligned}
\tau &= \operatorname{Alt}(\tau) = \sum_{\substack{1 \leq i_j \leq n \\ 1 \leq j \leq k}} a_{i_1, i_2, \dots, i_k} \operatorname{Alt}(\phi_{i_1} \otimes \phi_{i_2} \otimes \cdots \otimes \phi_{i_k}) \\
&= k! \sum_{\substack{1 \leq i_j \leq n \\ 1 \leq j \leq k}} a_{i_1, i_2, \dots, i_k} (\phi_{i_1} \wedge \phi_{i_2} \wedge \cdots \wedge \phi_{i_k}) \\
&= \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} \left(k! \sum_{\sigma \in S_k} a_{\sigma(i_1), \sigma(i_2), \dots, \sigma(i_k)} \phi_{\sigma(i_1)} \wedge \phi_{\sigma(i_2)} \wedge \cdots \wedge \phi_{\sigma(i_k)} \right) \\
&= \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} \left(k! \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) a_{\sigma(i_1), \sigma(i_2), \dots, \sigma(i_k)} \right) \phi_{i_1} \wedge \phi_{i_2} \wedge \cdots \wedge \phi_{i_k}.
\end{aligned}$$

ii) The proof that the elements of $\mathcal{N}$ are linearly independent is similar to the proof that the elements of $\mathcal{M}$ are linearly independent in Theorem 4.1.3. We first note that

$$\begin{aligned} & (\phi_{i_1} \wedge \phi_{i_2} \wedge \dots \wedge \phi_{i_k})(\mathbf{v}_{j_1}, \mathbf{v}_{j_2}, \dots, \mathbf{v}_{j_k}) \\ &= \sum_{\sigma \in S_k} \text{sgn}(\sigma) \phi_{i_1}(\mathbf{v}_{\sigma(j_1)}) \otimes \phi_{i_2}(\mathbf{v}_{\sigma(j_2)}) \otimes \dots \otimes \phi_{i_k}(\mathbf{v}_{\sigma(j_k)}) \\ &= \begin{cases} \text{sgn}(\sigma) & \text{if } \sigma(j_m) = i_m \text{ for } 1 \leq m \leq k \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

for $1 \leq i_m, j_m \leq n$ and $1 \leq m \leq k$. In particular, for $1 \leq i_1 < i_2 < \dots < i_k \leq n$ and $1 \leq j_1 < j_2 < \dots < j_k \leq n$, we get

$$(\phi_{i_1} \wedge \phi_{i_2} \wedge \dots \wedge \phi_{i_k})(\mathbf{v}_{j_1}, \mathbf{v}_{j_2}, \dots, \mathbf{v}_{j_k}) = \begin{cases} 1 & \text{if } j_m = i_m \text{ for } 1 \leq m \leq k \\ 0 & \text{otherwise} \end{cases} \quad (4.1.3)$$

Suppose that

$$\sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} a_{i_1, i_2, \dots, i_k} \phi_{i_1} \wedge \phi_{i_2} \wedge \dots \wedge \phi_{i_k} = 0 .$$

If we evaluate this expression at $(\mathbf{v}_{m_1}, \mathbf{v}_{m_2}, \dots, \mathbf{v}_{m_k})$ with $1 \leq m_1 < m_2 < \dots < m_k \leq n$ and use (4.1.3), then we get $a_{m_1, m_2, \dots, m_k} = 0$.

iii) To determine the number of elements in the basis $\mathcal{N}$, we first note that there are $n(n-1)(n-2)\dots(n-k+1) = n!/(n-k)!$ ways to select k distinct ϕ_i from $\mathcal{B}^*$. However, for each set of indices $1 \leq i_1 < i_2 < \dots < i_k \leq n$, all $k!$ combinations $\{\phi_{i_{\sigma(j)}}\}_{j=1}^k$ for $\sigma \in S_k$ are included in the previous count while only the ordering $i_1 < i_2 < \dots < i_k$ matters for the basis $\mathcal{N}$. Therefore, the number of elements of $\mathcal{N}$ is $n!/((n-k)!k!)$. ■

Theorem 4.1.16

Let $\mathcal{B} = \{\mathbf{v}_i\}_{i=1}^n$ be a basis of the vector space H over $\mathbb{R}$. If $\tau \in \Omega^n(H)$ and $\mathbf{w}_i = \sum_{j=1}^n a_{j,i} \mathbf{v}_j$ for $1 \leq i \leq n$ are n vectors in H , then

$$\tau(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n) = \det(A) \tau(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) , \quad (4.1.4)$$

where the matrix A has entries $a_{j,i}$ for $1 \leq i, j \leq n$.

Proof.

The crux of this proof is the basic result from linear algebra that all alternating multilinear maps

$$\rho : \underbrace{\mathbb{R}^n \times \mathbb{R}^n \times \dots \times \mathbb{R}^n}_{n \text{ times}} \rightarrow \mathbb{R}^n ,$$

namely $\rho \in \Omega^n(\mathbb{R}^n)$, are of the form $\rho(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) = b \det(\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_n)$ for some $b \in \mathbb{R}$.

Consider the alternating multilinear map $\rho : \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}$ defined by

$$\rho \left(\begin{pmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{n,1} \end{pmatrix}, \begin{pmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{n,2} \end{pmatrix}, \dots, \begin{pmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{n,n} \end{pmatrix} \right) = \tau \left(\sum_{i=1}^n a_{i,1} \mathbf{v}_i, \sum_{i=1}^n a_{i,2} \mathbf{v}_i, \dots, \sum_{i=1}^n a_{i,n} \mathbf{v}_i \right).$$

Then, from our initial remark,

$$\tau \left(\sum_{i=1}^n a_{i,1} \mathbf{v}_i, \sum_{i=1}^n a_{i,2} \mathbf{v}_i, \dots, \sum_{i=1}^n a_{i,n} \mathbf{v}_i \right) = b \det \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{pmatrix}$$

for some $b \in \mathbb{R}$. With $a_{j,i} = \delta_{j,i}$, we get $\tau(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) = b \det \text{Id}_n = b$. ■

Corollary 4.1.17

Suppose that $\tau \in \Omega^k(H)$ and $\mathbf{w}_i \in H$ for $1 \leq i \leq n$. Then

1. $\tau(\mathbf{w}_{\sigma(1)}, \mathbf{w}_{\sigma(2)}, \dots, \mathbf{w}_{\sigma(n)}) = \text{sgn}(\sigma) \tau(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n)$ for all $\sigma \in S_n$, and
2. $\tau(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n) = 0$ if $\mathbf{w}_p = \mathbf{w}_q$ for some $1 \leq p < q \leq n$.

Proof.

1) A permutation on $J_n = \{1, 2, \dots, n\}$ may be represented by a matrix with only one 1 per row and column, and all the other entries being null. For instance, the permutation $\sigma \in S_4$ defined by $\sigma(1) = 2$, $\sigma(2) = 4$, $\sigma(3) = 3$ and $\sigma(4) = 1$ is represented by the

matrix $M_\sigma = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$ ³. The transposition $\sigma \in S_4$ defined by $\sigma(1) = 3$, $\sigma(2) = 2$,

$\sigma(3) = 1$ and $\sigma(4) = 4$ is represented by the matrix $M_\sigma = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$. We have that

$\text{sgn}(\sigma) = \det(M_\sigma)$ ⁴. With this information, the proof of (1) is now a simple consequence of the previous theorem.

³The matrix representation of σ may also be represented by the transpose of M_σ . It all depends if the multiplication with a vector is from the left or the right. We have chosen from the right.

⁴The reader may wonder why we did not define the sign of a permutation using the determinant of its matrix representation. But, by doing so, we will have created a vicious circle because the sign of permutations is used to define the determinant.

Let $\mathcal{B} = \{\mathbf{v}_i\}_{i=1}^n$ be a basis of the vector space H over $\mathbb{R}$. We may write $\mathbf{w}_i = \sum_{j=1}^n a_{j,i} \mathbf{v}_j$ for $1 \leq i \leq n$ and some $a_{j,i} \in \mathbb{R}$. Let A be the matrix with entry in position (j, i) given by $a_{j,i}$ for $1 \leq i, j \leq n$. Since $\mathbf{w}_{\sigma(i)} = \sum_{j=1}^n a_{j,\sigma(i)} \mathbf{v}_j$ for $1 \leq i \leq n$, we have that the matrix $\tilde{A}$ with entry in position (j, i) given by $a_{j,\sigma(i)}$ is $\tilde{A} = AM_\sigma$.

Hence, we get from Theorem 4.1.16 that

$$\begin{aligned} \tau(\mathbf{w}_{\sigma(1)}, \mathbf{w}_{\sigma(2)}, \dots, \mathbf{w}_{\sigma(n)}) &= \det(\tilde{A}) \tau(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) = \det(A) \det(M_\sigma) \tau(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) \\ &= \operatorname{sgn} \sigma \det(A) \tau(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) = \operatorname{sgn}(\sigma) \tau(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n). \end{aligned}$$

2) If $\sigma = \sigma_{p,q}$ is the transposition of p and q , then we get from (1) that

$$\begin{aligned} \tau(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n) &= \tau(\mathbf{w}_{\sigma(1)}, \mathbf{w}_{\sigma(2)}, \dots, \mathbf{w}_{\sigma(n)}) = \operatorname{sgn}(\sigma) \tau(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n) \\ &= -\tau(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n), \end{aligned}$$

where the first equality comes from $\mathbf{w}_p = \mathbf{w}_q$ and $\sigma(i) = i$ for $i \neq p, q$. ■

4.2 Orientation

If $\mathcal{C} = \{\mathbf{w}_i\}_{i=1}^n$ in Theorem 4.1.16 is a basis of H , then A is the matrix of change of bases from $\mathcal{C}$ to $\mathcal{B}$. In particular, $\det(A) \neq 0$. Hence (4.1.4) divides the bases of H into two groups. Two bases like $\mathcal{B}$ and $\mathcal{C}$ are in the same group if (4.1.4) is satisfied with $\det(A) > 0$.

Definition 4.2.1

We define an **orientation** on a vector space H by selecting one of the two groups of bases described above.

If $\mathcal{B} = \{\mathbf{v}_i\}_{i=1}^n$ is a basis belonging to the chosen group for the orientation on H , then we denote this orientation by $[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n]$. The order of the vector is fundamental. The group of bases not belonging to the orientation $[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n]$ on H is the opposite orientation and is denoted $-[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n]$.

Definition 4.2.2

The **standard orientation** on $\mathbb{R}^n$ is given by $[\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n]$ where $\mathcal{E} = \{\mathbf{e}_i\}_{i=1}^n$ is the canonical basis of $\mathbb{R}^n$.

If $\mathcal{B} = \{\mathbf{v}_i\}_{i=1}^n$ and $\mathcal{C} = \{\mathbf{w}_i\}_{i=1}^n$ are two orthonormal bases with respect to an inner product $\tau \in \mathcal{T}^2(H)$, then A in Theorem 4.1.16 is an orthogonal matrix, namely $A^\top A = \operatorname{Id}_n$. This follows from

$$\delta_{i,j} = \tau(\mathbf{w}_i, \mathbf{w}_j) = \tau\left(\sum_{k_1=1}^n a_{k_1,i} \mathbf{v}_{k_1}, \sum_{k_2=1}^n a_{k_2,j} \mathbf{v}_{k_2}\right) = \sum_{k_1, k_2=1}^n a_{k_1,i} a_{k_2,j} \tau(\mathbf{v}_{k_1}, \mathbf{v}_{k_2})$$

$$= \sum_{k_1, k_2=1}^n a_{k_1, i} a_{k_2, j} \delta_{k_1, k_2} = \sum_{k=1}^n a_{k, i} a_{k, j}.$$

Hence, if $\mathcal{B} = \{\mathbf{v}_i\}_{i=1}^n$ and $\mathcal{C} = \{\mathbf{w}_i\}_{i=1}^n$ are two orthonormal bases with respect to the inner product $\tau \in \mathcal{T}^2(H)$, then $\det A = \pm 1$ in Theorem 4.1.16.

Consider the orientation $[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n]$ associated to a given orthonormal basis $\mathcal{B} = \{\mathbf{v}_i\}_{i=1}^n$ on H with respect to the inner product $\tau \in \mathcal{T}^2(H)$. Since $\Omega^n(H)$ is of dimension one, there exists a unique $\nu \in \Omega^n(H)$ such that $\nu(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) = 1$. It follows from the previous paragraph that $\nu(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n) = 1$ for all orthonormal basis $\mathcal{C} = \{\mathbf{w}_i\}_{i=1}^n$ belonging to the orientation $[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n]$.

Definition 4.2.3

Let $\mathcal{B} = \{\mathbf{v}_i\}_{i=1}^n$ be an orthonormal basis on S with respect to the inner product $\tau \in \mathcal{T}^2(H)$. The unique $\nu \in \Omega^n(H)$ such that $\nu(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) = 1$ is called the **volume element** on H determined by the inner product $\tau \in \mathcal{T}^2(H)$ and the orientation $[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n]$.

A particular case of volume element is given by $\nu = \det \in \Omega^n(\mathbb{R}^n)$ seen in the proof of Theorem 4.1.16. We have that $\det$ is the volume element on $\mathbb{R}^n$ determined by the standard inner product $\tau \equiv \langle, \rangle$ on $\mathbb{R}^n$ and the orientation $[\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n]$, where $\mathcal{E} = \{\mathbf{e}_i\}_{i=1}^n$ is the canonical basis of $\mathbb{R}^n$.

To conclude this section, we extend the classical cross product in $\mathbb{R}^3$ to $\mathbb{R}^n$ for $n > 3$. Given $\mathbf{w}_i \in \mathbb{R}^n$ for $1 \leq i \leq n-1$, let $g \in \Omega^1(\mathbb{R}^n)$ be the linear functional defined by

$$g(\mathbf{v}) = \det \begin{pmatrix} \mathbf{w}_1 & \mathbf{w}_2 & \dots & \mathbf{w}_{n-1} & \mathbf{v} \end{pmatrix}$$

for $\mathbf{v} \in \mathbb{R}^n$. From the theory of representation of linear functionals on $\mathbb{R}^n$, there exists a unique $\mathbf{z} \in \mathbb{R}^n$ such that $g(\mathbf{v}) = \langle \mathbf{v}, \mathbf{z} \rangle$ for all $\mathbf{v} \in \mathbb{R}^n$.

Definition 4.2.4

The **cross product** $\mathbf{w}_1 \times \mathbf{w}_2 \times \dots \times \mathbf{w}_{n-1}$ is the unique vector $\mathbf{z}$ such that $\langle \mathbf{v}, \mathbf{z} \rangle = \det \begin{pmatrix} \mathbf{w}_1 & \mathbf{w}_2 & \dots & \mathbf{w}_{n-1} & \mathbf{v} \end{pmatrix}$ for all $\mathbf{v} \in \mathbb{R}^n$.

We have by construction that $\mathbf{z}$ is orthogonal to $\mathbf{w}_i$ for $1 \leq i \leq n-1$. The cross product is a multilinear map. The reader may be interested to prove that $\mathbf{w}_{\sigma(1)} \times \mathbf{w}_{\sigma(2)} \times \dots \times \mathbf{w}_{\sigma(n-1)} = \text{sgn}(\sigma) \mathbf{w}_1 \times \mathbf{w}_2 \times \dots \times \mathbf{w}_{n-1}$ for $\sigma \in S_{n-1}$.

4.3 Tangent Spaces

Definition 4.3.1

Let $V \subset \mathbb{R}^n$ be an open set. Given $\mathbf{v} \in V$, the **tangent space** of V at $\mathbf{v}$ is the set

$$T_{\mathbf{v}}V = \{(\mathbf{v}, \mathbf{x}) : \mathbf{x} \in \mathbb{R}^n\}.$$

The tangent space $T_{\mathbf{v}}V$ is given the structure of a vector space by setting

$$a_1(\mathbf{v}, \mathbf{x}_1) + a_2(\mathbf{v}, \mathbf{x}_2) = (\mathbf{v}, a_1\mathbf{x}_1 + a_2\mathbf{x}_2)$$

for all $a_i \in \mathbb{R}$ and $\mathbf{x}_i \in \mathbb{R}^n$. For each $\mathbf{v} \in V$, we have that $T_{\mathbf{v}}V$ is linearly isomorphic to $\mathbb{R}^n$. The canonical basis $\{\mathbf{e}_i\}_{i=1}^n$ of $\mathbb{R}^n$ yields a **canonical basis** $\mathcal{E}$ on $T_{\mathbf{v}}V$ defined by $\mathcal{E} = \{(\mathbf{v}, \mathbf{e}_i)\}_{i=1}^n$.

This definition of tangent space for an open subset of $\mathbb{R}^n$ may seem to be unnecessary and cumbersome but we will see later, when working on manifolds, that it is very useful.

Definition 4.3.2

Let $V \subset \mathbb{R}^n$ be an open set and $\mathbf{v} \in V$. We define an inner product $\langle \cdot, \cdot \rangle_{\mathbf{v}}$ on $T_{\mathbf{v}}V$ by

$$\langle (\mathbf{v}, \mathbf{x}_1), (\mathbf{v}, \mathbf{x}_2) \rangle_{\mathbf{v}} = \tau(\mathbf{x}_1, \mathbf{x}_2)$$

for all $\mathbf{x}_i \in \mathbb{R}^n$, where $\tau \in \mathcal{T}^2(\mathbb{R}^n)$ is an inner product.

The most frequent situation in this chapter is when τ is the standard inner product on $\mathbb{R}^n$.

The concept of tensors can be naturally extended to $T_{\mathbf{v}}V$.

Definition 4.3.3

Let $V \subset \mathbb{R}^n$ be an open set and $\mathbf{v} \in V$. A k -tensors $\tau_{\mathbf{v}} \in \mathcal{T}^k(T_{\mathbf{v}}V)$ is defined by

$$\tau_{\mathbf{v}}((\mathbf{v}, \mathbf{x}_1), (\mathbf{v}, \mathbf{x}_2), \dots, (\mathbf{v}, \mathbf{x}_k)) = \tau(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k)$$

for all $\mathbf{x}_i \in \mathbb{R}^n$, where $\tau \in \mathcal{T}^k(\mathbb{R}^n)$. We say that $\tau_{\mathbf{v}}$ is generated by τ .

The space $\mathcal{T}^k(T_{\mathbf{v}}V)$ is a vector space over $\mathbb{R}$. If $\tau_{\mathbf{v}}^{[1]}, \tau_{\mathbf{v}}^{[2]} \in \mathcal{T}^k(T_{\mathbf{v}}V)$ are generated by $\tau^{[1]}, \tau^{[2]} \in \mathcal{T}^k(\mathbb{R}^n)$ respectively, and $a, b \in \mathbb{R}$, then $a\tau_{\mathbf{v}}^{[1]} + b\tau_{\mathbf{v}}^{[2]} \in \mathcal{T}(T_{\mathbf{v}}V)$ is generated by $a\tau^{[1]} + b\tau^{[2]} \in \mathcal{T}(\mathbb{R}^n)$. This follows from

$$\begin{aligned} & (a\tau_{\mathbf{v}}^{[1]} + b\tau_{\mathbf{v}}^{[2]})((\mathbf{v}, \mathbf{x}_1), (\mathbf{v}, \mathbf{x}_2), \dots, (\mathbf{v}, \mathbf{x}_k)) \\ &= a\tau_{\mathbf{v}}^{[1]}((\mathbf{v}, \mathbf{x}_1), (\mathbf{v}, \mathbf{x}_2), \dots, (\mathbf{v}, \mathbf{x}_k)) + b\tau_{\mathbf{v}}^{[2]}((\mathbf{v}, \mathbf{x}_1), (\mathbf{v}, \mathbf{x}_2), \dots, (\mathbf{v}, \mathbf{x}_k)) \\ &= a\tau^{[1]}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k) + b\tau^{[2]}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k) = (a\tau^{[1]} + b\tau^{[2]})(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k) \end{aligned}$$

for all $\mathbf{x}_i \in \mathbb{R}^n$ and $1 \leq i \leq k$.

We can also extend the product of tensors to $T_v V$ as it follows.

Definition 4.3.4

Let $V \subset \mathbb{R}^n$ be an open set and $\mathbf{v} \in V$. The **product of two tensors** $\tau_v^{[1]} \in \mathcal{T}^{k_1}(T_v V)$ and $\tau_v^{[2]} \in \mathcal{T}^{k_2}(T_v V)$ is the tensor $\tau_v^{[1]} \otimes \tau_v^{[2]} \in \mathcal{T}^{k_1+k_2}(T_v V)$ defined by

$$\begin{aligned} & (\tau_v^{[1]} \otimes \tau_v^{[2]})((\mathbf{v}, \mathbf{x}_1), (\mathbf{v}, \mathbf{x}_2), \dots, (\mathbf{v}, \mathbf{x}_{k_1+k_2})) \\ &= \tau_v^{[1]}((\mathbf{v}, \mathbf{x}_1), (\mathbf{v}, \mathbf{x}_2), \dots, (\mathbf{v}, \mathbf{x}_{k_1})) \tau_v^{[2]}((\mathbf{v}, \mathbf{x}_{k_1+1}), (\mathbf{v}, \mathbf{x}_{k_1+2}), \dots, (\mathbf{v}, \mathbf{x}_{k_1+k_2})) \end{aligned}$$

for all $\mathbf{x}_i \in \mathbb{R}^n$.

If $\tau_v^{[i]} \in \mathcal{T}^{k_i}(T_v V)$ is generated by $\tau^{[i]} \in \mathcal{T}^{k_i}(\mathbb{R}^n)$ for $1 \leq i \leq 2$, then $\tau_v^{[1]} \otimes \tau_v^{[2]} \in \mathcal{T}^{k_1+k_2}(T_v V)$ is generated by $\tau^{[1]} \otimes \tau^{[2]} \in \mathcal{T}^{k_1+k_2}(\mathbb{R}^n)$. This follows from

$$\begin{aligned} & (\tau_v^{[1]} \otimes \tau_v^{[2]})((\mathbf{v}, \mathbf{x}_1), (\mathbf{v}, \mathbf{x}_2), \dots, (\mathbf{v}, \mathbf{x}_{k_1+k_2})) \\ &= \tau_v^{[1]}((\mathbf{v}, \mathbf{x}_1), (\mathbf{v}, \mathbf{x}_2), \dots, (\mathbf{v}, \mathbf{x}_{k_1})) \tau_v^{[2]}((\mathbf{v}, \mathbf{x}_{k_1+1}), (\mathbf{v}, \mathbf{x}_{k_1+2}), \dots, (\mathbf{v}, \mathbf{x}_{k_1+k_2})) \\ &= \tau^{[1]}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{k_1}) \tau^{[2]}(\mathbf{x}_{k_1+1}, \mathbf{x}_{k_1+2}, \dots, \mathbf{x}_{k_1+k_2}) = (\tau^{[1]} \otimes \tau^{[2]})(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{k_1+k_2}) \end{aligned}$$

for all $\mathbf{x}_i \in \mathbb{R}^n$.

With our definition of tensors on $T_v V$, we may extend the wedge product of alternating tensors on $T_v V$.

Definition 4.3.5

Let $V \subset \mathbb{R}^n$ be an open set and $\mathbf{v} \in V$. The **wedge product of two alternating tensors** $\tau_v^{[1]} \in \Omega^{k_1}(T_v V)$ and $\tau_v^{[2]} \in \Omega^{k_2}(T_v V)$ is the alternating tensor $\tau_v^{[1]} \wedge \tau_v^{[2]} \in \Omega^{k_1+k_2}(T_v V)$ defined by

$$\begin{aligned} & (\tau_v^{[1]} \wedge \tau_v^{[2]})((\mathbf{v}, \mathbf{x}_1), (\mathbf{v}, \mathbf{x}_2), \dots, (\mathbf{v}, \mathbf{x}_{k_1+k_2})) \\ &= \frac{(k_1 + k_2)!}{k_1! k_2!} \text{Alt}(\tau_v^{[1]} \otimes \tau_v^{[2]})((\mathbf{v}, \mathbf{x}_1), (\mathbf{v}, \mathbf{x}_2), \dots, (\mathbf{v}, \mathbf{x}_{k_1+k_2})) \end{aligned}$$

for all $\mathbf{x}_i \in \mathbb{R}^n$.

Suppose that $\tau_v^{[i]} \in \Omega^{k_i}(T_v V)$ is generated by $\tau^{[i]} \in \Omega^{k_i}(\mathbb{R}^n)$ for $1 \leq i \leq 2$. Then $\tau_v^{[1]} \wedge \tau_v^{[2]} \in \Omega^{k_1+k_2}(T_v V)$ is generated by $\tau^{[1]} \wedge \tau^{[2]} \in \Omega^{k_1+k_2}(\mathbb{R}^n)$. This follows from

$$\begin{aligned} & (\tau_v^{[1]} \wedge \tau_v^{[2]})((\mathbf{v}, \mathbf{x}_1), (\mathbf{v}, \mathbf{x}_2), \dots, (\mathbf{v}, \mathbf{x}_{k_1+k_2})) \\ &= \frac{(k_1 + k_2)!}{k_1! k_2!} \text{Alt}(\tau_v^{[1]} \otimes \tau_v^{[2]})((\mathbf{v}, \mathbf{x}_1), (\mathbf{v}, \mathbf{x}_2), \dots, (\mathbf{v}, \mathbf{x}_{k_1+k_2})) \\ &= \frac{1}{k_1! k_2!} \sum_{\sigma \in S_{k_1+k_2}} (\tau_v^{[1]} \otimes \tau_v^{[2]})((\mathbf{v}, \mathbf{x}_{\sigma(1)}), (\mathbf{v}, \mathbf{x}_{\sigma(2)}), \dots, (\mathbf{v}, \mathbf{x}_{\sigma(k_1+k_2)})) \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{k_1! k_2!} \sum_{\sigma \in S_{k_1+k_2}} \tau_{\mathbf{v}}^{[1]}((\mathbf{v}, \mathbf{x}_{\sigma(1)}), (\mathbf{v}, \mathbf{x}_{\sigma(2)}), \dots, (\mathbf{v}, \mathbf{x}_{\sigma(k_1)})) \\
&\quad \tau_{\mathbf{v}}^{[2]}((\mathbf{v}, \mathbf{x}_{\sigma(k_1+1)}), (\mathbf{v}, \mathbf{x}_{\sigma(k_1+2)}), \dots, (\mathbf{v}, \mathbf{x}_{\sigma(k_1+k_2)})) \\
&= \frac{1}{k_1! k_2!} \sum_{\sigma \in S_{k_1+k_2}} \tau^{[1]}(\mathbf{x}_{\sigma(1)}, \mathbf{x}_{\sigma(2)}, \dots, \mathbf{x}_{\sigma(k_1)}) \tau^{[2]}(\mathbf{x}_{\sigma(k_1+1)}, \mathbf{x}_{\sigma(k_1+2)}, \dots, \mathbf{x}_{\sigma(k_1+k_2)}) \\
&= \frac{(k_1+k_2)!}{k_1! k_2!} \text{Alt}(\tau^{[1]} \otimes \tau^{[2]})(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{k_1+k_2}) = (\tau^{[1]} \wedge \tau^{[2]})(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{k_1+k_2})
\end{aligned}$$

for all $\mathbf{x}_i \in \mathbb{R}^n$.

Suppose that $V \subset \mathbb{R}^n$ and $W \subset \mathbb{R}^m$ are two open sets. If $f : V \rightarrow W$ is a function of class C^1 and $\mathbf{v} \in V$, then $Df(\mathbf{v})$ can be used to define the following linear map.

$$\begin{aligned}
f_* : T_{\mathbf{v}}V &\rightarrow T_{f(\mathbf{v})}W \\
(\mathbf{v}, \mathbf{x}) &\mapsto (f(\mathbf{v}), Df(\mathbf{v})\mathbf{x})
\end{aligned}$$

for each $\mathbf{v} \in V$. This linear map induces a map $f^* : \mathcal{T}^k(T_{f(\mathbf{v})}W) \rightarrow \mathcal{T}^k(T_{\mathbf{v}}V)$ defined by

$$f^*(\tau_{f(\mathbf{v})})((\mathbf{v}, \mathbf{x}_1), (\mathbf{v}, \mathbf{x}_2), \dots, (\mathbf{v}, \mathbf{x}_k)) = \tau_{f(\mathbf{v})}(f_*(\mathbf{v}, \mathbf{x}_1), f_*(\mathbf{v}, \mathbf{x}_2), \dots, f_*(\mathbf{v}, \mathbf{x}_k)).$$

for all $\mathbf{x}_i \in \mathbb{R}^n$. We say that $f^*(\tau_{f(\mathbf{v})})$ is the **pull-back** of $\tau_{f(\mathbf{v})}$ by f .

Proposition 4.3.6

Suppose that $V \subset \mathbb{R}^n$ and $W \subset \mathbb{R}^m$ are two open sets. If $f : V \rightarrow W$ is a function of class C^1 and $\mathbf{v} \in V$, then

1. $f^*(\tau_{f(\mathbf{v})}^{[1]} + \tau_{f(\mathbf{v})}^{[2]}) = f^*(\tau_{f(\mathbf{v})}^{[1]}) + f^*(\tau_{f(\mathbf{v})}^{[2]})$ for all $\tau_{f(\mathbf{v})}^{[i]} \in \mathcal{T}^k(T_{f(\mathbf{v})}W)$.
2. $f^*(a \tau_{f(\mathbf{v})}) = a f^*(\tau_{f(\mathbf{v})})$ for all $\tau_{f(\mathbf{v})} \in \mathcal{T}^k(T_{f(\mathbf{v})}W)$ and $a \in \mathbb{R}$.
3. $f^*(\tau_{f(\mathbf{v})}^{[1]} \otimes \tau_{f(\mathbf{v})}^{[2]}) = f^*(\tau_{f(\mathbf{v})}^{[1]}) \otimes f^*(\tau_{f(\mathbf{v})}^{[2]})$ for all $\tau_{f(\mathbf{v})}^{[i]} \in \mathcal{T}^{k_i}(T_{f(\mathbf{v})}W)$.
4. $f^*(\tau_{f(\mathbf{v})}^{[1]} \wedge \tau_{f(\mathbf{v})}^{[2]}) = f^*(\tau_{f(\mathbf{v})}^{[1]}) \wedge f^*(\tau_{f(\mathbf{v})}^{[2]})$ for all $\tau_{f(\mathbf{v})}^{[i]} \in \Omega^{k_i}(T_{f(\mathbf{v})}W)$.

Proof.

We prove (2) and (4) and leave the proof of (1) and (3) to the reader.

2) For $\mathbf{v} \in V$ fixed, we have

$$\begin{aligned}
f^*(a \tau_{f(\mathbf{v})})((\mathbf{v}, \mathbf{x}_1), (\mathbf{v}, \mathbf{x}_2), \dots, (\mathbf{v}, \mathbf{x}_k)) &= (a \tau_{f(\mathbf{v})})(f_*(\mathbf{v}, \mathbf{x}_1), f_*(\mathbf{v}, \mathbf{x}_2), \dots, f_*(\mathbf{v}, \mathbf{x}_k)) \\
&= a (\tau_{f(\mathbf{v})}(f_*(\mathbf{v}, \mathbf{x}_1), f_*(\mathbf{v}, \mathbf{x}_2), \dots, f_*(\mathbf{v}, \mathbf{x}_k))) \\
&= a (f^*(\tau_{f(\mathbf{v})})((\mathbf{v}, \mathbf{x}_1), (\mathbf{v}, \mathbf{x}_2), \dots, (\mathbf{v}, \mathbf{x}_k)))
\end{aligned}$$

for all $\mathbf{x}_i \in \mathbb{R}^n$.

4) For $\mathbf{v} \in V$ fixed, we have

$$\begin{aligned}
 & f^*(\tau_{f(\mathbf{v})}^{[1]} \wedge \tau_{f(\mathbf{v})}^{[2]})((\mathbf{v}, \mathbf{x}_1), (\mathbf{v}, \mathbf{x}_2), \dots, (\mathbf{v}, \mathbf{x}_{k_1+k_2})) \\
 &= (\tau_{f(\mathbf{v})}^{[1]} \wedge \tau_{f(\mathbf{v})}^{[2]})(f_*(\mathbf{v}, \mathbf{x}_1), f_*(\mathbf{v}, \mathbf{x}_2), \dots, f_*(\mathbf{v}, \mathbf{x}_{k_1+k_2})) \\
 &= \frac{1}{k_1! k_2!} \sum_{\sigma \in S_{k_1+k_2}} (\tau_{f(\mathbf{v})}^{[1]} \otimes \tau_{f(\mathbf{v})}^{[2]})(f_*(\mathbf{v}, \mathbf{x}_{\sigma(1)}), f_*(\mathbf{v}, \mathbf{x}_{\sigma(2)}), \dots, f_*(\mathbf{v}, \mathbf{x}_{\sigma(k_1+k_2)})) \\
 &= \frac{1}{k_1! k_2!} \sum_{\sigma \in S_{k_1+k_2}} \tau_{f(\mathbf{v})}^{[1]}(f_*(\mathbf{v}, \mathbf{x}_{\sigma(1)}), f_*(\mathbf{v}, \mathbf{x}_{\sigma(2)}), \dots, f_*(\mathbf{v}, \mathbf{x}_{\sigma(k_1)})) \\
 &\quad \tau_{f(\mathbf{v})}^{[2]}(f_*(\mathbf{v}, \mathbf{x}_{\sigma(k_1+1)}), f_*(\mathbf{v}, \mathbf{x}_{\sigma(k_1+2)}), \dots, f_*(\mathbf{v}, \mathbf{x}_{\sigma(k_1+k_2)})) \\
 &= \frac{1}{k_1! k_2!} \sum_{\sigma \in S_{k_1+k_2}} f^*(\tau_{f(\mathbf{v})}^{[1]})((\mathbf{v}, \mathbf{x}_{\sigma(1)}), (\mathbf{v}, \mathbf{x}_{\sigma(2)}), \dots, (\mathbf{v}, \mathbf{x}_{\sigma(k_1)})) \\
 &\quad f^*(\tau_{f(\mathbf{v})}^{[2]})((\mathbf{v}, \mathbf{x}_{\sigma(k_1+1)}), (\mathbf{v}, \mathbf{x}_{\sigma(k_1+2)}), \dots, (\mathbf{v}, \mathbf{x}_{\sigma(k_1+k_2)})) \\
 &= (f^*(\tau_{f(\mathbf{v})}^{[1]}) \wedge f^*(\tau_{f(\mathbf{v})}^{[2]}))((\mathbf{v}, \mathbf{x}_1), (\mathbf{v}, \mathbf{x}_2), \dots, (\mathbf{v}, \mathbf{x}_{\sigma(k_1+k_2)}))
 \end{aligned}$$

for all $\mathbf{x}_i \in \mathbb{R}^n$. ■

4.4 Vector Fields

Definition 4.4.1

A **vector field** on an open set $V \subset \mathbb{R}^n$ is a map

$$\begin{aligned}
 F : V &\rightarrow \bigcup_{\mathbf{v} \in V} T_{\mathbf{v}}V \\
 \mathbf{v} &\mapsto (\mathbf{v}, f(\mathbf{v})) \in T_{\mathbf{v}}V
 \end{aligned}$$

where $f : V \rightarrow \mathbb{R}^n$.

It is important to note that $F(\mathbf{v}) \in T_{\mathbf{v}}V$ for each $\mathbf{v} \in V$. We have

$$F(\mathbf{v}) = \left(\mathbf{v}, \sum_{i=1}^n f_i(\mathbf{v}) \mathbf{e}_i \right) = \sum_{i=1}^n f_i(\mathbf{v}) (\mathbf{v}, \mathbf{e}_i)$$

for all $\mathbf{v} \in V$, where $f_i : V \rightarrow \mathbb{R}$ for $1 \leq i \leq n$. We say that the vector field F is of class C^k if each function f_i is of class C^k .

The sum of vector fields, the product of a vector field by a real valued function on V , the inner product of vector fields and the cross product of vector fields are defined by the respective operation in $T_{\mathbf{v}}V$ for every $\mathbf{v} \in V$. For instance, if $F, G : V \rightarrow \bigcup_{\mathbf{v} \in V} T_{\mathbf{v}}V$ are two

vector fields, then $\langle F, G \rangle$ is defined by $\langle F, G \rangle(\mathbf{v}) = \langle F(\mathbf{v}), G(\mathbf{v}) \rangle_{\mathbf{v}} = \tau(f(\mathbf{v}), g(\mathbf{v}))$ for all $\mathbf{v} \in V$, where $\tau \in \mathcal{T}^2(\mathbb{R}^n)$ is an inner product.

We have introduced this notation for vector fields because it will extend naturally to vector fields on manifolds as we will see later.

4.5 Differential Forms

Definition 4.5.1

Let V be an open subset of $\mathbb{R}^n$. A **differential k -form** on V is a map

$$\begin{aligned} \omega : V &\rightarrow \bigcup_{\mathbf{v} \in V} \Omega^k(\mathbf{T}_{\mathbf{v}}V) \\ \mathbf{v} &\mapsto \omega(\mathbf{v}) \in \Omega^k(\mathbf{T}_{\mathbf{v}}V) \end{aligned}$$

It is important to note that $\omega(\mathbf{v}) \in \Omega^k(\mathbf{T}_{\mathbf{v}}V)$ for each $\mathbf{v} \in V$.

Let $\{\phi_i\}_{i=1}^n \subset (\mathbb{R}^n)^*$ be the dual basis associated to a basis $\{\mathbf{u}_j\}_{j=1}^n$ of $\mathbb{R}^n$. For each i , we define a differential 1-form $\tilde{\phi}_i$ on V by $\tilde{\phi}_i(\mathbf{v})((\mathbf{v}, \mathbf{x})) = \phi_i(\mathbf{x})$ for each $\mathbf{v} \in V$ and all $\mathbf{x} \in \mathbb{R}^n$. For each $\mathbf{v} \in V$, the set $\{\tilde{\phi}_i(\mathbf{v})\}_{i=1}^n \subset (\mathbf{T}_{\mathbf{v}}V)^* = \Omega^1(\mathbf{T}_{\mathbf{v}}V)$ is the dual basis associated to the basis $\{(\mathbf{v}, \mathbf{u}_j)\}_{j=1}^n$ of $\mathbf{T}_{\mathbf{v}}V$ because $\tilde{\phi}_i(\mathbf{v})((\mathbf{v}, \mathbf{u}_j)) = \phi_i(\mathbf{u}_j) = \delta_{i,j}$ for $1 \leq i, j \leq n$. It follows from Theorem 4.1.15 that every differential k -form ω on V can be expressed as

$$\omega = \sum_{i_1 < i_2 < \dots < i_k} \omega_{i_1, i_2, \dots, i_k} \tilde{\phi}_{i_1} \wedge \tilde{\phi}_{i_2} \wedge \dots \wedge \tilde{\phi}_{i_k},$$

where $\omega_{i_1, i_2, \dots, i_k} : V \rightarrow \mathbb{R}$ for all indices. We say that the differential k -form is of class C^m if each $\omega_{i_1, i_2, \dots, i_k}$ is of class C^m .

The sum of differential k -forms, the product of a differential k -form by a real valued function on V and the wedge product of a differential k_1 -form with a differential k_2 -form are defined by the respective operation on $\mathbf{T}_{\mathbf{v}}V$ for each $\mathbf{v} \in V$. For instance, if ω_i is a differential k_i -form on V for $1 \leq i \leq 2$, then $\omega_1 \wedge \omega_2$ is the differential $(k_1 + k_2)$ -form on V defined by $(\omega_1 \wedge \omega_2)(\mathbf{v}) = \omega_1(\mathbf{v}) \wedge \omega_2(\mathbf{v})$ for all $\mathbf{v} \in V$.

If $f : V \rightarrow \mathbb{R}$ is differentiable, then $Df(\mathbf{v}) \in (\mathbb{R}^n)^*$. As we have done with ϕ_i above, we may define the following differential 1-form on V .

Definition 4.5.2

Given $f : V \rightarrow \mathbb{R}$ is of class C^1 , the differential 1-form $df : V \rightarrow \bigcup_{\mathbf{v} \in V} \Omega^1(\mathbf{T}_{\mathbf{v}}V)$ is defined by $df(\mathbf{v})((\mathbf{v}, \mathbf{x})) = Df(\mathbf{v})\mathbf{x}$ for each $\mathbf{v} \in V$ and all $\mathbf{x} \in \mathbb{R}^n$.

If $\pi_i : \mathbb{R}^n \rightarrow \mathbb{R}$ is the projection defined by $\pi_i(\mathbf{x}) = x_i$ for all $\mathbf{x} \in \mathbb{R}^n$, then the differential 1-form $d\pi_i$ is defined by $d\pi_i(\mathbf{v})((\mathbf{v}, \mathbf{x})) = D\pi_i(\mathbf{v})\mathbf{x} = \pi_i(\mathbf{x}) = x_i$ for each $\mathbf{v} \in V$ and all $\mathbf{x} \in \mathbb{R}^n$. Thus $d\pi_i(\mathbf{v}) \in \Omega^1(T_{\mathbf{v}}V)$ is generated by $\pi_i \in \Omega^1(\mathbb{R}^n)$ for all $\mathbf{v} \in V$.

For each $\mathbf{v} \in V$, the dual basis of the canonical basis $\mathcal{E} = \{(\mathbf{v}, \mathbf{e}_i)\}_{i=1}^n \subset T_{\mathbf{v}}V$ is $\{d\pi_i(\mathbf{v})\}_{i=1}^n \subset (T_{\mathbf{v}}V)^*$. By tradition, the differential 1-form $d\pi_i$ is denoted dx_i . Hence, every differential k -form ω on V can be expressed as

$$\omega = \sum_{i_1 < i_2 < \dots < i_k} \omega_{i_1, i_2, \dots, i_k} dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k},$$

where $\omega_{i_1, i_2, \dots, i_k} : V \rightarrow \mathbb{R}$ for all indices.

It is important to note that

$$dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k} = 0 \quad (4.5.1)$$

if there are two identical indices, and that

$$dx_{\sigma(i_1)} \wedge dx_{\sigma(i_2)} \wedge \dots \wedge dx_{\sigma(i_k)} = \text{sgn}(\sigma) dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k} \quad (4.5.2)$$

for all permutations $\sigma \in S_k$. These results follow from Proposition 4.1.14. For instance,

$$\begin{aligned} & (dx_{\sigma(i_1)} \wedge dx_{\sigma(i_2)} \wedge \dots \wedge dx_{\sigma(i_k)})(\mathbf{v})((\mathbf{v}, \mathbf{w}_1), (\mathbf{v}, \mathbf{w}_2), \dots, (\mathbf{v}, \mathbf{w}_k)) \\ &= (\pi_{\sigma(i_1)} \wedge \pi_{\sigma(i_2)} \wedge \dots \wedge \pi_{\sigma(i_k)})(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k) \\ &= \text{sgn } \sigma (\pi_{i_1} \wedge \pi_{i_2} \wedge \dots \wedge \pi_{i_k})(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k) \\ &= \text{sgn}(\sigma) (dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k})(\mathbf{v})((\mathbf{v}, \mathbf{w}_1), (\mathbf{v}, \mathbf{w}_2), \dots, (\mathbf{v}, \mathbf{w}_k)) \end{aligned}$$

for all $\mathbf{v} \in V$ and $\mathbf{w}_i \in \mathbb{R}^n$, where we have used Proposition 4.1.14 to obtain the second equality.

Theorem 4.5.3

Suppose that $V \subset \mathbb{R}^n$ is an open subset and $f : V \rightarrow \mathbb{R}$ is of class C^1 , then

$$df = \sum_{j=1}^n \frac{\partial f}{\partial x_j} dx_j.$$

Proof.

By definition, for every $\mathbf{v} \in V$,

$$df(\mathbf{v})((\mathbf{v}, \mathbf{x})) = Df(\mathbf{v})\mathbf{x} = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\mathbf{v})x_j = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\mathbf{v}) dx_j(\mathbf{v})((\mathbf{v}, \mathbf{x}))$$

for all $\mathbf{x} \in \mathbb{R}^n$. ■

Suppose that $V \subset \mathbb{R}^n$ and $W \subset \mathbb{R}^m$ are two open sets. If $f : V \rightarrow W$ is a function of class C^1 , then we can extend f^* to a map from differential k -forms on W to differential k -forms on V by $f^*(\omega)(\mathbf{v}) = f^*(\omega(f(\mathbf{v})))$ for every differential k -forms ω on W and $\mathbf{v} \in V$. We say that $f^*(\omega)$ is the **pull-back** of the differential k -form ω by f .

Proposition 4.5.4

Suppose that $V \subset \mathbb{R}^n$ and $W \subset \mathbb{R}^m$ are two open sets. If $f : V \rightarrow W$ is a function of class C^1 , then

1. $f^*(dx_i) = \sum_{j=1}^n \frac{\partial f_i}{\partial x_j} dx_j$ for $1 \leq i \leq m$.
2. $f^*(\omega_1 + \omega_2) = f^*(\omega_1) + f^*(\omega_2)$ for all differential k -forms ω_i on W .
3. $f^*(g\omega) = (g \circ f)f^*(\omega)$ for all differential k -form ω on W and functions $g : W \rightarrow \mathbb{R}$ of class C^1 .
4. $f^*(\omega_1 \wedge \omega_2) = f^*(\omega_1) \wedge f^*(\omega_2)$ for all differential k_1 -forms ω_1 and differential k_2 -form ω_2 on W .

Proof.

1) For every $\mathbf{v} \in V$, we have

$$\begin{aligned} f^*(dx_i)(\mathbf{v})(\mathbf{v}, \mathbf{x}) &= dx_i(f(\mathbf{v}))(\mathbf{f}_*(\mathbf{v}, \mathbf{x})) = dx_i(f(\mathbf{v}))(f(\mathbf{v}), Df(\mathbf{v})\mathbf{x}) = \pi_i(Df(\mathbf{v})\mathbf{x}) \\ &= \pi_i\left(\sum_{r=1}^m \left(\sum_{j=1}^n \frac{\partial f_r}{\partial x_j}(\mathbf{v})x_j\right) \mathbf{e}_r\right) = \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(\mathbf{v})x_j = \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(\mathbf{v}) dx_j(\mathbf{v})(\mathbf{v}, \mathbf{x}) \end{aligned}$$

for all $\mathbf{x} \in \mathbb{R}^n$.

2) This proof is left to the reader.

3) We have

$$\begin{aligned} f^*(g\omega)(\mathbf{v}) &= f^*((g\omega)(f(\mathbf{v}))) = f^*(g(f(\mathbf{v}))\omega(f(\mathbf{v}))) = g(f(\mathbf{v}))f^*(\omega(f(\mathbf{v}))) \\ &= g(f(\mathbf{v}))(f^*\omega)(\mathbf{v}) \end{aligned}$$

for all $\mathbf{v} \in V$, where we have used (2) of Proposition 4.3.6 for the third equality.

4) We have

$$\begin{aligned} f^*(\omega_1 \wedge \omega_2)(\mathbf{v}) &= f^*((\omega_1 \wedge \omega_2)(f(\mathbf{v}))) = f^*(\omega_1(f(\mathbf{v})) \wedge \omega_2(f(\mathbf{v}))) \\ &= f^*(\omega_1(f(\mathbf{v}))) \wedge f^*(\omega_2(f(\mathbf{v}))) = (f^*\omega_1)(\mathbf{v}) \wedge (f^*\omega_2)(\mathbf{v}) \end{aligned}$$

for all $\mathbf{v} \in V$, where we have used (4) of Proposition 4.3.6 for the third equality. ■

Theorem 4.5.5

Suppose that V and W are two open subsets of $\mathbb{R}^n$. Let $h : W \rightarrow \mathbb{R}$ and $f : V \rightarrow W$ be functions of class C^1 . Then

$$f^*(h dx_1 \wedge dx_2 \wedge \dots \wedge dx_n) = (h \circ f)(\det Df) dx_1 \wedge dx_2 \wedge \dots \wedge dx_n.$$

Proof.

We have from (3) of Proposition 4.5.4 that

$$f^*(h dx_1 \wedge dx_2 \wedge \dots \wedge dx_n) = (h \circ f) f^*(dx_1 \wedge dx_2 \wedge \dots \wedge dx_n). \quad (4.5.3)$$

Let $Q = Df(v)$; namely, the components $q_{i,j}$ of Q for $1 \leq i, j \leq n$ are given by $q_{i,j} = \frac{df_i}{dx_j}(v)$. Moreover, let $\mathcal{E}_v = \{(v, e_i)\}_{i=1}^n$ be the canonical basis of $T_v V$. Recall that $dx_i(v) = d\pi_i(v) \in \Omega^1(T_v V)$ for $1 \leq i \leq n$ is generated by the projection $\pi_i : \mathbb{R}^n \rightarrow \mathbb{R}$ defined by $\pi_i(x) = x_i$ for all $x \in \mathbb{R}^n$.

Given $w_k \in \mathbb{R}^n$ for $1 \leq k \leq n$, we may write $w_k = \sum_{j=1}^n a_{j,k} e_j$ for $1 \leq k \leq n$ and some $a_{j,k} \in \mathbb{R}$ where $\{e_i\}_{i=1}^n$ is the canonical basis on $\mathbb{R}^n$. Let A be the matrix with entries $a_{j,k}$.

We get from Theorem 4.1.16 that

$$\begin{aligned} & (f^*(dx_1 \wedge dx_2 \wedge \dots \wedge dx_n))(v) ((v, w_1), (v, w_2), \dots, (v, w_n)) \\ &= f^*((dx_1 \wedge dx_2 \wedge \dots \wedge dx_n)(f(v))) ((v, w_1), (v, w_2), \dots, (v, w_n)) \\ &= ((dx_1 \wedge dx_2 \wedge \dots \wedge dx_n)(f(v))) (f_*(v, w_1), f_*(v, w_2), \dots, f_*(v, w_n)) \\ &= ((dx_1 \wedge dx_2 \wedge \dots \wedge dx_n)(f(v))) \\ & \quad ((f(v), Df(v)w_1), (f(v), Df(v)w_2), \dots, (f(v), Df(v)w_n)) \\ &= (\pi_1 \wedge \pi_2 \wedge \dots \wedge \pi_n) \left(\sum_{j=1}^n a_{j,1} Df(v)e_j, \sum_{j=1}^n a_{j,2} Df(v)e_j, \dots, \sum_{j=1}^n a_{j,n} Df(v)e_j \right) \\ &= (\pi_1 \wedge \pi_2 \wedge \dots \wedge \pi_n) \left(\sum_{i=1}^n \left(\sum_{j=1}^n a_{j,1} q_{i,j} \right) e_i, \sum_{i=1}^n \left(\sum_{j=1}^n a_{j,2} q_{i,j} \right) e_i, \dots, \right. \\ & \quad \left. \sum_{i=1}^n \left(\sum_{j=1}^n a_{j,n} q_{i,j} \right) e_i \right) \\ &= \det(QA) (\pi_1 \wedge \pi_2 \wedge \dots \wedge \pi_n) (e_1, e_2, \dots, e_n) \\ &= \det(Q) \det(A) (\pi_1 \wedge \pi_2 \wedge \dots \wedge \pi_n) (e_1, e_2, \dots, e_n) \\ &= \det(Q) (\pi_1 \wedge \pi_2 \wedge \dots \wedge \pi_n) \left(\sum_{j=1}^n a_{j,1} e_j, \sum_{j=1}^n a_{j,2} e_j, \dots, \sum_{j=1}^n a_{j,n} e_j \right) \\ &= \det(Q) (\pi_1 \wedge \pi_2 \wedge \dots \wedge \pi_n) (w_1, w_2, \dots, w_n) \\ &= \det(Q) ((dx_1 \wedge dx_2 \wedge \dots \wedge dx_n)(v)) ((v, w_1), (v, w_2), \dots, (v, w_n)) \end{aligned}$$

for all $v \in V$ and $w_i \in \mathbb{R}^n$. Thus

$$f^*(dx_1 \wedge dx_2 \wedge \dots \wedge dx_n) = (\det(Df)) dx_1 \wedge dx_2 \wedge \dots \wedge dx_n. \quad \blacksquare$$

We now define an operator d that will map differential k -forms to differential $(k+1)$ -forms. We have already defined this operator for function $f : V \rightarrow \mathbb{R}$ in Definition 4.5.2 and Theorem 4.5.3.

Definition 4.5.6

Suppose that

$$\omega = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \omega_{i_1, i_2, \dots, i_k} dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k}$$

is a differential k -form of class C^1 on an open set $V \subset \mathbb{R}^n$. We define a differential $k+1$ -form $d\omega$ on V by

$$\begin{aligned} d\omega &= \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} d\omega_{i_1, i_2, \dots, i_k} \wedge dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k} \\ &= \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \sum_{j=1}^n \frac{\partial \omega_{i_1, i_2, \dots, i_k}}{\partial x_j} dx_j \wedge dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k}. \end{aligned}$$

Theorem 4.5.7

Suppose that $V \subset \mathbb{R}^n$ and $W \subset \mathbb{R}^m$ are two open subsets.

1. $d(\omega + \mu) = d\omega + d\mu$ for all differential k -forms ω and μ of class C^1 on V .
2. $d(\omega \wedge \mu) = d\omega \wedge \mu + (-1)^p \omega \wedge d\mu$ for all differential p -forms ω and differential q -form μ of class C^1 on V .
3. $d^2\omega \equiv d(d\omega) = 0$ for all differential k -forms ω of class C^2 on V .
4. $f^*(d\omega) = df^*(\omega)$ for all differential k -forms ω of class C^1 on W and all functions $f : V \rightarrow W$ of class C^1 .

Proof.

1) This proof is left to the reader.

2) We have that $\omega = \sum_{1 \leq i_1 < i_2 < \dots < i_p \leq n} \omega_{i_1, i_2, \dots, i_p} dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_p}$ and

$\mu = \sum_{1 \leq j_1 < j_2 < \dots < j_q \leq n} \mu_{j_1, j_2, \dots, j_q} dx_{j_1} \wedge dx_{j_2} \wedge \dots \wedge dx_{j_q}$. Hence

$$\begin{aligned} \omega \wedge \mu &= \sum_{\substack{1 \leq i_1 < i_2 < \dots < i_p \leq n \\ 1 \leq j_1 < j_2 < \dots < j_q \leq n}} \omega_{i_1, i_2, \dots, i_p} \mu_{j_1, j_2, \dots, j_q} \\ &\quad dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_p} \wedge dx_{j_1} \wedge dx_{j_2} \wedge \dots \wedge dx_{j_q}. \end{aligned}$$

From the definition of the operator d , we get

$$d\omega = \sum_{1 \leq i_1 < i_2 < \dots < i_p \leq n} \left(\sum_{m=1}^n \frac{\partial \omega_{i_1, i_2, \dots, i_p}}{\partial x_m} dx_m \right) \wedge dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_p}$$

and

$$d\mu = \sum_{1 \leq j_1 < j_2 < \dots < j_q \leq n} \left(\sum_{m=1}^n \frac{\partial \mu_{j_1, j_2, \dots, j_q}}{\partial x_m} dx_m \right) \wedge dx_{j_1} \wedge dx_{j_2} \wedge \dots \wedge dx_{j_q} .$$

Using (1), we get

$$\begin{aligned} d(\omega \wedge \mu) &= \sum_{\substack{1 \leq i_1 < i_2 < \dots < i_p \leq n \\ 1 \leq j_1 < j_2 < \dots < j_q \leq n}} d(\omega_{i_1, i_2, \dots, i_p} \mu_{j_1, j_2, \dots, j_q} \\ &\quad dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_p} \wedge dx_{j_1} \wedge dx_{j_2} \wedge \dots \wedge dx_{j_q}) \\ &= \sum_{\substack{1 \leq i_1 < i_2 < \dots < i_p \leq n \\ 1 \leq j_1 < j_2 < \dots < j_q \leq n}} \left(\sum_{m=1}^n \left(\frac{\partial \omega_{i_1, i_2, \dots, i_p}}{\partial x_m} \mu_{j_1, j_2, \dots, j_q} + \omega_{i_1, i_2, \dots, i_p} \frac{\partial \mu_{j_1, j_2, \dots, j_q}}{\partial x_m} \right) \right. \\ &\quad \left. dx_m \wedge dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_p} \wedge dx_{j_1} \wedge dx_{j_2} \wedge \dots \wedge dx_{j_q} \right) \\ &= \sum_{\substack{1 \leq i_1 < i_2 < \dots < i_p \leq n \\ 1 \leq j_1 < j_2 < \dots < j_q \leq n}} \left(\sum_{m=1}^n \frac{\partial \omega_{i_1, i_2, \dots, i_p}}{\partial x_m} dx_m \wedge dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_p} \right) \\ &\quad \wedge (\mu_{j_1, j_2, \dots, j_q} dx_{j_1} \wedge dx_{j_2} \wedge \dots \wedge dx_{j_q}) \\ &\quad + (-1)^p \sum_{\substack{1 \leq i_1 < i_2 < \dots < i_p \leq n \\ 1 \leq j_1 < j_2 < \dots < j_q \leq n}} (\omega_{i_1, i_2, \dots, i_p} \wedge dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_p}) \\ &\quad \wedge \left(\sum_{m=1}^n \frac{\partial \mu_{j_1, j_2, \dots, j_q}}{\partial x_m} dx_m \wedge dx_{j_1} \wedge dx_{j_2} \wedge \dots \wedge dx_{j_q} \right) \\ &= d\omega \wedge \mu + (-1)^p \omega \wedge d\mu , \end{aligned}$$

where the factor of $(-1)^p$ comes from $dx_m \wedge dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_p} \wedge dx_{j_1} \wedge dx_{j_2} \wedge \dots \wedge dx_{j_q} = (-1)^p dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_p} \wedge dx_m \wedge dx_{j_1} \wedge dx_{j_2} \wedge \dots \wedge dx_{j_q}$ according to (4.5.2).

3) If $\omega = \sum_{1 \leq i_1 < i_2 < \dots < i_p \leq n} \omega_{i_1, i_2, \dots, i_k} dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k}$, then

$$d\omega = \sum_{1 \leq i_1 < i_2 < \dots < i_p \leq n} \sum_{j=1}^n \frac{\partial \omega_{i_1, i_2, \dots, i_k}}{\partial x_j} dx_j \wedge dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k}$$

and

$$\begin{aligned} d(d\omega) &= \sum_{1 \leq i_1 < i_2 < \dots < i_p \leq n} \sum_{m=1}^n \sum_{j=1}^n \frac{\partial^2 \omega_{i_1, i_2, \dots, i_k}}{\partial x_j \partial x_m} dx_m \wedge dx_j \wedge dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k} \\ &= \sum_{1 \leq i_1 < i_2 < \dots < i_p \leq n} \left(\sum_{m=1}^n \sum_{j=1}^n \frac{\partial^2 \omega_{i_1, i_2, \dots, i_k}}{\partial x_j \partial x_m} dx_m \wedge dx_j \right) \wedge dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k} = 0 \end{aligned}$$

because $\frac{\partial^2 \omega_{i_1, i_2, \dots, i_k}}{\partial x_j \partial x_m} = \frac{\partial^2 \omega_{i_1, i_2, \dots, i_k}}{\partial x_m \partial x_j}$ for all $0 \leq j, m \leq n$ since $\omega_{i_1, i_2, \dots, i_k} \in C^2(V)$ implies that

$$\frac{\partial^2 \omega_{i_1, i_2, \dots, i_k}}{\partial x_j \partial x_m} dx_m \wedge dx_j = -\frac{\partial^2 \omega_{i_1, i_2, \dots, i_k}}{\partial x_m \partial x_j} dx_j \wedge dx_m$$

for all $1 \leq j, m \leq n$ according to (4.5.2); in particular, both sides are null when $m = j$ according to (4.5.1).

4) We prove this result using a proof by induction on the order of the differential form. Using the chain rule, it is easy to verify that the result is true if ω is a 0-form; namely, if $\omega : V \rightarrow \mathbb{R}$ is a function of class C^1 . Suppose that the result is true for all differential k -form.

If ω is a differential $(k+1)$ -form, then we can write $\omega = \sum_{j=1}^n \mu_j \wedge dx_j$ for some k -forms μ_j .

From (1), (2) and (3) above, we get

$$d\omega = \sum_{j=1}^n d(\mu_j \wedge dx_j) = \sum_{j=1}^n (d\mu_j \wedge dx_j + (-1)^k \mu_j \wedge d^2 x_j) = \sum_{j=1}^n (d\mu_j \wedge dx_j) .$$

It follows from (2) and (4) of Proposition 4.5.4 that $f^* d\omega = \sum_{j=1}^n (f^* d\mu_j \wedge f^* dx_j)$. Recall that $dx_j = d\pi_j$ where $\pi_j : \mathbb{R}^n \rightarrow \mathbb{R}$ is a projection. Since π_j is a 0-form, we have that $d(f^* dx_j) = d(f^* d\pi_j) = d(d(f^*(\pi_j))) = d^2(\pi_j \circ f) = d^2 f_j = 0$ for all j because of (3). Hence

$$d(f^* \mu_j \wedge f^* dx_j) = df^* \mu_j \wedge f^* dx_j + (-1)^k f^* \mu_j \wedge d(f^* dx_j) = df^* \mu_j \wedge f^* dx_j .$$

Finally, using this relation and our hypothesis of induction, we get

$$\begin{aligned} f^* d\omega &= \sum_{j=1}^n (f^* d\mu_j \wedge f^* dx_j) = \sum_{j=1}^n (df^* \mu_j \wedge f^* dx_j) = \sum_{j=1}^n d(f^* \mu_j \wedge f^* dx_j) \\ &= d\left(\sum_{j=1}^n f^* \mu_j \wedge f^* dx_j\right) = d\left(f^* \left(\sum_{j=1}^n \mu_j \wedge dx_j\right)\right) = d(f^* \omega) , \end{aligned}$$

where we have used (1) above to get the fourth equality, and (2) and (4) of Proposition 4.5.4 to get the second to last equality. This complete the proof by induction. ■

From now on, we will assume that each differential form that we consider is of class C^j for j large enough to meet the requirements of the statement about this differential form.

Definition 4.5.8

A differential k -form ω on the open set $V \subset \mathbb{R}^n$ is **closed** if $d\omega = 0$ on V . A differential k -form ω on the open set $V \subset \mathbb{R}^n$ is **exact** if there exists a differential $(k-1)$ -form μ on V such that $d\mu = \omega$.

It follows from item (3) of Theorem 4.5.7 that exact differential k -forms are closed.

Example 4.5.9

Are closed differential k -forms exact? The following example is from [26] though it is a classical example that can be found in almost all textbooks on differential geometry.

In general, the answer is negative. Consider the differential 1-form

$$\omega = \frac{-x_2}{x_1^2 + x_2^2} dx_1 + \frac{x_1}{x_1^2 + x_2^2} dx_2$$

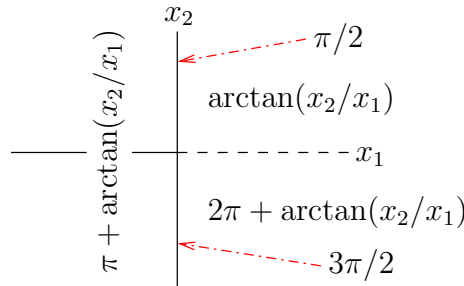
defined on $V = \mathbb{R}^2 \setminus \{\mathbf{0}\}$ (see warning below).

We have

$$\begin{aligned} d\omega &= \left(\frac{2x_1x_2}{(x_1^2 + x_2^2)^2} dx_1 + \frac{x_2^2 - x_1^2}{(x_1^2 + x_2^2)^2} dx_2 \right) \wedge dx_1 \\ &\quad + \left(\frac{x_2^2 - x_1^2}{(x_1^2 + x_2^2)^2} dx_1 - \frac{2x_1x_2}{(x_1^2 + x_2^2)^2} dx_2 \right) \wedge dx_2 \\ &= \frac{x_2^2 - x_1^2}{(x_1^2 + x_2^2)^2} dx_2 \wedge dx_1 + \frac{x_2^2 - x_1^2}{(x_1^2 + x_2^2)^2} dx_1 \wedge dx_2 \\ &= -\frac{x_2^2 - x_1^2}{(x_1^2 + x_2^2)^2} dx_1 \wedge dx_2 + \frac{x_2^2 - x_1^2}{(x_1^2 + x_2^2)^2} dx_1 \wedge dx_2 = 0. \end{aligned}$$

However, there does not exist any function μ on V such that $d\mu = \omega$. To prove this statement, let $W = \mathbb{R}^2 \setminus \{(x_1, 0) : x_1 \geq 0\}$ and consider the function $\theta : W \rightarrow \mathbb{R}$ defined by

$$\theta(\mathbf{x}) = \begin{cases} \arctan(x_2/x_1) & \text{if } x_1, x_2 > 0 \\ \pi/2 & \text{if } x_1 = 0, x_2 > 0 \\ \pi + \arctan(x_2/x_1) & \text{if } x_1 < 0 \\ 3\pi/2 & \text{if } x_1 = 0, x_2 < 0 \\ 2\pi + \arctan(x_2/x_1) & \text{if } x_1 > 0, x_2 < 0 \end{cases}$$



The image of θ is $]0, 2\pi[$. We have that $d\theta = \omega$ on W . If there was a function $\mu : V \rightarrow \mathbb{R}$ such that $d\mu = \omega$ on V , then we would have that $\frac{\partial \mu}{\partial x_1} = \frac{\partial \theta}{\partial x_1} = \frac{-x_2}{x_1^2 + x_2^2}$ and $\frac{\partial \mu}{\partial x_2} = \frac{\partial \theta}{\partial x_2} = \frac{x_1}{x_1^2 + x_2^2}$ on W . Thus $\mu = \theta + C$ on W for some constant C . However, at $\mathbf{x}_0 = (x_1, 0)$ with

$x_1 > 0$, we have that $\mu(\mathbf{x}) = \theta(\mathbf{x}) + C \rightarrow C$ as $\mathbf{x} \rightarrow \mathbf{x}_0$ with $x_2 > 0$ and $\mu(\mathbf{x}) = \theta(\mathbf{x}) + C \rightarrow 2\pi + C$ as $\mathbf{x} \rightarrow \mathbf{x}_0$ with $x_2 < 0$. This is a contradiction that μ is continuous at $\mathbf{x}_0 \in V$. Thus ω cannot be exact on V .

We provide another proof that ω cannot be exact in V in Example 4.7.4 later. ♣

Remark 4.5.10 (Warning)

The expression

$$\omega = \frac{-x_2}{x_1^2 + x_2^2} dx_1 + \frac{x_1}{x_1^2 + x_2^2} dx_2$$

is an abbreviation for $\omega = f_1 dx_1 + f_2 dx_2$ where $f_1 : \mathbb{R}^2 \setminus \{\mathbf{0}\} \rightarrow \mathbb{R}$ is the function defined by $f_1(\mathbf{x}) = -x_2/(x_1^2 + x_2^2)$ and $f_2 : \mathbb{R}^2 \setminus \{\mathbf{0}\} \rightarrow \mathbb{R}$ is the function defined by $f_2(\mathbf{x}) = x_1/(x_1^2 + x_2^2)$. To introduce the variable $\mathbf{x}$ and be rigorously correct, we should have written instead

$$\omega(\mathbf{x}) = \frac{-x_2}{x_1^2 + x_2^2} dx_1(\mathbf{x}) + \frac{x_1}{x_1^2 + x_2^2} dx_2(\mathbf{x})$$

but that would have been a uglier expression. We will regularly use this type of abbreviation from now on. ♠

Suppose that $\omega = \sum_{i=1}^n \omega_i dx_i$ is a closed differential 1-form on $\mathbb{R}^n$. Let

$$\mu(\mathbf{x}) = \sum_{j=1}^n \int_0^1 \omega_j(t\mathbf{x}) x_j dt \quad (4.5.4)$$

for $1 \leq j \leq n$ and $\mathbf{x} \in \mathbb{R}^n$. Then

$$\frac{\partial \mu}{\partial x_i}(\mathbf{x}) = \sum_{\substack{j=1 \\ i \neq j}}^n \int_0^1 \left(tx_j \frac{\partial \omega_j}{\partial x_i}(t\mathbf{x}) \right) dt + \int_0^1 \left(tx_i \frac{\partial \omega_i}{\partial x_i}(t\mathbf{x}) + \omega_i(t\mathbf{x}) \right) dt. \quad (4.5.5)$$

Since ω is a closed differential form, we get that $d\omega = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial \omega_i}{\partial x_j} dx_j \wedge dx_i = 0$. Therefore, with $(\mathbf{x}, \mathbf{y}) \in T_{\mathbf{x}}\mathbb{R}^n$ where $y_i = 0$ for $i \neq p, q$ and $y_i = 1$ for $i = p, q$, we get from $d\omega(\mathbf{x})(\mathbf{x}, \mathbf{y}) = 0$ that $\frac{\partial \omega_p}{\partial x_q}(\mathbf{x}) = \frac{\partial \omega_q}{\partial x_p}(\mathbf{x})$ for $1 \leq p < q \leq n$. Thus (4.5.5) yields

$$\begin{aligned} \frac{\partial \mu}{\partial x_i}(\mathbf{x}) &= \sum_{\substack{j=1 \\ i \neq j}}^n \int_0^1 \left(tx_j \frac{\partial \omega_i}{\partial x_j}(t\mathbf{x}) \right) dt + \int_0^1 \left(tx_i \frac{\partial \omega_i}{\partial x_i}(t\mathbf{x}) + \omega_i(t\mathbf{x}) \right) dt \\ &= \int_0^1 \left(\sum_{j=1}^n tx_j \frac{\partial \omega_i}{\partial x_j}(t\mathbf{x}) + \omega_i(t\mathbf{x}) \right) dt = \int_0^1 \frac{d}{dt} (t\omega_i(t\mathbf{x})) dt = \omega_i(\mathbf{x}) \end{aligned}$$

for $\mathbf{x} \in \mathbb{R}^n$. Hence $d\mu = \omega$. This method to find a differential 0-form μ such that $d\mu = \omega$ works because the line from $\mathbf{0}$ to any $\mathbf{x} \in \mathbb{R}^n$ is included in the domain of ω . This is required in the definition of μ in (4.5.4). To extend this method, we consider domains with a similar property.

Definition 4.5.11

A set $V \subset \mathbb{R}^n$ is **star-shaped** if there exists a point $\mathbf{y} \in V$ such that the line segment $\{\mathbf{y} + t(\mathbf{x} - \mathbf{y}) : 0 \leq t \leq 1\} \subset V$ for all $\mathbf{x} \in V$.

If ω is a closed differential 1-form defined on a star-shaped domain V , then the integral formula in (4.5.4) (with appropriate shifting) can be used to find $\mu : V \rightarrow \mathbb{R}$ such that $d\mu = \omega$. A general form of this integral formula is used to prove the following theorem.

Theorem 4.5.12 (Poincaré Lemma)

If $V \subset \mathbb{R}^n$ is an open set which is star-shaped and ω is a closed differential k -form on V with $k > 0$, then ω is exact on V .

Proof.

Without loss of generality, we may assume that $\mathbf{0} \in V$ and every point $\mathbf{x} \in V$ can be joined by a straight line in V to the origin. If it is not, then we consider $\tilde{V} = V - \{\mathbf{y}\} = \{\mathbf{x} - \mathbf{y} : \mathbf{x} \in V\}$ and $\tilde{\omega}(\mathbf{x}) = \omega(\mathbf{x} + \mathbf{y})$ for $\mathbf{x} \in \tilde{V}$ where $\mathbf{y}$ is the "centre" of the star-shaped set described in Definition 4.5.11.

For every positive integer k , we prove the existence of a function L mapping differential k -forms on V to differential $(k-1)$ -forms on V such that $L(0) = 0$ and

$$\tau = [L, d]\tau \equiv L(d\tau) + dL(\tau) \quad (4.5.6)$$

for all differential k -forms τ on V ⁵. Assuming that L exists, (4.5.6) with $\tau = \omega$ yields $\omega = dL(\omega)$ because $d\omega = 0$. Thus ω is exact.

Suppose that

$$\tau = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \tau_{i_1, i_2, \dots, i_k} dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k}.$$

Let $g_{i_1, i_2, \dots, i_k, j} : V \rightarrow \mathbb{R}$ for $1 \leq i_1 < i_2 < \dots < i_k \leq n$ and $1 \leq j \leq k$ be the functions defined by

$$g_{i_1, i_2, \dots, i_k, j}(\mathbf{x}) = (-1)^{j-1} x_{i_j} \int_0^1 t^{k-1} \tau_{i_1, i_2, \dots, i_k}(t\mathbf{x}) dt$$

for $\mathbf{x} \in V$. We define $L(\tau)$ by

$$L(\tau) = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \left(\sum_{j=1}^k \left(g_{i_1, i_2, \dots, i_k, j} dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge \widehat{dx_{i_j}} \wedge \dots \wedge dx_{i_k} \right) \right)$$

where the hat over a dx_{i_j} indicates that this factor has been removed from the wedge product. The integral is well defined because $t\mathbf{x} \in V$ for $0 \leq t \leq 1$. For the clarity of the proof, it is

⁵The expression $[L, d]$ is called a bracket of L and d . Expressions of that form are frequent in differential geometry and, in fact, in many other fields of mathematics.

probably better to use the abbreviation mentioned in Remark 4.5.10 and simply write

$$L(\tau) = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \left(\sum_{j=1}^k \left((-1)^{j-1} x_{i_j} \left(\int_0^1 t^{k-1} \tau_{i_1, i_2, \dots, i_k}(t\mathbf{x}) dt \right) \right. \right. \\ \left. \left. dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge \widehat{dx_{i_j}} \wedge \dots \wedge dx_{i_k} \right) \right).$$

We have

$$\begin{aligned} dL(\tau) &= d(L(\tau)) \\ &= \sum_{i_1 < i_2 < \dots < i_k} \left(\sum_{j=1}^k \left((-1)^{j-1} \sum_{m=1}^n \left(\frac{\partial}{\partial x_m} \left(x_{i_j} \int_0^1 t^{k-1} \tau_{i_1, i_2, \dots, i_k}(t\mathbf{x}) dt \right) \right. \right. \right. \\ &\quad \left. \left. dx_m \wedge dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge \widehat{dx_{i_j}} \wedge \dots \wedge dx_{i_k} \right) \right) \right) \\ &= \sum_{i_1 < i_2 < \dots < i_k} \left(\sum_{j=1}^k \left((-1)^{j-1} \left(\int_0^1 t^{k-1} \tau_{i_1, i_2, \dots, i_k}(t\mathbf{x}) dt \right) \right. \right. \\ &\quad \left. \left. dx_{i_j} \wedge dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge \widehat{dx_{i_j}} \wedge \dots \wedge dx_{i_k} \right) \right) \\ &\quad + \sum_{i_1 < i_2 < \dots < i_k} \left(\sum_{j=1}^k \left((-1)^{j-1} x_{i_j} \sum_{m=1}^n \left(\left(\int_0^1 t^k \frac{\partial \tau_{i_1, i_2, \dots, i_k}}{\partial x_m}(t\mathbf{x}) dt \right) \right. \right. \right. \\ &\quad \left. \left. dx_m \wedge dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge \widehat{dx_{i_j}} \wedge \dots \wedge dx_{i_k} \right) \right) \right) \\ &= \sum_{i_1 < i_2 < \dots < i_k} \left(\left(\int_0^1 k t^{k-1} \tau_{i_1, i_2, \dots, i_k}(t\mathbf{x}) dt \right) dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k} \right) \\ &\quad + \sum_{i_1 < i_2 < \dots < i_k} \left(\sum_{j=1}^k \left((-1)^{j-1} x_{i_j} \sum_{m=1}^n \left(\left(\int_0^1 t^k \frac{\partial \tau_{i_1, i_2, \dots, i_k}}{\partial x_m}(t\mathbf{x}) dt \right) \right. \right. \right. \\ &\quad \left. \left. dx_m \wedge dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge \widehat{dx_{i_j}} \wedge \dots \wedge dx_{i_k} \right) \right) \right), \quad (4.5.7) \end{aligned}$$

where we have used $dx_{i_j} \wedge dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge \widehat{dx_{i_j}} \wedge \dots \wedge dx_{i_k} = (-1)^{j-1} dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_j} \wedge \dots \wedge dx_{i_k}$ to get the last equality.

Moreover, if we apply L to the differential $(k+1)$ -form

$$d\tau = \sum_{i_1 < i_2 < \dots < i_k} \sum_{m=1}^n \frac{\partial \tau_{i_1, i_2, \dots, i_k}}{\partial x_m} dx_m \wedge dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k},$$

then we get

$$L(d\tau) = \sum_{i_1 < i_2 < \dots < i_k} \left(\sum_{m=1}^n \left(x_m \int_0^1 t^k \frac{\partial \tau_{i_1, i_2, \dots, i_k}}{\partial x_m}(t\mathbf{x}) dt \right) dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k} \right)$$

$$+ \sum_{i_1 < i_2 < \dots < i_k} \left(\sum_{j=1}^k \left((-1)^j x_{i_j} \sum_{m=1}^n \left(\left(\int_0^1 t^k \frac{\partial \tau_{i_1, i_2, \dots, i_k}}{\partial x_m}(t\mathbf{x}) dt \right) dx_m \wedge dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge \widehat{dx_{i_j}} \wedge \dots \wedge dx_{i_k} \right) \right) \right). \quad (4.5.8)$$

Adding (4.5.7) and (4.5.8) gives

$$\begin{aligned} dL(\tau) + L(d\tau) &= \sum_{i_1 < i_2 < \dots < i_k} \left(\int_0^1 k t^{k-1} \tau_{i_1, i_2, \dots, i_k}(t\mathbf{x}) dt \right. \\ &\quad \left. + \sum_{m=1}^n x_m \int_0^1 t^k \frac{\partial \tau_{i_1, i_2, \dots, i_k}}{\partial x_m}(t\mathbf{x}) dt \right) dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k} \\ &= \sum_{i_1 < i_2 < \dots < i_k} \left(\int_0^1 \frac{d}{dt} (t^k \tau_{i_1, i_2, \dots, i_k}(t\mathbf{x})) dt \right) dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k} \\ &= \sum_{i_1 < i_2 < \dots < i_k} \tau_{i_1, i_2, \dots, i_k}(\mathbf{x}) dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k} = \tau(\mathbf{x}). \quad \blacksquare \end{aligned}$$

Remark 4.5.13

The previous theorem can be generalized to domains V that are **(smoothly) contractible to a point**; namely, there exist a point $\mathbf{y} \in V$ and a continuous differentiable map $H : V \times [0, 1] \rightarrow V$ such that $H(\mathbf{x}, 1) = \mathbf{x}$ and $H(\mathbf{x}, 0) = \mathbf{y}$ for all $\mathbf{x} \in V$. A star-shaped set V as given in Definition 4.5.11 is contractible to the point $\mathbf{y} \in V$. It suffices to use the map $H : V \times [0, 1] \rightarrow V$ defined by $H(\mathbf{x}, t) = \mathbf{y} + t(\mathbf{x} - \mathbf{y})$ for $\mathbf{x} \in V$ and $t \in [0, 1]$.

The proof of the generalized version of Theorem 4.5.12 is given in Section 7.1. ♠

4.6 Domains of Integration and their Boundary

Before stating the general form of the fundamental theorem of calculus, we have to introduce the domains of integration that will be considered.

Definition 4.6.1

A **singular m-cube** in $\mathbb{R}^n$ is a continuous function $\sigma : I_m \rightarrow \mathbb{R}^n$, where I_m is the unit cube $I_m = \{(x_1, x_2, \dots, x_m) : 0 \leq x_j \leq 1 \text{ for } 1 \leq j \leq m\} \subset \mathbb{R}^m$. For $m = 0$, we define $\mathbb{R}^m$ and I_m as $\{0\}$.

A singular 0-cube is simply a map $\sigma : \{0\} \rightarrow \mathbb{R}^n$. The image of a singular 0-cube σ is a point in $\mathbb{R}^n$. A singular 1-cube is a map $\sigma : [0, 1] \rightarrow \mathbb{R}^n$; namely, a **curve** in $\mathbb{R}^n$.

Definition 4.6.2

The **standard m-cube** is the function $i_m : I_m \rightarrow \mathbb{R}^m$ defined by $i_m(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in I_m$.

Definition 4.6.3

An **m-chain** in $\mathbb{R}^n$ is an expression of the form $\sum_{j=1}^J a_j \sigma_j$, where $a_j \in \mathbb{Z}$ and σ_j is a singular m -cube in $\mathbb{R}^n$ for $1 \leq j \leq J$ with J a positive integer.

A singular m -cube σ can be seen as the m -chain 1σ . The addition of m -chain is defined by

$$\sum_{j=1}^J a_j \sigma_j + \sum_{j=1}^J b_j \sigma_j = \sum_{j=1}^J (a_j + b_j) \sigma_j .$$

We have only added the coefficients associated to the same singular m -cubes. When adding m -chains, we may assume that they are sums over the same singular m -cubes by adding terms of the form $0\sigma_j$ if necessary. The multiplication of an m -chain by an integer c is defined by

$$c \sum_{j=1}^J a_j \sigma_j = \sum_{j=1}^J (ca_j) \sigma_j .$$

We use the standard m -cube i_m to define the singular $(m-1)$ -cubes $i_{m,j,i} : I_{m-1} \rightarrow \mathbb{R}^m$ for $1 \leq j \leq m$ and $i = 0$ or $i = 1$ by $i_{m,j,i}(\mathbf{x}) = i_m \left(\begin{pmatrix} x_1 & x_2 & \dots & x_{j-1} & i & x_j & \dots & x_{m-1} \end{pmatrix}^\top \right)$ for all $\mathbf{x} \in I_{m-1}$.

Definition 4.6.4

For $1 \leq j \leq m$ and $0 \leq i \leq 1$, the singular $(m-1)$ -cube $i_{m,j,i}$ is called the **(j, i) face** of the standard m -cube i_m . The **boundary** of the standard m -cube i_m , denoted ∂i_m , is defined by the $(m-1)$ -chain

$$\partial i_m = \sum_{j=1}^m (-1)^{j-1} (i_{m,j,1} - i_{m,j,0}) .$$

We have drawn the boundary of the standard m -cube for $m = 1$ and $m = 2$ in Figure 4.1. In particular, $i_{1,1,1} : \{0\} \rightarrow \mathbb{R}^n$ is defined by $i_{1,1,1}(0) = i_1(1)$ and $i_{1,1,0} : \{0\} \rightarrow \mathbb{R}^n$ is defined by $i_{1,1,0}(0) = i_1(0)$.

The (j, i) face of a general singular m -cube σ is defined by $\sigma_{j,i} = \sigma \circ i_{m,j,i}$ for $1 \leq j \leq m$ and $i = 0$ or $i = 1$.

Definition 4.6.5

Let σ be a singular m -cube. The **boundary** of σ , denoted $\partial\sigma$, is defined by the $(m-1)$ -chain

$$\partial\sigma = \sum_{j=1}^m (-1)^{j-1} (\sigma_{j,1} - \sigma_{j,0}) .$$

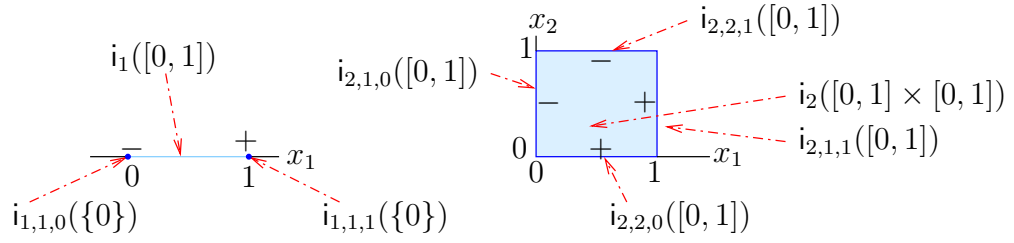


Figure 4.1: We have drawn a standard m -cube with its boundary for $m = 1$ on the left and $m = 2$ on the right. We have that $\partial i_1 = i_{1,1,1} - i_{1,1,0}$ and $\partial i_2 = i_{2,1,1} - i_{2,1,0} + i_{2,2,0} - i_{2,2,1}$. We have added the sign of the faces to the standard m -cube.

Let σ be a m -chain given by $\sigma = \sum_{j=1}^J a_j \sigma_j$. The **boundary** of σ , also denote $\partial\sigma$, is defined by the $(m-1)$ -chain $\partial\sigma = \sum_{j=1}^J a_j \partial\sigma_j$.

We have drawn the boundary of a singular 2-cube in Figure 4.2.

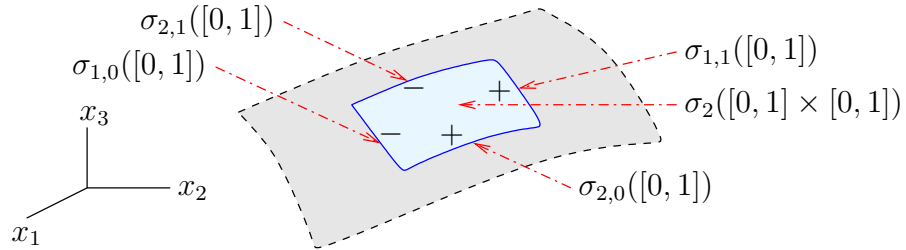


Figure 4.2: We have drawn a singular 2-cube in $\mathbb{R}^3$ with its boundary. We have that $\partial\sigma_2 = \sigma_{1,1} - \sigma_{1,0} + \sigma_{2,0} - \sigma_{2,1}$. We have added the sign of the faces to the singular 2-cube.

Theorem 4.6.6

if σ is a m -chain, then $\partial(\partial\sigma) = 0$ where 0 denotes the empty $(m-2)$ -chain.

Proof.

Since the boundary of m -chain is the sum of the boundary of the m -cubes that are parts of the chain, it is enough to prove that $\partial(\partial\sigma) = 0$ for a m -cube σ .

Since $\partial\sigma = \sum_{j=1}^m (-1)^{j-1} (\sigma_{j,1} - \sigma_{j,0})$, we get

$$\begin{aligned} \partial(\partial\sigma) &= \sum_{j=1}^m (-1)^{j-1} \left(\sum_{i=1}^{m-1} (-1)^{i-1} ((\sigma_{j,1})_{i,1} - (\sigma_{j,1})_{i,0}) - \sum_{i=1}^{m-1} (-1)^{i-1} ((\sigma_{j,0})_{i,1} - (\sigma_{j,0})_{i,0}) \right) \\ &= \sum_{j=1}^m \sum_{i=1}^{m-1} (-1)^{j+i-2} ((\sigma_{j,1})_{i,1} - (\sigma_{j,1})_{i,0} - (\sigma_{j,0})_{i,1} + (\sigma_{j,0})_{i,0}) \\ &= \sum_{j=1}^m \sum_{i=1}^{m-1} \sum_{s_1, s_2=0}^1 (-1)^{j+i+s_1+s_2} (\sigma_{j,s_2})_{i,s_1} . \end{aligned}$$

If $1 \leq i < j \leq m$, then

$$\begin{aligned} (\sigma_{j,s_2})_{i,s_1}(\mathbf{y}) &= \sigma_{j,s_2} \left(\begin{pmatrix} y_1 & y_2 & \cdots & y_{i-1} & s_1 & y_i & \cdots & y_{m-2} \end{pmatrix}^\top \right) \\ &= \left(\begin{pmatrix} y_1 & y_2 & \cdots & y_{i-1} & s_1 & y_i & \cdots & y_{j-2} & s_2 & y_{j-1} & \cdots & y_{m-2} \end{pmatrix}^\top \right) \\ &= \sigma_{i,s_1} \left(\begin{pmatrix} y_1 & y_2 & \cdots & y_{j-2} & s_2 & y_{j-1} & \cdots & y_{m-2} \end{pmatrix}^\top \right) = (\sigma_{i,s_1})_{j-1,s_2}(\mathbf{y}) \quad (4.6.1) \end{aligned}$$

for all $\mathbf{y} \in I_{m-2}$. If $i = j - 1$, then the section $y_i \dots y_{j-2}$ is removed from the second row of (4.6.1). The terms of the form $(\sigma_{j,s_2})_{i,s_1}(\mathbf{y})$ with $j = i$ are included in (4.6.1) because, when $i = k$ and $j = k + 1$ for $1 \leq k < m$, we get $(\sigma_{k,s_2})_{k,s_1}(\mathbf{y})$ for $1 \leq k < m$ on the right side of (4.6.1). Similarly, the terms of the form $(\sigma_{j,s_2})_{i,s_1}(\mathbf{y})$ with $1 \leq j < i < m$ are also included in (4.6.1) because we have $1 \leq i \leq j - 1 < m$ on the right side of (4.6.1).

Since the coefficient of $(\sigma_{j,s_2})_{i,s_1}(\mathbf{y})$ is $(-1)^{j+i+s_1+s_2}$ and the coefficient of $(\sigma_{i,s_1})_{j-1,s_2}(\mathbf{y})$ is $(-1)^{j-1+i+s_1+s_2}$, they have opposite signs and thus cancel each other. Therefore $\partial(\partial\sigma) = 0$. ■

Remark 4.6.7

It is worth emphasizing the difference between the topological boundary of a cube and the boundary of a m -cube. For instance, the topological boundary of the cube $I_3 = \{(x_1, x_2, x_3) : 0 \leq x_j \leq 1 \text{ for } 1 \leq j \leq 3\}$ is the set

$$\begin{aligned} \partial I_3 &= \bigcup_{m=0}^1 \left(\{(m, x_2, x_3) : 0 \leq x_j \leq 1 \text{ for } j = 2, 3\} \cup \{(x_1, m, x_3) : 0 \leq x_j \leq 1 \text{ for } j = 1, 3\} \right. \\ &\quad \left. \cup \{(x_1, x_2, m) : 0 \leq x_j \leq 1 \text{ for } j = 1, 2\} \right) \end{aligned}$$

and the topological boundary of ∂I_3 is $\partial(\partial I_3) = \partial I_3$. The issue about boundary will get even more confusing when we will consider manifolds later. The boundary of the manifold I_3 is still ∂I_3 defined above and denoted ∂I_3 . However, the boundary of ∂I_3 as a manifold, still denoted $\partial(\partial I_3)$, is the empty set $\emptyset$. As a manifold, ∂I_3 does not have a boundary. Therefore, the reader should be very careful of not mixing all these definitions of boundary. The context should determine which definition of boundary is used. It is unfortunate that by tradition we use the same symbol ∂ for all definitions of boundary. ♠

4.7 Integration on Chains

We begin with the definition of the integral of a differential k -forms on an k -chain. From now on all differential k -forms ω on an open set $V \subset \mathbb{R}^n$ and all singular k -cubes $\sigma : I_k \rightarrow V$ are assumed to be of class C^j for j as sufficiently large. Recall that $\sigma : I_k \rightarrow V$ of class C^j means that there exist an open set $U \supset I_k$ and a function $\tilde{\sigma} : U \rightarrow \mathbb{R}^n$ of class C^j such that $\tilde{\sigma}|_{I_k} = \sigma$.

Definition 4.7.1

Let k be a positive integer and $\omega = f dx_1 \wedge dx_2 \wedge \dots \wedge dx_k$ be a differential k -form on $\mathbb{R}^k$. The **integral of ω on I_k** is defined by $\int_{I_k} \omega = \int_{I_k} f(\mathbf{x}) d\mathbf{x}$. Let ω be a differential k -form defined on an open set $V \subset \mathbb{R}^n$ and σ be a singular k -cube with image in V . The **integral of ω on σ** is defined by $\int_{\sigma} \omega = \int_{I_k} \sigma^*(\omega)$.

Obviously, the previous definition include the case where $\omega = f dx_1 \wedge dx_2 \wedge \dots \wedge dx_k$ is a differential k -form on $\mathbb{R}^k$ and σ is the standard k -cube i_m . We then have that $\int_{i_k} \omega =$

$$\int_{I_k} i_k^*(\omega) = \int_{I_k} \omega = \int_{I_k} f(\mathbf{x}) d\mathbf{x}.$$

Definition 4.7.2

Let ω be a 0-form on an open set $V \subset \mathbb{R}^n$ and σ is a singular 0-cube with image in V . Namely, $\omega : V \rightarrow \mathbb{R}$ and $\sigma(0) \in V$. The **integral of ω on σ** is defined by $\int_{\sigma} \omega = \omega(\sigma(0))$.

Definition 4.7.3

Let ω be a differential k -form on an open set $V \subset \mathbb{R}^n$ and $\sigma = \sum_{j=1}^J a_j \sigma_j$ be a k -chain in V ; namely, the image of each singular k -cube σ_j is in V . The **integral of ω on σ** is defined by $\int_{\sigma} \omega = \sum_{j=1}^J a_j \int_{\sigma_j} \omega = \sum_{j=1}^J a_j \int_{I_k} \sigma_j^*(\omega)$.

Example 4.7.4

We give a second proof that the differential 1-form

$$\omega = \frac{-x_2}{x_1^2 + x_2^2} dx_1 + \frac{x_1}{x_1^2 + x_2^2} dx_2$$

defined on $V = \mathbb{R}^2 \setminus \{0\}$ and introduced in Example 4.5.9 cannot be exact.

Suppose that $\omega = df$ where $f : V \rightarrow \mathbb{R}$ is a differentiable function on V . Let σ be the singular 1-cube defined by $\sigma(\theta) = (\cos(2\pi\theta), \sin(2\pi\theta))$ for $\theta \in [0, 1]$. It follows that

$$\begin{aligned} \int_{\sigma} \omega &= \int_{\sigma} df = \int_{[0,1]} \sigma^*(df) = \int_{[0,1]} d(\sigma^*(f)) = \int_{[0,1]} d(f(\sigma)) \\ &= \int_{[0,1]} \frac{d}{d\theta}(f(\sigma(\theta))) d\theta = f(\sigma(1)) - f(\sigma(0)) = 0 \end{aligned}$$

because $\sigma(0) = \sigma(2\pi) = (1, 0)$. However, if we compute directly the integral $\int_{\sigma} \omega$, then we get

$$\int_{\sigma} \omega = \int_{[0,1]} \sigma^*(\omega) = 2\pi \int_{[0,1]} d\theta = 2\pi \neq 0.$$

Therefore, we cannot have that $\omega = df$ with $f : V \rightarrow \mathbb{R}$. ♣

The next theorem is often called the **fundamental theorem of calculus** in higher dimension. In fact, the classical fundamental theorem of calculus is the case $m = 1$ and $\omega = i_1$ in the next theorem.

Before stating and proving the next theorem, it is going to be useful to evaluate $i_{k,j,s}^* dx_i$ for $1 \leq i, j \leq k$ and $s = 0$ or $s = 1$. Recall that $dx_i = d\pi_i$. Hence

$$\begin{aligned} (i_{k,j,s}^* dx_i)(\mathbf{w})(\mathbf{w}, \mathbf{y}) &= (i_{k,j,s}^* d\pi_i)(\mathbf{w})(\mathbf{w}, \mathbf{y}) = (d(\pi_i \circ i_{k,j,s}))(\mathbf{w})(\mathbf{w}, \mathbf{y}) \\ &= (D(\pi_i \circ i_{k,j,s})(\mathbf{w}))(\mathbf{y}) \\ &= (\pi_i \circ i_{k,j,0})(\mathbf{y}) = \pi_i(i_{k,j,0}(\mathbf{y})) = \begin{cases} y_i & \text{if } i < j \\ 0 & \text{if } i = j \\ y_{i-1} & \text{if } i > j \end{cases} \end{aligned}$$

for all $\mathbf{w} \in I_{k-1}$ and $\mathbf{y} \in \mathbb{R}^{k-1}$. Therefore,

$$i_{k,j,s}^* dx_i = \begin{cases} dy_i & \text{if } i < j \\ 0 & \text{if } i = j \\ dy_{i-1} & \text{if } i > j \end{cases}$$

A more computational but less intuitive approach to get the previous relation would have been to use (1) of Proposition 4.5.4. We will use the computational approach in the future but it is wise to also keep in mind the intuitive approach above.

As a consequence of this, we have

$$\begin{aligned} i_{k,j,s}^* (dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_k) &= i_{k,j,s}^* dx_1 \wedge \dots \wedge \widehat{i_{k,j,s}^* dx_i} \wedge \dots \wedge i_{k,j,s}^* dx_k \\ &= \begin{cases} 0 & \text{if } i \neq j \\ dy_1 \wedge dy_2 \wedge \dots \wedge dy_{k-1} & \text{if } i = j \end{cases} \end{aligned}$$

where, as we stated before, the hat over dx_i for $1 \leq i \leq k$ means that this differential 1-form has been removed from the wedge product.

Theorem 4.7.5 (Stokes)

Suppose that ω is a differential $(k-1)$ -form on an open set $V \subset \mathbb{R}^n$ and that σ is an k -chain in V with $k \in \mathbb{N}^+$. Then

$$\int_{\sigma} d\omega = \int_{\partial\sigma} \omega .$$

Proof.

i) We first prove the theorem for $\sigma = i_k$ and ω a differential $(k-1)$ -form on $\mathbb{R}^k$. We have that ω is the sum of terms of the form

$$f dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_k , \quad (4.7.1)$$

By linearity of the integral, it suffices to prove the theorem for (4.7.1).

Since $\partial i_k = \sum_{j=1}^k \sum_{s=0}^1 (-1)^{j+s} i_{k,j,s}$ and

$$\begin{aligned} & \int_{I_{k-1}} i_{k,j,s}^* \left(f dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_k \right) \\ &= \begin{cases} 0 & \text{if } i \neq j \\ \int_{I_{k-1}} f \circ i_{k,j,s} dy_1 \wedge dy_2 \wedge \dots \wedge dy_{k-1} & \text{if } i = j \end{cases} \\ &= \begin{cases} 0 & \text{if } i \neq j \\ \int_{I_{k-1}} f(y_1, \dots, y_{i-1}, s, y_i, \dots, y_{k-1}) dy_1 dy_2 \dots dy_{k-1} & \text{if } i = j \end{cases} \end{aligned}$$

for $s = 0$ or $s = 1$, we get

$$\begin{aligned} \int_{\partial i_k} f dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_k &= \sum_{j=1}^k \sum_{s=0}^1 (-1)^{j+s} \int_{i_{k,j,s}} f dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_k \\ &= \sum_{j=1}^k \sum_{s=0}^1 (-1)^{j+s} \int_{I_{k-1}} i_{k,j,s}^* \left(f dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_k \right) \\ &= \sum_{s=0}^1 (-1)^{i+s} \int_{I_{k-1}} f(y_1, \dots, y_{i-1}, s, y_i, \dots, y_{k-1}) dy_1 dy_2 \dots dy_{k-1} . \end{aligned} \quad (4.7.2)$$

Moreover

$$\begin{aligned} \int_{i_k} d \left(f dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_k \right) &= \int_{i_k} df \wedge dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_k \\ &= \int_{i_k} \sum_{j=1}^k \frac{\partial f}{\partial x_j} dx_j \wedge dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_k \end{aligned}$$

$$= (-1)^{i-1} \int_{i_k} \frac{\partial f}{\partial x_i} dx_1 \wedge dx_2 \wedge \dots \wedge dx_k = (-1)^{i-1} \int_{I_k} \frac{\partial f}{\partial x_i}(\mathbf{x}) d\mathbf{x} .$$

It follows from Fubini's theorem and from the classical fundamental theorem of calculus in one dimension that

$$\begin{aligned} & \int_{i_k} d \left(f dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_k \right) \\ &= (-1)^{i-1} \int_0^1 \dots \int_0^1 \left(\int_0^1 \frac{\partial f}{\partial x_i}(x_1, x_2, \dots, x_k) dx_i \right) dx_1 \dots \widehat{dx_i} \dots dx_k \\ &= (-1)^{i-1} \int_0^1 \dots \int_0^1 f(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_k) \Big|_{x_i=0}^{x_i=1} dx_1 \dots \widehat{dx_i} \dots dx_k \\ &= \sum_{s=0}^1 (-1)^{i+s} \int_{I_{k-1}} f(y_1, \dots, y_{i-1}, s, y_i, \dots, y_{k-1}) dy_1 dy_2 \dots dy_{k-1} . \end{aligned} \quad (4.7.3)$$

If we combine (4.7.2) and (4.7.3), then we get

$$\int_{\partial i_k} f dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_k = \int_{i_k} d \left(f dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_k \right) .$$

Hence,

$$\int_{\partial i_k} \omega = \int_{i_k} d\omega$$

for all differential $(k-1)$ -forms ω on $\mathbb{R}^k$.

ii) Suppose that σ is a general singular k -cube in an open set $V \subset \mathbb{R}^n$ and that ω is a differential $(k-1)$ -form on V . Since

$$\partial\sigma = \sum_{j=1}^k \sum_{s=0}^1 (-1)^{i+s} \sigma_{j,s} = \sum_{j=1}^k \sum_{s=0}^1 (-1)^{i+s} \sigma \circ i_{k,j,s} ,$$

we have

$$\begin{aligned} \int_{\partial\sigma} \omega &= \sum_{j=1}^k \sum_{s=0}^1 (-1)^{i+s} \int_{\sigma_{j,s}} \omega = \sum_{j=1}^k \sum_{s=0}^1 (-1)^{i+s} \int_{I_{k-1}} \sigma_{j,s}^*(\omega) \\ &= \sum_{j=1}^k \sum_{s=0}^1 (-1)^{i+s} \int_{I_{k-1}} (\sigma \circ i_{k,j,s})^*(\omega) = \sum_{j=1}^k \sum_{s=0}^1 (-1)^{i+s} \int_{I_{k-1}} i_{k,j,s}^*(\sigma^*(\omega)) \\ &= \sum_{j=1}^k \sum_{s=0}^1 (-1)^{i+s} \int_{i_{k,j,s}} \sigma^*(\omega) = \int_{\partial i_k} \sigma^*(\omega) . \end{aligned}$$

It follows from (i) applied to $\sigma^*(\omega)$ and (4) of Proposition 4.5.4 that

$$\int_{\sigma} d\omega = \int_{I_k} \sigma^*(d\omega) = \int_{I_k} d(\sigma^*(\omega)) = \int_{i_k} d(\sigma^*(\omega)) = \int_{\partial i_k} \sigma^*(\omega) = \int_{\partial\sigma} \omega .$$

iii) Finally, if $\sigma = \sum_{j=1}^J a_j \sigma_j$ is a k -chain in an open set $V \subset \mathbb{R}^n$ and ω is a differential $(k-1)$ -form in V , then

$$\int_{\mu} d\omega = \sum_{j=1}^J a_j \int_{\sigma_j} d\omega = \sum_{j=1}^J a_j \int_{\partial\sigma_j} \omega = \int_{\partial\sigma} \omega . \quad \blacksquare$$

Chapter 5

Manifolds

Intuitively, if $V \subset \mathbb{R}^3$ is a small enough open neighbourhood of a point $\mathbf{x}$ on a nice surface S in $\mathbb{R}^3$, then the set $V \cap S$ looks like a piece of $\mathbb{R}^2$. How can we give a meaningful sense to this last sentence? The answer to this question comes from the introduction of a local system of coordinates at each point of S .

5.1 Preliminaries

Before defining what is a local system of coordinates, we need to define what is an open set in the half space

$$H_k = \{\mathbf{u} \in \mathbb{R}^k : u_k \geq 0\}$$

and what is a differentiable function $f : H_k \rightarrow \mathbb{R}^n$.

Definition 5.1.1

Let S be a subset of $\mathbb{R}^n$. An open set in S is a set of the form $U = S \cap V$ where V is an open set in $\mathbb{R}^n$. The collections of all the open sets on S defines a topology on S called the **induced topology** on S by $\mathbb{R}^n$.

We equip H_k with the induced topology from $\mathbb{R}^k$. It follows from the previous definition that a set $U \subset H_k$ is an open subset of H_k if there exists an open set $V \subset \mathbb{R}^k$ such that $U = H_k \cap V$.

Definition 5.1.2

The points $\mathbf{x} \in H_k$ such that $x_k > 0$ are the **interior points** of H_k and those with $x_k = 0$ are the **boundary points** of H_k . The set of all interior points of H_k is the **interior** of H_k and is denoted H_k° . The set of all boundary points of H_k is the **boundary** of H_k and is denoted ∂H_k .

The interior H_k° is open in the induced topology because $H_k^\circ = \{\mathbf{x} \in \mathbb{R}^k : x_k > 0\} \cap H_k$. The definition of interior and boundary points of H_k correspond to the topological definition

of interior and boundary points. This correspondence will be broken for manifolds as we will see shortly.

We need to extend our definition of differentiable functions to functions that may not be defined on an open subset of $\mathbb{R}^n$.

Definition 5.1.3

Suppose that U is an open subset of H_k . A function $f : U \rightarrow \mathbb{R}^m$ is **continuously differentiable** at $\mathbf{x} \in U$ if there exist an open neighbourhood $V_{\mathbf{x}}$ of $\mathbf{x}$ in $\mathbb{R}^k$ and a continuously differentiable function $F : V_{\mathbf{x}} \rightarrow \mathbb{R}^m$ such that $F = f$ on $V_{\mathbf{x}} \cap U$. The function F is an **extension** of f to an open neighbourhood of $\mathbf{x}$ in $\mathbb{R}^k$. We define $Df(\mathbf{x})$ as $DF(\mathbf{x})$. We say that $f : U \rightarrow \mathbb{R}^m$ is **continuously differentiable** on U if f is continuously differentiable at every $\mathbf{x} \in U$.

At the interior points of H_k , our previous definition of open neighbourhood and differentiable functions correspond to the classical definitions.

Proposition 5.1.4

If U is an open subset of H_k and $f : U \rightarrow \mathbb{R}^m$ is continuously differentiable at $\mathbf{x} \in U$, then the definition of $Df(\mathbf{x})$ is independent of the extension chosen.

Proof.

For $i = 1, 2$, suppose that $V_{\mathbf{x}}^{[i]}$ is an open neighbourhood of $\mathbf{x}$ in $\mathbb{R}^k$ and that $F^{[i]} : V_{\mathbf{x}}^{[i]} \rightarrow \mathbb{R}^m$ is a continuously differentiable function such that $F^{[i]} = f$ on $V_{\mathbf{x}}^{[i]} \cap H_k$. Let $V = V_{\mathbf{x}}^{[1]} \cap V_{\mathbf{x}}^{[2]} \cap U \cap H_k^\circ$.

If $\mathbf{x} \in H_k^\circ$, then V is an open neighbourhood of $\mathbf{x}$ in $\mathbb{R}^k$. We have that $F^{[1]} = F^{[2]} = f$ in V . Since the derivative at a point is a local property, we get $DF^{[1]}(\mathbf{x}) = DF^{[2]}(\mathbf{x}) = Df(\mathbf{x})$, where Df denotes the standard derivative of f .

Suppose that $\mathbf{x} \in U \cap \partial H_k$. Choose $\{\mathbf{x}_i\}_{i=1}^\infty \subset V$ such that $\{\mathbf{x}_i\}_{i=1}^\infty$ converges to $\mathbf{x}$. Since $F^{[1]} = F^{[2]} = f$ in V , we have as in the previous paragraph that $DF^{[1]}(\mathbf{x}_i) = DF^{[2]}(\mathbf{x}_i)$ for all i . Thus $DF^{[1]}(\mathbf{x}) = \lim_{i \rightarrow \infty} DF^{[1]}(\mathbf{x}_i) = \lim_{i \rightarrow \infty} DF^{[2]}(\mathbf{x}_i) = DF^{[2]}(\mathbf{x})$ by continuity of $DF^{[1]}$ and $DF^{[2]}$ on V . ■

Proposition 5.1.5

Let U , V and W be open subsets of H_k . Suppose that $f : U \rightarrow V$ and $g : V \rightarrow W$ are two differentiable functions, then $D(g \circ f)(\mathbf{x}) = Dg(f(\mathbf{x})) Df(\mathbf{x})$ for all $\mathbf{x} \in U$.

Proof.

Given $\mathbf{x} \in U$, suppose that U_1 is an open neighbourhood of $\mathbf{x}$ in $\mathbb{R}^k$ and that $F : U_1 \rightarrow \mathbb{R}^k$ is a continuously differentiable function such that $F = f$ on $U_1 \cap U$. Similarly, suppose that V_1 is an open neighbourhood of $f(\mathbf{x})$ in $\mathbb{R}^k$ and that $G : V_1 \rightarrow \mathbb{R}^k$ is a continuously differentiable function such that $G = g$ on $V_1 \cap V$.

Since F is continuous, $U_2 = F^{-1}(V_1) \cap U_1$ is an open neighbourhood of $\mathbf{x}$ in $\mathbb{R}^k$. Hence $G \circ F : U_2 \rightarrow \mathbb{R}^k$ is a continuously differentiable function such that $G \circ F = g \circ f$ on $U_2 \cap U$.

We have that $D(G \circ F)(\mathbf{x}) = D G(F(\mathbf{x})) D F(\mathbf{x})$. Since $D f(\mathbf{x})$, $D g(f(\mathbf{x}))$ and $D(g \circ f)(\mathbf{x})$ are defined by $D F(\mathbf{x})$, $D G(F(\mathbf{x}))$ and $D(G \circ f)(\mathbf{x})$ respectively, we get that $D(g \circ f)(\mathbf{x}) = D g(f(\mathbf{x})) D f(\mathbf{x})$. ■

The next definition is well known but it is worth repeating it here since it is going to play a major role from now on.

Definition 5.1.6

Suppose that V and W are two open subsets of H_k . The function $h : V \rightarrow W$ is a **diffeomorphism** if h is a homeomorphism, and h and h^{-1} are continuously differentiable.

5.2 Local Charts, Atlases and Manifolds

Definition 5.2.1

Let S be a subset of $\mathbb{R}^n$. We assume that the topology on S is the induced topology from $\mathbb{R}^n$. A **local chart** is defined by three objects.

1. An open subset U of S .
2. An open subset W of H_k for some integer k such that $0 \leq k \leq n$.
3. A homeomorphism $\phi : W \rightarrow U$.

This local chart is denoted (W, U, ϕ) (Figures 5.1 and 5.2).

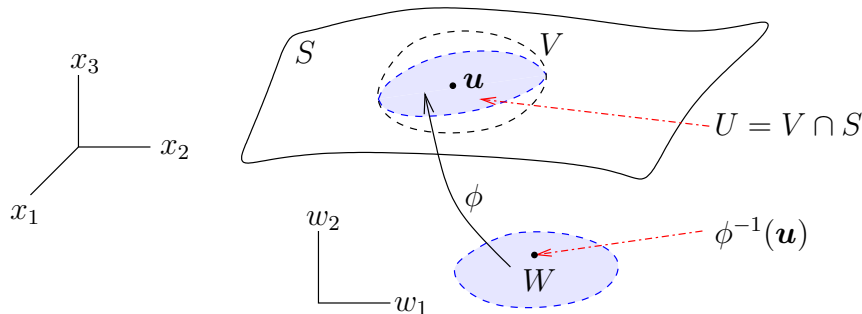


Figure 5.1: (W, U, ϕ) is a local chart about an interior point of the manifold S .

Remark 5.2.2

For $k = 0$, we define H_k and $\mathbb{R}^k$ as $\{0\} \subset \mathbb{R}$. Hence, a local chart is given by $W = \{0\}$ and

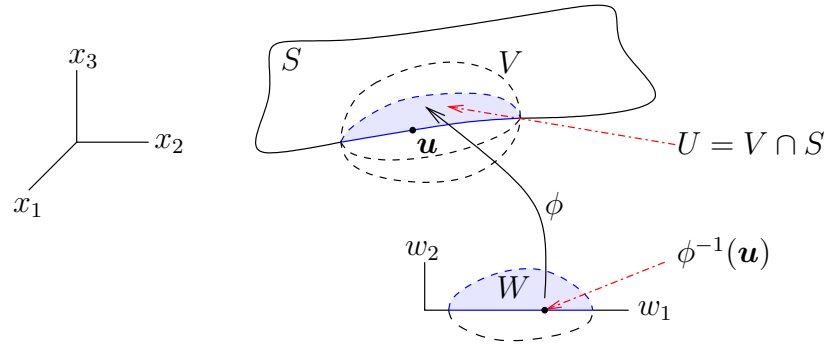


Figure 5.2: (W, U, ϕ) is a local chart about a boundary point of the manifold S .

the function $\phi : W \rightarrow U$ maps 0 to a single point $\phi(0) \in S$. The conditions for continuity and differentiability are meaningless in this case and therefore they are ignored. This special notation for $k = 0$ will be useful later to simplify the statements of some results. ♠

Definition 5.2.3

An **atlas** of class C^j , denoted $\mathcal{A}(k, j)$, for a set $S \subset \mathbb{R}^n$ is a collection of local charts $\{(W_i, U_i, \phi_i)\}_{i \in I}$, where I is an index set, such that the following conditions are satisfied.

1. $S = \bigcup_{i \in I} U_i$.
2. $W_i \subset H_k$ for all $i \in I$.
3. If $(W_{i_1}, U_{i_1}, \phi_{i_1})$ and $(W_{i_2}, U_{i_2}, \phi_{i_2})$ are two local charts such that $U_{i_1} \cap U_{i_2} \neq \emptyset$, then

$$\phi_{i_2}^{-1} \circ \phi_{i_1} : M_1 = \phi_{i_1}^{-1}(U_{i_1} \cap U_{i_2}) \rightarrow M_2 = \phi_{i_2}^{-1}(U_{i_1} \cap U_{i_2}) \quad (5.2.1)$$

is a C^j -diffeomorphism (Figure 5.3).

For $j > 0$, the condition that $\phi_{i_2}^{-1} \circ \phi_{i_1} : M_1 \rightarrow M_2$ be a C^j -diffeomorphism for all indices $i_1, i_2 \in I$ implies that $D\phi_i(\mathbf{w})$ is of rank k for all $\mathbf{w} \in W_i$ and all $i \in I$. We will come back on this subject in Subsection 5.2.1.

Note that the set M_1 and M_2 in the previous definition are open subsets of H_k because the functions ϕ_i are homeomorphisms.

In the previous definition, we could have considered only homeomorphisms $\phi_{i_2}^{-1} \circ \phi_{i_1}$ instead of C^j -diffeomorphism. However, since our goal is integration on manifold, we will need differentiable local charts anyway. In most of our examples, the local charts will even be of class C^j with $j > 1$ or $j = \infty$.

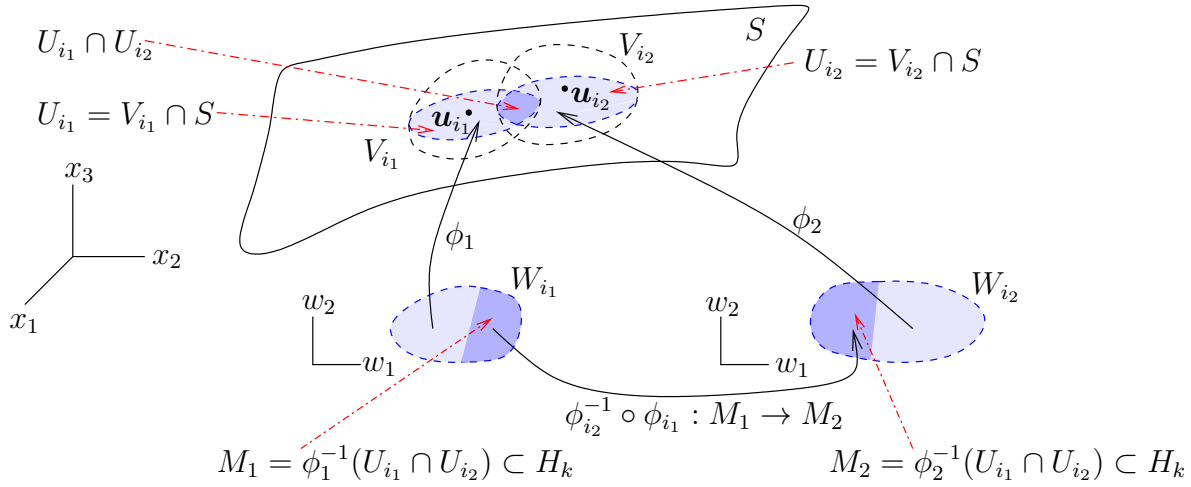


Figure 5.3: The map $\phi_2^{-1} \circ \phi_1$ obtained from two local charts provides a diffeomorphism from $M_1 = \phi_1^{-1}(U_{i_1} \cap U_{i_2})$ to $M_2 = \phi_2^{-1}(U_{i_1} \cap U_{i_2})$.

Definition 5.2.4

Two atlases $\mathcal{A}_1(k, j)$ and $\mathcal{A}_2(k, j)$ for a set $S \subset \mathbb{R}^n$ are **equivalent** if $\mathcal{A}_1(k, j) \cup \mathcal{A}_2(k, j)$ yields an atlas $\mathcal{A}_3(k, j)$ for S . For k and j fixed, a **differential structure** $\mathcal{S}(k, j)$ for a set S is an equivalent class of atlases $\mathcal{A}(k, j)$ for S . The union of all the atlases of an equivalent class of atlases for S forms a **maximal atlas** for S . An **admissible local chart** is a local chart in the chosen maximal atlas.

Definition 5.2.5

A **k-dimensional manifold** of class C^j is a subset S of $\mathbb{R}^n$ with a differential structure $\mathcal{S}(k, j)$ for S .

From now on, when we talk about a local chart of the manifold S , we mean an admissible local chart from the chosen maximal atlas for S . Moreover, a local chart (W, U, ϕ) about a point $\mathbf{u}$ of a manifold S is simply a local chart (W, U, ϕ) with $\mathbf{u} \in U$.

To determine if a set $S \subset \mathbb{R}^n$ is a manifold, it suffices to find one atlas for S .

Definition 5.2.6

Let S be a manifold. The point $\mathbf{u} \in S$ is an **interior point** of S if there exists a local chart (W, U, ϕ) about $\mathbf{u}$ such that $W \subset H_k^\circ$. The set of all interior points of S is the **interior** of S . The interior of the manifold S is denoted $\text{Int } S$. The point $\mathbf{u} \in S$ is a **boundary point** of S if there exists a local chart (W, U, ϕ) about $\mathbf{u}$ such that $\phi^{-1}(\mathbf{u}) \in \partial H_k$. The set of all boundary points of S is the **boundary** of S . The boundary of the manifold S is denoted ∂S . If $\partial S = \emptyset$, then we say that we have a

manifold without boundary. If $\partial S \neq \emptyset$, then we say that we have a **manifold with boundary**.

The interior and boundary for a manifold differ from the definition of the interior and boundary given in topology. The context will determine the meaning given to interior and boundary.

Remark 5.2.7

For the manifold without boundary, we may obviously replace H_k° in the definition of a manifold by $\mathbb{R}^k$ without any consequence. This is often very convenient when defining local charts to simplify the definition of the maps ϕ_i . ♠

Example 5.2.8

Let S be the circle of radius 2 centred at the origin in $\mathbb{R}^2$. We have that S is a 1-dimensional manifold. A possible atlas for S is given by the local charts:

$$\begin{aligned} & \left(] - 2, 2[, \{(x_1, x_2) \in S : x_2 > 0\}, \phi_1(w) = (w - 2, \sqrt{4 - w^2}) \right), \\ & \left(] - 2, 2[, \{(x_1, x_2) \in S : x_2 < 0\}, \phi_2(w) = (w - 2, -\sqrt{4 - w^2}) \right), \\ & \left(] - 2, 2[, \{(x_1, x_2) \in S : x_1 > 0\}, \phi_3(w) = (\sqrt{4 - w^2}, w - 2) \right) \end{aligned}$$

and

$$\left(\left(] - 2, 2[, \{(x_1, x_2) \in S : x_1 < 0\}, \phi_4(w) = (-\sqrt{4 - w^2}, w - 2) \right) \right).$$

We have that $] - 2, 2[\subset \mathbb{R}$.

The open upper-hemisphere $S_o = \{\mathbf{x} \in \mathbb{R}^3 : \|\mathbf{x}\| = 2 \text{ and } x_3 > 0\}$ is a 2-dimension manifold. A possible atlas for S_o is given by the single local chart

$$\left(\{(w_1, w_2) : w_1^2 + w_2^2 < 4\}, S_o, \phi(\mathbf{w}) = (w_1, w_2 - 2, \sqrt{4 - w_1^2 - w_2^2}) \right).$$

We have that $\{(w_1, w_2) : w_1^2 + w_2^2 < 4\} \subset \mathbb{R}^2$.

Both S and S_o are manifolds of class C^∞ without boundary.

The closed upper-hemisphere $S_c = \{\mathbf{x} \in \mathbb{R}^3 : \|\mathbf{x}\| = 2 \text{ and } x_3 \geq 0\}$ is a 2-dimension manifold with boundary. Let $B = \{(w_1, w_2) : w_1^2 + (w_2 - 2)^2 < 4\} \subset H_2^\circ$ and $B_o = \{(w_1, w_2) : w_1^2 + w_2^2 < 4, w_2 \geq 0\} \subset H_2$. A possible atlas for S_c is given by the local charts:

$$\begin{aligned} & \left(B, \{(x_1, x_2, x_3) \in S_c : x_3 > 0\}, \phi_1(\mathbf{w}) = (w_1, w_2 - 2, \sqrt{4 - w_1^2 - (w_2 - 2)^2}) \right), \\ & \left(B_o, \{(x_1, x_2, x_3) \in S_c : x_2 > 0\}, \phi_2(\mathbf{w}) = (w_1, \sqrt{4 - w_1^2 - w_2^2}, w_2) \right), \\ & \left(B_o, \{(x_1, x_2, x_3) \in S_c : x_2 < 0\}, \phi_3(\mathbf{w}) = (w_1, -\sqrt{4 - w_1^2 - w_2^2}, w_2) \right), \end{aligned}$$

$$\left(B_o, \{(x_1, x_2, x_3) \in S_c : x_1 > 0\}, \phi_4(\mathbf{w}) = (\sqrt{4 - w_1^2 - w_2^2}, w_1, w_2) \right)$$

and

$$\left(B_o, \{(x_1, x_2, x_3) \in S_c : x_1 < 0\}, \phi_5(\mathbf{w}) = (-\sqrt{4 - w_1^2 - w_2^2}, w_1, w_2) \right),$$

where $B = \{(w_1, w_2) : w_1^2 + (w_2 - 2)^2 < 4\} \subset H_2^\circ$ and $B_o = \{(w_1, w_2) : w_1^2 + w_2^2 < 4, w_2 \geq 0\} \subset H_2$.

Some readers may have thought that a possible atlas for S_c could be given by the single local chart

$$\left(\{(w_1, w_2) : w_1^2 + (w_2 - 2)^2 \leq 4\}, S_c, \phi(\mathbf{w}) = (w_1, w_2 - 2, \sqrt{4 - w_1^2 - (w_2 - 2)^2}) \right).$$

However, this is not a local chart according to our definition. The set $\{(w_1, w_2) : w_1^2 + (w_2 - 2)^2 \leq 4\}$ is not an open subset of H_2 . Moreover, ϕ is not differentiable at (w_1, w_2) with $w_1^2 + (w_2 - 2)^2 = 4$; the map ϕ cannot be extended to a differentiable function in an open neighbourhood of these points.

The points on the circle $S_1 = \{\mathbf{x} \in \mathbb{R}^3 : \|\mathbf{x}\| = 2 \text{ and } x_3 = 0\}$ form the boundary of S_c . The set S_1 is a 1-dimensional manifold without boundary. ♣

If $\mathbf{u}$ is a boundary point of a manifold S according to one local chart (W_1, U_1, ϕ_1) with $\mathbf{u} \in U_1$, is $\mathbf{u}$ a boundary point according to another equivalent local chart (W_2, U_2, ϕ_2) with $\mathbf{u} \in U_2$? The following propositions will be used to answer this question affirmatively. Therefore, our definition of boundary for a manifold is consistent.

Proposition 5.2.9

Suppose that W is an open subset of H_k and $\phi : W \rightarrow H_k$ is a differentiable map. Moreover, suppose that $\phi(\tilde{\mathbf{w}}) \in \partial H_k$ for some $\tilde{\mathbf{w}} \in W \cap H_k^\circ$. Then $D\phi(\tilde{\mathbf{w}})(\mathbb{R}^k) \subset \partial H_k$.

Proof.

Let $\pi_k : \mathbb{R}^k \rightarrow \mathbb{R}$ be the canonical projection on the k^{th} coordinate. Choose $\mathbf{x} \in \mathbb{R}^k$. Since ϕ is differentiable at $\tilde{\mathbf{w}} \in W \cap H_k^\circ$, we have that $\phi(\tilde{\mathbf{w}} + t\mathbf{x}) = \phi(\tilde{\mathbf{w}}) + D\phi(\tilde{\mathbf{w}})(t\mathbf{x}) + g(t\mathbf{x})$ for $|t|$ small enough where $g : \mathbb{R}^k \rightarrow \mathbb{R}^k$ a function such that $\lim_{\mathbf{y} \rightarrow \mathbf{0}} \frac{\|g(\mathbf{y})\|}{\|\mathbf{y}\|} = 0$.

Since $\phi(\tilde{\mathbf{w}}) \in \partial H_0$ and $\pi_k(\phi(\mathbf{w})) \geq 0$ for all $\mathbf{w} \in \mathbb{R}^k$ by hypothesis, we have

$$\begin{aligned} 0 &\leq \pi_k(\phi(\tilde{\mathbf{w}} + t\mathbf{x})) = \pi_k(\phi(\tilde{\mathbf{w}})) + (\pi_k \circ D\phi(\tilde{\mathbf{w}}))(t\mathbf{x}) + (\pi_k \circ g)(t\mathbf{x}) \\ &= (\pi_k \circ D\phi(\tilde{\mathbf{w}}))(t\mathbf{x}) + (\pi_k \circ g)(t\mathbf{x}). \end{aligned} \tag{5.2.2}$$

Hence, for $t > 0$ small enough, we get from (5.2.2) that

$$0 \leq (\pi_k \circ D\phi(\tilde{\mathbf{w}}))(\mathbf{x}) + \frac{1}{t} (\pi_k \circ g)(t\mathbf{x}).$$

If we let t converges to 0, then we get $0 \leq (\pi_k \circ D\phi(\tilde{\mathbf{w}}))(\mathbf{x})$ because

$$\left| \frac{1}{t} (\pi_k \circ g)(t\mathbf{x}) \right| \leq \|\mathbf{x}\| \frac{\|g(t\mathbf{x})\|}{\|t\mathbf{x}\|} \rightarrow 0$$

as $t \rightarrow 0$. Similarly, for $t < 0$ closed to 0, we get from (5.2.2) that

$$0 \geq (\pi_k \circ D\phi(\tilde{\mathbf{w}}))(\mathbf{x}) + \frac{1}{t} (\pi_k \circ g)(t\mathbf{x}) .$$

Again, if we let t converges to 0, then we get $0 \geq (\pi_k \circ D\phi(\tilde{\mathbf{w}}))(\mathbf{x})$.

Hence $(\pi_k \circ D\phi(\tilde{\mathbf{w}}))(\mathbf{x}) = 0$. Since $\mathbf{x} \in \mathbb{R}^k$ is arbitrary, we get the conclusion of the proposition. ■

Proposition 5.2.10

Suppose that W_1 and W_2 are open subsets of H_k and $f : W_1 \rightarrow W_2$ is a diffeomorphism. Then $f(W_1 \cap H_k^\circ) \subset W_2 \cap H_k^\circ$ and $f(W_1 \cap \partial H_k) \subset W_2 \cap \partial H_k$. Moreover, $f|_{W_1 \cap H_k^\circ} : W_1 \cap H_k^\circ \rightarrow W_2 \cap H_k^\circ$ and $f|_{W_1 \cap \partial H_k} : W_1 \cap \partial H_k \rightarrow W_2 \cap \partial H_k$ are diffeomorphisms.

Proof.

i) We consider the case with $W_1 \subset H_k^\circ$. We prove that $W_2 \subset H_k^\circ$. Hence $f|_{W_1 \cap H_k^\circ} : W_1 \cap H_k^\circ \rightarrow W_2 \cap H_k^\circ$ will be a diffeomorphism because $W_1 \cap H_k^\circ = W_1$ and $W_2 \cap H_k^\circ = W_2$ are open subsets of $\mathbb{R}^k$, and $f : W_1 \rightarrow W_2$ is then a diffeomorphism in the traditional sense.

Suppose that $W_2 \cap \partial H_k \neq \emptyset$. Since f is onto W_2 , there exists $\tilde{\mathbf{w}} \in W_1$ such that $f(\tilde{\mathbf{w}}) \in W_2 \cap \partial H_k$. It follows from Proposition 5.1.5 that

$$Df(\tilde{\mathbf{w}}) Df^{-1}(f(\tilde{\mathbf{w}})) = Df(f^{-1}(f(\tilde{\mathbf{w}}))) Df^{-1}(f(\tilde{\mathbf{w}})) = D(f \circ f^{-1})(f(\tilde{\mathbf{w}})) = \text{Id}$$

and

$$Df^{-1}(f(\tilde{\mathbf{w}})) Df(\tilde{\mathbf{w}}) = D(f^{-1} \circ f)(\tilde{\mathbf{w}}) = \text{Id} .$$

Thus $Df(\tilde{\mathbf{w}})$ is invertible; in particular, $Df(\tilde{\mathbf{w}})$ maps $\mathbb{R}^k$ onto $\mathbb{R}^k$. This result contradicts the conclusion of Proposition 5.2.9 which says that $Df(\tilde{\mathbf{w}})(\mathbb{R}^k) \subset \partial H_k$. Thus $W_2 \cap \partial H_k = \emptyset$.

ii) We consider the case with $W_1 \cap \partial H_k \neq \emptyset$. We must have that $W_2 \cap \partial H_k \neq \emptyset$ because (i) applied to f^{-1} instead of f would then yield $W_1 \cap \partial H_k = \emptyset$.

Given $\mathbf{w} \in W_1 \cap H_k^\circ$, there exists an open neighbourhood $V_1 \subset W_1 \cap H_k^\circ$ of $\mathbf{w}$. If we apply (i) with W_1 replaced by V_1 , then we get that $f(V_1) \subset H_k^\circ$. Thus $f(V_1)$ is an open neighbourhood of $f(\mathbf{w})$ in $W_2 \cap H_k^\circ$. Since $\mathbf{w}$ is arbitrary, this shows that $f(W_1 \cap H_k^\circ) \subset W_2 \cap H_k^\circ$. The same reasoning applied to f^{-1} shows that $f^{-1}(W_2 \cap H_k^\circ) \subset W_1 \cap H_k^\circ$. Thus $f(W_1 \cap H_k^\circ) = W_2 \cap H_k^\circ$ and $f|_{W_1 \cap H_k^\circ} : W_1 \cap H_k^\circ \rightarrow W_2 \cap H_k^\circ$ is a diffeomorphism.

Since $f : W_1 \rightarrow W_2$ is one to one and onto, we then have that $f(W_1 \cap \partial H_k) = W_2 \cap \partial H_k$ and $f|_{W_1 \cap \partial H_k} : W_1 \cap \partial H_k \rightarrow W_2 \cap \partial H_k$ is a diffeomorphism. Note that $W_i \cap \partial H_k$ for $1 \leq i \leq 2$ are open subsets of $\partial H_k \cong \mathbb{R}^{k-1}$. So $f|_{W_1 \cap \partial H_k} : W_1 \cap \partial H_k \rightarrow W_2 \cap \partial H_k$ is a standard diffeomorphism between open sets of $\mathbb{R}^{k-1}$. ■

Theorem 5.2.11

Suppose that S is a k -dimensional manifold. Then $\text{Int } S$ is a k -dimensional manifold without boundary and ∂S is a $(k - 1)$ -dimensional manifold without boundary.

Proof.

To prove this theorem, it suffices to apply the results of Proposition 5.2.10 to the compositions $\phi_i^{-1} \circ \phi_j$, where (W_i, U_i, ϕ_i) and (W_j, U_j, ϕ_j) are two local charts of S such that $U_i \cap U_j \neq \emptyset$.

If $\{(W_i, U_i, \phi_i)\}_{i \in I}$ is a maximal atlas for S , then

$$\left\{ \left(W_i \cap H_k^\circ, U_i \cap \text{Int } S, \phi_i|_{W_i \cap H_k^\circ} \right) : i \in I \right\}$$

is an atlas for $\text{Int } S$. An atlas for ∂S is given by

$$\left\{ \left(W_i \cap \partial H_k, U_i \cap \partial S, \phi_i|_{W_i \cap \partial H_k} \right) : i \in I \text{ and } U_i \cap \partial H_k \neq \emptyset \right\}.$$

Recall that $W_i \cap \partial H_k$ is an open subset of $\{\mathbf{x} \in \mathbb{R}^k : x_k = 0\} \cong \mathbb{R}^{k-1}$ for all i . So ∂S is a manifold without boundary of dimension $k - 1$. ■

The boundary of a manifold S is in fact a “submanifold” of S .

Definition 5.2.12

Let S be a k -dimensional manifold. A **submanifold** of S is a subset M of S such that there exist a positive integer $q < k$, indices $1 \leq k_1 < k_2 < \dots < k_q \leq k$, and an atlas $\{(W_i, U_i, \phi_i)\}_{i \in I}$ for S such that $U_i \cap M = \phi_i(W \cap H_k^{[k_1, k_2, \dots, k_q]})$ for all $i \in I$, where $H_k^{[k_1, k_2, \dots, k_q]} = \{\mathbf{w} \in H_k : w_{k_i} = 0 \text{ for } 1 \leq i \leq q\}$.

If we set $\tilde{W}_i = W \cap H_k^{[k_1, k_2, \dots, k_q]}$ and $\tilde{U}_i = U_i \cap M$ for all $i \in I$, then $\{(\tilde{W}_i, \tilde{U}_i, \phi_i)\}_{i \in I}$ is an atlas for M . Hence M is itself a q -dimensional manifold.

5.2.1 Manifolds in $\mathbb{R}^n$

The next theorem gives another method to describe manifolds in $\mathbb{R}^n$.

Theorem 5.2.13

Let S be a subset of $\mathbb{R}^n$ and $\mathbf{u} \in S$. The following two statements are equivalent.

1. There exist an open neighbourhood $V_1 \subset \mathbb{R}^n$ of $\mathbf{u}$, an open set $V_2 \subset \mathbb{R}^n$, and a diffeomorphism $h : V_1 \rightarrow V_2$ such that $h(V_1 \cap S) = \{\mathbf{v} \in V_2 : v_{k+1} = v_{k+2} = \dots = v_n = 0\}$ (Figure 5.4).
2. There exist an open subset $W \subset \mathbb{R}^k$, an open neighbourhood $V \subset \mathbb{R}^n$ of $\mathbf{u}$ and a function $\phi : W \rightarrow V \cap S$ such that ϕ is a homeomorphism, ϕ is of class C^1 and

$D\phi(\mathbf{w})$ is of rank k for all $\mathbf{w} \in W$.

Moreover, (1) implies that $\phi^{-1} : V \cap S \rightarrow W$ in (2) can be extended to a continuously differentiable function in the neighbourhood V_1 of $\mathbf{u}$.

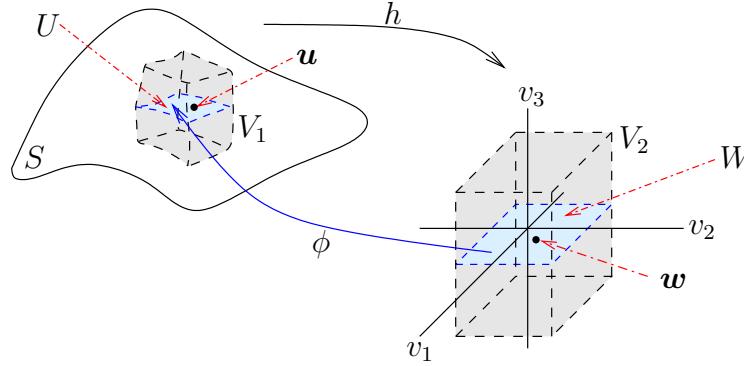


Figure 5.4: Local representation of a manifold S .

Proof.

i) Suppose that (1) is true. By definition of induced topology $U = V_1 \cap S$ is an open subset of S . Let

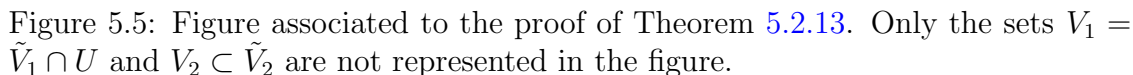
$$W = \left\{ \mathbf{w} \in \mathbb{R}^k : \begin{pmatrix} \mathbf{w} \\ \mathbf{0} \end{pmatrix} \in V_2 \right\} \cong V_2 \cap \left\{ \begin{pmatrix} \mathbf{w} \\ \mathbf{0} \end{pmatrix} \in \mathbb{R}^n : \mathbf{w} \in \mathbb{R}^k \right\}.$$

By definition of the induced topology, W is an open subset of $\mathbb{R}^k$. Let $\phi : W \rightarrow U$ be the function defined by $\phi(\mathbf{w}) = h^{-1}(\mathbf{w}, \mathbf{0})$ for $\mathbf{w} \in W$. We have that $\phi(W) = U$, $\phi = h^{-1}|_W$ and $\phi^{-1} = h|_U$ are continuous, and $\phi \in C^1(W)$ because $h^{-1} \in C^1(V_2)$ with $W \subset V_2$. Moreover, $\tilde{h} : V_1 \rightarrow \mathbb{R}^k$ defined by $\tilde{h}(\mathbf{x}) = (h_1(\mathbf{x}) \ h_2(\mathbf{x}) \ \dots \ h_k(\mathbf{x}))^\top$ for all $\mathbf{x} \in V_1$ is a continuously differentiable extension of ϕ^{-1} .

It follows from $\tilde{h} \circ \phi = \text{Id}$ on W that $D\tilde{h}(\phi(\mathbf{w})) D\phi(\mathbf{w}) = D(\tilde{h} \circ \phi)(\mathbf{w}) = \text{Id}_k$ for all $\mathbf{w} \in W$ where Id_k is the identity matrix on $\mathbb{R}^k$. Thus, the rank of $D\phi(\mathbf{w})$ is k .

ii) Suppose that (2) is true. The reader may want to refer to Figure 5.5 for an illustration of the proof. Choose $\mathbf{w} \in W$ such that $\phi(\mathbf{w}) = \mathbf{u}$. By changing the order of the coordinates in $\mathbb{R}^n$ if necessary, we may assume that the matrix

$$M = \begin{pmatrix} \frac{\partial \phi_1}{\partial w_1}(\mathbf{w}) & \frac{\partial \phi_1}{\partial w_2}(\mathbf{w}) & \dots & \frac{\partial \phi_1}{\partial w_k}(\mathbf{w}) \\ \frac{\partial \phi_2}{\partial w_1}(\mathbf{w}) & \frac{\partial \phi_2}{\partial w_2}(\mathbf{w}) & \dots & \frac{\partial \phi_2}{\partial w_k}(\mathbf{w}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \phi_k}{\partial w_1}(\mathbf{w}) & \frac{\partial \phi_k}{\partial w_2}(\mathbf{w}) & \dots & \frac{\partial \phi_k}{\partial w_k}(\mathbf{w}) \end{pmatrix}$$


$$D = \left\{ \mathbf{z} \in W : \begin{pmatrix} \mathbf{z} \\ \mathbf{0} \end{pmatrix} \in \tilde{V}_2 \right\} \cong \tilde{V}_2 \cap \left\{ \begin{pmatrix} \mathbf{z} \\ \mathbf{0} \end{pmatrix} : \mathbf{z} \in W \right\}$$

Let $V_1 = \tilde{V}_1 \cap U$, $V_2 = g^{-1}(V_1) \subset \tilde{V}_2$, and set $h = g^{-1}|_{V_1}$. Since

$$h(V_1 \cap S) = g^{-1}(V_1 \cap S) = g^{-1} \left(\left\{ g \left(\begin{pmatrix} z \\ \mathbf{0} \end{pmatrix} \right) : \begin{pmatrix} z \\ \mathbf{0} \end{pmatrix} \in V_2 \right\} \right) = V_2 \cap \left\{ \begin{pmatrix} z \\ \mathbf{0} \end{pmatrix} : z \in \mathbb{R}^k \right\} . \blacksquare$$

Condition (1) in the previous theorem is often used to define manifolds in $\mathbb{R}^n$ (e.g. [6, 26]).

The previous theorem provides local charts for a manifold without boundary. A slight modification of the previous proof (replace $\mathbb{R}^k$ by H_k) can also provide local charts for manifold with boundary.

Theorem 5.2.14

Let S be a subset of $\mathbb{R}^n$ and $\mathbf{u} \in S$. The following two statements are equivalent

1. There exist an open neighbourhood $V_1 \subset \mathbb{R}^n$ of $\mathbf{u}$, an open set $V_2 \subset \mathbb{R}^n$, and a diffeomorphism $h : V_1 \rightarrow V_2$ such that $h(V_1 \cap S) = \{\mathbf{w} \in V_2 : w_k \geq 0 \text{ and } w_{k+1} = w_{k+2} = \dots = w_n = 0\}$.
2. There exist an open subset $W \subset H_k$, an open neighbourhood $V \subset \mathbb{R}^n$ of $\mathbf{u}$ and a function $\phi : W \rightarrow V \cap S$ such that ϕ is a homeomorphism, ϕ is of class C^1 and $D\phi(\mathbf{w})$ is of rank k for all $\mathbf{w} \in W$.

Moreover, (1) implies that $\phi^{-1} : V \cap S \rightarrow W$ in (2) can be extended to a continuously differentiable function in the neighbourhood V_1 of $\mathbf{u}$.

A consequence of the Theorem 5.2.13 is the following result.

Theorem 5.2.15

If $V \subset \mathbb{R}^n$ is open and $g : V \rightarrow \mathbb{R}^{n-k}$ is a function of class C^1 such that $Dg(\mathbf{x})$ is of rank $n - k$ for all $\mathbf{x} \in g^{-1}(\{\mathbf{0}\})$, then $S = g^{-1}(\{\mathbf{0}\}) \subset V$ is a k -dimensional manifold.

Proof.

We may apply Theorem 1.2.1 to each $\mathbf{v} \in V$ such that $g(\mathbf{v}) = \mathbf{0}$. We get an open set $W_{\mathbf{v}} \subset \mathbb{R}^n$, an open neighbourhood $U_{\mathbf{v}} \subset V$ of $\mathbf{v}$, and a diffeomorphism $h : W_{\mathbf{v}} \rightarrow U_{\mathbf{v}}$ such that $g(h(\mathbf{x})) = (x_{n-k+1} \ x_{n-k+2} \ \dots \ x_n)^\top$ for $\mathbf{x} \in W_{\mathbf{v}}$.

Permuting the coordinates in V if necessary, we may assume that

$$P = \begin{pmatrix} \frac{\partial g_1}{\partial x_{k+1}}(\mathbf{v}) & \frac{\partial g_1}{\partial x_{k+2}}(\mathbf{v}) & \dots & \frac{\partial g_1}{\partial x_n}(\mathbf{v}) \\ \frac{\partial g_2}{\partial x_{k+1}}(\mathbf{v}) & \frac{\partial g_2}{\partial x_{k+2}}(\mathbf{v}) & \dots & \frac{\partial g_2}{\partial x_n}(\mathbf{v}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_{n-k}}{\partial x_{k+1}}(\mathbf{v}) & \frac{\partial g_{n-k}}{\partial x_{k+2}}(\mathbf{v}) & \dots & \frac{\partial g_{n-k}}{\partial x_n}(\mathbf{v}) \end{pmatrix}$$

has rank $n - k$ and thus is invertible. Consider the function $f : W_{\mathbf{v}} \rightarrow \mathbb{R}^n$ defined by $f(\mathbf{x}) = (x_1 \ x_2 \ \dots \ x_k \ g(h(\mathbf{x})))^\top$ for $\mathbf{x} \in W_{\mathbf{v}}$. If $\mathbf{u} = h^{-1}(\mathbf{v})$, then we have

$$Df(\mathbf{u}) = \begin{pmatrix} \text{Id}_{k \times k} & 0 \\ * & P \end{pmatrix} Dh(\mathbf{u})$$

where the $*$ stands for the $(n - k) \times k$ matrix of derivatives $\frac{\partial g_i}{\partial x_j}(\mathbf{v})$ with $1 \leq i \leq n - k$ and $1 \leq j \leq k$. Since $\det(Df(\mathbf{u})) = \det(P) \det(Dh(\mathbf{u})) \neq 0$, we have that $Df(\mathbf{u})$ is invertible.

Therefore, by shrinking $W_{\mathbf{v}}$ if necessary, we may assume from the Inverse Function Theorem that $f : W_{\mathbf{v}} \rightarrow V_{\mathbf{v}} = f(W_{\mathbf{v}})$ is a diffeomorphism.

The function f maps $h^{-1}(S) \cap W_{\mathbf{v}}$ onto $\{\mathbf{x} \in V_{\mathbf{v}} : x_{n-k+1} = x_{n-k+2} = \dots = x_n = 0\}$. It follows from Theorem 5.2.13 that there exists a local chart (W, U, ϕ) of $h^{-1}(S) \cap W_{\mathbf{v}}$ about $\mathbf{w} = h^{-1}(\mathbf{v})$. Thus $(W, h(U), h \circ \phi)$ is a local chart of S about $\mathbf{v}$.

$$\begin{array}{ccc}
 U_{\mathbf{v}} \cap S & \xrightarrow{g} & \{\mathbf{0}\} \subset \mathbb{R}^{n-k} \\
 \uparrow h & \nearrow g \circ h & \\
 U \subset W_{\mathbf{v}} \cap h^{-1}(S) & & \\
 \uparrow \phi & & \\
 W & &
 \end{array}$$

Since we can find a local chart about each point of S , a structure of manifold can be defined on S . ■

5.3 Tangent Spaces and Oriented Manifolds

Let $W \subset \mathbb{R}^k$ be an open set. Given a vector $\mathbf{w} \in W$, we have defined in Section 4.3 the tangent space of W at $\mathbf{w}$ to be the set $T_{\mathbf{w}}W = \{(\mathbf{w}, \mathbf{v}) : \mathbf{v} \in \mathbb{R}^k\}$. We extend our definition of a tangent space of an open subset of $\mathbb{R}^k$ to a tangent space of an open subset of H_k in a natural way.

Definition 5.3.1

Suppose that $W \subset H_k$ is an open subset and $\mathbf{w} \in W$. By definition of induced topology, there exists an open subset V of $\mathbb{R}^k$ such that $W = H_k \cap V$. Then $T_{\mathbf{w}}W$ is defined as $T_{\mathbf{w}}V$.

There are many equivalent ways to define the tangent space of a manifold S at a point $\mathbf{w} \in S$. The approach that we have chosen is more visual and closer to the definition of the tangent space of an open subset $W \subset H_k$ at a point $\mathbf{w} \in W$ given above.

Definition 5.3.2

Let $S \subset \mathbb{R}^n$ be a k -dimensional manifold of class C^1 and $\mathbf{u} \in S$. Suppose that (W, U, ϕ) is a local chart such that $\mathbf{u} \in U$ and $\mathbf{w} = \phi^{-1}(\mathbf{u})$. The **tangent space** $T_{\mathbf{u}}S$ of S at $\mathbf{u}$ is the subspace $T_{\mathbf{u}}S = \phi_*(T_{\mathbf{w}}W) \subset T_{\mathbf{u}}\mathbb{R}^n$.

Recall that $\phi_* : T_{\mathbf{w}}W \rightarrow T_{\mathbf{u}}\mathbb{R}^n$ is defined by $\phi_*(\mathbf{w}, \mathbf{v}) = (\phi(\mathbf{w}), D\phi(\mathbf{w})\mathbf{v})$ for $\mathbf{v} \in \mathbb{R}^k$. By Proposition 5.1.4, $D\phi(\mathbf{w})$ is well defined even if $\mathbf{w} \in \partial W$.

If S is a k -dimensional manifold with boundary of class C^1 , then we have that $T_{\mathbf{u}}\partial S$ is a $(k-1)$ dimensional subspace of $T_{\mathbf{u}}S$. In fact, if (W, U, ϕ) is a local chart with $\mathbf{u} \in U \cap \partial S$ and $\mathbf{w} = \phi(\mathbf{u}) \in W \cap \partial H_k$, then it follows from the proof of Theorem 5.2.11 that

$T_{\mathbf{u}}\partial S = \phi_*(T_{\mathbf{w}}(W \cap \partial H_k))$ where $T_{\mathbf{w}}(W \cap \partial H_k) \cong \{(\mathbf{w}, \mathbf{y}) : \mathbf{y} \in \mathbb{R}^k \text{ and } y_k = 0\} \cong \mathbb{R}^{k-1}$. In particular, $T_{\mathbf{u}}\partial S$ can be seen as a $(k-1)$ -dimensional subspace of $T_{\mathbf{u}}S$.

The proposition below shows that the definition of the tangent space to a k -dimensional manifold at a point is independent of the local chart used. But before stating this proposition, suppose that U and V are open subsets of H_k , and $f : U \rightarrow V$ and $g : V \rightarrow W \subset \mathbb{R}^k$ are two differentiable functions. It follows from Proposition 5.1.5 that

$$\begin{aligned} (g \circ f)_*(\mathbf{u}, \mathbf{x}) &= ((g \circ f)(\mathbf{u}), D(g \circ f)(\mathbf{u})\mathbf{x}) = (g(f(\mathbf{u})), Dg(f(\mathbf{u})) Df(\mathbf{u})\mathbf{x}) \\ &= g_*(f(\mathbf{u}), Df(\mathbf{u})\mathbf{x}) = g_*(f_*(\mathbf{u}, \mathbf{x})) = (g_* \circ f_*)(\mathbf{u}, \mathbf{x}) \end{aligned}$$

for all $(\mathbf{u}, \mathbf{x}) \in T_{\mathbf{u}}U$. Thus $(g \circ f)_* = g_* \circ f_*$ is still true. Suppose that $f : U \rightarrow V$ is a diffeomorphism. Using $W = U$ and $g = f^{-1}$ in the previous relation, we get that $(f^{-1})_* \circ f_* = (f^{-1} \circ f)_* = \text{Id}_*$ and $f_* \circ (f^{-1})_* = (f \circ f^{-1})_* = \text{Id}_*$. Thus $(f_*)^{-1} = (f^{-1})_*$.

Proposition 5.3.3

Let $S \subset \mathbb{R}^n$ be a k -dimensional manifold of class C^1 and $\mathbf{u} \in S$. Suppose that (W_i, U_i, ϕ_i) for $i = 1, 2$ are local charts about $\mathbf{u} \in S$. If $\mathbf{w}_i = \phi_i^{-1}(\mathbf{u})$ for $i = 1, 2$, then

$$(\phi_1)_*(T_{\mathbf{w}_1}W_1) = (\phi_2)_*(T_{\mathbf{w}_2}W_2) . \quad (5.3.1)$$

Moreover $(\phi_2^{-1} \circ \phi_1)_* : T_{\mathbf{w}_1}W_1 \rightarrow T_{\mathbf{w}_2}W_2$ is a linear isomorphism.

Proof.

i) We begin by proving that $(\phi_2^{-1} \circ \phi_1)_* : T_{\mathbf{w}_1}W_1 \rightarrow T_{\mathbf{w}_2}W_2$ is a linear isomorphism.

Let $h = \phi_2^{-1} \circ \phi_1 : \phi_1^{-1}(U_1 \cap U_2) \rightarrow \phi_2^{-1}(U_1 \cap U_2)$. By definition of a local chart, h is a diffeomorphism of $\phi_1^{-1}(U_1 \cap U_2)$ onto $\phi_2^{-1}(U_1 \cap U_2)$. Since $h \circ h^{-1} = \text{Id}$ on $\phi_2^{-1}(U_1 \cap U_2)$, we get that $h_* \circ h_*^{-1} = (h \circ h^{-1})_* = \text{Id}_*$ on $T_{\mathbf{w}_2}W_2$. Similarly, since $h^{-1} \circ h = \text{Id}$ on $\phi_1^{-1}(U_1 \cap U_2)$, we get that $h_*^{-1} \circ h_* = (h^{-1} \circ h)_* = \text{Id}_*$ on $T_{\mathbf{w}_1}W_1$. Thus $h_* : T_{\mathbf{w}_1}W_1 \rightarrow T_{\mathbf{w}_2}W_2$ is invertible. The linearity of h on $T_{\mathbf{w}_1}W_1$ comes from the definition of h_* ; namely, $h_*(\mathbf{w}_1, \mathbf{v}) = (h(\mathbf{w}_1), Dh(\mathbf{w}_1)\mathbf{v}) = (\mathbf{w}_2, Dh(\mathbf{w}_1)\mathbf{v})$ for all $\mathbf{v} \in \mathbb{R}^k$.

ii) To prove (5.3.1), we note that $\phi_1 = \phi_2 \circ h$. Hence, $(\phi_1)_* = (\phi_2 \circ h)_* = (\phi_2)_* \circ h_*$. Since h_* maps $T_{\mathbf{w}_1}W_1$ onto $T_{\mathbf{w}_2}W_2$ by (i), we get (5.3.1). ■

From Proposition 5.3.3, we have that $(\phi_2^{-1} \circ \phi_1)_* : T_{\mathbf{w}_1}W_1 \rightarrow T_{\mathbf{w}_2}W_2$ is a linear isomorphism. Namely, $(\mathbf{w}_1, \mathbf{v}) \mapsto (\phi_2^{-1} \circ \phi_1)_*(\mathbf{w}_1, \mathbf{v}) = (\mathbf{w}_2, D(\phi_2^{-1} \circ \phi_1)(\mathbf{w}_1)\mathbf{v})$ is a linear isomorphism for $\mathbf{w}_1$ fixed. Hence $D(\phi_2^{-1} \circ \phi_1)(\mathbf{w}_1)$ is a linear isomorphism of $\mathbb{R}^k$ onto itself. Thus $D(\phi_2^{-1} \circ \phi_1)(\mathbf{w}_1)$ has rank k . Hence $D\phi_1(\mathbf{w}_1)$ has rank k and it is one-to-one because its domain is $\mathbb{R}^k$. Thus $T_{\mathbf{x}}S = (\phi_1)_*(T_{\mathbf{w}_1}W_1)$ is a linear subspace of $T_{\mathbf{x}}\mathbb{R}^n$ of dimension k .

Definition 5.3.4

Suppose that S_i is a k_i -dimensional manifold of class C^q for $i = 1, 2$. A function $f : S_1 \rightarrow S_2$ is of class C^q if $\phi_2^{-1} \circ f \circ \phi_1 : \phi_1^{-1}(f^{-1}(U_2)) \rightarrow W_2$ is of class C^q for all local charts (W_1, U_1, ϕ_1) of S_1 and (W_2, U_2, ϕ_2) of S_2 .

Suppose that $k_1 = k_2$ and $q > 0$. A function $f : S_1 \rightarrow S_2$ is a **diffeomorphism** if $f^{-1} : S_2 \rightarrow S_1$ exists, and both $f : S_1 \rightarrow S_2$ and $f^{-1} : S_2 \rightarrow S_1$ are functions of class C^1 .

The reader should prove that to determine if a map $f : S_1 \rightarrow S_2$ between two manifolds S_1 and S_2 is of class C^q , it suffices to prove that $\phi_2^{-1} \circ f \circ \phi_1 : \phi_1^{-1}(f^{-1}(U_2)) \rightarrow W_2$ is of class C^q for local charts (W_1, U_1, ϕ_1) of an atlas $\mathcal{A}_1(k, q)$ for S_1 and local charts (W_2, U_2, ϕ_2) of an atlas $\mathcal{A}_2(k, q)$ for S_2 .

The following theorem is an immediate consequence of Proposition 5.2.10.

Theorem 5.3.5

If $f : S_1 \rightarrow S_2$ is a diffeomorphism between two k -dimensional manifolds of class C^1 , then $f(\text{Int } S_1) \subset \text{Int } S_2$ and $f(\partial S_1) \subset \partial S_2$. Moreover, $f|_{\text{Int } S_1} : \text{Int } S_1 \rightarrow \text{Int } S_2$ and $f|_{\partial S_1} : \partial S_1 \rightarrow \partial S_2$ are diffeomorphisms.

Suppose that S_i is a k_i -dimensional manifold of class C^1 for $i = 1, 2$, and that $f : S_1 \rightarrow S_2$ is a map of class C^1 . Since $S_1^\circ = \emptyset$ when $k < n$, we cannot define the derivative of f from an extension of f to $\mathbb{R}^n$ as we did for the derivative of a function on the boundary points of H_k because there is no unique extension of f and we cannot use the interior of S_1 to ensure that the value of the derivative is unique.

Given $\mathbf{u}_1 \in S_1$, let $\mathbf{u}_2 = f(\mathbf{u}_1)$. We define $f_* : T_{\mathbf{u}_1}S_1 \rightarrow T_{\mathbf{u}_2}S_2$ as it follows. Let (W_1, U_1, ϕ_1) be a local chart of S_1 about $\mathbf{u}_1 \in S_1$ and (W_2, U_2, ϕ_2) be a local chart of S_2 about $\mathbf{u}_2 \in S_2$. We have that $\phi_2^{-1} \circ f \circ \phi_1 : \phi_1^{-1}(f^{-1}(U_2)) \rightarrow W_2$ and $(\phi_2^{-1} \circ f \circ \phi_1)_*(T_{\mathbf{w}_1}W_1) \subset T_{\mathbf{w}_2}W_2$ where $\mathbf{w}_1 = \phi_1^{-1}(\mathbf{u}_1)$ and $\mathbf{w}_2 = \phi_2^{-1}(\mathbf{u}_2)$. If $(\mathbf{u}_1, \mathbf{x}) \in T_{\mathbf{u}_1}S_1$, then $(\mathbf{u}_1, \mathbf{x}) = (\phi_1)_*(\mathbf{w}_1, \mathbf{y})$ for some unique $(\mathbf{w}_1, \mathbf{y}) \in T_{\mathbf{w}_1}W_1$. We set $f_*(\mathbf{u}_1, \mathbf{x}) = (\phi_2)_*((\phi_2^{-1} \circ f \circ \phi_1)_*(\mathbf{w}_1, \mathbf{y}))$. We get the following commutative diagram.

$$\begin{array}{ccc} T_{\mathbf{u}_1}S_1 & \xrightarrow{f_*} & T_{\mathbf{u}_2}S_2 \\ (\phi_1)_* \uparrow & & \uparrow (\phi_2)_* \\ T_{\mathbf{w}_1=\phi_1^{-1}(\mathbf{u}_1)}W_1 & \xrightarrow{(\phi_2^{-1} \circ f \circ \phi_1)_*} & T_{\mathbf{w}_2=\phi_2^{-1}(\mathbf{u}_2)}W_2 \end{array}$$

We note that $(\phi_i)_* : T_{\mathbf{w}_i}W_i \rightarrow T_{\mathbf{u}_i}S_i$ is a linear isomorphism because $D\phi_i(\mathbf{w}_i)$ is of rank k_i . Thus $D\phi_i(\mathbf{w}_i)$ is a linear isomorphism between $\mathbb{R}^{k_i}$ and a subspace of dimension k_i in $\mathbb{R}^n$.

Definition 5.3.6

Suppose that S is a k -dimensional manifold of class C^1 . The **tangent bundle of S** is the set

$$TS = \bigcup_{\mathbf{u} \in S} T_{\mathbf{u}}S.$$

The map $\pi_S : TS \rightarrow S$ defined by $\pi_S(\mathbf{u}, \mathbf{x}) = \mathbf{u}$ for all $(\mathbf{u}, \mathbf{x}) \in TS$ is the **tangent bundle projection** of S .

If S is a k -dimensional manifold, then we can give a structure of $2k$ -dimensional manifold to TS . Suppose that (W, U, ϕ) is a local chart of S . Let $\tilde{W} = W \times \mathbb{R}^k$ and $\tilde{U} = TU = \bigcup_{u \in U} T_u S$, and consider the map $\tilde{\phi} : \tilde{W} \rightarrow \tilde{U}$ defined by $\tilde{\phi}(\mathbf{w}, \mathbf{y}) = \phi_*(\mathbf{w}, \mathbf{y}) = (\phi(\mathbf{w}), D\phi(\mathbf{w})\mathbf{y})$ for all $(\mathbf{w}, \mathbf{y}) \in \tilde{W}$. Then $(\tilde{W}, \tilde{U}, \tilde{\phi})$ defines a local chart of TS . Note that $\tilde{W} = W \times \mathbb{R}^k$ can be seen as an open subset of H_{2k} because $H_k \times \mathbb{R}^k \cong H_{2k}$ by permuting coordinates.

Remark 5.3.7

The tangent bundle of a manifold S has the structure of what is called a **vector bundle**. The reader interested in learning more about this subject should consult [1, 27]. ♠

Suppose that S is a k -dimensional manifold. For each $\mathbf{u} \in S$, the tangent space $T_{\mathbf{u}}S$ is a subspace of $\mathbb{R}^n$ of dimension k . Recall that if (W, U, ϕ) is a local chart with $\mathbf{u} \in U$ and $\mathbf{w} \in W$ such that $\mathbf{u} = \phi(\mathbf{w})$, then $D\phi(\mathbf{w})(\mathbb{R}^k) \subset \mathbb{R}^n$ is a subspace of dimension k of $\mathbb{R}^n$. Therefore, we may associate an **orientation**

$$\mu_{\mathbf{u}} = [(\mathbf{u}, \mathbf{x}_{u,1}), (\mathbf{u}, \mathbf{x}_{u,2}), \dots, (\mathbf{u}, \mathbf{x}_{u,k})]$$

to $T_{\mathbf{u}}S$ where $\{\mathbf{x}_{u,i}\}_{i=1}^k$ is a basis of the subspace $D\phi(\mathbf{w})(\mathbb{R}^k)$.

Definition 5.3.8

Let S be a k -dimensional manifold of class C^1 . Suppose that we assign an orientation $\mu_{\mathbf{u}}$ on each $T_{\mathbf{u}}S$ for each $\mathbf{u} \in S$. Such choice of orientations is **consistent** if, for every local chart (W, U, ϕ) of S , and for every $\mathbf{w}_1$ and $\mathbf{w}_2$ in W , we have

$$[\phi_*(\mathbf{w}_1, \mathbf{e}_1), \phi_*(\mathbf{w}_1, \mathbf{e}_2), \dots, \phi_*(\mathbf{w}_1, \mathbf{e}_k)] = \mu_{\phi(\mathbf{w}_1)}$$

if and only if

$$[\phi_*(\mathbf{w}_2, \mathbf{e}_1), \phi_*(\mathbf{w}_2, \mathbf{e}_2), \dots, \phi_*(\mathbf{w}_2, \mathbf{e}_k)] = \mu_{\phi(\mathbf{w}_2)} .$$

As a little warning to the reader, we emphasize that, in the previous definition, we did not request that

$$[\phi_*(\mathbf{w}, \mathbf{e}_1), \phi_*(\mathbf{w}, \mathbf{e}_2), \dots, \phi_*(\mathbf{w}, \mathbf{e}_k)] = \mu_{\phi(\mathbf{w})}$$

for all $\mathbf{w} \in W$. We may as well have that

$$[\phi_*(\mathbf{w}, \mathbf{e}_1), \phi_*(\mathbf{w}, \mathbf{e}_2), \dots, \phi_*(\mathbf{w}, \mathbf{e}_k)] \neq \mu_{\phi(\mathbf{w})}$$

for all $\mathbf{w} \in W$. As we have seen in Section 4.2, we have for each $\mathbf{u} \in S$ two possible orientations on $T_{\mathbf{u}}S$. This will play an important role in the study of integration on manifolds later.

Definition 5.3.9

Let S be a k -dimensional manifold of class C^1 . Suppose that the choice of orientation $\mu_{\mathbf{u}}$ on each $\mathbf{u} \in S$ is consistent. A local chart (W, U, ϕ) of S is called **orientation**

preserving if

$$[\phi_*(\mathbf{w}, \mathbf{e}_1), \phi_*(\mathbf{w}, \mathbf{e}_2), \dots, \phi_*(\mathbf{w}, \mathbf{e}_k)] = \mu_{\phi(\mathbf{w})}$$

for one (and so all ¹) $\mathbf{w} \in W$.

If a local chart (W, U, ϕ) of a k -dimensional manifold S without boundary is not orientation preserving, then we can easily produce a local chart which will be orientation preserving. Choose a linear map $T : \mathbb{R}^k \rightarrow \mathbb{R}^k$ with $\det(T) = -1$ (e.g. a permutation of two coordinates), then $(T^{-1}(W), V, \phi \circ T)$ is a local chart which is orientation preserving.

Suppose that (W_i, U_i, ϕ_i) for $i = 1, 2$ are two orientation preserving local charts of a k -dimensional manifold S . If $\mathbf{u} \in U_1 \cap U_2$ and $\mathbf{u} = \phi_i(\mathbf{w}_i)$ for $\mathbf{w}_i \in W_i$, then we have

$$\begin{aligned} & [(\phi_1)_*(\mathbf{w}_1, \mathbf{e}_1), (\phi_1)_*(\mathbf{w}_1, \mathbf{e}_2), \dots, (\phi_1)_*(\mathbf{w}_1, \mathbf{e}_k)] \\ &= \mu_{\mathbf{u}} = [(\phi_2)_*(\mathbf{w}_2, \mathbf{e}_1), (\phi_2)_*(\mathbf{w}_2, \mathbf{e}_2), \dots, (\phi_2)_*(\mathbf{w}_2, \mathbf{e}_k)] . \end{aligned}$$

Thus

$$\begin{aligned} & [(\phi_2^{-1} \circ \phi_1)_*(\mathbf{w}_1, \mathbf{e}_1), (\phi_2^{-1} \circ \phi_1)_*(\mathbf{w}_1, \mathbf{e}_2), \dots, (\phi_2^{-1} \circ \phi_1)_*(\mathbf{w}_1, \mathbf{e}_k)] \\ &= [(\mathbf{w}_2, \mathbf{e}_1), (\mathbf{w}_2, \mathbf{e}_2), \dots, (\mathbf{w}_2, \mathbf{e}_k)] . \end{aligned}$$

This implies that $D(\phi_2^{-1} \circ \phi_1)(\mathbf{w}_1)$ preserves the orientation on $\mathbb{R}^k$. Hence $\det(D(\phi_2^{-1} \circ \phi_1)(\mathbf{w}_1)) > 0$.

Definition 5.3.10

A k -dimensional manifold S of class C^1 is **orientable** if there exists a consistent choice of orientations $\mu_{\mathbf{u}}$ on $T_{\mathbf{u}}S$ for each $\mathbf{u} \in S$. A consistent choice of orientation is called an **orientation** of S . A k -dimensional manifold S with an orientation is called an **oriented manifold**.

Remark 5.3.11

We will not consider non-orientable manifolds in this book. Many of the results about oriented manifolds presented in this book have versions for non-orientable manifolds. However, there are also many results about oriented manifolds that do not have counterparts for non-orientable manifolds and vice-versa. ♠

Example 5.3.12

We cannot mention non-orientable manifolds without providing the best known example of a non-orientable manifold; namely, the Möbius strip (Figure 5.6). It is not possible to provide a consistent orientation on the Möbius strip. Suppose that M is a Möbius strip and that $\tilde{\mathbf{u}} \in M$. It is easy to imagine a sequence of local charts (W_i, U_i, ϕ_i) for $1 \leq i \leq I$ such that $U_{i-1} \cap U_i \neq \emptyset$ for $1 < i \leq I$, (W_i, U_i, ϕ_i) preserves the orientation inherited from $(W_{i-1}, U_{i-1}, \phi_{i-1})$ for $1 < i \leq I$ ², and $\tilde{\mathbf{u}} \in U_1 \cap U_I$. If $\tilde{\mathbf{u}} = \phi_1(\tilde{\mathbf{w}}_1) = \phi_I(\tilde{\mathbf{w}}_I)$, then $[(\phi_1)_*(\tilde{\mathbf{w}}_1, \mathbf{e}_1), (\phi_1)_*(\tilde{\mathbf{w}}_1, \mathbf{e}_2)] \neq [(\phi_I)_*(\tilde{\mathbf{w}}_I, \mathbf{e}_1), (\phi_I)_*(\tilde{\mathbf{w}}_I, \mathbf{e}_2)]$. ♣

¹Because the choice of orientation is consistent.

²Namely, if $\mathbf{u} = \phi_i(\mathbf{w}_i) = \phi_{i-1}(\mathbf{w}_{i-1})$, then $[(\phi_i)_*(\mathbf{w}_i, \mathbf{e}_1), (\phi_i)_*(\mathbf{w}_i, \mathbf{e}_2)] = [(\phi_{i-1})_*(\mathbf{w}_{i-1}, \mathbf{e}_1), (\phi_{i-1})_*(\mathbf{w}_{i-1}, \mathbf{e}_2)] = \mu_{\mathbf{u}}$.

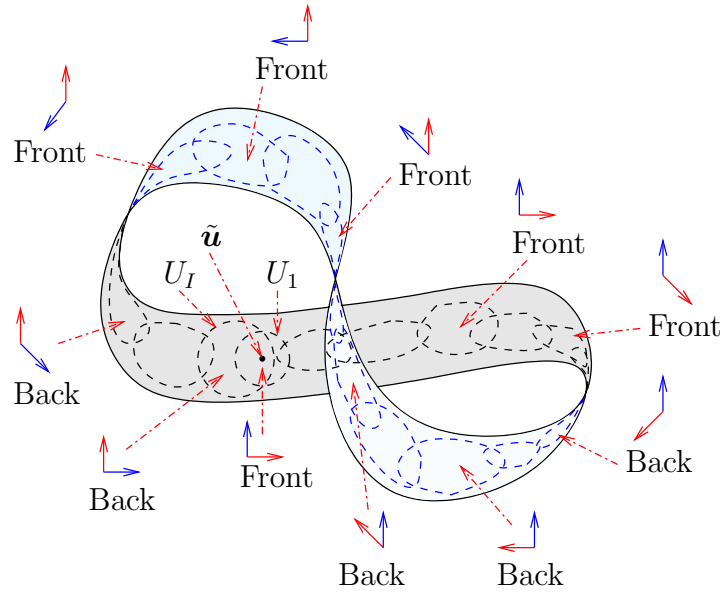


Figure 5.6: Möbius strip.

Definition 5.3.13

Let S and $\tilde{S}$ be two k -dimensional manifolds of class C^1 . Suppose that $\mu_{\mathbf{u}}$ for each $\mathbf{u} \in S$ is the orientation on S and $\tilde{\mu}_{\mathbf{u}}$ for each $\mathbf{u} \in \tilde{S}$ is the orientation on $\tilde{S}$. We say that a diffeomorphism $g : S \rightarrow \tilde{S}$ is **orientation preserving** if

$$[g_*(\mathbf{u}, \mathbf{x}_1), g_*(\mathbf{u}, \mathbf{x}_2), \dots, g_*(\mathbf{u}, \mathbf{x}_k)] = \tilde{\mu}_{g(\mathbf{u})}$$

for all $\{(\mathbf{u}, \mathbf{x}_i)\}_{1 \leq i \leq k} \subset T_{\mathbf{u}}S$ such that $[(\mathbf{u}, \mathbf{x}_1), (\mathbf{u}, \mathbf{x}_2), \dots, (\mathbf{u}, \mathbf{x}_k)] = \mu_{\mathbf{u}}$ and all $\mathbf{u} \in S$.

The orientation on the k -dimensional manifold S with boundary induces an orientation on the $(k - 1)$ -dimensional manifold ∂S .

Suppose that S is an oriented k -dimensional manifold. To define the induced orientation on ∂S , we begin by defining the outward unit normal to the manifold ∂S at $\mathbf{u} \in \partial S$. Since $T_{\mathbf{u}}\partial S$ is a $(k - 1)$ dimensional subspace of $T_{\mathbf{u}}S$, there are two unit vectors orthogonal to $T_{\mathbf{u}}\partial S$ in $T_{\mathbf{u}}S$. Let $(\mathbf{u}, \mathbf{x}_1)$ and $(\mathbf{u}, \mathbf{x}_2)$ be these two vectors. One of the two is the outward unit normal. Choose an orientation preserving local chart (W, U, ϕ) of S such that $\mathbf{u} \in U \cap \partial S$, and let $\mathbf{w} \in \partial W \cap H_k$ be the vector such that $\mathbf{u} = \phi(\mathbf{w})$. Since $\phi_* : T_{\mathbf{w}}W \rightarrow T_{\mathbf{u}}S$ is a linear isomorphism, there exists a unique $(\mathbf{w}, \mathbf{q}) \in T_{\mathbf{w}}W$ such that $\mathbf{q} \in \mathbb{R}^k \setminus H_k$ and $\phi_*(\mathbf{w}, \mathbf{q})$ is one of $(\mathbf{u}, \mathbf{x}_1)$ or $(\mathbf{u}, \mathbf{x}_2)$.

Definition 5.3.14

The vector $\mathbf{n}_u = (\mathbf{u}, \mathbf{n}) = \phi_*(\mathbf{w}, \mathbf{q}) \in T_u S$ is called the **outward unit normal** to ∂S at $\mathbf{u} \in \partial S$ (Figure 5.7).

One can show using Proposition 5.2.10 that the definition of the outward unit normal is independent of the local chart used.

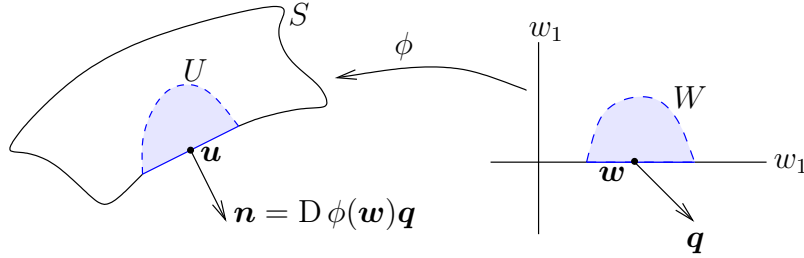


Figure 5.7: The outward unit normal $\mathbf{n}_u \in T_u S$ to the manifold ∂S at the point $\mathbf{u} \in \partial S$ is given by $\mathbf{n}_u = (\mathbf{u}, \mathbf{n})$ where $\mathbf{n} = D\phi(\mathbf{w})(\mathbf{q})$.

Definition 5.3.15

Suppose that S is a k -dimensional manifold of class C^1 with boundary and that μ_u for $\mathbf{u} \in S$ is an orientation on S . For each $\mathbf{u} \in \partial S$, choose $\{(\mathbf{u}, \mathbf{x}_{u,i})\}_{i=1}^{k-1} \subset T_u \partial S$ such that $[\mathbf{n}_u, (\mathbf{u}, \mathbf{x}_{u,1}), (\mathbf{u}, \mathbf{x}_{u,2}), \dots, (\mathbf{u}, \mathbf{x}_{u,k-1})] = \mu_u$. The **induced orientation on ∂S** is defined as $(\partial\mu)_u \equiv [(\mathbf{u}, \mathbf{x}_{u,1}), (\mathbf{u}, \mathbf{x}_{u,2}), \dots, (\mathbf{u}, \mathbf{x}_{u,k-1})]$ for $\mathbf{u} \in \partial S$.

Using the local charts of ∂S provided in Theorem 5.2.11, it is easy to prove that if μ_u is a consistent choice of orientations on $T_u S$ for $\mathbf{u} \in S$, then $(\partial\mu)_u$ is a consistent choice of orientations on $T_u \partial S$ for $\mathbf{u} \in \partial S$.

Remark 5.3.16

For the special case where $S = H_k$ with the standard orientation, then the induced orientation on $\partial S \cong \mathbb{R}^{k-1}$ is $(-1)^k$ times the standard orientation on $\mathbb{R}^{k-1}$. We have that $\mathbf{n}_w = (\mathbf{w}, -\mathbf{e}_k)$ for $\mathbf{w} \in \partial H_k$. Hence

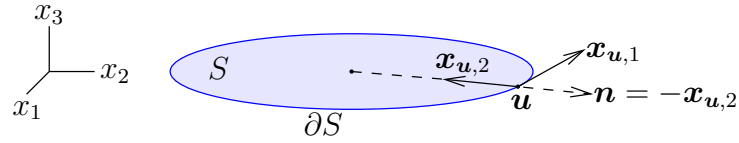
$$\begin{aligned} & [(\mathbf{w}, -\mathbf{e}_k), (\mathbf{w}, \mathbf{e}_1), (\mathbf{w}, \mathbf{e}_2), \dots, (\mathbf{w}, \mathbf{e}_{k-1})] \\ &= (-1)^{k-1} [(\mathbf{w}, \mathbf{e}_1), (\mathbf{w}, \mathbf{e}_2), \dots, (\mathbf{w}, \mathbf{e}_{k-1}), (\mathbf{w}, -\mathbf{e}_k)] \\ &= (-1)^k [(\mathbf{w}, \mathbf{e}_1), (\mathbf{w}, \mathbf{e}_2), \dots, (\mathbf{w}, \mathbf{e}_{k-1}), (\mathbf{w}, \mathbf{e}_k)] . \end{aligned}$$

The two orientations are the same when k is even and are of opposite signs when k is odd. ♠

Example 5.3.17

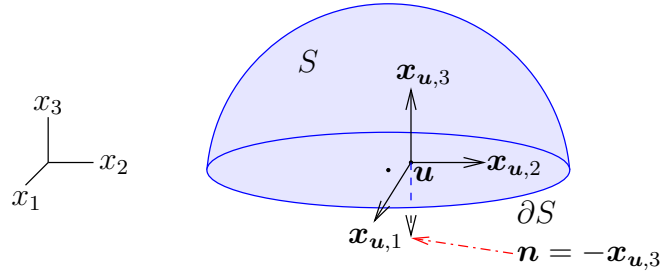
We illustrate the previous concept of induced orientation and outward unit normal for two of the simplest manifolds in $\mathbb{R}^3$.

For the first example, S is a disk in the x_1, x_2 plane. Thus S is an orientable 2-dimensional manifold.



The chosen orientation on $T_u S$ is given by $\mu_u = [(\mathbf{u}, \mathbf{x}_{u,1}), (\mathbf{u}, \mathbf{x}_{u,2})]$. The outward unit normal to S at $\mathbf{u} \in \partial S$ is $\mathbf{n}_u = (\mathbf{u}, -\mathbf{x}_{u,2})$. The orientation on $T_u \partial S$ is $(\partial\mu)_u = [(\mathbf{u}, \mathbf{x}_{u,1})]$ because $[\mathbf{n}_u, (\mathbf{u}, \mathbf{x}_{u,1})] = [(\mathbf{u}, -\mathbf{x}_{u,2}), (\mathbf{u}, \mathbf{x}_{u,1})] = -[(\mathbf{u}, \mathbf{x}_{u,1}), (\mathbf{u}, -\mathbf{x}_{u,2})] = \mu_u$.

For our second example, S is half a ball. Thus S is an orientable 3-dimensional manifold.



We consider $\mathbf{u} \in \partial S$ on the bottom face of S . The chosen orientation on $T_u S$ is given by $\mu_u = [(\mathbf{u}, \mathbf{x}_{u,1}), (\mathbf{u}, \mathbf{x}_{u,2}), (\mathbf{u}, \mathbf{x}_{u,3})]$. The outward unit normal to S at $\mathbf{u} \in \partial S$ is $\mathbf{n}_u = (\mathbf{u}, -\mathbf{x}_{u,3})$. The orientation on $T_u \partial S$ is $(\partial\mu)_u = [(\mathbf{u}, \mathbf{x}_{u,2}), (\mathbf{u}, \mathbf{x}_{u,1})]$ because

$$\begin{aligned} [\mathbf{n}_u, (\mathbf{u}, \mathbf{x}_{u,2}), (\mathbf{u}, \mathbf{x}_{u,1})] &= [(\mathbf{u}, -\mathbf{x}_{u,3}), (\mathbf{u}, \mathbf{x}_{u,2}), (\mathbf{u}, \mathbf{x}_{u,1})] \\ &= -[(\mathbf{u}, \mathbf{x}_{u,1}), (\mathbf{u}, \mathbf{x}_{u,2}), -(\mathbf{u}, \mathbf{x}_{u,3})] = \mu_u . \end{aligned}$$



For $(n-1)$ -dimensional manifold S in $\mathbb{R}^n$, we may define an outward unit normal to S at every point of S even if S is not the boundary of a n -dimensional manifold.

Definition 5.3.18

Suppose that S is a $(n-1)$ -dimensional manifold of class C^1 . Suppose that $\mu_u = [(\mathbf{u}, \mathbf{x}_{u,1}), (\mathbf{u}, \mathbf{x}_{u,2}), \dots, (\mathbf{u}, \mathbf{x}_{u,n-1})]$ is the selected orientation on $T_u S$ for $\mathbf{u} \in S$. The **outward unit normal** to S at $\mathbf{u} \in S$ is the unit vector $\mathbf{n}_u = (\mathbf{u}, \mathbf{n}(\mathbf{u})) \in T_u \mathbb{R}^n \cong \mathbb{R}^n$ such that $[(\mathbf{u}, \mathbf{n}(\mathbf{u})), (\mathbf{u}, \mathbf{x}_{u,1}), (\mathbf{u}, \mathbf{x}_{u,2}), \dots, (\mathbf{u}, \mathbf{x}_{u,n-1})] = [(\mathbf{u}, \mathbf{e}_1), (\mathbf{u}, \mathbf{e}_2), \dots, (\mathbf{u}, \mathbf{e}_n)]$, the standard orientation on $\mathbb{R}^n$.

5.3.1 Tangent Spaces as Equivalence Classes of Curves

We find in the literature several equivalent definitions of the tangent space to a manifold at a point. The choice of the definition is usually determined by how we plan to use them. We

have chosen the definition which is closer to the idea that the reader may have of a tangent space. We briefly present in this subsection and the next subsection two other definitions of the tangent space to a manifold at a point.

Suppose that S is a k -dimensional manifold of class C^1 and $\mathbf{u} \in S$. A differentiable curve at $\mathbf{u} \in S$ is a function $\rho :]-a, a[\rightarrow S$ of class C^1 such that $\rho(0) = \mathbf{u}$ and $a > 0$. Two differentiable curves $\rho_i :]-a_i, a_i[\rightarrow S$ for $1 \leq i \leq 2$ at $\mathbf{u}$ are **equivalent** if $\rho'_1(0) = \rho'_2(0)$.

The **tangent space** $\tilde{T}_{\mathbf{u}}S$ is defined as the set of all equivalent class of differentiable curves $\rho :]-a, a[\rightarrow S$ at $\mathbf{u} \in S$. If $[\rho]_{\mathbf{u}}$ denotes the equivalent class for the differentiable curve $\rho :]-a, a[\rightarrow S$ at $\mathbf{u} \in S$, then

$$\tilde{T}_{\mathbf{u}}S = \{[\rho]_{\mathbf{u}} : \rho :]-a, a[\rightarrow S \text{ is a differentiable curve at } \mathbf{u}\} .$$

We now explain why this definition of a tangent space is equivalent to Definition 5.3.2.

Suppose that (W, U, ϕ) is a local chart such that $\mathbf{u} \in U$ and let $\mathbf{w} = \phi^{-1}(\mathbf{u})$. The function

$$\begin{aligned} \tilde{\phi} : \tilde{T}_{\mathbf{w}}W &\rightarrow \tilde{T}_{\mathbf{u}}S \\ [\rho]_{\mathbf{w}} &\mapsto [\phi \circ \rho]_{\mathbf{u}} \end{aligned}$$

defines a bijective map.

i) The function $\tilde{\phi}$ is well defined.

1. If $\rho :]-a, a[\rightarrow W$ is a differentiable curve at $\mathbf{w}$, then $\phi \circ \rho :]-a, a[\rightarrow S$ is a differentiable curve at $\mathbf{u}$ because $\phi \circ \rho$ is of class C^1 (composition of functions of class C^1) and $(\phi \circ \rho)(0) = \phi(\rho(0)) = \phi(\mathbf{w}) = \mathbf{u}$.
2. Let ρ_1 and ρ_2 be two equivalent differentiable curves at $\mathbf{w}$. Then $\phi \circ \rho_1$ and $\phi \circ \rho_2$ are two equivalent differentiable curves at $\mathbf{u}$ because $\rho'_1(0) = \rho'_2(0)$ implies that

$$\begin{aligned} (\phi \circ \rho_1)'(0) &= D\phi(\rho_1(0))\rho'_1(0) = D\phi(\mathbf{w})\rho'_1(0) = D\phi(\mathbf{w})\rho'_2(0) \\ &= D\phi(\rho_2(0))\rho'_2(0) = (\phi \circ \rho_2)'(0) . \end{aligned}$$

ii) The function $\tilde{\phi} : \tilde{T}_{\mathbf{w}}W \rightarrow \tilde{T}_{\mathbf{u}}S$ is one-to-one. If ρ_1 and ρ_2 are two differentiable curves at $\mathbf{w}$ such that $\phi \circ \rho_1$ and $\phi \circ \rho_2$ are equivalent differentiable curves at $\mathbf{u}$, then

$$\begin{aligned} D\phi(\mathbf{w})\rho'_1(0) &= D\phi(\rho_1(0))\rho'_1(0) = (\phi \circ \rho_1)'(0) = (\phi \circ \rho_2)'(0) \\ &= D\phi(\rho_2(0))\rho'_2(0) = D\phi(\mathbf{w})\rho'_2(0) . \end{aligned}$$

Since $D\phi(\mathbf{w})$ is one-to-one, $\rho'_1(0) = \rho'_2(0)$ and so ρ_1 and ρ_2 are two equivalent differentiable curves at $\mathbf{w}$.

iii) To prove that $\tilde{\phi} : \tilde{T}_{\mathbf{w}}W \rightarrow \tilde{T}_{\mathbf{u}}S$ is onto, it suffices to note that if $\rho :]-a, a[\rightarrow U$ is a differentiable curve at $\mathbf{u}$, then $\eta :]-a, a[\rightarrow W$ defined by $\rho(t) = \phi(\eta(t))$ for $t \in]-a, a[$ is a differentiable curve at $\mathbf{w}$. To prove that the curve is η is of class C^1 , we use the map $\tilde{h}$ provided in the proof of Theorem 5.2.13 and Theorem 5.2.14. We have that $\tilde{h} : \mathbb{R}^n \rightarrow \mathbb{R}^k$ is

of class C^1 and $\tilde{h}|_{S \cap U} = \phi^{-1} : U \rightarrow W$. So $\eta = \tilde{h} \circ \rho :] - a, a[\rightarrow W$ is a differentiable curve at $\mathbf{w}$.

We define a map ψ from $T_{\mathbf{w}}W$ to $\tilde{T}_{\mathbf{w}}W$ as it follows. We assign to each $(\mathbf{w}, \mathbf{y}) \in T_{\mathbf{w}}W$ a unique equivalence class of differentiable curves at $\mathbf{w}$ such that the derivative at 0 is $\mathbf{y}$; namely, the class $[\rho_{\mathbf{w}, \mathbf{y}}]_{\mathbf{w}}$ of all differentiable curves equivalent to $\rho_{\mathbf{w}, \mathbf{y}}(t) = \mathbf{w} + t\mathbf{y}$ for $t \in] - a, a[$ and a small enough. We leave it to the reader to convince themselves that the function

$$\begin{aligned} \psi : T_{\mathbf{w}}W &\rightarrow \tilde{T}_{\mathbf{w}}W \\ (\mathbf{w}, \mathbf{y}) &\mapsto [\rho_{\mathbf{w}, \mathbf{y}}]_{\mathbf{w}} \end{aligned}$$

is a bijection.

Finally, since $D\phi(\mathbf{w})$ is a linear isomorphism onto $D\phi(\mathbf{w})(\mathbb{R}^k)$ because $D\phi(\mathbf{w})$ is of rank k , we have that $\phi_* : T_{\mathbf{w}}W \rightarrow T_{\mathbf{u}}S$ is a linear isomorphism. Thus $\tilde{\phi} \circ \psi \circ (\phi_*)^{-1} : T_{\mathbf{u}}S \rightarrow \tilde{T}_{\mathbf{u}}S$ is an isomorphism. Hence $T_{\mathbf{u}}S$ and $\tilde{T}_{\mathbf{u}}S$ are isomorphic. In fact, it follows from the definition of this isomorphism that the equivalent class of curves at $\mathbf{u}$ associated to $(\mathbf{u}, \mathbf{x}) \in T_{\mathbf{u}}S$ is the class of all differentiable curves $\rho :] - a, a[\rightarrow S$ such that $\rho(0) = \mathbf{u}$ and $\rho'(0) = \mathbf{x}$. To prove this claim, suppose that $(\mathbf{u}, \mathbf{x}) \in T_{\mathbf{u}}S$. Then $(\phi_*)^{-1}(\mathbf{u}, \mathbf{x}) = (\mathbf{w}, \mathbf{y})$ where $\phi(\mathbf{w}) = \mathbf{u}$ and $(D\phi(\mathbf{w}))(\mathbf{y}) = \mathbf{x}$. Hence $(\psi \circ (\phi_*)^{-1})(\mathbf{u}, \mathbf{x}) = \psi(\mathbf{w}, \mathbf{y}) = [\rho_{\mathbf{w}, \mathbf{y}}]_{\mathbf{w}}$. Thus $(\tilde{\phi} \circ \psi \circ (\phi_*)^{-1})(\mathbf{u}, \mathbf{x}) = \tilde{\phi}([\rho_{\mathbf{w}, \mathbf{y}}]_{\mathbf{w}}) = [(\phi \circ \rho_{\mathbf{w}, \mathbf{y}})]_{\mathbf{u}}$ because $(\phi \circ \rho_{\mathbf{w}, \mathbf{y}})(0) = \phi(\rho_{\mathbf{w}, \mathbf{y}}(0)) = \phi(\mathbf{w}) = \mathbf{u}$. We also have that $(\phi \circ \rho_{\mathbf{w}, \mathbf{y}})'(0) = (D\phi(\mathbf{w}))(\mathbf{y}) = \mathbf{x}$.

In the previous discussion, we have assumed that $\mathbf{u} \in \text{Int } S$. We leave it to the reader to study the case where $\mathbf{u} \in \partial S$. We then have to consider the intervals $[0, a[$ or $] - a, 0]$ to get a curve in S .

5.3.2 Tangent Spaces as Spaces of differential Linear Operators

Suppose that S is a k -dimensional manifold of class C^1 and that $\mathbf{u} \in S$. A **tangent vector** to S at $\mathbf{u}$ is a map $q : C^\infty(S) \rightarrow \mathbb{R}$ with the following property. If (W, U, ϕ) is a local chart of S about $\mathbf{u} \in S$, then there exists $\tilde{q} \in \mathbb{R}^k$ such that

$$\tilde{q} = \sum_{i=1}^k y_i \frac{\partial}{\partial w_i} \Big|_{\mathbf{w}}$$

with $\mathbf{w} = \phi^{-1}(\mathbf{u})$ is the local representation of q ; namely,

$$q(g) = \phi^*(\tilde{q})(g) = \sum_{i=1}^k y_i \frac{\partial}{\partial w_i} (g \circ \phi) \Big|_{\mathbf{w}=\phi^{-1}(\mathbf{u})} \quad (5.3.2)$$

for all real valued functions $g \in C^\infty(S)$.

The **tangent space** $\tilde{T}_{\mathbf{u}}S$ is the set of all the tangent vectors to S at $\mathbf{u}$.

Uniqueness of $\mathbf{y}$: The vector $\mathbf{y}$ satisfying (5.3.2) is unique. Suppose that $V \subset \mathbb{R}^n$ is an open set such that $U = S \cap V$. Choose an open set B such that $\mathbf{u} \in B \subset \overline{B} \subset V$ and a smooth function $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$ such that

1. $0 \leq \psi(\mathbf{x}) \leq 1$ for all $\mathbf{x} \in \mathbb{R}^n$,
2. $\psi(\mathbf{x}) = 1$ for all $\mathbf{x} \in B$ and
3. $\psi(\mathbf{x}) = 0$ for all $\mathbf{x} \in \mathbb{R}^n \setminus B$.

We can get the function ψ from Theorem 3.4.1, Partition of Unity, with the open cover $\{V, \mathbb{R}^n \setminus \overline{B}\}$. If

$$g(\mathbf{u}) = \begin{cases} \pi_j(\phi^{-1}(\mathbf{u}))\psi(\mathbf{u}) & \text{if } \mathbf{u} \in U \\ 0 & \text{if } \mathbf{u} \in \mathbb{R}^n \setminus U \end{cases}$$

for $1 \leq j \leq k$ fixed, then $g \in C^\infty(S)$. The projection $\pi_j : \mathbb{R}^k \rightarrow \mathbb{R}$ defined by $\pi_j(\mathbf{w}) = w_j$ is used to obtain the j^{th} component of ϕ^{-1} . Since $(g \circ \phi)(\mathbf{w}) = \pi_j(\phi^{-1}(\phi(\mathbf{w})))\psi(\phi(\mathbf{w})) = w_j$ for all $\mathbf{w} \in \phi^{-1}(S \cap B) \subset W$, an open neighbourhood of $\phi^{-1}(\mathbf{u})$, we get

$$q(g) = \sum_{i=1}^k y_i \frac{\partial}{\partial w_i} (g \circ \phi)(\phi^{-1}(\mathbf{u})) = y_j. \quad (5.3.3)$$

Thus y_j is uniquely determinate by the value of $q(g)$.

Dependence on the local chart: The existence of the vector $\mathbf{y} \in \mathbb{R}^k$ is independent of the local chart used. Suppose that (W_m, U_m, ϕ_m) for $m = 1, 2$ are two local charts of S and $\mathbf{u} \in U_1 \cap U_2$. Moreover, suppose that there exists $\mathbf{y}_1 \in \mathbb{R}^k$ such that

$$q(g) = \sum_{i=1}^k y_{1,i} \frac{\partial}{\partial w_i} (g \circ \phi_1) \Big|_{\mathbf{w}=\phi_1^{-1}(\mathbf{u})}$$

for all $g \in C^\infty(S)$ is the local representation of q with respect to the local chart (W, U, ϕ) . If we work on the open subset $U = U_1 \cap U_2$ of S containing $\mathbf{u}$, then we have

$$\begin{aligned} q(g) &= \sum_{i=1}^k y_{1,i} \frac{\partial}{\partial w_i} (g \circ \phi_1) \Big|_{\mathbf{w}=\phi_1^{-1}(\mathbf{u})} = \sum_{i=1}^k y_{1,i} \frac{\partial}{\partial w_i} ((g \circ \phi_2) \circ (\phi_2^{-1} \circ \phi_1)) \Big|_{\mathbf{w}=\phi_1^{-1}(\mathbf{u})} \\ &= \sum_{i=1}^k y_{1,i} \left(\sum_{j=1}^k \frac{\partial}{\partial w_j} (g \circ \phi_2) \Big|_{\mathbf{w}=\phi_2^{-1}(\mathbf{u})} \frac{\partial}{\partial w_i} (\pi_j(\phi_2^{-1} \circ \phi_1)) \Big|_{\mathbf{w}=\phi_1^{-1}(\mathbf{u})} \right) \\ &= \sum_{j=1}^k \left(\sum_{i=1}^k y_{1,i} \frac{\partial}{\partial w_i} (\pi_j(\phi_2^{-1} \circ \phi_1)) \Big|_{\mathbf{w}=\phi_1^{-1}(\mathbf{u})} \right) \frac{\partial}{\partial w_j} (g \circ \phi_2) \Big|_{\mathbf{w}=\phi_2^{-1}(\mathbf{u})} \\ &= \sum_{j=1}^k y_{2,j} \frac{\partial}{\partial w_j} (g \circ \phi_2) \Big|_{\mathbf{w}=\phi_2^{-1}(\mathbf{u})} \end{aligned}$$

for all $g \in C^\infty(S)$ if and only if

$$\mathbf{y}_2 = (D(\phi_2^{-1} \circ \phi_1)(\phi_1^{-1}(\mathbf{u}))) \mathbf{y}_1. \quad (5.3.4)$$

The projection $\pi_j : \mathbb{R}^k \rightarrow \mathbb{R}$ defined by $\pi_j(\mathbf{w}) = w_j$ was used to obtain the j^{th} component of $\phi_2^{-1} \circ \phi_1$. This relation (5.3.4) is exactly the relation provided by $\phi_2^{-1} \circ \phi_1$ in Proposition 5.3.3.

$\check{T}_{\mathbf{u}}S$ is a Vector Space over $\mathbb{R}$: If $q_m \in \check{T}_{\mathbf{u}}S$ for $m = 1, 2$ and

$$q_m(g) = \sum_{i=1}^k y_{m,i} \frac{\partial}{\partial w_i} (g \circ \phi) \Big|_{\mathbf{w}=\phi^{-1}(\mathbf{u})}$$

for $m = 1, 2$, then it follows from (5.3.3) that $q_1 + q_2 \in \check{T}_{\mathbf{u}}S$ and

$$(q_1 + q_2)(g) = \sum_{i=1}^k (y_{1,i} + y_{2,i}) \frac{\partial}{\partial w_i} (g \circ \phi) \Big|_{\mathbf{w}=\phi^{-1}(\mathbf{u})}.$$

If $q \in \check{T}_{\mathbf{u}}S$ and (5.3.2) is satisfied, then it again follows from (5.3.3) that $\lambda q \in \check{T}_{\mathbf{u}}S$ for all $\lambda \in \mathbb{R}$ and

$$\lambda q(g) = \sum_{i=1}^k (\lambda y_i) \frac{\partial}{\partial w_i} (g \circ \phi) \Big|_{\mathbf{w}=\phi^{-1}(\mathbf{u})}$$

for all $\lambda \in \mathbb{R}$.

Product rule: Let (W, U, ϕ) be a local chart of S about $\mathbf{u} \in S$. Suppose that $g_1, g_2 \in C^\infty(S)$ and that $q \in \check{T}_{\mathbf{u}}S$ is given by (5.3.2). Since $g_1 g_2 \in C^\infty(S)$, we may compute

$$\begin{aligned} q(g_1 g_2) &= \sum_{i=1}^k y_i \frac{\partial}{\partial w_i} ((g_1 \circ \phi)(g_2 \circ \phi)) \Big|_{\mathbf{w}=\phi^{-1}(\mathbf{u})} \\ &= \left(\sum_{i=1}^k y_i \frac{\partial}{\partial w_i} (g_1 \circ \phi) \Big|_{\mathbf{w}=\phi^{-1}(\mathbf{u})} \right) (g_2 \circ \phi)(\phi^{-1}(\mathbf{u})) \\ &\quad + (g_1 \circ \phi)(\phi^{-1}(\mathbf{u})) \sum_{i=1}^k y_i \left(\frac{\partial}{\partial w_i} (g_2 \circ \phi) \Big|_{\mathbf{w}=\phi^{-1}(\mathbf{u})} \right) \\ &= q(g_1) g_2(\mathbf{u}) + g_1(\mathbf{u}) q(g_2). \end{aligned}$$

Since this relation is independent of the local chart, it is true on $\check{T}_{\mathbf{u}}S$ for all $\mathbf{u} \in S$.

It can be proved that the definition given in (5.3.2) of a tangent vector q to S at $\mathbf{u}$ is uniquely determined by the three conditions: (1) $q(g_1 + g_2) = q(g_1) + q(g_2)$ for all $g_1, g_2 \in C^\infty(S)$, (2) $q(\lambda g) = \lambda q(g)$ for all $g \in C^\infty(S)$ and $\lambda \in \mathbb{R}$, and (3) $q(g_1 g_2) = q(g_1) g_2(\mathbf{u}) + g_1(\mathbf{u}) q(g_2)$ for all $f, g \in C^\infty(S)$ and $\mathbf{u} \in S$.

Isomorphism between $T_{\mathbf{u}}S$ and $\check{T}_{\mathbf{u}}S$: We may identify $\check{T}_{\mathbf{u}}S$ to $T_{\mathbf{u}}S$ by associating $\mathbf{y} \in \mathbb{R}^k$ given by (5.3.2) to $(\mathbf{u}, \mathbf{x}) = \phi_*(\mathbf{w}, \mathbf{y}) \in T_{\mathbf{u}}S$.

Suppose that (W, U, ϕ) is a local chart of a k -dimensional manifold S with $\mathbf{u} \in U$, the tangent vectors to S at $\mathbf{u}$ with the local representation (5.3.2) given by $\mathbf{y} = \mathbf{e}_j \in \mathbb{R}^k$ for $1 \leq j \leq k$ are denoted $\frac{\partial}{\partial w_j} \Big|_{\mathbf{w}=\phi^{-1}(\mathbf{u})}$. It follows that $\left\{ \frac{\partial}{\partial w_j} \Big|_{\mathbf{w}=\phi^{-1}(\mathbf{u})} \right\}_{1 \leq j \leq k}$ is a basis of $\check{T}_{\mathbf{u}}S$.

To learn more about this definition of a tangent plane to a manifold at a point using the approach in this subsection, the reader should consult [25, 27]. We will make use of this notation when studying Riemann manifolds in Chapter 9 where this notation proves to be more convenient.

5.4 Vector Fields

Definition 5.4.1

A **vector field** on a k -dimensional manifold $S \subset \mathbb{R}^n$ of class C^1 is a map $F : S \rightarrow \mathbf{T}S = \bigcup_{\mathbf{u} \in S} \mathbf{T}_{\mathbf{u}}S$ defined by $F(\mathbf{u}) = (\mathbf{u}, f(\mathbf{u})) \in \mathbf{T}_{\mathbf{u}}S$ for all $\mathbf{u} \in S$ where $f : S \rightarrow \mathbb{R}^n$.

Let (W, U, ϕ) be a local chart of S . There exists a unique vector field $\tilde{F} : W \rightarrow \bigcup_{\mathbf{w} \in W} \mathbf{T}_{\mathbf{w}}W$ such that $\phi_*(\tilde{F}(\mathbf{w})) = F(\phi(\mathbf{w}))$ for all $\mathbf{w} \in W$. Moreover, if $\tilde{F}(\mathbf{w}) = (\mathbf{w}, \tilde{f}(\mathbf{w})) \in \mathbf{T}_{\mathbf{w}}W$ for all $\mathbf{w} \in W$, then $\tilde{f}$ satisfies $(D\phi(\mathbf{w}))(\tilde{f}(\mathbf{w})) = f(\phi(\mathbf{w}))$ for all $\mathbf{w} \in W$. An equivalent statement is that $f(\mathbf{u}) = (D_{\mathbf{w}}\phi(\phi^{-1}(\mathbf{u})))\tilde{f}(\phi^{-1}(\mathbf{u}))$ for all $\mathbf{u} \in U$. We say that F is of class C^j on S if $\tilde{F}$ is of class C^j on W for all local charts (W, U, ϕ) of S .

The theory of ordinary differential equations can be applied to vector fields on manifolds.

Definition 5.4.2

Suppose that F is a vector field on a k -dimensional manifold S of class C^1 . An **integral curve** of F at $\mathbf{u} \in S$ is a function $\sigma : I \rightarrow S$ where I is an open neighbourhood of the origin such that $\sigma(0) = \mathbf{u}$ and $\sigma_*(t, 1) = F(\sigma(t))$ for all $t \in I$.

Suppose that (W, U, ϕ) is a local chart of the k -dimensional manifold S about $\mathbf{u} \in S$. Suppose that $\tilde{\sigma} : I \rightarrow W$ is the local representation of $\sigma : I \rightarrow S$; namely, $\tilde{\sigma} = \phi^{-1} \circ \sigma$ and so $\sigma = \phi \circ \tilde{\sigma}$. Since

$$\phi_*(\tilde{\sigma}_*(t, 1)) = \sigma_*(t, 1) = F(\sigma(t)) = \phi_*(\tilde{F}(\phi^{-1}(\sigma(t)))) = \phi_*(\tilde{F}(\tilde{\sigma}(t)))$$

for all $t \in I$, we get the following local representation of $\sigma_*(t, 1) = F(\sigma(t))$.

$$(\tilde{\sigma}(t), \tilde{\sigma}'(t)) = \tilde{\sigma}_*(t, 1) = \tilde{F}(\tilde{\sigma}(t)) = (\tilde{\sigma}(t), \tilde{f}(\tilde{\sigma}(t)))$$

for all $t \in I$. Therefore, solving $\sigma_*(t, 1) = F(\sigma(t))$ for t near 0 is equivalent to solving $\tilde{\sigma}'(t) = \tilde{f}(\tilde{\sigma}(t))$ for t near 0. We have reduce the problem to an ordinary differential equation in $\mathbb{R}^k$.

Suppose that $g : S_1 \rightarrow S_2$ is a diffeomorphism between two k -dimensional manifolds of class C^1 . The **pull-back** of a vector field $F : S_2 \rightarrow \mathbf{T}S_2$ is the vector field on S_1 defined by $g^*(F)(\mathbf{u}) = (g^{-1})_*(F(g(\mathbf{u})))$ for all $\mathbf{u} \in S_1$.

For the readers who would like to learn more about vector fields and differential equations, they can find a full chapter on these fabulous topics in [1, 27].

Remark 5.4.3

Suppose that $F : S \rightarrow \mathbf{T}S$ is a vector field on a k -dimensional manifold $S \subset \mathbb{R}^n$ of class C^1 , and that (W, U, ϕ) is a local chart of S . The local representation of the vector field F with

respect to the description of tangent spaces given in Subsection 5.3.2 is given by

$$\tilde{F}(\mathbf{w}) = \sum_{i=1}^k \tilde{f}_i(\mathbf{w}) \frac{\partial}{\partial w_i} \Big|_{\mathbf{w}}$$

for $\mathbf{w} \in W$ where $\tilde{f}_j : W \rightarrow \mathbb{R}$ for $1 \leq j \leq k$. The vector field F is of class C^j on S if, for each local chart (W, U, ϕ) , the function $\tilde{f}_i : W \rightarrow \mathbb{R}$ are of class C^j .

For all $g \in C^\infty(S)$, we have

$$F(\mathbf{u})(g) = \tilde{F}(\phi^{-1}(\mathbf{u}))(g \circ \phi) = \sum_{i=1}^k \tilde{f}_i(\phi^{-1}(\mathbf{u})) \frac{\partial}{\partial w_i} (g \circ \phi) \Big|_{\mathbf{w}=\phi^{-1}(\mathbf{u})}$$

for all $\mathbf{u} \in U$. ♠

5.5 Differential Forms

Definition 5.5.1

A **differential p-form** ω on a k -dimensional manifold S of class C^1 is a map $\omega : S \rightarrow \Omega^p(S) = \bigcup_{\mathbf{u} \in S} \Omega^p(T_{\mathbf{u}}S)$ defined by $\omega(\mathbf{u}) \in \Omega^p(T_{\mathbf{u}}S)$ for all $\mathbf{u} \in S$.

Let ω be a differential p -form on a k -dimensional manifold S of class C^1 . It is important to note that $\omega(\mathbf{u}) \in \Omega^p(T_{\mathbf{u}}S)$ for each $\mathbf{u} \in S$; namely, for each $\mathbf{u} \in S$, $\omega(\mathbf{u})$ is an alternating p -tensor on the k -dimensional subspace $T_{\mathbf{u}}S$ of $T_{\mathbf{u}}\mathbb{R}^n \cong \mathbb{R}^n$.

If (W, U, ϕ) is a local chart of S , then $\phi^*(\omega)$ is a differential p -form on W defined by

$$\begin{aligned} \phi^*(\omega)(\mathbf{w})((\mathbf{w}, \mathbf{y}_1), \dots, (\mathbf{w}, \mathbf{y}_p)) &= \phi^*(\omega(\phi(\mathbf{w})))((\mathbf{w}, \mathbf{y}_1), \dots, (\mathbf{w}, \mathbf{y}_p)) \\ &= \omega(\phi(\mathbf{w}))(\phi_*(\mathbf{w}, \mathbf{y}_1), \dots, \phi_*(\mathbf{w}, \mathbf{y}_p)) \\ &= \omega(\phi(\mathbf{w}))((\phi(\mathbf{w}), (D\phi(\mathbf{w}))\mathbf{y}_1), \dots, (\phi(\mathbf{w}), (D\phi(\mathbf{w}))\mathbf{y}_p)) \end{aligned}$$

for all $\mathbf{w} \in W$ and $(\mathbf{w}, \mathbf{y}_i) \in T_{\mathbf{w}}W$ with $1 \leq i \leq p$. Thus $\phi^*(\omega) : W \rightarrow \bigcup_{\mathbf{w} \in W} \Omega^p(T_{\mathbf{w}}W)$ and

$\phi^*(\omega)(\mathbf{w}) \in \Omega^p(T_{\mathbf{w}}W)$ for all $\mathbf{w} \in W$ as expected.

Remark 5.5.2

If S is a k -dimensional manifold of class C^1 , then the set $T^*S = \bigcup_{\mathbf{u} \in S} (T_{\mathbf{u}}S)^* = \bigcup_{\mathbf{u} \in S} \Omega^1(T_{\mathbf{u}}S)$ is called the **cotangent bundle of S**. As the tangent bundle TS , it has the structure of a vector bundle. ♠

Definition 5.5.3

We say that a differential p -form ω is of class C^j on a k -dimensional manifold S of class C^j if, for every local chart (W, U, ϕ) of S , the differential p -form $\phi^*(\omega)$ is of class C^j on W .

More explicitly, let ω be a differential p -form on a k -dimensional manifold S of class C^j and (W, U, ϕ) be a local chart of S . We have seen in Section 4.5 that we can express $\phi^*(\omega)$ as

$$\phi^*(\omega) = \sum_{1 \leq i_1 < i_2 < \dots < i_p \leq k} \omega_{i_1, i_2, \dots, i_p} dw_{i_1} \wedge dw_{i_2} \wedge \dots \wedge dw_{i_p}$$

where $\omega_{i_1, i_2, \dots, i_p} : W \rightarrow \mathbb{R}$ for all $1 \leq i_1 < i_2 < \dots < i_p \leq k$. If, for all local chart (W, U, ϕ) , the functions $\omega_{i_1, i_2, \dots, i_p} : W \rightarrow \mathbb{R}$ are of class C^j , then ω is a differential p -form of class C^j on S .

Suppose that S_1 and S_2 are two manifolds of class C^1 and that $f : S_1 \rightarrow S_2$ is a function of class C^1 . Let ω is a differential p -form on S_2 . As expected, the **pull-back** of ω by f is defined by

$$\begin{aligned} f^*(\omega)(\mathbf{u}_1)((\mathbf{u}_1, \mathbf{x}_1), \dots, (\mathbf{u}_1, \mathbf{x}_p)) &= f^*(\omega(f(\mathbf{u}_1)))(\mathbf{u}_1, \mathbf{x}_1, \dots, (\mathbf{u}_1, \mathbf{x}_p)) \\ &= \omega(f(\mathbf{u}_1))(f_*(\mathbf{u}_1, \mathbf{x}_1), \dots, f_*(\mathbf{u}_1, \mathbf{x}_p)) \end{aligned}$$

for all $\mathbf{u}_1 \in S_1$ and $(\mathbf{u}_1, \mathbf{x}_i) \in T_{\mathbf{u}_1}S_1$, where f_* is defined in the paragraph following Theorem 5.3.5.

The properties that we have listed for differential forms on open subsets of $\mathbb{R}^n$ in Section 4.5 are also true for differential forms on manifolds. The proofs are generally almost identical since they are often reduced to computation with alternating tensors on the tangent space at each point of the manifold. For instance, we have the following result that generalize (4) of Proposition 4.5.4 to manifolds.

Proposition 5.5.4

Suppose that S_1 and S_2 are two manifolds of class C^1 and that $f : S_1 \rightarrow S_2$ is a function of class C^1 . If ω_i is a differential p_i -form on S_2 for $i = 1, 2$, then $f^*(\omega_1 \wedge \omega_2) = f^*(\omega_1) \wedge f^*(\omega_2)$.

Proof.

We have

$$\begin{aligned} f^*(\omega_1 \wedge \omega_2)(\mathbf{u})((\mathbf{u}, \mathbf{x}_1), \dots, (\mathbf{u}, \mathbf{x}_{p_1+p_2})) &= (\omega_1(f(\mathbf{u})) \wedge \omega_2(f(\mathbf{u}))) (f_*(\mathbf{u}, \mathbf{x}_1), \dots, f_*(\mathbf{u}, \mathbf{x}_{p_1+p_2})) \\ &= \frac{1}{p_1! p_2!} \sum_{\sigma \in S_{p_1+p_2}} (\omega_1(f(\mathbf{u})) \otimes \omega_2(f(\mathbf{u}))) (f_*(\mathbf{u}, \mathbf{x}_{\sigma(1)}), \dots, f_*(\mathbf{u}, \mathbf{x}_{\sigma(p_1+p_2)})) \\ &= \frac{1}{p_1! p_2!} \sum_{\sigma \in S_{p_1+p_2}} \omega_1(f(\mathbf{u})) (f_*(\mathbf{u}, \mathbf{x}_{\sigma(1)}), \dots, f_*(\mathbf{u}, \mathbf{x}_{\sigma(p_1)})) \otimes \end{aligned}$$

$$\begin{aligned}
& \omega_2(f(\mathbf{u})) (f_*(\mathbf{u}, \mathbf{x}_{\sigma(p_1+1)}), \dots, f_*(\mathbf{u}, \mathbf{x}_{\sigma(p_1+p_2)})) \\
&= \frac{1}{p_1! p_2!} \sum_{\sigma \in S_{p_1+p_2}} (f^*(\omega_1))(\mathbf{u})((\mathbf{u}, \mathbf{x}_{\sigma(1)}), \dots, (\mathbf{u}, \mathbf{x}_{\sigma(p_1)})) \otimes \\
& \quad (f^*(\omega_2))(\mathbf{u})((\mathbf{u}, \mathbf{x}_{\sigma(p_1+1)}), \dots, (\mathbf{u}, \mathbf{x}_{\sigma(p_1+p_2)})) \\
&= (f^*(\omega_1) \wedge f^*(\omega_2))(\mathbf{u})((\mathbf{u}, \mathbf{x}_1), \dots, (\mathbf{u}, \mathbf{x}_{p_1+p_2}))
\end{aligned}$$

for all $\mathbf{u} \in S_1$ and $(\mathbf{u}, \mathbf{x}_i) \in T_{\mathbf{u}}S_1$ with $1 \leq i \leq p_1 + p_2$. ■

If ω is a differential p -form is of class C^j on a k -dimensional manifold S of class C^j , how are we going to define $d\omega$? We cannot simply derive the ω as we did for differential p -forms on $\mathbb{R}^n$ because ω is only defined on S which is not an open subset of $\mathbb{R}^n$ when $k < n$. The following theorem will answer this question.

Theorem 5.5.5

Let ω be a p -form of class C^1 on a k -dimensional manifold S of class C^1 . There exists a unique differential $(p+1)$ -form μ on S such that $\phi^*(\mu) = d(\phi^*(\omega))$ on W for every local chart (W, U, ϕ) of S . This unique differential $(p+1)$ -form is denoted $d\omega$.

Proof.

Let (W, U, ϕ) be a local chart of S . The differential $(p+1)$ -form μ on S is defined locally by

$$\begin{aligned}
& \mu|_U(\mathbf{u})((\mathbf{u}, \mathbf{x}_1), (\mathbf{u}, \mathbf{x}_2), \dots, (\mathbf{u}, \mathbf{x}_{p+1})) \\
&= d(\phi^*(\omega))(\mathbf{w})((\mathbf{w}, \mathbf{y}_1), (\mathbf{w}, \mathbf{y}_2), \dots, (\mathbf{w}, \mathbf{y}_{p+1}))
\end{aligned}$$

for $\mathbf{u} \in U$ and $(\mathbf{u}, \mathbf{x}_i) \in T_{\mathbf{u}}S$, where $(\mathbf{u}, \mathbf{x}_i) = \phi_*(\mathbf{w}, \mathbf{y}_i)$ for $\mathbf{w} \in W$ and $(\mathbf{w}, \mathbf{y}_i) \in T_{\mathbf{w}}W$ with $1 \leq i \leq p+1$. Since $\phi_* : T_{\mathbf{w}}W \rightarrow T_{\phi(\mathbf{w})}S$ is a linear isomorphism for every $\mathbf{w} \in W$, there is a unique $\mathbf{y}_i$ for each $\mathbf{x}_i$.

i) The previous definition is independent of the local chart used. Suppose that (W_i, U_i, ϕ_i) for $i = 1, 2$ are two local charts of S . Since $\phi_2^{-1} \circ \phi_1 : \phi_1(U_1 \cap U_2) \rightarrow \phi_2(U_1 \cap U_2)$ is a diffeomorphism, the map $(\phi_2^{-1} \circ \phi_1)_* : T_{\mathbf{w}}W_1 \rightarrow T_{\phi_2^{-1}(\phi_1(\mathbf{w}))}W_2$ is a linear isomorphism for every $\mathbf{w} \in \phi_1(U_1 \cap U_2)$. Suppose that $\phi_1(\mathbf{w}_1) = \phi_2(\mathbf{w}_2) = \mathbf{u} \in U_1 \cap U_2$. Then $\mathbf{w}_2 = \phi_2^{-1}(\phi_1(\mathbf{w}_1))$.

If $(\phi_1)_*(\mathbf{w}_1, \mathbf{y}_j) = (\mathbf{u}, \mathbf{x}_j)$ for $1 \leq j \leq p+1$, then

$$\begin{aligned}
& \mu|_{U_1}(\mathbf{u})((\mathbf{u}, \mathbf{x}_1), (\mathbf{u}, \mathbf{x}_2), \dots, (\mathbf{u}, \mathbf{x}_{p+1})) \\
&= d(\phi_1^*(\omega))(\mathbf{w}_1)((\mathbf{w}_1, \mathbf{y}_1), (\mathbf{w}_1, \mathbf{y}_2), \dots, (\mathbf{w}_1, \mathbf{y}_{p+1})) \\
&= d((\phi_2 \circ \phi_2^{-1} \circ \phi_1)^*(\omega))(\mathbf{w}_1)((\mathbf{w}_1, \mathbf{y}_1), (\mathbf{w}_1, \mathbf{y}_2), \dots, (\mathbf{w}_1, \mathbf{y}_{p+1})) \\
&= d((\phi_2^{-1} \circ \phi_1)^*(\phi_2^*(\omega)))(\mathbf{w}_1)((\mathbf{w}_1, \mathbf{y}_1), (\mathbf{w}_1, \mathbf{y}_2), \dots, (\mathbf{w}_1, \mathbf{y}_{p+1})) \\
&= (\phi_2^{-1} \circ \phi_1)^*(d(\phi_2^*(\omega)))(\mathbf{w}_1)((\mathbf{w}_1, \mathbf{y}_1), (\mathbf{w}_1, \mathbf{y}_2), \dots, (\mathbf{w}_1, \mathbf{y}_{p+1})) \\
&= d(\phi_2^*(\omega))(\mathbf{w}_2)((\mathbf{w}_2, D(\phi_2^{-1} \circ \phi_1)(\mathbf{w}_1)\mathbf{y}_1), (\mathbf{w}_2, D(\phi_2^{-1} \circ \phi_1)(\mathbf{w}_1)\mathbf{y}_2), \dots, \\
& \quad (\mathbf{w}_2, D(\phi_2^{-1} \circ \phi_1)(\mathbf{w}_1)\mathbf{y}_{p+1}))
\end{aligned}$$

$$= \mu|_{U_2}(\mathbf{u})((\mathbf{u}, \mathbf{x}_1), (\mathbf{u}, \mathbf{x}_2), \dots, (\mathbf{u}, \mathbf{x}_{p+1})) \quad (5.5.1)$$

because

$$\begin{aligned} (\phi_2)_* (\mathbf{w}_2, D(\phi_2^{-1} \circ \phi_1)(\mathbf{w}_1) \mathbf{y}_j) &= (\phi_2(\mathbf{w}_2), D(\phi_2)(\mathbf{w}_2) D(\phi_2^{-1} \circ \phi_1)(\mathbf{w}_1) \mathbf{y}_j) \\ &= (\mathbf{u}, D\phi_1(\mathbf{w}_1) \mathbf{y}_j) = (\mathbf{u}, \mathbf{x}_j) \end{aligned}$$

for $1 \leq j \leq p+1$ as required. Note that we have used Theorem 4.5.7 to get the fourth equality in (5.5.1).

ii) Finally, for each local chart (W, U, ϕ) of S , we have by construction that

$$\begin{aligned} \phi^*(\mu)(\mathbf{w})((\mathbf{w}, \mathbf{y}_1), (\mathbf{w}, \mathbf{y}_2), \dots, (\mathbf{w}, \mathbf{y}_{p+1})) \\ &= \mu|_U(\phi(\mathbf{w}))(\phi_*(\mathbf{w}, \mathbf{y}_1), \phi_*(\mathbf{w}, \mathbf{y}_2), \dots, \phi_*(\mathbf{w}, \mathbf{y}_{p+1})) \\ &= d(\phi^*(\omega))(\mathbf{w})((\mathbf{w}, \mathbf{y}_1), (\mathbf{w}, \mathbf{y}_2), \dots, (\mathbf{w}, \mathbf{y}_{p+1})) \end{aligned}$$

for all $\mathbf{w} \in W$ and $(\mathbf{w}, \mathbf{y}_i) \in T_{\mathbf{w}}W$ with $1 \leq i \leq p+1$. Thus $\phi^*(\mu) = d(\phi^*(\omega))$ on W . ■

Since the derivative of a differential form is defined locally, we may generalize (4) of Theorem 4.5.7.

Proposition 5.5.6

Suppose that S_1 and S_2 are two manifolds of class C^1 and that $f : S_1 \rightarrow S_2$ is a function of class C^1 . If ω be a p -form of class C^1 on S_2 . then $f^*(d\omega) = d(f^*(\omega))$.

Proof.

Let (W_i, U_i, ϕ_i) be a local chart of S_i for $1 \leq i \leq 2$. Without loss of generality, we may assume that $f(U_1) \subset U_2$. We have by definition of the derivative of a differential form given in Theorem 5.5.5 that $d(\phi_1^*(f^*(\omega))) = \phi_1^*(d(f^*(\omega)))$ on W_1 . Moreover, we have

$$\begin{aligned} d(\phi_1^*(f^*(\omega))) &= d((f \circ \phi_1)^*(\omega)) = d((\phi_2^{-1} \circ f \circ \phi_1)^*(\phi_2^*(\omega))) \\ &= (\phi_2^{-1} \circ f \circ \phi_1)^* d(\phi_2^*(\omega)) = (\phi_2^{-1} \circ f \circ \phi_1)^*(\phi_2^*(d\omega)) \\ &= (f \circ \phi_1)^*(d\omega) = \phi_1^*(f^*(d\omega)), \end{aligned}$$

where the third equality comes from (4) of Theorem 4.5.7 and the fourth equality from the definition of the derivative of a differential form given in Theorem 5.5.5. Thus $\phi_1^*(d(f^*(\omega))) = \phi_1^*(f^*(d\omega))$. Since (W_1, U_1, ϕ_1) is an arbitrary local chart of S_1 , we get that $d(f^*(\omega)) = f^*(d\omega)$. ■

Remark 5.5.7

1. We would like to warn the reader that a different notation is often used in the literature on differential geometry. Some authors write

$$\omega = \sum_{1 \leq i_1 < i_2 < \dots < i_p \leq k} \omega_{i_1, i_2, \dots, i_p} dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_p},$$

where $\omega_{i_1, i_2, \dots, i_p} : S \rightarrow \mathbb{R}$ as a differential q -form on a k -dimensional manifold S of class S^1 . This in fact represents the local representation of the differential q -form ω . Their definition of a local chart (W, U, ϕ) of S is slightly different than our definition. Instead of $\phi : W \rightarrow U$, they use $\phi : U \rightarrow W$. The x_i for $1 \leq i \leq k$ represent the local coordinates in $\mathbb{R}^k$ instead of the w_i in our definition. Therefore, to match our theory to their, all the local formulae that we have presented have to be reformulated according to this new convention.

2. As we have seen in Section 4.5, we may write a differential p -form ω on $\mathbb{R}^n$ as

$$\omega = \sum_{1 \leq i_1 < i_2 < \dots < i_p \leq n} \omega_{i_1, i_2, \dots, i_p} dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_p}, \quad (5.5.2)$$

where $\omega_{i_1, i_2, \dots, i_p} : \mathbb{R}^n \rightarrow \mathbb{R}$ for all indices. If S is a k -dimensional manifold of class C^1 in $\mathbb{R}^n$, then it is also frequent in the literature to use this notation to talk about $\omega|_S$. Recall that dx_i for $1 \leq i \leq n$ are differential 1-form on $\mathbb{R}^n$ defined by $dx_i(\mathbf{u})(\mathbf{u}, \mathbf{x}) = \pi_i(\mathbf{x}) = x_i$ for all $(\mathbf{u}, \mathbf{x}) \in T_{\mathbf{u}}\mathbb{R}^n$, $\mathbf{u} \in \mathbb{R}^n$ and $1 \leq i \leq n$. Since $T_{\mathbf{u}}S$ is a linear subspace of $T_{\mathbf{u}}\mathbb{R}^n$ for all $\mathbf{u} \in S$, we have that $dx_i|_S$ is a differential 1-form on S where $dx_i(\mathbf{u})(\mathbf{u}, \mathbf{x}) = \pi_i(\mathbf{x}) = x_i$ for all $(\mathbf{u}, \mathbf{x}) \in T_{\mathbf{u}}S$, $\mathbf{u} \in S$ and $1 \leq i \leq n$. Thus, when we write dx_{i_j} in (5.5.2), we really mean $dx_{i_j}|_S$. As a warning, when $k < n$, the set $\{dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_p} : 1 \leq i_1 < i_2 < \dots < i_p \leq n\}$ is a set of generators for $\Omega^p(T_{\mathbf{u}}S)$ but not a basis.

The fact that we do not have a precise action of the differential 1-form dx_i on S is not going to be an issue in the integrals that we will consider later because differential p -form as in (5.5.2) are reinterpreted in terms of the local coordinates given by the local charts on S using substitution.



Remark 5.5.8

Suppose that $F : S \rightarrow TS$ is a vector field on a k -dimensional manifold $S \subset \mathbb{R}^n$ of class C^1 . Moreover, suppose that $g : S \rightarrow \mathbb{R}$ is a function of class C^∞ . Then dg is a differential 1-form on S . We may therefore compute $dg(\mathbf{u})(F(\mathbf{u}))$ for all $\mathbf{u} \in S$.

Suppose that (W, U, ϕ) is a local chart of S . If we work on the local chart (W, U, ϕ) and assume that $\mathbf{u} = \phi(\mathbf{w})$ for $\mathbf{w} \in W$, then we get

$$\begin{aligned} dg(\mathbf{u})(F(\mathbf{u})) &= dg(\phi(\mathbf{w}))(\phi_*(\tilde{F}(\mathbf{w}))) = \phi^*(dg)(\mathbf{w})(\tilde{F}(\mathbf{w})) = d(\phi^*(g))(\mathbf{w})(\tilde{F}(\mathbf{w})) \\ &= d(g \circ \phi)(\mathbf{w})(\tilde{F}(\mathbf{w})) = D(g \circ \phi)(\mathbf{w})(\tilde{F}(\mathbf{w})) \\ &= \sum_{i=1}^k \tilde{f}_i(\mathbf{w}) \frac{\partial}{\partial w_i} (g \circ \phi) \Big|_{\mathbf{w}} = \sum_{i=1}^k \tilde{f}_i(\phi^{-1}(\mathbf{u})) \frac{\partial}{\partial w_i} (g \circ \phi) \Big|_{\mathbf{w}=\phi^{-1}(\mathbf{u})} \end{aligned}$$

for all $\mathbf{u} \in U$, where the fifth equality comes from Definition 4.5.2. With respect to the description of tangent spaces given in Subsection 5.3.2, we have that F has the locale representation

$\tilde{F}(\mathbf{w}) = \sum_{i=1}^k \tilde{f}_i(\mathbf{w}) \frac{\partial}{\partial w_i} \Big|_{\mathbf{w}}$ for $\mathbf{w} \in W$. Therefore $dg(\mathbf{u})(F(\mathbf{u})) = F(\mathbf{u})(g)$ for all

$\mathbf{u} \in U$. The expression $dg(\mathbf{u})(F(\mathbf{u}))$ is called the **directional derivative** or **Lie derivative** of g along F at $\mathbf{u}$.



5.6 Integration on Manifolds

We first expand a little Definition 4.7.1.

Definition 5.6.1

Let k be a positive integer and $\omega = f dx_1 \wedge dx_2 \wedge \dots \wedge dx_k$ be a differential k -form defined on an open set $V \subset \mathbb{R}^k$. The **integral of ω** on V is defined by $\int_V \omega = \int_V f(\mathbf{x}) d\mathbf{x}$.

Definition 5.6.2

Let S be a k -dimensional oriented manifold of class C^1 and (W, U, ϕ) be an orientation preserving local chart of S . Suppose that ω is a differential k -form on S with compact support in U . We define the **integral of ω** on S as

$$\int_S \omega = \int_W \phi^*(\omega) .$$

The next proposition shows that this definition is independent of the local chart containing the support of ω that is used.

Proposition 5.6.3

Let V and W be two open subsets of H_k and suppose that $g : W \rightarrow V$ is an orientation preserving diffeomorphism. If ω is a differential k -form with compact support in V , then $g^*(\omega)$ is a differential k -form with compact support in W and

$$\int_V \omega = \int_W g^*(\omega) .$$

Proof.

Since $g : W \rightarrow V$ is a diffeomorphism, it follows that $g^*(\omega)$ is a differential k -form on W . Since $\text{supp } \omega$ is a compact set and $g^{-1} : V \rightarrow W$ is continuous, we have that $\text{supp } g^*(\omega) = g^{-1}(\text{supp } \omega)$ is a compact set.

If $\omega = f dx_1 \wedge dx_2 \wedge \dots \wedge dx_k$, then we have

$$\begin{aligned} \int_W g^*(\omega) &= \int_W (f \circ g)(\det Dg) dx_1 \wedge dx_2 \wedge \dots \wedge dx_k \\ &= \int_W (f \circ g)(\det Dg) dx_1 dx_2 \dots dx_k = \int_V f dx_1 dx_2 \dots dx_k = \int_V \omega . \end{aligned}$$

The first equality is a consequence of the definition of ω combined with Theorem 4.5.5 and the second equality is the definition of the integral of a differential k -form over an open

subset of H_k . The third equality comes from the change of variable theorem and the fact that $\det Dg > 0$ on W since g is orientation preserving. Finally, the last equality is again the definition of the integral of a differential k -form over an open subset of H_k .

Note that we used the extended definition of diffeomorphism between open subsets of H_k when $\text{supp } \omega \cap \partial H_k \neq \emptyset$. ■

It follows from the previous proposition that if (W_i, U_i, ϕ_i) for $1 \leq i \leq 2$ are two orientation preserving local charts of a k -dimensional oriented manifold S and that ω is a differential k -form on S with $\text{supp } \omega \subset U_1 \cap U_2$, then $\int_S \omega = \int_{W_1} \phi_1^*(\omega) = \int_{W_2} \phi_2^*(\omega)$. To be more precise, let $\tilde{W}_i = \phi_i^{-1}(U_1 \cap U_2)$ for $i = 1, 2$ and $g = \phi_2^{-1} \circ \phi_1$. Then g is an orientation preserving diffeomorphism between $\tilde{W}_1$ and $\tilde{W}_2$. Hence

$$\int_{W_2} \phi_2^*(\omega) = \int_{\tilde{W}_2} \phi_2^*(\omega) = \int_{\tilde{W}_1} g^*(\phi_2^*(\omega)) = \int_{\tilde{W}_1} (\phi_2 \circ g)^*(\omega) = \int_{\tilde{W}_1} \phi_1^*(\omega) = \int_{W_1} \phi_1^*(\omega).$$

To define the integral of a differential form on a manifold, we need the following definition.

Definition 5.6.4

Let S be a k -dimensional oriented manifold of class C^1 and $\mathcal{A} = \{(W_\alpha, U_\alpha, \phi_\alpha)\}_{\alpha \in A}$ be an atlas of orientation preserving local charts for S . A **partition of unity** subordinate to the atlas $\mathcal{A}$ is a collection $\{\psi_j\}_{j \in \mathbb{N}^+}$ of functions of class C^∞ such that:

1. $0 \leq \psi_j(\mathbf{x}) \leq 1$ for all $\mathbf{x} \in S$ and all $j \in \mathbb{N}^+$.
2. For each $\mathbf{x} \in S$, there exists an open neighbourhood $U_{\mathbf{x}} \subset S$ of $\mathbf{x}$ such that $\psi_j|_{U_{\mathbf{x}}} = 0$ for all but a finite number of ψ_j .
3. For each $\mathbf{x} \in S$, $\sum_{j \in \mathbb{N}^+} \psi_j(\mathbf{x}) = 1$. This sum is finite according to (2).
4. For each $j \in \mathbb{N}^+$, we have that $\text{supp } \psi_j$ is a compact set with $\text{supp } \psi_j \subset U_{\alpha_j}$ for some $\alpha_j \in A$.

It follows from Theorem 3.4.1 that we can always find a partition of unity subordinate to the atlas $\mathcal{A}$. It suffices to take a partition of unity $\{\psi_j\}_{j \in \mathbb{N}^+}$ subordinate to the open cover $\{V_\alpha\}_{\alpha \in A}$ of S , where V_α is an open set in $\mathbb{R}^n$ such that $U_\alpha = S \cap V_\alpha$ for $\alpha \in A$.

The following result is a direct consequence of Proposition 3.4.4.

Proposition 5.6.5

Let S be a k -dimensional oriented manifold of class C^1 . Let $\mathcal{A} = \{(W_\alpha, U_\alpha, \phi_\alpha)\}_{\alpha \in A}$ be an atlas of orientation preserving local charts for S and $\{\psi_j\}_{j \in \mathbb{N}^+}$ be a partition of unity subordinate to $\mathcal{A}$. If $K \subset S$ is a compact set, then there exist only a finite number of functions ψ_j such that $\phi_j|_K \neq 0$.

Definition 5.6.6

Let S be a k -dimensional oriented manifold of class C^1 and $\mathcal{A} = \{(W_\alpha, U_\alpha, \phi_\alpha)\}_{\alpha \in A}$ be an atlas of orientation preserving local charts for S . Let $\{\psi_j\}_{j \in \mathbb{N}^+}$ be a partition of unity subordinate to $\mathcal{A}$. Suppose that ω is a differential k -form on S with compact support. For each $j \in \mathbb{N}^+$, the differential k -forms $\omega_j = \psi_j \omega$ has a compact support in U_{α_j} for some $\alpha_j \in A$. The **integral** of ω on S is defined by

$$\int_S \omega = \sum_{j \in \mathbb{N}^+} \int_S \omega_j = \sum_{j \in \mathbb{N}^+} \int_{W_{\alpha_j}} \phi_j^*(\omega_j) . \quad (5.6.1)$$

The next proposition shows that the definition of the integral of a differential k -form on a k -dimensional oriented manifold of class C^1 is not ambiguous.

Proposition 5.6.7

1. The sum (5.6.1) contains only a finite number of non-zero terms.
2. The value of the sum in (5.6.1) is independent of the atlas for S chosen.

Proof.

The proof used the information provided in Definition 5.6.6.

- 1) According to Proposition 5.6.5, we have that $\omega_j = \psi_j \omega = 0$ for all $j \in \mathbb{N}^+$ but a finite number of $j \in \mathbb{N}^+$ because ω has a compact support.
- 2) Suppose that $\mathcal{A} = \{(W_\alpha, U_\alpha, \phi_\alpha)\}_{\alpha \in A}$ and $\tilde{\mathcal{A}} = \{(\tilde{W}_{\tilde{\alpha}}, \tilde{U}_{\tilde{\alpha}}, \tilde{\phi}_{\tilde{\alpha}})\}_{\tilde{\alpha} \in \tilde{A}}$ are two atlases of orientation preserving local charts for S . Let $\{\psi_j\}_{j \in \mathbb{N}^+}$ be a partition of unity subordinate to $\mathcal{A}$ and $\{\tilde{\psi}_{\tilde{j}}\}_{\tilde{j} \in \mathbb{N}^+}$ be a partition of unity subordinate to $\tilde{\mathcal{A}}$. Then

$$\begin{aligned} \sum_{j_1 \in \mathbb{N}^+} \int_S \psi_{j_1} \omega &= \sum_{j_1 \in \mathbb{N}^+} \int_S \psi_{j_1} \left(\sum_{j_2 \in \mathbb{N}^+} \tilde{\psi}_{j_2} \right) \omega = \sum_{j_1 \in \mathbb{N}^+} \int_S \left(\sum_{j_2 \in \mathbb{N}^+} \tilde{\psi}_{j_2} (\psi_{j_1} \omega) \right) \\ &= \sum_{j_1 \in \mathbb{N}^+} \sum_{j_2 \in \mathbb{N}^+} \int_S \tilde{\psi}_{j_2} \psi_{j_1} \omega = \sum_{j_2 \in \mathbb{N}^+} \sum_{j_1 \in \mathbb{N}^+} \int_S \tilde{\psi}_{j_2} \psi_{j_1} \omega = \sum_{j_2 \in \mathbb{N}^+} \int_S \left(\sum_{j_1 \in \mathbb{N}^+} \tilde{\psi}_{j_2} \psi_{j_1} \omega \right) \\ &= \sum_{j_2 \in \mathbb{N}^+} \int_S \tilde{\psi}_{j_2} \left(\sum_{j_1 \in \mathbb{N}^+} \psi_{j_1} \right) \omega = \sum_{j_2 \in \mathbb{N}^+} \int_S \tilde{\psi}_{j_2} \omega . \end{aligned}$$

The first equality comes from $\sum_{j_2 \in \mathbb{N}^+} \tilde{\psi}_{j_2}(\mathbf{x}) = 1$ for all $\mathbf{x}$. Since the summation with respect

to j_2 is finite because of (1), we may pull the summation with respect to j_2 out of the integral to get the third equality. Because $\text{supp}(\psi_{j_1} \tilde{\psi}_{j_2}) \cap \text{supp} \omega \neq \emptyset$ for only a finite number of $\psi_{j_1} \tilde{\psi}_{j_2}$, the summations may be inverted to give the fourth equality. Since the summation with respect to j_1 is finite because of (1), we may bring the summation with respect to j_1 in

the integral to get the fifth equality. The last equality comes from $\sum_{j_1 \in \mathbb{N}^+} \psi_{j_1}(\mathbf{x}) = 1$ for all $\mathbf{x}$. ■

There is a change of variable formula for integration of differential forms on manifolds.

Proposition 5.6.8

Let S and $\tilde{S}$ be two k -dimensional manifolds of class C^1 . Suppose that ω is a differential k -form on $\tilde{S}$ with compact support and that $g : S \rightarrow \tilde{S}$ is an orientation preserving diffeomorphism. Then $\int_{\tilde{S}} \omega = \int_S g^*(\omega)$.

Proof.

Let $K = \text{supp } \omega$. Then $\text{supp } g^*(\omega) = g^{-1}(K)$. This is a compact set because g^{-1} is continuous and the image of a compact set by a continuous function is a compact set. Therefore, our definition of integral of differential forms on manifolds can be used for both ω and $g^*(\omega)$.

Let $\mathcal{A} = \{(W_\alpha, U_\alpha, \phi_\alpha)\}_{\alpha \in A}$ be an atlas of orientation preserving local charts for S , and $\{\psi_j\}_{j \in \mathbb{N}^+}$ be a partition of unity subordinate to $\mathcal{A}$. Then $\mathcal{B} = \{(W_\alpha, g(U_\alpha), g \circ \phi_\alpha)\}_{\alpha \in A}$ is an atlas of orientation preserving local charts for $\tilde{S}$ and $\{\psi_j \circ g^{-1}\}_{j \in \mathbb{N}^+}$ is a partition of unity subordinate to $\mathcal{B}$. If $\alpha_j \in A$ is an index such that $\text{supp } \psi_j \subset U_{\alpha_j}$, then $\text{supp}(\psi_j \circ g^{-1}) \subset g(U_{\alpha_j})$. Hence

$$\begin{aligned} \int_S g^*(\omega) &= \sum_{j \in \mathbb{N}^+} \int_{U_{\alpha_j}} \psi_j g^*(\omega) = \sum_{j \in \mathbb{N}^+} \int_{W_{\alpha_j}} \phi_{\alpha_j}^*(\psi_j g^*(\omega)) = \sum_{j \in \mathbb{N}^+} \int_{W_{\alpha_j}} (\psi_j \circ \phi_{\alpha_j}) \phi_{\alpha_j}^*(g^*(\omega)) \\ &= \sum_{j \in \mathbb{N}^+} \int_{W_{\alpha_j}} (\psi_j \circ g^{-1} \circ (g \circ \phi_{\alpha_j})) (g \circ \phi_{\alpha_j})^*(\omega) = \sum_{j \in \mathbb{N}^+} \int_{W_{\alpha_j}} (g \circ \phi_{\alpha_j})^*((\psi_j \circ g^{-1}) \omega) \\ &= \sum_{j \in \mathbb{N}^+} \int_{g(U_{\alpha_j})} (\psi_j \circ g^{-1}) \omega = \int_{\tilde{S}} \omega. \end{aligned}$$

Before stating and proving the main theorem of this chapter, we need to clarify some technical details.

Lemma 5.6.9

Suppose that v is a differential k -form on an open set $W \subset H_k$ and $\text{supp } v$ is a compact subset of W . There exists a finite collection $\{\sigma_m\}_{1 \leq m \leq M}$ of singular k -cubes of class C^1 such that:

1. $\sigma_m(I_k)$ is a closed cube in W with sides parallel to the axes in H_k ,
2. $\sigma_{m_1}(I_k) \cap \sigma_{m_2}(I_k)$ for $m_1 \neq m_2$ is either the empty set or a common side of these two closed cubes (namely, $(\sigma_{m_1})_{j,1}(I_{k-1}) = (\sigma_{m_2})_{j,0}(I_{k-1})$ or $(\sigma_{m_1})_{j,0}(I_{k-1}) = (\sigma_{m_2})_{j,1}(I_{k-1})$ for some $1 \leq j \leq k$), and

3. (a) $\text{supp } v \subset \left(\bigcup_{m=1}^M \sigma_m(I_k) \right)^\circ$ if $\text{supp } v \cap \partial H_k = \emptyset$ or
- (b) $\text{supp } v \subset \bigcup_{m=1}^M \sigma_m(I_k)$ and $\text{supp } v \cap \partial \left(\bigcup_{m=1}^M \sigma_m(I_k) \right) \subset \partial H_k$ (the topological boundaries) if $\text{supp } v \cap \partial H_k \neq \emptyset$.

Proof.

Since $\text{supp } v$ is a compact subset of the open set W of H_k with the induced topology from $\mathbb{R}^k$, there exists an open set $V \subset H_k$ such that $\text{supp } v \subset V \subset \bar{V} \subset W$. For each $\mathbf{w} \in \text{supp } v$, choose an open cube $B_{\mathbf{w}} \subset V$ in H_k such that $\mathbf{w} \in B_{\mathbf{w}}$. The collection $\{B_{\mathbf{w}}\}_{\mathbf{w} \in \text{supp } v}$ is an open cover of the compact set $\text{supp } v$. Therefore, there exists a finite subcover $\{B_{\mathbf{w}_j}\}_{1 \leq j \leq J}$ of $\text{supp } v$. By subdividing the cubes $\{\bar{B}_{\mathbf{w}_j}\}_{1 \leq j \leq J}$, we may get a finite collections of closed cubes $\{B_m\}_{1 \leq m \leq M}$ such that $\text{supp } v \subset \bigcup_{m=1}^M B_m \subset \bar{V} \subset W$ and $B_{m_1} \cap B_{m_2}$ for $m_1 \neq m_2$ is either the empty set or a common side of these two closed cubes (Figure 5.8 and 5.9). Each cube B_m is the image of an affine transformation $\sigma_m : I_k \rightarrow B_m$. These σ_m are obviously singular k -cubes of class C^1 . ■

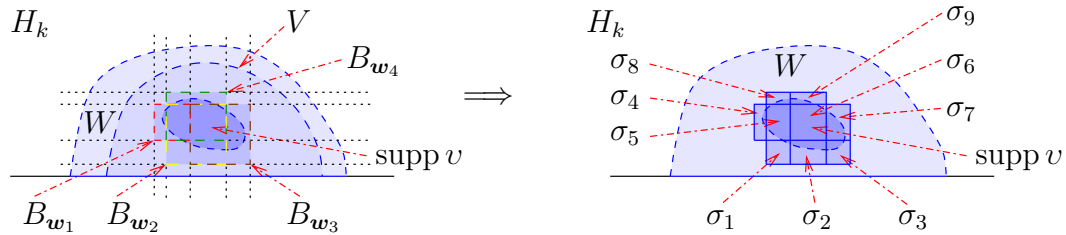


Figure 5.8: The collection of open boxes $\{B_{\mathbf{w}_i}\}_{1 \leq i \leq 4}$ on the left cover $\text{supp } v$. These boxes can be subdivided to give the singular k -cubes $\{\sigma_i\}_{1 \leq i \leq 9}$ on the right whose images cover $\text{supp } v$.

Lemma 5.6.10

Let S be an orientable k -dimensional manifold of class C^1 with boundary. Suppose that $\mu_{\mathbf{u}}$ is the chosen orientation on $T_{\mathbf{u}}S$ for every $\mathbf{u} \in S$, that (W, U, ϕ) is an orientation preserving local chart on S such that $U \cap \partial S \neq \emptyset$ and that ω is a differential $(k-1)$ -form on S with compact support in U such that $\text{supp } \omega \cap \partial S \neq \emptyset$ (Figure 5.10). If $\{\sigma_m\}_{1 \leq m \leq M}$ is a finite collection of singular k -cubes satisfying (3(b)) of Lemma 5.6.9 with $v = \phi^*(\omega)$, then $\int_{\partial S} \omega = \int_{\partial \sigma} \phi^*(\omega)$.

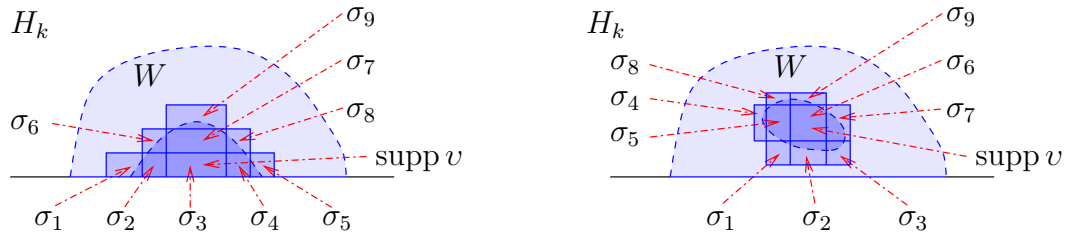


Figure 5.9: Covering of the support of the differential k -form v defined on W by singular k -cubes whose sides are parallel to the axes in H_k .

Proof.

Since $\text{supp } \phi^*(\omega)$ is a compact subset of W , we may effectively apply Lemma 5.6.9 to get the finite collection $\{\sigma_m\}_{1 \leq m \leq M}$.

We have that $\sigma = \sum_{m=1}^M \sigma_m$ is a k -chain in W whose image contains the support of $\phi^*(\omega)$ and such that the image of the $(k-1)$ -chain $\sum_{\sigma_m(I_k) \cap \partial H_k \neq \emptyset} (\sigma_m)_{k,0}$ contains the support of the differential $(k-1)$ -form $\phi^*(\omega|_{U \cap \partial S}) = \phi^*(\omega)|_{W \cap \partial H_k}$ on $W \cap \partial H_k$.

Since ϕ is orientation preserving, we have

$$\begin{aligned} \mu_{\mathbf{u}} &= [\phi_*(\mathbf{w}, -\mathbf{e}_k), \phi_*(\mathbf{w}, \mathbf{e}_1), \phi_*(\mathbf{w}, \mathbf{e}_2), \dots, \phi_*(\mathbf{w}, \mathbf{e}_{k-1})] \\ &= (-1)^k [\phi_*(\mathbf{w}, \mathbf{e}_1), \phi_*(\mathbf{w}, \mathbf{e}_2), \dots, \phi_*(\mathbf{w}, \mathbf{e}_{k-1}), \phi_*(\mathbf{w}, \mathbf{e}_k)] \end{aligned}$$

for $\mathbf{u} \in U \cap \partial S$ and $\mathbf{w} \in W \cap \partial H_k$ such that $\mathbf{u} = \phi(\mathbf{w})$. Therefore, the induced orientation on ∂S differs from the orientation provided by the standard orientation on H_k by a factor of $(-1)^k$. The two orientations match when k is even and are of opposite signs when k is odd. See also Remark 5.3.16. Thus, we have

$$\int_{\partial S} \omega = \int_{U \cap \partial S} \omega = (-1)^k \int_{W \cap \partial H_k} \phi^*(\omega) = (-1)^k \sum_{\sigma_m(I_k) \cap \partial H_k \neq \emptyset} \int_{(\sigma_m)_{k,0}} \phi^*(\omega) .$$

Since $\partial \sigma = \sum_{m=1}^M \left(\sum_{i=1}^k (-1)^i ((\sigma_m)_{i,0} - (\sigma_m)_{i,1}) \right)$, we get

$$\begin{aligned} \int_{\partial S} \omega &= \sum_{\sigma_m(I_k) \cap \partial H_k \neq \emptyset} (-1)^k \int_{(\sigma_m)_{k,0}} \phi^*(\omega) = \sum_{\sigma_m(I_k) \cap \partial H_k \neq \emptyset} \int_{(-1)^k (\sigma_m)_{k,0}} \phi^*(\omega) \\ &= \sum_{m=1}^M \sum_{i=1}^k \left(\int_{(-1)^i (\sigma_m)_{i,0}} \phi^*(\omega) - \int_{(-1)^i (\sigma_m)_{i,1}} \phi^*(\omega) \right) = \int_{\partial \sigma} \phi^*(\omega) , \end{aligned}$$

where we have used the fact that $\int_{(\sigma_m)_{i,j}} \phi^*(\omega) = 0$ if $\text{supp } \phi^*(\omega) \cap (\sigma_m)_{i,j}(I_{k-1}^\circ) = \emptyset$ with $j = 0, 1$ and $1 \leq i \leq k$, and $\int_{(\sigma_{m_1})_{i,1}} \phi^*(\omega) = - \int_{(\sigma_{m_2})_{i,0}} \phi^*(\omega)$ if $(\sigma_{m_1})_{i,1}(I_{k-1}) = (\sigma_{m_2})_{i,0}(I_{k-1})$ for some $m_1 \neq m_2$ and $1 \leq i \leq k$. ■

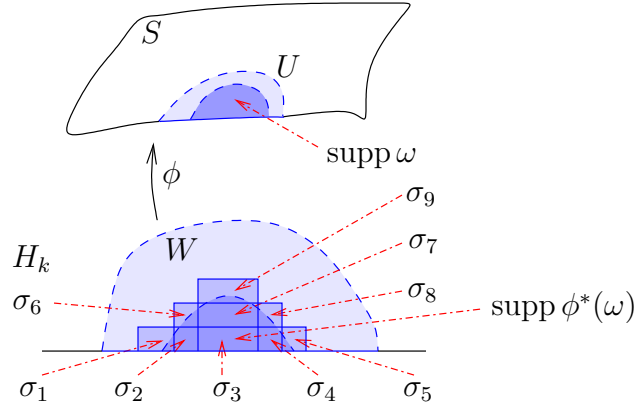


Figure 5.10: Integration along a chain on a manifold

The following theorem is the general form of the fundamental theorem of calculus.

Theorem 5.6.11 (Stokes' Theorem)

Suppose that S is a compact and oriented k -dimensional manifold of class C^1 , and ω is a differential $(k-1)$ -form on S . Then,

$$\int_S d\omega = \int_{\partial S} \omega. \quad (5.6.2)$$

Proof.

Let $\mathcal{A} = \{(W_\alpha, U_\alpha, \phi_\alpha)\}_{\alpha \in A}$ be an atlas of orientation preserving local charts for S . Since S is compact, we can find a finite open cover $\{U_{\alpha_i}\}_{1 \leq i \leq I}$ of S . Therefore

$\mathcal{A}_I = \{(W_{\alpha_i}, U_{\alpha_i}, \phi_{\alpha_i})\}_{1 \leq i \leq I}$ is a finite atlas of orientation preserving local charts for S . Let $\{\psi_i\}_{1 \leq i \leq I}$ be a partition of unity subordinate to $\mathcal{A}_I$ such that $\text{supp } \psi_i \subset V_{\alpha_i}$ for $1 \leq i \leq I$ (see Remark 3.4.2).

Let $\omega_i = \psi_i \omega$ for $1 \leq i \leq I$. Then

$$\int_S \omega = \sum_{i=1}^I \int_S \omega_i.$$

There are two cases to consider.

i) If $(\text{supp } \omega_i) \cap \partial U_{\alpha_i} = \emptyset$ (topological neighbourhood of U_{α_i}), then there exists a finite number of orientation preserving singular k -cubes satisfying (3(a)) of Lemma 5.6.9 with U and W replaced by U_{α_i} and W_{α_i} respectively, and $v = \phi_i^*(\omega_i)$. Consider the k -chain

$\sigma = \sum_{m=1}^M \sigma_m$ in W_{α_i} . It follows from Theorems 4.7.5 that

$$\int_S d\omega_i = \int_{U_{\alpha_i}} d\omega_i = \int_{W_{\alpha_i}} \phi_i^*(d\omega_i) = \int_{W_{\alpha_i}} d\phi_i^*(\omega_i) = \int_{\sigma} d\phi_i^*(\omega_i) = \int_{\partial\sigma} \phi_i^*(\omega_i) = 0$$

because $\phi_i^*(\omega_i)|_{\partial\sigma} = 0$. Since $\omega_i|_{\partial S} = 0$, we also get $\int_{\partial S} \omega_i = 0$. Thus $\int_S d\omega_i = \int_{\partial S} \omega_i$.

ii) If $(\text{supp } \omega_i) \cap \partial U_{\alpha_i} \neq \emptyset$, then we may use Lemma 5.6.10 with U and W replaced by U_{α_i} and W_{α_i} respectively, and $v = \phi_i^*(\omega_i)$, to conclude that $\int_{\partial S} \omega_i = \int_{\partial\sigma} \phi_i^*(\omega_i)$. It again follows from Theorems 4.7.5 that

$$\int_{\partial S} \omega_i = \int_{\partial\sigma} \phi_i^*(\omega_i) = \int_{\sigma} d\phi_i^*(\omega_i) = \int_{\sigma} \phi_i^*(d\omega_i) = \int_{W_{\alpha_i}} \phi_i^*(d\omega_i) = \int_{U_{\alpha_i}} d\omega_i = \int_S d\omega_i .$$

iii) Since $\sum_{i=1}^I \psi_i = 1$, we get that $0 = d\left(\sum_{i=1}^I \psi_i\right) = \sum_{i=1}^I d\psi_i$. Therefore,

$$\sum_{i=1}^I \int_S d\psi_i \wedge \omega = \int_S \left(\sum_{i=1}^I d\psi_i\right) \wedge \omega = 0 . \quad (5.6.3)$$

iv) Finally, it follows from (i), (ii) and (iii) that

$$\begin{aligned} \int_S d\omega &= \sum_{i=1}^I \int_S \psi_i d\omega = \sum_{i=1}^I \int_S d\psi_i \wedge \omega + \sum_{i=1}^I \int_S \psi_i d\omega = \sum_{i=1}^I \int_S (d\psi_i \wedge \omega + \psi_i d\omega) \\ &= \sum_{i=1}^I \int_S d(\psi_i \omega) = \sum_{i=1}^I \int_S d\omega_i = \sum_{i=1}^I \int_{\partial S} \omega_i = \int_{\partial S} \omega . \end{aligned}$$

The first equality comes from the definition of the integral of a differential k -form since $\{\psi_i\}_{1 \leq i \leq I}$ is a partition of unity subordinate to the open cover $\{V_{\alpha_i}\}_{1 \leq i \leq I}$ of S , the second equality comes from (5.6.3), the last equality comes from the definition of the integral of a differential $k-1$ -form since $\{\psi_i\}_{1 \leq i \leq I}$ is a partition of unity subordinate to the open cover $\{V_{\alpha_i}\}_{1 \leq i \leq I}$ of ∂S . ■

The next proposition is almost a corollary of Stokes' theorem and in some textbooks it is stated as Stokes' theorem.

Proposition 5.6.12

Suppose that S is a oriented k -dimensional manifold of class C^1 and ω is a differential $(k-1)$ -form on S with compact support in S . Then,

$$\int_S d\omega = \int_{\partial S} \omega .$$

Proof.

Let $\mathcal{A} = \{(W_\alpha, U_\alpha, \phi_\alpha)\}_{\alpha \in A}$ be an atlas of orientation preserving local charts for S , and $\{\psi_j\}_{j \in \mathbb{N}^+}$ be a partition of unity subordinate to the atlas $\mathcal{A}$.

Since $\text{supp } \omega$ is compact, it follows from Proposition 5.6.5 that there exist only a finite number of ψ_j that are not null on $\text{supp } \omega$. Without loss of generality, we may assume that they are ψ_j for $1 \leq j \leq J$. Suppose that $\alpha_j \in A$ is an index such that $\text{supp } \psi_j \subset U_{\alpha_j}$.

Let $\omega_j = \psi_j \omega$ for $1 \leq j \leq J$. Then

$$\int_S \omega = \sum_{j=1}^J \int_S \omega_j .$$

The rest of the proof is almost identical to the proof of Theorem 5.6.11. ■

In Theorem 5.6.11, we assumed that S was a compact manifold. The proof of that theorem used the fact that the support of the differential $(k-1)$ -form ω on S and ∂S were covered by local charts because $\bar{S} = S$ and $\partial S = \bar{S} \setminus \text{Int } S$.

In proposition 5.6.12, we did not assume that S was compact. Therefore, ∂S might not be equal to $\bar{S} \setminus \text{Int } S$. Hence, there might be points of $\bar{S} \setminus \text{Int } S$ that were not elements of a local chart. However, we assumed that the support of the differential $(k-1)$ -form ω on S was a subset of S . This ensures that $\text{supp } \omega$ and ∂S were covered by local charts.

What can we do if S is not compact and the support of a differential $(k-1)$ -form ω on S is not a subset of S ? Namely, what can we do if $\text{supp } \omega \cap (\bar{S} \setminus S) \neq \emptyset$? This is one of the issues that we address in the next subsection.

5.6.1 Generalization of Stokes' Theorem

The Stokes' theorem stated above does not apply to simple objects like cones because they are not manifolds. They have “corners” that cannot be smoothly mapped to some open subsets of H_k . Nevertheless, they are almost like manifolds except at some points.

There are two possible approaches to generalize Stokes' theorem. We can either generalize our definition of manifolds to include the “manifolds with corners.” This is the approach adopted in [1]. The second approach is to show that the set of points where there are “corners” is negligible and that the integrals can be computed by ignoring these points. This is the approach used in [14]. Both approach however approximate the integral of a

differential form ω on a “manifold with corners” by the limit of integrals of differential forms ω_j for $j \in \mathbb{N}$ that do not involve the “corners.”

We present the approaches in [14] and sketch the approach in [1]. Both approaches require some notions of measure theory and functional analysis that the reader may not have. So, the next subsections are for readers with a more advanced background in mathematics than the background that is required for the rest of the book. Nevertheless, we believe that the reader without the proper background may still be able to follow the main steps of the proofs.

First Approach

We need to introduce some new concepts before stating and proving a generalized version of Stokes’ theorem. The next results heavily relied on measure theory and functional analysis.

Theorem 5.6.13

Let S be a oriented k -dimensional manifold of class C^1 without boundary and ω be a continuous differential k -form on S .

1. There exists a unique linear functional η on $C_c(S)$, the space of continuous functions on S with compact supports, such that, if (W, U, ϕ) is an orientation preserving local chart of S and $\phi^*\omega = f dw_1 \wedge dw_2 \wedge \dots \wedge dw_k$ is the local representation of ω on W , then $\eta(g) = \int_W g(\phi(\mathbf{w})) f(\mathbf{w}) d\mathbf{w}$ for all $g \in C_c(S)$ with support in U .
2. The linear functional η is continuous on $C_c(S)$ with respect to the supremum norm.

If in addition ω has compact support, then

3. η can be extended to a continuous linear functional on the space $C_0(S)$ of continuous functions that vanish at infinity in S ³, and
4. there exists a unique real measure η_ω on S such that $\eta(g) = \int_S g d\eta_\omega$ for all $g \in C_0(S)$.

Proof (Sketch).

1) Let $\mathcal{A} = \{(W_\alpha, U_\alpha, \phi_\alpha)\}_{\alpha \in A}$ be an atlas of orientation preserving local charts for S and $\{\psi_j\}_{j \in \mathbb{N}^+}$ be a partition of unity subordinate to the atlas $\mathcal{A}$. Suppose that $\alpha_j \in A$ is an index such that $\text{supp } \psi_j \subset V_{\alpha_j}$ for $j \in \mathbb{N}^+$. We also assume that $\phi_\alpha^*\omega = f_\alpha dw_1 \wedge dw_2 \wedge \dots \wedge dw_k$ on W_α for all $\alpha \in A$ where $f_\alpha : W_\alpha \rightarrow \mathbb{R}$ is continuous.

³A continuous function f on S **vanish at infinity** if for every $\epsilon > 0$ there exists a compact set $K \subset S$ such that $|f(\mathbf{x})| < \epsilon$ for all $\mathbf{x} \in S \setminus K$.

Given $g \in C_c(S)$, it follows from Proposition 5.6.5 that there exists a finite subset $J \subset \mathbb{N}^+$ such that $g = \sum_{j \in J} \psi_j g$ because g has compact support. We define $\eta(g)$ as

$$\eta(g) = \eta \left(\sum_{j \in J} \psi_j g \right) = \sum_{j \in J} \int_{W_{\alpha_j}} \psi_j(\phi_{\alpha_j}(\mathbf{w})) g(\phi_{\alpha_j}(\mathbf{w})) f_{\alpha_j}(\mathbf{w}) d\mathbf{w} . \quad (5.6.4)$$

Note that $\psi_j g$ is a continuous function on S with compact support in U_{α_j} . It is easy to show that this defines a linear function η on $C_c(S)$.

Using the change of variable formula, it is not too difficult to prove that this definition is independent of the chosen atlas. Suppose that $\tilde{\mathcal{A}} = \left\{ (\tilde{W}_{\tilde{\alpha}}, \tilde{U}_{\tilde{\alpha}}, \tilde{\phi}_{\tilde{\alpha}}) \right\}_{\tilde{\alpha} \in \tilde{A}}$ is another atlas of orientation preserving local charts for S . We assume that $\tilde{\phi}_{\tilde{\alpha}}^* \omega = \tilde{f}_{\tilde{\alpha}} dw_1 \wedge dw_2 \wedge \dots \wedge dw_k$ on $\tilde{W}_{\tilde{\alpha}}$ for all $\tilde{\alpha} \in \tilde{A}$ where $\tilde{f}_{\tilde{\alpha}} : \tilde{W}_{\tilde{\alpha}} \rightarrow \mathbb{R}$ is continuous. It is enough to prove that

$$\int_{W_{\alpha}} g(\phi_{\alpha}(\mathbf{w})) f_{\alpha}(\mathbf{w}) d\mathbf{w} = \int_{\tilde{W}_{\tilde{\alpha}}} g(\tilde{\phi}_{\tilde{\alpha}}(\mathbf{w})) \tilde{f}_{\tilde{\alpha}}(\mathbf{w}) d\mathbf{w}$$

for all $g \in C_c(S)$ with $\text{supp } g \subset U_{\alpha} \cap \tilde{U}_{\tilde{\alpha}}$. We get from Theorem 4.5.5 that

$$\begin{aligned} \tilde{f}_{\tilde{\alpha}} dw_1 \wedge dw_2 \wedge \dots \wedge dw_k &= (\phi_{\alpha}^{-1} \circ \tilde{\phi}_{\tilde{\alpha}})^* f_{\alpha} dw_1 \wedge dw_2 \wedge \dots \wedge dw_k \\ &= f_{\alpha}((\phi_{\alpha}^{-1} \circ \tilde{\phi}_{\tilde{\alpha}})(\mathbf{w})) \det D(\phi_{\alpha}^{-1} \circ \tilde{\phi}_{\tilde{\alpha}})(\mathbf{w}) dw_1 \wedge dw_2 \wedge \dots \wedge dw_k \end{aligned}$$

on $\tilde{\phi}_{\tilde{\alpha}}^{-1}(U_{\alpha} \cap \tilde{U}_{\tilde{\alpha}})$. It then follows from Theorem 3.6.1 about change of variables that

$$\begin{aligned} &\int_{W_{\alpha}} g(\phi_{\alpha}(\mathbf{w})) f_{\alpha}(\mathbf{w}) d\mathbf{w} \\ &= \int_{\tilde{W}_{\tilde{\alpha}}} g(\phi_{\alpha} \circ (\phi_{\alpha}^{-1} \circ \tilde{\phi}_{\tilde{\alpha}})(\mathbf{w})) f_{\alpha}((\phi_{\alpha}^{-1} \circ \tilde{\phi}_{\tilde{\alpha}})(\mathbf{w})) |\det D(\phi_{\alpha}^{-1} \circ \tilde{\phi}_{\tilde{\alpha}})(\mathbf{w})| d\mathbf{w} \\ &= \int_{\tilde{W}_{\tilde{\alpha}}} g(\tilde{\phi}_{\tilde{\alpha}}(\mathbf{w})) f_{\alpha}((\phi_{\alpha}^{-1} \circ \tilde{\phi}_{\tilde{\alpha}})(\mathbf{w})) \det D(\phi_{\alpha}^{-1} \circ \tilde{\phi}_{\tilde{\alpha}})(\mathbf{w}) d\mathbf{w} \\ &= \int_{\tilde{W}_{\tilde{\alpha}}} g(\tilde{\phi}_{\tilde{\alpha}}(\mathbf{w})) \tilde{f}_{\tilde{\alpha}}(\mathbf{w}) d\mathbf{w} . \end{aligned}$$

Note that the second identity comes from the fact that the determinant is positive since we have orientation preserving local charts. For similar reason, the definition of η in (5.6.4) is independent of the chosen partition of unity.

The uniqueness of the linear functional η comes from (5.6.4).

2) This follows from the definition of η in (5.6.4) and the fact that integration is continuous with respect to uniform convergence.

3) It suffices to note that there exists a finite set $J \subset \mathbb{N}^+$ such that $\text{supp } \omega \cap \text{supp } \psi_j \neq \emptyset$ because $\text{supp } \omega$ is compact. We may choose this set J in (5.6.4) for all $g \in C_c(S)$. It is then easy to prove that η is uniformly continuous using (5.6.4) or equivalently that η is a bounded linear functional.

Since $C_c(S)$ is dense with respect to the supremum norm in the space $C_0(S)$ (see [24]), we can extend η continuously to $C_0(S)$.

4) It follows from Riesz Representation Theorem (see Theorem 6.19 in [24]) that there exists a unique real measure η_ω on S such that $\eta(g) = \int_S g d\eta_\omega$ for all $g \in C_0(S)$. ■

Thus, if ω is a continuously differential k -form with compact support, then we may then define $\int_S g\omega$ for all (Lebesgue) integrable function g on S as $\int_S g d\eta_\omega$. Note that one of the properties of a real measure μ on a set X is that $\mu(M) \in \mathbb{R}$ for all measurable sets M of X . In particular, $\mu(X) \in \mathbb{R}$. Thus, $\eta_\omega(S) \in \mathbb{R}$ and it makes sense to set $\int_S \omega = \int_S d\eta_\omega$. In fact, if we take the set J defined in (3) of the previous proof and a function $g \in C_c(S)$ such that $g = 1$ on K , then

$$\int_S \omega = \int_S g\omega = \eta(g) = \int_S g d\eta_\omega = \int_S d\eta_\omega$$

because $\text{supp } \eta_\omega \subset \text{supp } \omega$ as it can be proved (see (4) or Proposition 5.6.18 below). We will see shortly that Theorem 5.6.13 is also true even if $\text{supp } \omega$ is not a compact subset of S . Thus, defining $\int_S \omega = \int_S d\eta_\omega$ is justified in that case. In fact, if we use a sequence $\{g_j\}_{j \in \mathbb{N}^+}$ of functions in $C_c(S)$ such that $g_{j+1} \leq g_j$ for all j and $\{g_j\}_{j \in \mathbb{N}^+}$ converges pointwise to χ_S , then we get from the Lebesgue dominated convergence theorem ⁴ that

$$\int_S g_j \omega = \eta(g_j) = \int_S g_j d\eta_\omega \rightarrow \int_S \chi_S d\eta_\omega = \int_S d\eta_\omega = \eta_\omega(S) < \infty.$$

Definition 5.6.14

The linear functional η defined in Theorem 5.6.13 is called the **functional associated to ω** on S . The measure η_ω is called the **measure associated to ω** .

Theorem 5.6.15

Let S be a oriented k -dimensional manifold of class C^1 without boundary and ω be a continuous differential k -form on S .

1. There exists a unique positive linear functional η on $C_c(S)$ ⁵ such that, if (W, U, ϕ) is an orientation preserving local chart of S and $\phi^*\omega = f dw_1 \wedge dw_2 \wedge \dots \wedge dw_k$ is the local representation of ω on W , then $\eta(g) = \int_W g(\phi(\mathbf{w})) |f(\mathbf{w})| d\mathbf{w}$ for all $g \in C_c(S)$ with support in U .
2. There exists a unique positive (countably additive and Borel) measure $\eta_{|\omega|}$ on S

⁴Since we are working with a real measure, we cannot use the Lebesgue monotone convergence theorem as if we had a positive measure.

such that $\eta(g) = \int_S g \, d\eta_{|\omega|}$ for all $g \in C_c(S)$.

Proof (Sketch).

1) Let $\mathcal{A} = \{(W_\alpha, U_\alpha, \phi_\alpha)\}_{\alpha \in A}$ be an atlas of orientation preserving local charts for S and $\{\psi_j\}_{j \in \mathbb{N}^+}$ be a partition of unity subordinate to the atlas $\mathcal{A}$. Suppose that $\alpha_j \in A$ is an index such that $\text{supp } \psi_j \subset V_{\alpha_j}$ for $j \in \mathbb{N}^+$. We also assume that $\phi_\alpha^* \omega = f_\alpha \, dw_1 \wedge dw_2 \wedge \dots \wedge dw_k$ on W_α for all $\alpha \in A$ where $f_\alpha : W_\alpha \rightarrow \mathbb{R}$ is continuous.

Given $g \in C_c(S)$, it follows from Proposition 5.6.5 that there exists a finite subset $J \subset \mathbb{N}^+$ such that $g = \sum_{j \in J} \psi_j g$ because g has a compact support. We define $\eta(g)$ as

$$\eta(g) = \eta \left(\sum_{j \in J} \psi_j g \right) = \sum_{j \in J} \int_{W_{\alpha_j}} \psi_j(\phi_{\alpha_j}(\mathbf{w})) g(\phi_{\alpha_j}(\mathbf{w})) |f_{\alpha_j}(\mathbf{w})| \, d\mathbf{w} . \quad (5.6.5)$$

Note that $\psi_j g$ is continuous function on S with compact support in U_{α_j} . It is easy to show that this defines a positive linear function η on $C_c(S)$. Using the change of variable formulae, we can prove as we did in the proof of Theorem 5.6.13 that this definition is independent of the chosen atlas and satisfies item (1) in the statement of the theorem. Likewise, the definition of η in (5.6.5) is independent of the chosen partition of unity.

3) It follows from Riesz Representation Theorem (see Theorem 2.24 in [24]) that there exists a positive (countably additive and Borel) measure $\eta_{|\omega|}$ on S such that $\eta(g) = \int_S g \, d\eta_{|\omega|}$ for all $g \in C_c(S)$. ■

In Theorem 5.6.15, the assumption that S is an oriented manifold is not needed but we still stated it because we will only work with oriented manifolds. There is a version of Stokes' theorem for non-orientable manifolds but we will not present it.

Definition 5.6.16

The positive measure $\eta_{|\omega|}$ given in Theorem 5.6.15 is called the **measure associated to $|\omega|$** .

We say that a positive measure μ on a manifold S is a **finite** if $\mu(S) < \infty$.

Remark 5.6.17

1. Since $C_c(S)$ is dense with respect to the supremum norm in the space $C_0(S)$, we have that $\eta(g) = \int_S g \, d\eta_{|\omega|}$ for all $g \in C_0(S)$.

⁵Namely, $\eta(g) \geq 0$ for all $g \in C_c(S)$ such that $g \geq 0$ on S .

2. Theorems 5.6.13 is true if $\eta_{|\omega|}$ given in Theorem 5.6.15 is a finite measure.

In the proof of Theorems 5.6.13, the only issue is to prove that η is uniformly continuous on $C_c(S)$ so that it can be extended continuously to $C_0(S)$.

Suppose that g_1 and g_2 are two functions of $C_c(S)$. Since they both have compact support, there exists a finite set $J \subset \mathbb{N}^+$ such that $\text{supp } \phi_j \cap \text{supp } g_1 = \emptyset$ and $\text{supp } \phi_j \cap \text{supp } g_2 = \emptyset$ for $j \notin J$. Using (5.6.4), we get

$$\begin{aligned} |\eta(g_1) - \eta(g_2)| &= \left| \eta \left(\sum_{j \in J} \psi_j(g_1 - g_2) \right) \right| \\ &\leq \sum_{j \in J} \int_{W_{\alpha_j}} \psi_j(\phi_{\alpha_j}(\mathbf{w})) |g_1(\phi_{\alpha_j}(\mathbf{w})) - g_2(\phi_{\alpha_j}(\mathbf{w}))| |f_{\alpha_j}(\mathbf{w})| d\mathbf{w} \\ &\leq \|g_1 - g_2\|_{\infty} \sum_{j \in J} \int_{W_{\alpha_j}} \psi_j(\phi_{\alpha_j}(\mathbf{w})) |f_{\alpha_j}(\mathbf{w})| d\mathbf{w} \leq \|g_1 - g_2\|_{\infty} \eta_{|\omega|}(S). \end{aligned}$$

It follows that η is uniformly continuous on $C_c(S)$.



Proposition 5.6.18

We assume that

- S is an oriented k -dimensional manifold of class C^1 without boundary,
- ω is a continuous differential k -form on S ,
- $\eta_{|\omega|}$ is the positive measure provide in Theorem 5.6.15,
- $\mathcal{A} = \{(W_{\alpha}, U_{\alpha}, \phi_{\alpha})\}_{\alpha \in A}$ is an atlas of orientation preserving local charts for S ,
- $\{\psi_j\}_{j \in \mathbb{N}^+}$ is a partition of unity subordinate to the atlas $\mathcal{A}$,
- $\alpha_j \in A$ is an index such that $\text{supp } \psi_j \subset V_{\alpha_j}$ for $j \in \mathbb{N}^+$, and
- $\phi_{\alpha}^* \omega = f_{\alpha} dw_1 \wedge dw_2 \wedge \dots \wedge dw_k$ on W_{α} for all $\alpha \in A$ where $f_{\alpha} : W_{\alpha} \rightarrow \mathbb{R}$ is continuous.

Then

1. $\left| \int_S g \omega \right| \leq \int_S |g| d\eta_{|\omega|}$ for all $g \in C_c(S)$,
2. $\left| \int_S \omega \right| \leq \int_S d\eta_{|\omega|}$ if ω has compact support,
3. given a set $V \subset S$ (open or closed) with a compact closure in S ,

$$\eta_{|\omega|}(V) = \sum_{j \in J} \int_{W_{\alpha_j} \cap \phi_{\alpha_j}^{-1}(V)} \psi_j(\phi_{\alpha_j}(\mathbf{w})) |f_{\alpha_j}(\mathbf{w})| d\mathbf{w}$$

for some finite subset $J \subset \mathbb{N}^+$, and

4. $\text{supp } \eta_{|\omega|} \subset \text{supp } \omega$ if ω has compact support.

Proof.

1) Suppose that $g \in C_c(S)$. It follows from Proposition 5.6.5 that there exists a finite subset $J \subset \mathbb{N}^+$ such that $g = \sum_{j \in J} \psi_j g$ because g has a compact support. We then get from (5.6.5) that

$$\begin{aligned} \left| \int_S g \omega \right| &= \left| \sum_{j \in J} \int_S \psi_j g \omega \right| = \left| \sum_{j \in J} \int_{U_{\alpha_j}} \psi_j g \omega \right| \\ &= \left| \sum_{j \in J} \int_{W_{\alpha_j}} \psi_j(\phi_{\alpha_j}(\mathbf{w})) g(\phi_{\alpha_j}(\mathbf{w})) f_{\alpha_j}(\mathbf{w}) d\mathbf{w} \right| \\ &\leq \sum_{j \in J} \int_{W_{\alpha_j}} \psi_j(\phi_{\alpha_j}(\mathbf{w})) |g(\phi_{\alpha_j}(\mathbf{w}))| |f_{\alpha_j}(\mathbf{w})| d\mathbf{w} = \eta(|g|) = \int_S |g| d\eta_{|\omega|}. \end{aligned}$$

2) This result is true even if ω does not have a compact support⁶. We only prove (2) for ω with compact support so that we do not have to refer to real measure. Choose $\mathbf{u} \in S$. For $m > 0$, choose $g_m \in C_c(S)$ such that $0 \leq g_m(\mathbf{x}) \leq 1$ for all $\mathbf{x} \in S$, $g_m(\mathbf{x}) = 1$ for all $\mathbf{x} \in S \cap \overline{B_m(\mathbf{u})}$ and $g_m(\mathbf{x}) = 0$ for all $\mathbf{x} \in S \cap B_{m+1}(\mathbf{u})$ ⁷. This is a classical result. In fact, given a compact set K and an open set $V \supset K$ in $\mathbb{R}^n$, we can use a partition of unity for instance to prove that there exists a function g of class C^∞ with compact support such that $0 \leq g(\mathbf{x}) \leq 1$ for all $\mathbf{x}$, $g(\mathbf{x}) = 1$ for all $\mathbf{x} \in K$ and $g(\mathbf{x}) = 0$ for $\mathbf{x} \notin V$. We have already used this type of arguments before.

We have that $g_m(\mathbf{x}) \rightarrow 1$ for all $\mathbf{x} \in S$ as $m \rightarrow \infty$ and $g_m(\mathbf{x}) \leq g_{m+1}(\mathbf{x})$ for all $m > 0$. It follows that

$$\left| \int_S \omega \right| = \left| \lim_{m \rightarrow \infty} \int_S g_m \omega \right| \leq \lim_{m \rightarrow \infty} \left| \int_S g_m \omega \right| \leq \lim_{m \rightarrow \infty} \int_S |g_m| d\eta_{|\omega|} = \int_S d\eta_{|\omega|}, \quad (5.6.6)$$

where the Lebesgue dominated convergence theorem (see [24]) is used to get the first equality and the Lebesgue monotone convergence theorem (see [24]) is used to get the last equality. The use of the Lebesgue dominated convergence theorem to get the first equality in (5.6.6) is explained in detail below.

Since $K = \text{supp } \omega$ is a compact set, it follows from Proposition 5.6.5 that $\psi_j|_K = 0$ for all but a finite number of $j \in \mathbb{N}^+$. Let $J \subset \mathbb{N}^+$ be the finite set of indices such that $\psi_j|_K \neq 0$. As usual, we assume that $\alpha_j \in A$ is an index such that $\text{supp } \psi_j \subset U_{\alpha_j}$ for $j \in \mathbb{N}^+$. Then

⁶If $\eta_{|\omega|}(S) = \infty$, then it is obviously true. If $\eta_{|\omega|}(S) < \infty$, then it follows from the previous remark that the measure η_ω on S exists. The result in (2) can be proved using the total variation $|\eta_\omega|$ for η_ω and $|\eta_\omega| = \eta_{|\omega|}$.

⁷Recall that $B_m(\mathbf{u})$ is the open ball of radius m centred at $\mathbf{u}$.

$\text{supp } \omega \subset \bigcup_{j \in J} U_{\alpha_j}$. Hence

$$\int_S \omega = \sum_{j \in J} \int_{W_{\alpha_j}} \psi_j(\phi_{\alpha_j}(\mathbf{w})) f_{\alpha_j}(\mathbf{w}) \, d\mathbf{w}.$$

Since $(\psi_j(\phi_{\alpha_j}(\mathbf{w})) g_m(\phi_{\alpha_j}(\mathbf{w})) f_{\alpha_j}(\mathbf{w})) \rightarrow \psi_j(\phi_{\alpha_j}(\mathbf{w})) f_{\alpha_j}(\mathbf{w})$ as $m \rightarrow \infty$ for all $\mathbf{w} \in W_{\alpha_j}$, $|\psi_j(\phi_{\alpha_j}(\mathbf{w})) g_m(\phi_{\alpha_j}(\mathbf{w})) f_{\alpha_j}(\mathbf{w})| \leq |\psi_j(\phi_{\alpha_j}(\mathbf{w})) f_{\alpha_j}(\mathbf{w})|$ for all $\mathbf{w} \in W_{\alpha_j}$ and $m > 0$, and $|(\psi_j \circ \phi_{\alpha_j}) f_{\alpha_j}|$ has a bounded integral on W_{α_j} for all $j \in J$, we may use the Lebesgue dominated convergence theorem to conclude that

$$\int_{W_{\alpha_j}} \psi_j(\phi_{\alpha_j}(\mathbf{w})) f_{\alpha_j}(\mathbf{w}) \, d\mathbf{w} = \lim_{m \rightarrow \infty} \int_{W_{\alpha_j}} \psi_j(\phi_{\alpha_j}(\mathbf{w})) g_m(\phi_{\alpha_j}(\mathbf{w})) f_{\alpha_j}(\mathbf{w}) \, d\mathbf{w}$$

for all $j \in J$. Hence

$$\begin{aligned} \int_S \omega &= \sum_{j \in J} \left(\lim_{m \rightarrow \infty} \int_{W_{\alpha_j}} \psi_j(\phi_{\alpha_j}(\mathbf{w})) g_m(\phi_{\alpha_j}(\mathbf{w})) f_{\alpha_j}(\mathbf{w}) \, d\mathbf{w} \right) \\ &= \lim_{m \rightarrow \infty} \left(\sum_{j \in J} \int_{W_{\alpha_j}} \psi_j(\phi_{\alpha_j}(\mathbf{w})) g_m(\phi_{\alpha_j}(\mathbf{w})) f_{\alpha_j}(\mathbf{w}) \, d\mathbf{w} \right) = \lim_{m \rightarrow \infty} \int_S g_m \omega. \end{aligned}$$

3) Since $K = \bar{V}$ is a compact subset of S , it follows from Proposition 5.6.5 that $\psi_j|_K = 0$ for all but a finite number of $j \in \mathbb{N}^+$. Let $J \subset \mathbb{N}^+$ be the finite set of indices such that $\psi_j|_K \neq 0$.

For $m > 0$, choose $g_m \in C_c(S)$ such that $0 \leq g_m(\mathbf{x}) \leq 1$ for all $\mathbf{x} \in S$, $g_m(\mathbf{x}) = 1$ for all $\mathbf{x} \in \{\mathbf{x} \in S : \text{dist}(\mathbf{x}, S \setminus V) \geq 1/m\}$ and $g_m(\mathbf{x}) = 0$ for all $\mathbf{x} \in S \setminus V$. We have that $g_m(\mathbf{x}) \rightarrow \chi_V(\mathbf{x})$ for all $\mathbf{x} \in S$ as $m \rightarrow \infty$ and $|g_m(\mathbf{x})| \leq \chi_V(\mathbf{x})$ for all $m > 0$. Note that $\text{supp } g_m \subset K \subset \bigcup_{j \in J} U_{\alpha_j}$ for all m .

Using the Lebesgue dominated convergence theorem, we get

$$\begin{aligned} \eta_{|\omega|}(V) &= \int_S \chi_V \, d\eta_{|\omega|} = \lim_{m \rightarrow \infty} \int_S g_m \, d\eta_{|\omega|} \\ &= \lim_{m \rightarrow \infty} \left(\sum_{j \in J} \int_{W_{\alpha_j}} \psi_j(\phi_{\alpha_j}(\mathbf{w})) g_m(\phi_{\alpha_j}(\mathbf{w})) |f_{\alpha_j}(\mathbf{w})| \, d\mathbf{w} \right) \\ &= \sum_{j \in J} \int_{W_{\alpha_j}} \psi_j(\phi_{\alpha_j}(\mathbf{w})) \chi_V(\phi_{\alpha_j}(\mathbf{w})) |f_{\alpha_j}(\mathbf{w})| \, d\mathbf{w} \\ &= \sum_{j \in J} \int_{W_{\alpha_j} \cap \phi_{\alpha_j}^{-1}(V)} \psi_j(\phi_{\alpha_j}(\mathbf{w})) |f_{\alpha_j}(\mathbf{w})| \, d\mathbf{w}. \end{aligned}$$

4) The support of $\eta_{|\omega|}$ is the complement of the largest open set G such that $\eta_{|\omega|}(G) = 0$.

It follows from (3) that $\eta_{|\omega|}(V) = 0$ for all open subsets V of S such that $\bar{V}$ is a compact subset of S and $V \cap \text{supp } \omega = \emptyset$. Since the countable union of sets of measure zero is a set of measure zero, we have that the open set $S \setminus \text{supp } \omega$ has measure zero. The set $S \setminus \text{supp } \omega$ is open because $\text{supp } \omega$ is compact and thus closed. Hence $\text{supp } \eta_{|\omega|} \subset \text{supp } \omega$. ■

Definition 5.6.19

Let S be a k -dimensional manifold of class C^1 without boundary. Given $\mathbf{u} \in \bar{S} \setminus S$, if there exists a local chart (W, U, ϕ) about $\mathbf{u}$ with U an open subset of $\bar{S}$ according to the induced topology on $\bar{S}$ from $\mathbb{R}^n$, then $\mathbf{u}$ is a **regular point** of S . The set of regular points is denoted S_r . The points in the set $S_s = \bar{S} \setminus (S \cup S_r)$ are called **singular points** of S .

This definition requires some explanation. We want to emphasize the difference between the boundary ∂S of a manifold S as we have defined it and the set $\bar{S} \setminus \text{Int } S$. We have that $S \cup S_r$ is a manifold with boundary and $\partial(S \cup S_r) = S_r$. Moreover $\bar{S} = S \cup S_r \cup S_s$. The set S_r is an open subset of $\bar{S} \setminus S$ because, given a local chart $(W_{\mathbf{u}}, U_{\mathbf{u}}, \phi_{\mathbf{u}})$ about $\mathbf{u} \in S_r$, all points in $U_{\mathbf{u}} \cap (\bar{S} \setminus S)$ are themselves in S_r since they belong to the local chart $(W_{\mathbf{u}}, U_{\mathbf{u}}, \phi_{\mathbf{u}})$. The set S_s is a closed subset of $\bar{S} \setminus S$ because $S_s = \bar{S} \setminus (S \cup S_r)$. Moreover, the set S_s is also a closed subset of $\bar{S}$ because $S_s = \bar{S} \setminus \left(S \cup \bigcup_{\mathbf{u} \in S_r} U_{\mathbf{u}} \right)$ with $S = \bigcup_{\alpha \in A} U_{\alpha}$ where $\{(W_{\alpha}, U_{\alpha}, \phi_{\alpha})\}_{\alpha \in A}$ is an atlas for S .

Example 5.6.20

Consider the cone $C = \{\mathbf{x} \in \mathbb{R}^3 : x_1^2 + x_2^2 \leq x_3^2 \text{ with } 0 \leq x_3 \leq 1\}$. This is not a manifold because of the corners. Let $S = \{\mathbf{x} \in \mathbb{R}^3 : x_1^2 + x_2^2 < x_3^2 \text{ with } 0 < x_3 < 1\}$. The set S is a manifold without boundary and $C = \bar{S}$. The set of regular points is

$$S_r = \{\mathbf{x} \in \mathbb{R}^3 : x_1^2 + x_2^2 = x_3^2 \text{ with } 0 < x_3 < 1 \text{ or } x_1^2 + x_2^2 < 1 \text{ with } x_3 = 0\}.$$

The set of singular points is the set

$$S_s = \{\mathbf{x} \in \mathbb{R}^3 : \mathbf{x} = \mathbf{e}_3 \text{ or } x_1^2 + x_2^2 = 1 \text{ with } x_3 = 0\}.$$

**Definition 5.6.21**

A **fundamental sequence** for a set $R \subset \mathbb{R}^n$ is a collection of open sets $\{V_j\}_{j \in \mathbb{N}^+}$ of $\mathbb{R}^n$ such that $R \subset V_j$ for all $j \in \mathbb{N}^+$ and, if $V \subset \mathbb{R}^n$ is an open set containing R , then there exists $J \in \mathbb{N}^+$ such that $V_j \subset V$ for $j \geq J$.

Definition 5.6.22

Let S be an oriented k -dimensional manifold of class C^1 without boundary. A closed subset $R \subset \mathbb{R}^n$ is **negligible** in S if there exist

1. an open set $G \subset \mathbb{R}^n$ such that $R \subset G$,
2. a fundamental sequence $\{G_j\}_{j \in \mathbb{N}^+}$ of R in G with $\overline{G_j} \subset G$ for all j , and
3. a sequence of functions $\{g_j\}_{j \in \mathbb{N}^+}$ such that
 - a. $g_j : G \rightarrow \mathbb{R}$ is of class C^1 for all $j \in \mathbb{N}^+$,

- b. $0 \leq g_j(\mathbf{x}) \leq 1$ for all $\mathbf{x} \in G$ and $j \in \mathbb{N}^+$,
- c. $g_j(\mathbf{x}) = 1$ for all $\mathbf{x} \in G \setminus G_j$ and $j \in \mathbb{N}^+$,
- d. there exists for each $j \in \mathbb{N}^+$ an open set $\tilde{G}_j \subset G_j$ such that $R \subset \tilde{G}_j$ and $g_j(\mathbf{x}) = 0$ for all $\mathbf{x} \in \tilde{G}_j$, and,
- e. for any $(k-1)$ -form ω of class C^1 on G , the positive measures $\eta_{|dg_j \wedge \omega|}$ associated with $dg_j \wedge \omega$ on $G \cap S$ for $j \in \mathbb{N}^+$ are finite measures satisfying $\lim_{j \rightarrow \infty} \eta_{|dg_j \wedge \omega|}(G \cap S) = 0$.

This previous definition is scarring but we will shortly demonstrate how to work with it in most of the situations that we encounter in practice.

Theorem 5.6.23 (Stokes' Theorem with Singular Points)

Let S be an oriented k -dimensional manifold of class C^1 without boundary. Let $O \subset \mathbb{R}^n$ be an open set such that $\bar{S} \subset O$, and ω be a differential $(k-1)$ -form of class C^1 on O with compact support. If the set of singular points S_s of S is such that $S_s \cap \text{supp } \omega$ is negligible in S , and the measures $\mu_{|d\omega|}$ on S associated to $|d\omega|$ and $\mu_{|\omega|}$ on S_r associated to $|\omega|$ are finite, then

$$\int_S d\omega = \int_{S_r} \omega. \quad (5.6.7)$$

We will show later that the sets S_s of “manifolds with corners.” satisfy the statement of the previous theorem.

Note that the version of Stokes' theorem above is more general than the previous versions of Stokes' theorem. First, it does not require that S be a compact set as in Theorem 5.6.11. So, we may not have that $S_r = \bar{S} \setminus \text{Int } S$ as in Theorem 5.6.11. It also does not require that ω has a compact support in S as in Proposition 5.6.12. Thus, the intersection of $\text{supp } \omega$ and S_s may not be empty.

Remark 5.6.24

Some authors define ∂S as the set S_r even if the point $\mathbf{u} \in S_r$ does not belong to S as we have for our definition of ∂S . In that case, the right hand side of (5.6.7) is $\int_{\partial S} \omega$.

The previous theorem is also true if S is a manifold with boundary. It suffices to apply the theorem to $\text{Int } S$ and note that $\int_S \omega = \int_{\text{Int } S} \omega$. To justify this equality, suppose that S is a k -dimensional manifold and (W, U, ϕ) is a local chart about $\mathbf{u} \in \partial S$. Then $W \cap \partial H_k$ is a set of measure zero in W and the integral over $W \cap \partial H_k$ is equal to the integral over $W \cap \{\mathbf{y} \in \mathbb{R}^k : y_k > 0\}$.

Note that S_r is often the disjoint union of the interior of some manifolds. For instance, in Example 5.6.20, we have that $S_r = \text{Int } M_1 \cup \text{Int } M_2$ where $M_1 = \{\mathbf{x} \in \mathbb{R}^3 : x_1^2 + x_2^2 = x_3^2 \text{ with } 0 < x_3 \leq 1\}$ and $M_2 = \{\mathbf{x} \in \mathbb{R}^3 : x_1^2 + x_2^2 \leq 1 \text{ with } x_3 = 0\}$ are both manifolds

with boundary. For a reason similar to the one given in the previous paragraph, if $\partial S = \bigcup_{m=1}^M \text{Int } M_m$ for some manifolds M_m , then we write $\int_{S_r} \omega = \sum_{m=1}^M \int_{M_m} \omega$. ♠

Example 5.6.25

We consider our cone C in Example 5.6.20 and assume that ω is a differential $(k-1)$ -form with compact support defined in an open set containing the cone C . Our new version of Stokes' theorem states that $\int_S d\omega = \int_{S_r} \omega$.

Despite the fact that C is not a manifold, we often write $\int_C d\omega$ instead of $\int_S d\omega$. We accept this informal terminology because S_s is a set of measure zero in $\mathbb{R}^3$. Moreover, we write $\int_{S_r} \omega = \int_{M_1} \omega + \int_{M_2} \omega$ despite the fact that $S_r = \text{Int } M_1 \cup \text{Int } M_2$ as we explained in the previous remark. We also informally write $\int_{\partial C}$ instead of $\int_{S_r}$ though ∂C is the topological boundary of C and is different from the manifold boundary S_r . ♣

Proof (of Theorem 5.6.23).

Since $S_s \cap \text{supp } \omega$ is negligible, there exist an open set $G \subset O$, a fundamental sequence $\{G_j\}_{j \in \mathbb{N}^+}$ of S_s in G and sequence of functions $\{g_j\}_{j \in \mathbb{N}^+}$ satisfying Definition 5.6.22. Moreover, there exists a collection of open sets $\{\tilde{G}_j\}_{j \in \mathbb{N}^+}$ such that $S_s \cap \text{supp } \omega \subset \tilde{G}_j \subset G_j$ and $g_j = 0$ on $\tilde{G}_j$ for all $j \in \mathbb{N}^+$.

Let $\omega_j = g_j \omega$ for all $j \in \mathbb{N}^+$ ⁸. We have that ω_j has a compact support because $\text{supp } \omega_j \subset \text{supp } \omega$ and ω has a compact support. Since $\omega_j = 0$ on the open set $\tilde{G}_j \supset S_s \cap \text{supp } \omega$, we have that $S_s \cap \text{supp } \omega_j = \emptyset$. Thus $(S \cup S_r) \cap \text{supp } \omega_j = \bar{S} \cap \text{supp } \omega_j$ is a compact set for all $j \in \mathbb{N}^+$ because $\bar{S} \cap \text{supp } \omega_j$ is a closed subset of the compact set $\text{supp } \omega_j$. We may therefore use Proposition 5.6.12 to obtain

$$\int_{S_r} \omega_j = \int_{S \cup S_r} d\omega_j = \int_S d\omega_j = \int_S dg_j \wedge \omega + \int_S g_j d\omega. \quad (5.6.8)$$

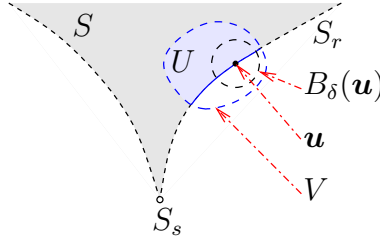
i) If η_ω is the measure associated to ω given in Theorem 5.6.13 with S_r instead of S , then we get

$$\begin{aligned} \left| \int_{S_r} \omega_j - \int_{S_r} \omega \right| &= \left| \int_{S_r} g_j d\eta_\omega - \int_{S_r} d\eta_\omega \right| = \left| \int_{S_r} (g_j - 1) d\eta_\omega \right| = \left| \int_{S_r} (g_j - 1) \omega \right| \\ &\leq \int_{S_r} |g_j - 1| d\eta_{|\omega|} = \int_{G_j \cap S_r} |g_j - 1| d\eta_{|\omega|} \\ &\leq \int_{G_j \cap S_r} d\eta_{|\omega|} = \eta_{|\omega|}(G_j \cap S_r), \end{aligned} \quad (5.6.9)$$

⁸Since $g_j = 1$ on $G \setminus G_j$ we may assume that $g_j = 1$ on $O \setminus G_j$ and so ω_j is also defined on O .

where we have used the fact that $g_j - 1 \in C_c(S_r)$ to get the third equality, Proposition 5.6.18 (2) to get the first inequality, $g_j(\mathbf{x}) - 1 = 0$ for $\mathbf{x} \in O \setminus G_j$ to get the fourth equality, and $0 \leq g_j(\mathbf{x}) \leq 1$ for all $\mathbf{x} \in O$ to get the last inequality.

We now show that $\bigcap_{j \in \mathbb{N}^+} G_j \cap S_r = \emptyset$. If $\mathbf{u} \in S_r$, then there exists a local chart (W, U, ϕ) such that $\mathbf{u} \in U$. Recall that $U = V \cap \bar{S}$ where $V \in \mathbb{R}^n$ is an open set. We have that $U \cap S_s = \emptyset$. Choose δ small enough such that $\bar{B}_\delta(\mathbf{u}) \subset V$.



Thus $\bar{B}_\delta(\mathbf{u}) \cap \bar{S} \subset U$ and $\bar{B}_\delta(\mathbf{u}) \cap (\bar{S} \setminus S) \subset S_r$. Hence $O \setminus \bar{B}_\delta(\mathbf{u})$ is an open set in $\mathbb{R}^n$ containing S_s . By definition of fundamental sequence for S_s , there exists $J \in \mathbb{N}^+$ such that $G_j \subset O \setminus \bar{B}_\delta(\mathbf{u})$ for $j \geq J$. Thus $\mathbf{u} \notin G_j \cap S_r$ for $j \geq J$.

Since $\bigcap_{j \in \mathbb{N}^+} G_j \cap S_r = \emptyset$, we have from measure theory⁹ that $\eta_{|\omega|}(G_j \cap S_r) \rightarrow 0$ as $j \rightarrow \infty$.

It follows from (5.6.9) that

$$\lim_{j \rightarrow \infty} \int_{S_r} \omega_j = \int_{S_r} \omega. \quad (5.6.10)$$

ii) The proof that $\int_S g_j d\omega$ converge to $\int_S d\omega$ as $j \rightarrow \infty$ is similar to proof given in (i).

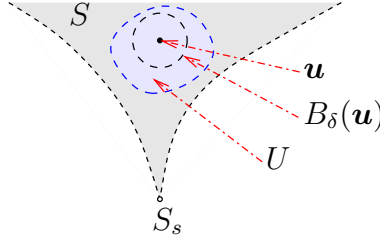
If $\eta_{d\omega}$ is the measure associated to $d\omega$ given in Theorem 5.6.13 with S , then we get

$$\begin{aligned} \left| \int_S g_j d\omega - \int_S d\omega \right| &= \left| \int_{S_r} g_j d\eta_{d\omega} - \int_{S_r} d\eta_{d\omega} \right| = \left| \int_{S_r} (g_j - 1) d\eta_{d\omega} \right| = \left| \int_S (g_j - 1) d\omega \right| \\ &\leq \int_S |g_j - 1| d\eta_{|d\omega|} = \int_{G_j \cap S} |g_j - 1| d\eta_{|d\omega|} \\ &\leq \int_{G_j \cap S} d\eta_{|d\omega|} = \eta_{|d\omega|}(G_j \cap S), \end{aligned} \quad (5.6.11)$$

where we have used the fact that $g_j - 1 \in C_c(S)$ to get the third equality, Proposition 5.6.18 (2) to get the first inequality, $g_j(\mathbf{x}) - 1 = 0$ for $\mathbf{x} \in O \setminus G_j$ to get the fourth equality and $0 \leq g_j(\mathbf{x}) \leq 1$ for all $\mathbf{x} \in O$ to get the last inequality.

⁹Use Theorem 1.19 in [24] with $A_n = \bigcap_{1 \leq j \leq n} G_j \cap S_r$ for $n \geq 0$. The fact that $\mu_{|\omega|}$ is a finite measure is also needed to apply Theorem 1.19.

We now show that $\bigcap_{j \in \mathbb{N}^+} G_j \cap S = \emptyset$. If $\mathbf{u} \in S$, then there exists a local chart (W, U, ϕ) such that $\mathbf{u} \in U$.



Since $U \cap (\bar{S} \setminus S) = \emptyset$ We have that $U \cap S_s = \emptyset$. Choose δ small enough such that $\overline{B_\delta(\mathbf{u})} \cap (\bar{S} \setminus S) = \emptyset$. Then $O \setminus \overline{B_\delta(\mathbf{u})}$ is an open set in $\mathbb{R}^n$ containing S_s . By definition of fundamental sequence for S_s , there exists $J \in \mathbb{N}^+$ such that $G_j \subset O \setminus \overline{B_\delta(\mathbf{u})}$ for $j \geq J$. Thus $\mathbf{u} \notin G_j \cap S$ for $j \geq J$.

Since $\bigcap_{j \in \mathbb{N}^+} G_j \cap S = \emptyset$, we have as in (i) that $\mu_{|\mathrm{d}\omega|}(G_j \cap S) \rightarrow 0$ as $j \rightarrow \infty$. It follows from (5.6.11) that

$$\lim_{j \rightarrow \infty} \int_S g_j \, \mathrm{d}\omega = \int_S \mathrm{d}\omega. \quad (5.6.12)$$

iii) We get from (3.e) of the definition of negligible sets, Definition 5.6.22, and (2) and (4) of Proposition 5.6.18 that

$$\left| \int_S \mathrm{d}g_j \wedge \omega \right| \leq \int_S \mathrm{d}\eta_{|\mathrm{d}g_j \wedge \omega|} = \eta_{|\mathrm{d}g_j \wedge \omega|}(S) = \eta_{|\mathrm{d}g_j \wedge \omega|}(G \cap S) \rightarrow 0 \quad (5.6.13)$$

as $j \rightarrow \infty$. We have used the fact that $\mathrm{supp} \eta_{|\mathrm{d}g_j \wedge \omega|} \subset \mathrm{supp} \mathrm{d}g_j \wedge \omega \subset G$ ¹⁰.

iv) If we let $j \rightarrow \infty$ in (5.6.8) and use (5.6.10), (5.6.12) and (5.6.13), then we get (5.6.7). ■

This generalized version of Stokes' theorem was not too hard to prove but it leaves a big question to answer. How can we determine if $S_s \cap \mathrm{supp} \omega$ is negligible in S ? The next two propositions will answer this question.

The next proposition is use to reduce the problem of determining if a set is negligible to a local problem.

Proposition 5.6.26

Let S be a oriented k -dimensional manifold of class C^1 . If R_1 and R_2 are negligible in S , then $R_1 \cup R_2$ is negligible in S .

¹⁰When restricted to measurable subsets of $G \cap S$, the measure $\eta_{|\mathrm{d}g_j \wedge \omega|}$ defined from the k -form $\mathrm{d}g_j \wedge \omega$ on the entire manifold S is equal to the measure $\eta_{|\mathrm{d}g_j \wedge \omega|}$ defined from the k -form $\mathrm{d}g_j \wedge \omega$ on $G \cap S$ because $\mathrm{supp} \mathrm{d}g_j \wedge \omega \subset G$. This justifies the use of the same name for the two measures.

Proof (Sketch).

Suppose that $\{G_{i,j}\}_{j \in \mathbb{N}^+}$ and $\{g_{i,j}\}_{j \in \mathbb{N}^+}$ are given by the definition of negligible sets applied to R_i in S for $i = 1, 2$. It suffices to show that $\{G_{1,j} \cup G_{2,j}\}_{j \in \mathbb{N}^+}$ and $\{g_{1,j} g_{2,j}\}_{j \in \mathbb{N}^+}$ satisfy the definition of negligible sets in S for $R_1 \cup R_2$.

Note that $d(g_{1,j} g_{2,j}) \wedge \omega = g_{1,j} dg_{2,j} \wedge \omega + g_{2,j} dg_{1,j} \wedge \omega$. Thus the measure $\eta_{|d(g_{1,j} g_{2,j}) \wedge \omega|}$ associated to $d(g_{1,j} g_{2,j}) \wedge \omega$ is the sum of the measures $\eta_{|g_{1,j} dg_{2,j} \wedge \omega|}$ and $\eta_{|g_{2,j} dg_{1,j} \wedge \omega|}$ associated to $g_{1,j} dg_{2,j} \wedge \omega$ and $g_{2,j} dg_{1,j} \wedge \omega$ respectively. ■

Proposition 5.6.27

Let S be a oriented n -dimensional manifold of class C^1 . A compact set R is negligible in S if R is a singular q -cube with $q \leq n - 2$.

Example 5.6.28

Before proving this proposition, let us use our two previous propositions to show that the set of singular points S_s for our famous cone C in Example 5.6.20 is negligible in S .

We have that $S_s = R_1 \cup R_2$ where $R_1 = \{\mathbf{e}_3\}$ and $R_2 = \{\mathbf{x} \in \mathbb{R}^3 : x_3 = 0 \text{ and } x_1^2 + x_2^2 = 1\}$. Since $R_1 = \sigma_1(\{0\})$ where σ_1 is defined by $\sigma_1(0) = \mathbf{e}_3$, we have that R_1 is the image of a singular 0-cube. Since $0 \leq 3 - 2 = 1$, it follows from Proposition 5.6.27 that R_1 is negligible in S . Similarly, since $R_2 = \sigma_2([0, 1])$ where σ_2 is defined by $\sigma_2(\theta) = (\cos(2\pi\theta), \sin(2\pi\theta), 0)$ for $0 \leq \theta \leq 1$, we have that R_2 is the image of a singular 1-cube. Since $1 \leq 1$, it follows from Proposition 5.6.27 that R_2 is negligible in S . Finally, it follows from Proposition 5.6.26 that $S_s = R_1 \cup R_2$ is negligible in S . ♣

Proof (of Proposition 5.6.27).

a) Suppose that $R = \sigma(I_q)$. We split I_q into p^q cubes of the form

$$I_{i_1, i_2, \dots, i_q} = \left[\frac{i_1 - 1}{p}, \frac{i_1}{p} \right] \times \left[\frac{i_2 - 1}{p}, \frac{i_2}{p} \right] \times \dots \times \left[\frac{i_q - 1}{p}, \frac{i_q}{p} \right]$$

for $1 \leq i_j \leq p$ with $1 \leq j \leq q$, and $p > 0$. Since σ is continuously differentiable on the compact set I_k , we have that $C_0 = \sup_{\mathbf{x} \in I_k} \|D\sigma(\mathbf{x})\| < \infty$. Hence, from the Mean Value

Theorem for functions of several variables, we have that $\|\sigma(\mathbf{x}) - \sigma(\mathbf{y})\| \leq \frac{C_0 \sqrt{q}}{p}$ for all

$\mathbf{x}, \mathbf{y} \in I_{i_1, i_2, \dots, i_q}$. Thus $\sigma(I_{i_1, i_2, \dots, i_q})$ is a subset of a closed cube $\tilde{I}_{i_1, i_2, \dots, i_q} \subset \mathbb{R}^n$ of volume $\left(\frac{C_0 \sqrt{q}}{p}\right)^n$ and $R = \sigma(I_q) \subset \bigcup_{\substack{1 \leq i_j \leq p \\ 1 \leq j \leq q}} \tilde{I}_{i_1, i_2, \dots, i_q}$.

b) Let $G_p = \{\mathbf{x} \in \mathbb{R}^n : \text{dist}(\mathbf{x}, R) < 2/p\}$. We construct a function $g_p : \mathbb{R}^n \rightarrow \mathbb{R}$ of class C^∞ ¹¹ such that:

- i. $0 \leq g_p(\mathbf{x}) \leq 1$ for all $\mathbf{x} \in \mathbb{R}^n$,

¹¹This is more than needed but it comes for free.

- ii. $g_p(\mathbf{x}) = 1$ for all $\mathbf{x} \in \mathbb{R}^n \setminus G_p$,
- iii. there exists an open set $\tilde{G}_p$ with $R \subset \tilde{G}_p \subset G_p$ and $g_p(\mathbf{x}) = 0$ for all $\mathbf{x} \in \tilde{G}_p$, and
- iv. there exists a constant C_1 that does not depend on p such that $\left| \frac{\partial g_p}{\partial x_s}(\mathbf{x}) \right| \leq C_1 p$ for all $\mathbf{x} \in \mathbb{R}^n$ and $1 \leq s \leq n$.

Choose a function $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$ of class C^∞ such that $0 \leq \psi(\mathbf{x}) \leq 1$ for all $\mathbf{x} \in \mathbb{R}^n$, $\psi(\mathbf{x}) = 0$ for $\|\mathbf{x}\| \leq 1/2$ and $\psi(\mathbf{x}) = 1$ for $\|\mathbf{x}\| \geq 1$ ¹². For $p > 0$, the function g_p is defined by

$$g_p(\mathbf{x}) = \prod_{\mathbf{z} \in R_p} \psi(p(\mathbf{x} - (2p)^{-1}\mathbf{z}))$$

for $\mathbf{x} \in \mathbb{R}^n$, where $R_p \equiv \{\mathbf{z} \in \mathbb{Z}^n : \text{dist}((2p)^{-1}\mathbf{z}, R) \leq 1/p\}$.

We prove that the product in the definition of g_p is finite. Since $R \subset \mathbb{R}^n$ is compact, there exist intervals $[a_i, b_i]$ for $1 \leq i \leq n$ such that $R \subset B \equiv [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$. We have that $R_p \subset B_p \equiv \{\mathbf{z} \in \mathbb{Z}^n : \text{dist}((2p)^{-1}\mathbf{z}, B) \leq 1/p\}$. Since

$$\text{dist}\left(\frac{1}{2p}\mathbf{z}, B\right) \leq \frac{1}{p} \Rightarrow a_i - \frac{1}{p} \leq \frac{1}{2p}z_i \leq b_i + \frac{1}{p} \Rightarrow 2pa_i - 2 \leq z_i \leq 2pb_i + 2$$

for $1 \leq i \leq n$, there is a finite number of points in B_p and therefore in R_p .

Therefore, the functions g_p are of class C^∞ and $0 \leq g_p(\mathbf{x}) \leq 1$ for all $\mathbf{x} \in \mathbb{R}^n$ because $0 \leq \psi(\mathbf{x}) \leq 1$ for all $\mathbf{x} \in \mathbb{R}^n$. Thus (i) is satisfied.

Given $\mathbf{x}$ such that $\text{dist}(\mathbf{x}, R) < (4p)^{-1}$, there exists at least one $\mathbf{z} \in R_p$ such that $\|\mathbf{x} - (2p)^{-1}\mathbf{z}\| \leq (2p)^{-1}$. For this value of $\mathbf{z}$, we have that $\|p(\mathbf{x} - (2p)^{-1}\mathbf{z})\| \leq 1/2$. It follows from the definition of ψ that $\psi(p(\mathbf{x} - (2p)^{-1}\mathbf{z})) = 0$ and so $g_p(\mathbf{x}) = 0$. Therefore, we may take $\tilde{G}_p = \{\mathbf{x} \in \mathbb{R}^n : \text{dist}(\mathbf{x}, R) < (4p)^{-1}\}$. Hence (iii) is satisfied.

Given $\mathbf{x}$ such that $\text{dist}(\mathbf{x}, R) \geq 2/p$ and $\mathbf{z} \in R_p$, we have that $\|\mathbf{x} - (2p)^{-1}\mathbf{z}\| \geq \|\mathbf{x} - \mathbf{y}\| - \|\mathbf{y} - (2p)^{-1}\mathbf{z}\|$ for all $\mathbf{y} \in R$ implies that $\|\mathbf{x} - (2p)^{-1}\mathbf{z}\| \geq 2/p - 1/p = 1/p$ because there is at least one $\mathbf{y} \in R$ with $\|\mathbf{y} - (2p)^{-1}\mathbf{z}\| \leq 1/p$ since $\mathbf{z} \in R_p$. Hence $\|p(\mathbf{x} - (2p)^{-1}\mathbf{z})\| \geq 1$ for all $\mathbf{z} \in R_p$. It follows from the definition of ψ that $\psi(p(\mathbf{x} - (2p)^{-1}\mathbf{z})) = 1$ for all $\mathbf{z} \in R_p$ and so $g_p(\mathbf{x}) = 1$. Therefore, we have that $g_p(\mathbf{x}) = 1$ for $\mathbf{x} \in \mathbb{R}^n \setminus G_p$. Hence (ii) is satisfied.

Finally, we prove that item (iv) is satisfied. We note that

$$\frac{\partial g_p}{\partial x_i}(\mathbf{x}) = \sum_{\mathbf{z} \in R_p} \left(\left(\prod_{\substack{\mathbf{w} \in R_p \\ \mathbf{w} \neq \mathbf{z}}} \psi(p(\mathbf{x} - (2p)^{-1}\mathbf{w})) \right) \frac{\partial}{\partial x_i} \psi(p(\mathbf{x} - (2p)^{-1}\mathbf{z})) \right). \quad (5.6.14)$$

Let $M = \max_{1 \leq i \leq n} \sup_{\mathbf{x} \in \mathbb{R}^n} \left| \frac{\partial \psi}{\partial x_i}(\mathbf{x}) \right|$. We have that $M < \infty$ because $\frac{\partial \psi}{\partial x_i}(\mathbf{x}) = 0$ for $\|\mathbf{x}\| \leq 1/2$ or $\|\mathbf{x}\| \geq 1$. We get

$$\left| \frac{\partial}{\partial x_i} \psi(p(\mathbf{x} - (2p)^{-1}\mathbf{z})) \right| = p \left| \frac{\partial \psi}{\partial x_i}(p(\mathbf{x} - (2p)^{-1}\mathbf{z})) \right| \leq pM$$

¹²As we have seen, we can use a partition of unity to construct such a function.

for all $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{z} \in \mathbb{Z}^n$ and $1 \leq i \leq n$. Hence

$$\begin{aligned} & \left| \left(\prod_{\substack{\mathbf{w} \in R_p \\ \mathbf{w} \neq \mathbf{z}}} \psi(p(\mathbf{x} - (2p)^{-1}\mathbf{w})) \right) \frac{\partial}{\partial x_i} \psi(p(\mathbf{x} - (2p)^{-1}\mathbf{z})) \right| \\ &= \prod_{\substack{\mathbf{w} \in R_p \\ \mathbf{w} \neq \mathbf{z}}} \left| \psi(p(\mathbf{x} - (2p)^{-1}\mathbf{w})) \right| \left| \frac{\partial}{\partial x_i} \psi(p(\mathbf{x} - (2p)^{-1}\mathbf{z})) \right| \leq pM \end{aligned}$$

for all $\mathbf{x} \in \mathbb{R}^n$ and all $\mathbf{z} \in R_p$ because $0 \leq \psi(\mathbf{y}) \leq 1$ for all $\mathbf{y} \in \mathbb{R}^n$.

Moreover, the number of terms in the sum in (5.6.14) does not grow with p . There are at most 9^n terms independently of the value of p . To understand why it is so, we first note that $\psi(p(\mathbf{y} - (2p)^{-1}\mathbf{z})) = 1$ for all $\mathbf{y} \in \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{y} - (2p)^{-1}\mathbf{z}\| > 1/p\}$. Thus, the partial derivatives of $\psi(p(\mathbf{y} - (2p)^{-1}\mathbf{z}))$ are null in the set $\mathbf{y} \in \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{y} - (2p)^{-1}\mathbf{z}\| > 1/p\}$. For $\mathbf{x}$ given and fixed, the only terms possibly non null in the sum in (5.6.14) are therefore those associated to the values of $\mathbf{z}$ such that $\|\mathbf{x} - (2p)^{-1}\mathbf{z}\| \leq 1/p$. Suppose that $\mathbf{z}_j \in \mathbb{Z}^n$ satisfies $\|\mathbf{x} - (2p)^{-1}\mathbf{z}_j\| \leq 1/p$ for $j = 1, 2$. Then $\|(2p)^{-1}\mathbf{z}_1 - (2p)^{-1}\mathbf{z}_2\| \leq \|\mathbf{x} - (2p)^{-1}\mathbf{z}_1\| + \|\mathbf{x} - (2p)^{-1}\mathbf{z}_2\| \leq 2/p$ implies that $\|\mathbf{z}_1 - \mathbf{z}_2\| \leq 4$. Thus $-4 \leq z_{1,i} - z_{2,i} \leq 4$ for $1 \leq i \leq n$. Since the $\mathbf{z}_j$ are in $\mathbb{Z}^n$, there are at most 9 acceptable values for each coordinates. In fact, we reach the maximum of 9 possible values only when $x_i \in \mathbb{Z}$ because we then have $x_i - 4 \leq z_{j,i} \leq x_i + 4$ for $j = 1, 2$. Even in that case, these $\mathbf{z}$ may not all be included in the sum because they may not be elements of R_p .

It follows from (5.6.14) that $\left| \frac{\partial g_p(\mathbf{x})}{\partial x_i} \right| \leq C_1 p$ for all $\mathbf{x} \in \mathbb{R}^n$ where $C_1 = 9^n M$.

c) We show that G_p can be covered by closed cubes whose total volume is at most $\frac{(4 + C_0\sqrt{q})^n}{p^2}$.

Given $\mathbf{x} \in G_p$, choose $\mathbf{y} \in \tilde{I}_{i_1, i_2, \dots, i_q}$ for some indices $1 \leq i_1, i_2, \dots, i_q \leq p$ such that $\|\mathbf{x} - \mathbf{y}\| < 2/p$. Then $\mathbf{x}$ is an element of a closed cube $\tilde{I}_{i_1, i_2, \dots, i_q} \supset \check{I}_{i_1, i_2, \dots, i_q}$ whose sides are of length less or equal to $\frac{4}{p} + \frac{C_0\sqrt{q}}{p}$. The volume of this cube is $\left(\frac{4 + C_0\sqrt{q}}{p} \right)^n$. Hence G_p is covered by closed cubes whose total volume is at most $p^q \left(\frac{4 + C_0\sqrt{q}}{p} \right)^n = \frac{(4 + C_0\sqrt{q})^n}{p^{n-q}} \leq \frac{(4 + C_0\sqrt{q})^n}{p^2}$.

d) Since G_1 is bounded because R is compact, we may choose a bounded open set G such that $G \supset \overline{G_1}$. It reminds to prove (3.e) of Definition 5.6.22 to complete the proof that $\{G_j\}_{j>0}$ and $\{g_j\}_{j>0}$ are associated to the definition of negligible sets in S .

Suppose that ω is a $(n-1)$ -form on the open set G . We may assume that

$$\omega = \sum_{s=1}^n f_s dx_1 \wedge dx_2 \wedge \widehat{dx_s} \wedge \dots \wedge dx_n,$$

where $f_s : G \rightarrow \mathbb{R}$ for $1 \leq s \leq n$ are continuous. We also have

$$dg_p = \sum_{s=1}^n \frac{\partial g_p}{\partial x_s} dx_s$$

on G . Hence

$$dg_p \wedge \omega = \sum_{s=1}^n (-1)^{s-1} \left(\frac{\partial g_p}{\partial x_s} f_s \right) dx_1 \wedge dx_2 \wedge \dots \wedge dx_n$$

on S . It follows from (3) of Proposition 5.6.18 that

$$\eta_{|dg_p \wedge \omega|}(S \cap G) = \int_{S \cap G} \left| \sum_{s=1}^n (-1)^{s-1} \frac{\partial g_p}{\partial x_s}(\mathbf{x}) f_s(\mathbf{x}) \right| d\mathbf{x} \quad (5.6.15)$$

because $\overline{G}$ is compact. We have that $\left| \frac{\partial g_p}{\partial x_s}(\mathbf{x}) \right| \leq C_1 p$ for all $\mathbf{x} \in G$ and $p > 0$. Since $f_s : G \rightarrow \mathbb{R}$ is continuous on the compact set $\overline{G_1} \subset G$ for $1 \leq s \leq n$, there exists a constant C_2 such that $|f_s(\mathbf{x})| \leq C_2$ for all $\mathbf{x} \in S \cap G_1$ and $1 \leq s \leq n$. Since $\text{supp } \frac{\partial g_p}{\partial x_s} \subset \overline{G_p} \subset \overline{G_1}$ for all $p > 0$, we get from (c) above that

$$\eta_{|dg_p \wedge \omega|}(S \cap G) = n C_1 C_2 p \int_{S \cap G_p} d\mathbf{x} \leq n C_1 C_2 p \frac{(4 + C_0 \sqrt{q})^n}{p^2} \rightarrow 0$$

as $p \rightarrow \infty$. ■

Second Approach

The first step is to expand the definition of local charts.

Definition 5.6.29

Let W be an open subset of $\mathbb{R}^{k-1}$ and $f : W \rightarrow \mathbb{R}$ be a continuous function. A point $\mathbf{p} \in G_{f,W} = \{(\mathbf{x}, f(\mathbf{x})) : \mathbf{x} \in W\}$ is a **regular point** if there exists an open set $V \subset W$ such that $G_{f,V}$ is a $k-1$ -dimensional manifold. The set of regular points is denoted $R_{f,W}$. The points in $S_{f,W} \equiv G_{f,W} \setminus R_{f,W}$ are called **singular points**.

The function f is said to be **piecewise smooth** if $f^{-1}(S_{f,W}) \subset \mathbb{R}^{k-1}$ is a set of measure zero and, for each compact set $K \subset W$, there exists a constant L such that $|f(\mathbf{x}_1) - f(\mathbf{x}_2)| \leq L \|\mathbf{x}_1 - \mathbf{x}_2\|$ for all $\mathbf{x}_1, \mathbf{x}_2 \in K$.

Definition 5.6.30

Let S be a subset of $\mathbb{R}^n$. We assume that the topology on S is the induced topology from $\mathbb{R}^n$. A **local chart** is either defined as in Definition 5.2.1 or defined by the following three objects:

1. an open subset U of S ,

2. an open subset $W = W_1 \times W_2 \subset \mathbb{R}^{k-1} \times \mathbb{R}$ with a piecewise smooth function $f : W_1 \rightarrow W_2$, and
3. a homeomorphism $\phi : W \rightarrow U$

such that $\phi(G_{f,W_1}) = U \cap \partial S$ and $\phi(\{(\mathbf{x}, y) \in W : y < f(\mathbf{x})\}) = U \cap \text{Int } S$. This local chart is denoted (W, U, ϕ) (Figures 5.11).

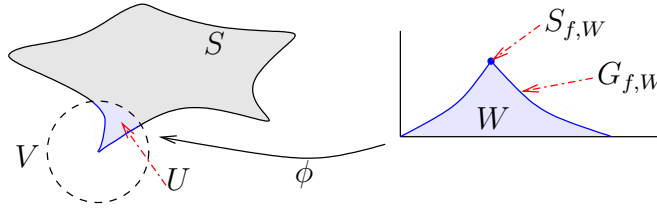


Figure 5.11: (W, U, ϕ) is a local chart of a manifold with a piecewise smooth boundary.

We can now define atlases for a set S exactly as we did in Definition 5.2.3 except that we now also accept local charts as defined above. Equivalent atlases, differential structures, maximal atlases and admissible local charts are defined as in Definition 5.2.4. Finally, the definition of manifold given in Definition 5.2.5 stays the same except that, as we said above, we now also accept local charts as in Definition 5.6.30. To distinguish the manifolds that accept local charts as in Definition 5.6.30 from those that only accept local charts as in Definition 5.2.1, we call the former **manifolds with piecewise smooth boundary**.

Theorem 5.6.11 can be generalized to manifolds with piecewise smooth boundary.

Theorem 5.6.31 (Stokes' Theorem with Piecewise Smooth Boundary)

Let S be a closed and oriented k -dimensional manifold with piecewise smooth boundary. Let ω be a differential $(k-1)$ -form of class C^1 defined in an open set of $\mathbb{R}^n$ containing $\bar{S}$. Then $\int_S d\omega = \int_{\partial S} \omega$.

Proof (Sketch).

The proof of the theorem used arguments similar to those introduced in the proofs of Theorem 5.6.23 and Proposition 5.6.27.

Since S is closed in $\mathbb{R}^n$, we have that $\bar{S} \setminus \text{Int } S = \partial S$.

Using an atlas for S and a partition of unity subordinate to this atlas, the problem is reduced to proving the theorem on a local chart (W, U, ϕ) of S as in Definition 5.6.30. Let K be a compact subset of the set $S_{f,W}$ of singular points. The first step is to construct functions g_p as in the proof of Proposition 5.6.27 with R replaced by K and to show that the

volume of $G_p \cap W$ times $\max_{1 \leq i \leq k} \sup_{\mathbf{x} \in \mathbb{R}^k} \left| \frac{\partial g_p}{\partial x_i}(\mathbf{x}) \right|$ is bounded by C/p for a constant C independent of p .

The second step is to prove the conclusion of the theorem on this local chart using techniques similar to those used in the proof of Theorem 5.6.23. ■

5.7 The Volume Element

To talk about the volume element, we have first to expand the definition of inner product to tangent spaces of manifolds.

Definition 5.7.1

Let S be a k -dimensional manifold of class C^1 . Given $\mathbf{u} \in S$, an **inner product** on $T_{\mathbf{u}}S$ is a symmetric and positive definite 2-tensor $\tau_{\mathbf{u}} \in \mathcal{T}^2(T_{\mathbf{u}}S)$

Consider $\tau : S \rightarrow \bigcup_{\mathbf{u} \in S} \mathcal{T}^2(T_{\mathbf{u}}S)$ where $\tau(\mathbf{u})$ is an inner product on $T_{\mathbf{u}}S$ for all $\mathbf{u} \in S$. If (W, U, ϕ) is a local chart of S , then $(\phi^*(\tau))(\mathbf{w})$ is an inner product on $T_{\mathbf{w}}W$ defined by

$$\begin{aligned} \phi^*(\tau)(\mathbf{w})((\mathbf{w}, \mathbf{y}_1), (\mathbf{w}, \mathbf{y}_2)) &= \phi^*(\tau(\phi(\mathbf{w})))((\mathbf{w}, \mathbf{y}_1), (\mathbf{w}, \mathbf{y}_2)) \\ &= \tau(\phi(\mathbf{w}))(\phi_*(\mathbf{w}, \mathbf{y}_1), \phi_*(\mathbf{w}, \mathbf{y}_2)) \end{aligned}$$

for all $(\mathbf{w}, \mathbf{y}_1), (\mathbf{w}, \mathbf{y}_2) \in T_{\mathbf{w}}W \cong \mathbb{R}^k$ and $\mathbf{w} \in W$. It follows from Theorem 4.1.3 that we can express τ locally as $\phi^*(\tau) = \sum_{1 \leq i_1 \leq i_2 \leq k} \tau_{i_1, i_2} dw_{i_1} \otimes dw_{i_2}$ where $\tau_{i_1, i_2} : W \rightarrow \mathbb{R}$ for all indices.

We say that τ is of class C^j on U if each τ_{i_1, i_2} is of class C^j on W . If this is true for all local chart (W, U, ϕ) , then we say that τ is of class C^j on S .

Definition 5.7.2

Let S be a k -dimensional manifold of class C^j , and $\tau : S \rightarrow \bigcup_{\mathbf{u} \in S} \mathcal{T}^2(T_{\mathbf{u}}S)$ be a map such that $\tau(\mathbf{u})$ is an inner product on $T_{\mathbf{u}}S$ for all $\mathbf{u} \in S$. If τ is of class C^j on S , then we say that the inner product on $T_{\mathbf{u}}S$ given by $\langle \cdot, \cdot \rangle_{\mathbf{u}} = \tau(\mathbf{u})$ is of class C^j on S .

Proposition 5.7.3

Let S be a k -dimensional manifold of class C^j . Then there exists a map $\tau : S \rightarrow \bigcup_{\mathbf{u} \in S} \mathcal{T}^2(T_{\mathbf{u}}S)$ of class C^j such that $\tau(\mathbf{u})$ is an inner product on $T_{\mathbf{u}}S$ for all $\mathbf{u} \in S$.

Proof.

Suppose that $\mathcal{A} = \{(W_{\alpha}, U_{\alpha}, \phi_{\alpha})\}_{\alpha \in A}$ is an atlas of orientation preserving local charts for

S and that $\{\psi_j\}_{j \in \mathbb{N}^+}$ is a partition of unity subordinate to the atlas $\mathcal{A}$. Given $j \in \mathbb{N}^+$, we assume that $\alpha_j \in A$ is an index such that $\text{supp } \psi_j \subset U_{\alpha_j}$.

For each local chart $(W_\alpha, U_\alpha, \phi_\alpha)$, we define $\tau_{U_\alpha} : U_\alpha \rightarrow \bigcup_{\mathbf{u} \in U_\alpha} \mathcal{T}^2(T_{\mathbf{u}}S)$ as it follows. For each $\mathbf{u} \in U_\alpha$, $\tau_{U_\alpha}(\mathbf{u})$ is the inner product on $T_{\mathbf{u}}S$ defined by

$$\langle (\mathbf{u}, \mathbf{x}_1), (\mathbf{u}, \mathbf{x}_2) \rangle_{\mathbf{u}} = \tau_{U_\alpha}(\mathbf{u})((\mathbf{u}, \mathbf{x}_1), (\mathbf{u}, \mathbf{x}_2)) = \langle (\mathbf{w}, \mathbf{y}_1), (\mathbf{w}, \mathbf{y}_2) \rangle_{\mathbf{w}} = \langle \mathbf{y}_1, \mathbf{y}_2 \rangle,$$

where $(\phi_\alpha)_*(\mathbf{w}, \mathbf{y}_i) = (\mathbf{u}, \mathbf{x}_i)$ for $i = 1, 2$, and the last inner product is the standard inner product on $\mathbb{R}^k$. Note that τ_{U_α} is of class C^j because it is defined from ϕ_α . In fact, $\phi_\alpha^*(\tau_{U_\alpha}) =$

$$\sum_{i=1}^k dw_i \otimes dw_i.$$

We have that $\tau = \sum_{j \in \mathbb{N}} \psi_j \tau_{U_{\alpha_j}}$ satisfies the conclusion of the proposition. Since the sum is finite on every compact subset of S , we have that τ is also of class C^j . In particular, the sum is finite at every point $\mathbf{u} \in S$. ■

Remark 5.7.4

A smooth k -dimensional manifold S with a smooth inner product on $T_{\mathbf{u}}S$ is called a **Riemannian manifold**. We will study this type of manifolds in greater detail in Chapter 9. ♠

Let S be an oriented k -dimension manifold and let $\mu_{\mathbf{u}}$ be the selected orientation on $T_{\mathbf{u}}S$ for each $\mathbf{u} \in S$. Given an inner product $\langle \cdot, \cdot \rangle_{\mathbf{u}}$ on $T_{\mathbf{u}}S$ of class C^1 , we have seen in Section 4.2 that there exists a unique $\nu(\mathbf{u}) \in \Omega^k(T_{\mathbf{u}}S)$ such that

$$\nu(\mathbf{u})((\mathbf{u}, \mathbf{x}_1), (\mathbf{u}, \mathbf{x}_2), \dots, (\mathbf{u}, \mathbf{x}_k)) = 1 \quad (5.7.1)$$

for all orthonormal basis $\{(\mathbf{u}, \mathbf{x}_i)\}_{i=1}^k$ of $T_{\mathbf{u}}S$ such that $[(\mathbf{u}, \mathbf{x}_1), (\mathbf{u}, \mathbf{x}_2), \dots, (\mathbf{u}, \mathbf{x}_k)] = \mu_{\mathbf{u}}$.

Definition 5.7.5

Let S be an oriented k -dimensional manifold of class C^1 . The unique continuous differential k -form on S satisfying (5.7.1) is called the **volume element on S** and is often denoted dV . The **volume** of the manifold S is defined as $\int_S dV$.

The notation dV is confusing because it designate the differential k -form ν that may not be the derivative of any differential $(k-1)$ -form.

When S is a surface (a 2-dimensional manifold), we usually replace the word volume by **surface** and we write dA or dS instead of dV . Similarly, when S is a curve (a 1-dimensional manifold), we replace the word volume by **length** and we write ds instead of dV .

We will not prove the existence of volume elements for oriented k -dimensional manifolds in general. Instead, we will provide the volume element for some of the low dimensional manifolds.

Example 5.7.6

Suppose that S is an oriented 1-dimensional manifold of class C^1 in $\mathbb{R}^n$, and that the orientation on S is given by $\mu_{\mathbf{u}} = [(\mathbf{u}, \mathbf{t}(\mathbf{u}))]$ where $(\mathbf{u}, \mathbf{t}(\mathbf{u})) \in T_{\mathbf{u}}S$ is a unit vector with respect to the inner product $\langle \cdot, \cdot \rangle_{\mathbf{u}}$ on $T_{\mathbf{u}}S$ defined by $\langle (\mathbf{u}, \mathbf{x}), (\mathbf{u}, \mathbf{y}) \rangle_{\mathbf{u}} = \langle \mathbf{x}, \mathbf{y} \rangle$ for all $(\mathbf{u}, \mathbf{x}), (\mathbf{u}, \mathbf{y}) \in T_{\mathbf{u}}S$ and the standard inner product $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^n$.

Let $\nu(\mathbf{u}) \in \Omega^1(T_{\mathbf{u}}S)$ be the alternating 1-tensor defined by

$$\nu(\mathbf{u})((\mathbf{u}, \mathbf{x})) = \langle (\mathbf{u}, \mathbf{x}), (\mathbf{u}, \mathbf{t}(\mathbf{u})) \rangle_{\mathbf{u}} = \langle \mathbf{x}, \mathbf{t}(\mathbf{u}) \rangle$$

for all $(\mathbf{u}, \mathbf{x}) \in T_{\mathbf{u}}S$. Then $\nu(\mathbf{u})((\mathbf{u}, \mathbf{t}(\mathbf{u}))) = \langle \mathbf{t}(\mathbf{u}), \mathbf{t}(\mathbf{u}) \rangle = 1$ for all $\mathbf{u} \in S$.

To prove that ν is a continuous differential 1-form on S , we consider an orientation preserving local chart (W, U, ϕ) of S . Since ϕ is orientation preserving, we have that $\mathbf{t}(\mathbf{u}) = \|\phi'(w)\|^{-1}\phi'(w)$ for $\mathbf{u} = \phi(w)$ with $w \in W$. Hence

$$\begin{aligned} \phi^*(\nu)(w)((w, y)) &= \phi^*(\nu(\phi(w)))((w, y)) = \nu(\phi(w))(\phi_*(w, y)) \\ &= \nu(\phi(w))(\phi(w), y\phi'(w)) = \langle y\phi'(w), \mathbf{t}(\phi(w)) \rangle \\ &= y\|\phi'(w)\| = ((\|\phi'\| dw)(w))(w, y) \end{aligned} \quad (5.7.2)$$

for all $(w, y) \in T_w W \cong \mathbb{R}$ and $w \in W$. Thus $\phi^*(\nu)$ is continuous on W . Note that $\phi'(w) \neq \mathbf{0}$ for all $w \in W$ since $\phi'(w)$ is a rank one by definition of local charts.

Therefore, by uniqueness of the volume element, $ds = \nu$. It also follows from (5.7.2) that

$$\phi^* ds = \|\phi'\| dw. \quad (5.7.3)$$

The element $(\mathbf{u}, \mathbf{t}(\mathbf{u})) \in T_{\mathbf{u}}S$, or simply $\mathbf{t}(\mathbf{u})$ when there is no confusion about the base $\mathbf{u}$, is called a **tangent unit vector** to S at $\mathbf{u}$. ♣

Before going on with reading the rest of this section, the reader should review the second part of Remark 5.5.7. The next results and examples will be useful for our study of the classical vector calculus in the next chapter.

Theorem 5.7.7

Let C be a smooth curve in $\mathbb{R}^3$ (i.e. an oriented 1-dimensional manifold of class C^1) and let $\mu_{\mathbf{u}} = [(\mathbf{u}, \mathbf{t}(\mathbf{u}))]$ be the chosen orientation on $T_{\mathbf{u}}S$ for all $\mathbf{u} \in C$ where $(\mathbf{u}, \mathbf{t}(\mathbf{u}))$ is of norm one; thus $\mathbf{t}(\mathbf{u})$ is a tangent unit vector to C at $\mathbf{u} \in C$. Then

$$ds = t_1 dx_1 + t_2 dx_2 + t_3 dx_3. \quad (5.7.4)$$

Moreover, on C (and only on C), we have

$$t_i ds = dx_i \quad (5.7.5)$$

for $1 \leq i \leq 3$.

Proof.

For all $(\mathbf{u}, \mathbf{x})$ in $T_{\mathbf{u}}C$, we have

$$\begin{aligned} ds(\mathbf{u})(\mathbf{u}, \mathbf{x}) &= \langle \mathbf{x}, \mathbf{t}(\mathbf{u}) \rangle = x_1 t_1(\mathbf{u}) + x_2 t_2(\mathbf{u}) + x_3 t_3(\mathbf{u}) \\ &= t_1(\mathbf{u}) dx_1(\mathbf{u})(\mathbf{u}, \mathbf{x}) + t_2(\mathbf{u}) dx_2(\mathbf{u})(\mathbf{u}, \mathbf{x}) + t_3(\mathbf{u}) dx_3(\mathbf{u})(\mathbf{u}, \mathbf{x}) \end{aligned}$$

where $\langle, \rangle$ is the standard inner product on $\mathbb{R}^3$. Since $\mathbf{u} \in C$ is arbitrary, we get (5.7.4).

To prove (5.7.5), we first note that $\mathbf{x} = \epsilon \|\mathbf{x}\| \mathbf{t}(\mathbf{u})$ for all vector $(\mathbf{u}, \mathbf{x}) \in T_{\mathbf{u}}C$ (vectors tangent to the curve C at $\mathbf{u}$) where $\epsilon = 1$ if $\mathbf{x}$ points in the same direction than $\mathbf{t}(\mathbf{u})$ and $\epsilon = -1$ otherwise. Hence, $ds(\mathbf{u})(\mathbf{u}, \mathbf{x}) = \langle \epsilon \|\mathbf{x}\| \mathbf{t}(\mathbf{u}), \mathbf{t}(\mathbf{u}) \rangle = \epsilon \|\mathbf{x}\|$ for all $(\mathbf{u}, \mathbf{x})$ in $T_{\mathbf{u}}C$. We already knew that from the previous example. It follows that

$$\langle \mathbf{q}, \mathbf{t}(\mathbf{u}) \rangle ds(\mathbf{u})(\mathbf{u}, \mathbf{x}) = \langle \mathbf{q}, \mathbf{t}(\mathbf{u}) \rangle \epsilon \|\mathbf{x}\| = \langle \mathbf{q}, \epsilon \|\mathbf{x}\| \mathbf{t}(\mathbf{u}) \rangle = \langle \mathbf{q}, \mathbf{x} \rangle$$

for all $(\mathbf{u}, \mathbf{x})$ in $T_{\mathbf{u}}C$ and $\mathbf{q} \in \mathbb{R}^3$. With $\mathbf{q} = \mathbf{e}_i$ for $1 \leq i \leq 3$, we get

$$t_i(\mathbf{u}) ds(\mathbf{u})(\mathbf{u}, \mathbf{x}) = \langle \mathbf{e}_i, \mathbf{x} \rangle = x_i = dx_i(\mathbf{u})(\mathbf{u}, \mathbf{x})$$

for all $(\mathbf{u}, \mathbf{x})$ in $T_{\mathbf{u}}C$. This yields (5.7.5). ■

Remark 5.7.8

If we use the local chart introduced in Example 5.7.6, then we can also prove (5.7.5). Recall that $\mathbf{t}(\mathbf{u}) = \|\phi'(w)\|^{-1} \phi'(w)$ for $\mathbf{u} \in U$ and $w \in W$ such that $\phi(w) = \mathbf{u}$.

We have from Proposition 4.5.4 (1) that $\phi^*(dx_i)(w) = \phi'_i(w) dw$. This can also be proved from

$$\begin{aligned} \phi^*(dx_i)(w)((w, r)) &= \phi^*(dx_i(\phi(w)))((w, r)) = dx_i(\phi(w))(\phi_*(w, r)) \\ &= dx_i(\phi(w))(\phi(w), r\phi'(w)) = r\phi'_i(w) = \phi'_i(w) dw(w)(w, r) \end{aligned}$$

for all $(w, r) \in T_w W$ ¹³. Moreover

$$\begin{aligned} \phi^*(t_i ds)(w)((w, r)) &= \phi^*(t_i(\phi(w)) ds(\phi(w)))((w, r)) = t_i(\phi(w)) ds(\phi(w))(\phi_*(w, r)) \\ &= t_i(\phi(w)) ds(\phi(w))(\phi(w), r\phi'(w)) = t_i(\phi(w)) \langle r\phi'(w), \mathbf{t}(\phi(w)) \rangle \\ &= t_i(\phi(w)) r \|\phi'(w)\| = r\phi'_i(w) = \phi'_i(w) dw(w)(w, r). \end{aligned}$$

Therefore, $t_i ds = dx_i$ because they have the same local representation. ♠

Remark 5.7.9

Suppose that S is an oriented 2-dimensional manifold of class C^1 in $\mathbb{R}^3$. The outward unit normal to S at $\mathbf{u} \in S$ (as defined in Definition 5.3.18) can be computed using the cross product.

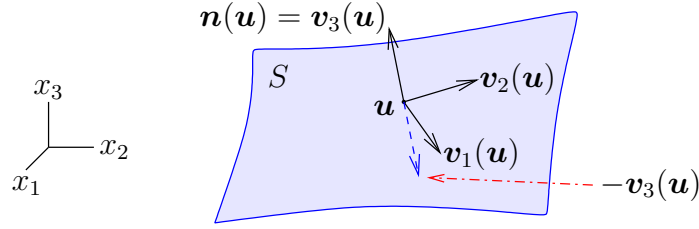
If $\mu_{\mathbf{u}} = [(\mathbf{u}, \mathbf{v}_1(\mathbf{u})), (\mathbf{u}, \mathbf{v}_2(\mathbf{u}))]$ is the selected orientation on $T_{\mathbf{u}}S$, where $\{(\mathbf{u}, \mathbf{v}_1(\mathbf{u})), (\mathbf{u}, \mathbf{v}_2(\mathbf{u}))\}$ is an orthonormal basis of $T_{\mathbf{u}}S$, then the outward unit normal is

¹³We have that $dw(w)((w, r)) = r$ for all $(w, r) \in T_w W \cong \mathbb{R}$ because the projection of $\mathbb{R}$ on itself is the identity map.

$\mathbf{n}_u = (\mathbf{u}, \mathbf{v}_1(\mathbf{u}) \times \mathbf{v}_2(\mathbf{u}))$. The inner product used to determine orthogonality in the present situation is the inner product $\langle \cdot, \cdot \rangle_u$ on $T_u S$ defined above. If $\mathbf{v}_3(\mathbf{u}) = \mathbf{v}_1(\mathbf{u}) \times \mathbf{v}_2(\mathbf{u})$, then we have in the following figure

$$\begin{aligned} [(\mathbf{u}, \mathbf{v}_3(\mathbf{u})), (\mathbf{u}, \mathbf{v}_1(\mathbf{u})), (\mathbf{u}, \mathbf{v}_2(\mathbf{u}))] &= [(\mathbf{u}, \mathbf{v}_1(\mathbf{u})), (\mathbf{u}, \mathbf{v}_2(\mathbf{u})), (\mathbf{u}, \mathbf{v}_3(\mathbf{u}))] \\ &= [(\mathbf{u}, \mathbf{e}_1), (\mathbf{u}, \mathbf{e}_2), (\mathbf{u}, \mathbf{e}_3)] . \end{aligned}$$

Thus $\mathbf{n}(\mathbf{u}) = \mathbf{v}_3(\mathbf{u})$ and $\mathbf{n}_u = (\mathbf{u}, \mathbf{v}_3(\mathbf{u}))$.



Example 5.7.10

Suppose that S is a 2-dimensional oriented manifold in $\mathbb{R}^3$. Let $\mathbf{n}_u = (\mathbf{u}, \mathbf{n}(\mathbf{u}))$ be the outward unit normal to S at $\mathbf{u}$ associated to the chosen orientation μ_u on $T_u S$ as defined in Definition 5.3.18. Let ν be the differential 2-form such that $\nu(\mathbf{u}) \in \Omega^2(T_x S)$ is the alternating 2-tensor defined by

$$\nu(\mathbf{u})((\mathbf{u}, \mathbf{x}_1), (\mathbf{u}, \mathbf{x}_2)) = \det \begin{pmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{n}(\mathbf{u}) \end{pmatrix}$$

for all $(\mathbf{u}, \mathbf{x}_i) \in T_u S$ with $i = 1, 2$. We have that $\nu(\mathbf{u})((\mathbf{u}, \mathbf{v}_1), (\mathbf{u}, \mathbf{v}_2)) = 1$ for all orthonormal bases $\{(\mathbf{u}, \mathbf{v}_1), (\mathbf{u}, \mathbf{v}_2)\}$ of $T_u S$ with $[(\mathbf{u}, \mathbf{v}_1), (\mathbf{u}, \mathbf{v}_2)] = \mu_u$; namely, such that $[\mathbf{n}_u, (\mathbf{u}, \mathbf{v}_1), (\mathbf{u}, \mathbf{v}_2)] = [(\mathbf{u}, \mathbf{e}_1), (\mathbf{u}, \mathbf{e}_2), (\mathbf{u}, \mathbf{e}_3)]$. By definition of the cross product, we also have

$$\nu(\mathbf{u})((\mathbf{u}, \mathbf{x}_1), (\mathbf{u}, \mathbf{x}_2)) = \langle \mathbf{x}_1 \times \mathbf{x}_2, \mathbf{n}(\mathbf{u}) \rangle$$

for all $(\mathbf{u}, \mathbf{x}_i) \in T_u S$ with for $i = 1, 2$, where $\langle \cdot, \cdot \rangle$ is the standard inner product on $\mathbb{R}^3$.

To prove that ν is a continuous differential 2-form on S , we consider an orientation preserving local chart (W, U, ϕ) of S . Since ϕ is orientation preserving, we have

$$[(\phi(\mathbf{w}), D\phi(\mathbf{w}) \mathbf{e}_1), (\phi(\mathbf{w}), D\phi(\mathbf{w}) \mathbf{e}_2))] = \left[\left(\phi(\mathbf{w}), \frac{\partial \phi}{\partial w_1}(\mathbf{w}) \right), \left(\phi(\mathbf{w}), \frac{\partial \phi}{\partial w_2}(\mathbf{w}) \right) \right] = \mu_{\phi(\mathbf{w})} .$$

Moreover

$$\mathbf{n}(\mathbf{u}) = \left\| \frac{\partial \phi}{\partial w_1}(\mathbf{w}) \times \frac{\partial \phi}{\partial w_2}(\mathbf{w}) \right\|^{-1} \left(\frac{\partial \phi}{\partial w_1}(\mathbf{w}) \times \frac{\partial \phi}{\partial w_2}(\mathbf{w}) \right)$$

for $\mathbf{u} = \phi(\mathbf{w})$ and $\mathbf{w} \in W$. Given $(\mathbf{w}, \mathbf{y}_i) \in T_{\mathbf{w}} W \cong \mathbb{R}^2$ for $i = 1, 2$. We have that $(\mathbf{w}, \mathbf{y}_i) = (\mathbf{w}, b_{i,1} \mathbf{e}_1 + b_{i,2} \mathbf{e}_2)$ for $i = 1, 2$ and some $b_{i,j} \in \mathbb{R}$. Hence

$$\phi^*(\nu)(\mathbf{w})((\mathbf{w}, \mathbf{y}_1), (\mathbf{w}, \mathbf{y}_2)) = \nu(\phi(\mathbf{w}))(\phi_*(\mathbf{w}, \mathbf{y}_1), \phi_*(\mathbf{w}, \mathbf{y}_2))$$

$$\begin{aligned}
&= \nu(\phi(\mathbf{w}))((\phi(\mathbf{w}), D\phi(\mathbf{w}) \mathbf{y}_1), (\phi(\mathbf{w}), D\phi(\mathbf{w}) \mathbf{y}_2)) \\
&= \nu(\phi(\mathbf{w}))\left((\phi(\mathbf{w}), b_{1,1} D\phi(\mathbf{w}) \mathbf{e}_1 + b_{1,2} D\phi(\mathbf{w}) \mathbf{e}_2), \right. \\
&\quad \left. (\phi(\mathbf{w}), b_{2,1} D\phi(\mathbf{w}) \mathbf{e}_1 + b_{2,2} D\phi(\mathbf{w}) \mathbf{e}_2)\right) \\
&= \nu(\phi(\mathbf{w}))\left(\left(\phi(\mathbf{w}), b_{1,1} \frac{\partial \phi}{\partial w_1}(\mathbf{w}) + b_{1,2} \frac{\partial \phi}{\partial w_2}(\mathbf{w})\right), \left(\phi(\mathbf{w}), b_{2,1} \frac{\partial \phi}{\partial w_1}(\mathbf{w}) + b_{2,2} \frac{\partial \phi}{\partial w_2}(\mathbf{w})\right)\right) \\
&= \left\langle \left(b_{1,1} \frac{\partial \phi}{\partial w_1}(\mathbf{w}) + b_{1,2} \frac{\partial \phi}{\partial w_2}(\mathbf{w})\right) \times \left(b_{2,1} \frac{\partial \phi}{\partial w_1}(\mathbf{w}) + b_{2,2} \frac{\partial \phi}{\partial w_2}(\mathbf{w})\right), \mathbf{n}(\phi(\mathbf{w})) \right\rangle \\
&= \left\langle (b_{1,1}b_{2,2} - b_{1,2}b_{2,1}) \frac{\partial \phi}{\partial w_1}(\mathbf{w}) \times \frac{\partial \phi}{\partial w_2}(\mathbf{w}), \mathbf{n}(\phi(\mathbf{w})) \right\rangle \\
&= (b_{1,1}b_{2,2} - b_{1,2}b_{2,1}) \left\| \frac{\partial \phi}{\partial w_1}(\mathbf{w}) \times \frac{\partial \phi}{\partial w_2}(\mathbf{w}) \right\| \\
&= \left(\left\| \frac{\partial \phi}{\partial w_1} \times \frac{\partial \phi}{\partial w_2} \right\| dw_1 \wedge dw_2 \right) (\mathbf{w})((\mathbf{w}, \mathbf{y}_1), (\mathbf{w}, \mathbf{y}_2)) . \tag{5.7.6}
\end{aligned}$$

Thus $\phi^*(\nu)$ is continuous on W .

Therefore, by uniqueness of the volume element, $dA = \nu$. It also follows from (5.7.6) that

$$\phi^* dA = \left\| \frac{\partial \phi}{\partial w_1} \times \frac{\partial \phi}{\partial w_2} \right\| dw_1 \wedge dw_2 . \tag{5.7.7}$$



A proof similar to the proof of the previous theorem gives the following result.

Theorem 5.7.11

Let S be a 2-dimensional oriented manifold in $\mathbb{R}^3$ and let $\mathbf{n}_{\mathbf{u}} = (\mathbf{u}, \mathbf{n}(\mathbf{u}))$ be the outward unit normal to S at $\mathbf{u}$ (as defined in Definition 5.3.18) associated to the chosen orientation $\mu_{\mathbf{u}}$ on $T_{\mathbf{u}}S$ for all $\mathbf{u} \in S$. Then

$$dA = n_1 dx_2 \wedge dx_3 + n_2 dx_3 \wedge dx_1 + n_3 dx_1 \wedge dx_2 . \tag{5.7.8}$$

Moreover, on S (and only on S), we have

$$n_1 dA = dx_2 \wedge dx_3 , \quad n_2 dA = dx_3 \wedge dx_1 \quad \text{and} \quad n_3 dA = dx_1 \wedge dx_2 . \tag{5.7.9}$$

Proof.

For all $(\mathbf{u}, \mathbf{v}_1)$ and $(\mathbf{u}, \mathbf{v}_2)$ in $T_{\mathbf{u}}S$, we have

$$\begin{aligned}
dA(\mathbf{u})((\mathbf{u}, \mathbf{v}_1), (\mathbf{u}, \mathbf{v}_2)) &= \det \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{n}(\mathbf{u}) \end{pmatrix} \\
&= n_1(\mathbf{u})(v_{1,2}v_{2,3} - v_{1,3}v_{2,2}) - n_2(\mathbf{u})(v_{1,1}v_{2,3} - v_{1,3}v_{2,1}) + n_3(\mathbf{u})(v_{1,1}v_{2,2} - v_{1,2}v_{2,1}) \\
&= n_1(\mathbf{u})(dx_2 \wedge dx_3)(\mathbf{u})((\mathbf{u}, \mathbf{v}_1), (\mathbf{u}, \mathbf{v}_2)) - n_2(\mathbf{u})(dx_1 \wedge dx_3)(\mathbf{u})((\mathbf{u}, \mathbf{v}_1), (\mathbf{u}, \mathbf{v}_2)) \\
&\quad + n_3(\mathbf{u})(dx_1 \wedge dx_2)(\mathbf{u})((\mathbf{u}, \mathbf{v}_1), (\mathbf{u}, \mathbf{v}_2)) .
\end{aligned}$$

Since $\mathbf{u} \in S$ is arbitrary, we get (5.7.8).

To prove (5.7.9), we note that

$$\begin{aligned} \langle \mathbf{q}, \mathbf{n}(\mathbf{u}) \rangle dA(\mathbf{u})((\mathbf{u}, \mathbf{v}_1), (\mathbf{u}, \mathbf{v}_2)) &= \langle \mathbf{q}, \mathbf{n}(\mathbf{u}) \rangle \langle \mathbf{v}_1 \times \mathbf{v}_2, \mathbf{n}(\mathbf{u}) \rangle \\ &= \langle \mathbf{q}, \mathbf{n}(\mathbf{u}) \rangle \|\mathbf{v}_1 \times \mathbf{v}_2\| = \langle \mathbf{q}, \|\mathbf{v}_1 \times \mathbf{v}_2\| \mathbf{n}(\mathbf{u}) \rangle = \langle \mathbf{q}, \mathbf{v}_1 \times \mathbf{v}_2 \rangle \end{aligned}$$

for all $(\mathbf{x}, \mathbf{v}_i) \in T_{\mathbf{u}}S$ with $i = 1, 2$ and all $\mathbf{q} \in \mathbb{R}^3$. With $\mathbf{q} = \mathbf{e}_i$, we get

$$\begin{aligned} n_i(\mathbf{u}) dA(\mathbf{u})((\mathbf{u}, \mathbf{v}_1), (\mathbf{x}, \mathbf{v}_2)) &= \langle \mathbf{e}_i, \mathbf{v}_1 \times \mathbf{v}_1 \rangle = \begin{cases} v_{1,2}v_{2,3} - v_{1,3}v_{2,2} & \text{if } i = 1 \\ -v_{1,1}v_{2,3} + v_{1,3}v_{2,1} & \text{if } i = 2 \\ v_{1,1}v_{2,2} - v_{1,2}v_{2,1} & \text{if } i = 3 \end{cases} \\ &= \begin{cases} (dx_2 \wedge dx_3)((\mathbf{u}, \mathbf{v}_1), (\mathbf{u}, \mathbf{v}_2)) & \text{if } i = 1 \\ -(dx_1 \wedge dx_3)((\mathbf{u}, \mathbf{v}_1), (\mathbf{u}, \mathbf{v}_2)) & \text{if } i = 2 \\ (dx_1 \wedge dx_2)((\mathbf{u}, \mathbf{v}_1), (\mathbf{u}, \mathbf{v}_2)) & \text{if } i = 3 \end{cases} \end{aligned}$$

for all $(\mathbf{x}, \mathbf{v}_i) \in T_{\mathbf{u}}S$ with $i = 1, 2$. This proves the relations in (5.7.9). ■

Example 5.7.12

Let S be a compact and oriented k -dimensional manifold without boundary. There exists a closed and non-exact differential k -form on S . One such differential form is the volume element ν on S . It is obviously closed because the derivative of a dimensional k -form on a manifold of dimension k is always null. Moreover, it follows from Stokes' theorem that ν cannot be exact. Suppose that there were a differential $(k-1)$ -form η on S such that $d\eta = \nu$. Then we would have from Stokes' theorem that $0 = \int_{\partial S} \eta = \int_S d\eta = \int_S \nu$. But $\int_S \nu$ is a positive real number because ν is the volume element on a compact manifold. ♣

5.7.1 Polar Coordinates

We begin by expressing in a convenient way the volume element on $S^{n-1} \equiv \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| = 1\}$.

Suppose that $\mu_{\mathbf{u}}$ is the orientation on $T_{\mathbf{u}}S^{n-1}$ for $\mathbf{u} \in S^{n-1}$ and that $(\mathbf{u}, \mathbf{n}(\mathbf{u})) \in T_{\mathbf{u}}\mathbb{R}^n$ is the outward unit normal to S at $\mathbf{u} \in S$ as in Definition 5.3.18. We note that $\mathbf{n}(\mathbf{u}) = \mathbf{u}$ for all $\mathbf{u} \in S^{n-1}$. If $\{(\mathbf{u}, \mathbf{v}_{u,i})\}_{1 \leq i \leq n-1} \subset T_{\mathbf{u}}S^{n-1}$ is such that $\mu_{\mathbf{u}} = [(\mathbf{u}, \mathbf{v}_{u,1}), \dots, (\mathbf{u}, \mathbf{v}_{u,n-1})]$, then $[(\mathbf{u}, \mathbf{n}(\mathbf{u})), (\mathbf{u}, \mathbf{v}_{u,1}), \dots, (\mathbf{u}, \mathbf{v}_{u,n-1})] = [(\mathbf{u}, \mathbf{e}_1), (\mathbf{u}, \mathbf{e}_2), \dots, (\mathbf{u}, \mathbf{e}_n)]$.

Let

$$\nu(\mathbf{u})((\mathbf{u}, \mathbf{v}_1), \dots, (\mathbf{u}, \mathbf{v}_{n-1})) = \det \begin{pmatrix} \mathbf{u} & \mathbf{v}_1 & \dots & \mathbf{v}_{n-1} \end{pmatrix}$$

for all $(\mathbf{u}, \mathbf{v}_1) \in T_{\mathbf{u}}S^{n-1}$ and $\mathbf{u} \in S^{n-1}$. It is easy to prove that ν is a differential $(n-1)$ -form on S^{n-1} of class C^∞ since the standard local atlas for S^1 is of class C^∞ . Moreover, if $\{(\mathbf{u}, \mathbf{v}_{u,i})\}_{1 \leq i \leq n-1} \subset T_{\mathbf{u}}S^{n-1}$ is an orthonormal basis on $T_{\mathbf{u}}S^{n-1}$ with respect to the inner product $\langle (\mathbf{u}, \mathbf{v}_i), (\mathbf{u}, \mathbf{v}_j) \rangle_{\mathbf{u}} = \langle \mathbf{v}_i, \mathbf{v}_j \rangle$, where the inner product on the right is the standard

inner product on $\mathbb{R}^n$, then $\nu(\mathbf{u})((\mathbf{u}, \mathbf{v}_1), \dots, (\mathbf{u}, \mathbf{v}_{n-1})) = 1$. Therefore, ν is the volume element on S .

Example 5.7.13

We have shown in Examples 4.5.9 and 4.7.4 that

$$\omega = \frac{-x_2}{x_1^2 + x_2^2} dx_1 + \frac{x_1}{x_1^2 + x_2^2} dx_2$$

is a closed differential 1-form on $V = \mathbb{R}^2 \setminus \{\mathbf{0}\}$ that is not exact.

We now show that we can always find a closed differential $n-1$ -form on $\mathbb{R}^n \setminus \{\mathbf{0}\}$ for $n \geq 2$ that is not exact. For this purpose, we will use the volume element ν on S^{n-1} . We have that ν is closed because it is a differential $(n-1)$ -form on a $(n-1)$ -dimensional manifold.

Consider the retraction $\rho : \mathbb{R}^n \setminus \{\mathbf{0}\} \rightarrow S^{n-1}$ defined by $\rho(\mathbf{x}) = \|\mathbf{x}\|^{-1}\mathbf{x}$, and the inclusion $\iota : S^{n-1} \rightarrow \mathbb{R}^n \setminus \{\mathbf{0}\}$ defined by $\iota(\mathbf{u}) = \mathbf{u}$. We have that $\rho \circ \iota = \text{Id} : S^{n-1} \rightarrow S^{n-1}$.

The differential $(n-1)$ -form $\rho^*(\nu)$ on $\mathbb{R}^n \setminus \{\mathbf{0}\}$ is closed because $d(\rho^*(\nu)) = \rho^*(d\nu) = 0$ since $d\nu = 0$. However, $\rho^*(\nu)$ is not exact. This is proved by contradiction. Suppose that $\rho^*(\nu)$ is exact and $\rho^*(\nu) = d\eta$ for some differential $(n-2)$ -form on $\mathbb{R}^n \setminus \{\mathbf{0}\}$. Then $\nu = (\rho \circ \iota)^*\nu = \iota^*(\rho^*(\nu)) = \iota^*(d\eta) = d(\iota^*(\eta))$. However, this is a contradiction because we know from Example 5.7.12 that ν cannot be exact on S^{n-1} since $\partial S^{n-1} = \emptyset$. ♣

Let ν be the volume element on S^{n-1} , and let $\iota : S^{n-1} \rightarrow \mathbb{R}^n \setminus \{\mathbf{0}\}$ and $\rho : \mathbb{R}^n \setminus \{\mathbf{0}\} \rightarrow S^{n-1}$ be respectively the inclusion and retraction defined in Example 5.7.13. Consider the following differential $(n-1)$ -form on $\mathbb{R}^n$.

$$\tilde{\nu} = \sum_{j=1}^n (-1)^{j-1} x_j dx_1 \wedge \dots \wedge \widehat{dx_j} \wedge \dots \wedge dx_n \quad (5.7.10)$$

(see Remark 4.5.10). The reader should verify that

$$\tilde{\nu}(\mathbf{x})((\mathbf{x}, \mathbf{y}_1), \dots, (\mathbf{x}, \mathbf{y}_{n-1})) = \det \begin{pmatrix} \mathbf{x} & \mathbf{y}_1 & \dots & \mathbf{y}_{n-1} \end{pmatrix}$$

for all $\mathbf{x} \in \mathbb{R}^n$ and $(\mathbf{x}, \mathbf{y}_j) \in T_{\mathbf{x}}\mathbb{R}^n$ with $1 \leq j \leq n-1$. Thus $\nu = \iota^*(\tilde{\nu})$. However, $\rho^*(\nu)$ is not equal to $\tilde{\nu}$. Instead, we have the following result.

Proposition 5.7.14

Let ν be the volume element on S^{n-1} , $\tilde{\nu}$ be the differential $(n-1)$ -form given in (5.7.10), and $\rho : \mathbb{R}^n \setminus \{\mathbf{0}\} \rightarrow S^{n-1}$ be the retraction defined by $\rho(\mathbf{x}) = \|\mathbf{x}\|^{-1}\mathbf{x}$ for $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Then $\rho^*(\nu)(\mathbf{u}) = \|\mathbf{u}\|^{-n}\tilde{\nu}(\mathbf{u})$ in $\Omega^{n-1}(T_{\mathbf{u}}\mathbb{R}^n)$ for all $\mathbf{u} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$.

Proof.

An orthonormal basis of $T_{\mathbf{u}}\mathbb{R}^n$ is given by the outward unit normal $\mathbf{n}_{\mathbf{u}} = (\mathbf{u}, \mathbf{n}(\mathbf{u})) = (\mathbf{u}, \|\mathbf{u}\|^{-1}\mathbf{u})$ to $\partial B_{\|\mathbf{u}\|}(\mathbf{0})$ at $\mathbf{u} \in \partial B_{\|\mathbf{u}\|}(\mathbf{0})$ and an orthonormal basis $\{(\mathbf{u}, \mathbf{v}_i)\}_{1 \leq i \leq n-1}$ of $T_{\mathbf{u}}(\partial B_{\|\mathbf{u}\|}(\mathbf{0}))$.

Because of the multi-linearity of $\rho^*(\nu)$, it suffices to verify that

$$\rho^*(\nu)(\mathbf{u})((\mathbf{u}, \mathbf{y}_1), \dots, (\mathbf{u}, \mathbf{y}_{n-1})) = \|\mathbf{u}\|^{-n} \tilde{\nu}(\mathbf{u})((\mathbf{u}, \mathbf{y}_1), \dots, (\mathbf{u}, \mathbf{y}_{n-1})) \quad (5.7.11)$$

where each $(\mathbf{u}, \mathbf{y}_j)$ is $(\mathbf{u}, \|\mathbf{u}\|^{-1}\mathbf{u})$ or one of $(\mathbf{u}, \mathbf{v}_i)$ for $1 \leq i \leq n-1$.

We first note that $\rho_*(\mathbf{u}, \|\mathbf{u}\|^{-1}\mathbf{u}) = (\|\mathbf{u}\|^{-1}\mathbf{u}, \mathbf{0})$ because $D\rho(\mathbf{u})(\|\mathbf{u}\|^{-1}\mathbf{u})$ is the directional derivative of ρ at $\mathbf{u}$ in the radial direction $\mathbf{n}(\mathbf{u})$. But $\rho(\mathbf{u}) = \|\mathbf{u}\|^{-1}\mathbf{u}$ is constant along the radial direction $\mathbf{n}(\mathbf{u})$. Thus, the directional derivative is null. It follows that the left side of (5.7.11) is null if one of the $(\mathbf{u}, \mathbf{y}_j)$ is $(\mathbf{u}, \|\mathbf{u}\|^{-1}\mathbf{u})$. As for the right hand side of (5.7.11), if $(\mathbf{u}, \mathbf{y}_j) = (\mathbf{u}, \|\mathbf{u}\|^{-1}\mathbf{u})$ for $j = j_0 \in \{1, 2, \dots, n-1\}$, then

$$\begin{aligned} \tilde{\nu}(\mathbf{u})((\mathbf{u}, \mathbf{y}_1), \dots, (\mathbf{u}, \mathbf{y}_{n-1})) &= \det \begin{pmatrix} \mathbf{u} & \mathbf{y}_1 & \dots & \mathbf{y}_{n-1} \end{pmatrix} \\ &= \det \begin{pmatrix} \mathbf{u} & \mathbf{y}_1 & \dots & \mathbf{y}_{j_0-1} & \|\mathbf{u}\|^{-1}\mathbf{u} & \mathbf{y}_{j_0+1} & \dots & \mathbf{y}_{n-1} \end{pmatrix} \\ &= \|\mathbf{u}\|^{-1} \det \begin{pmatrix} \mathbf{u} & \mathbf{y}_1 & \dots & \mathbf{y}_{j_0-1} & \mathbf{u} & \mathbf{y}_{j_0+1} & \dots & \mathbf{y}_{n-1} \end{pmatrix} = 0. \end{aligned}$$

We now consider (5.7.11) when each $(\mathbf{u}, \mathbf{y}_j)$ is one of $(\mathbf{u}, \mathbf{v}_i)$ for $1 \leq i \leq n-1$. If two $(\mathbf{u}, \mathbf{y}_j)$ are equal to the same $(\mathbf{u}, \mathbf{v}_i)$, then both sides of (5.7.11) are null. Therefore, we may assume that the $(\mathbf{u}, \mathbf{y}_j)$ are distinct. We have

$$(D\rho(\mathbf{u}))\mathbf{v}_j = \begin{pmatrix} D\left(\frac{x_1}{\|\mathbf{x}\|}\right) \Big|_{\mathbf{x}=\mathbf{u}} \cdot \mathbf{v}_j \\ D\left(\frac{x_2}{\|\mathbf{x}\|}\right) \Big|_{\mathbf{x}=\mathbf{u}} \cdot \mathbf{v}_j \\ \vdots \\ D\left(\frac{x_n}{\|\mathbf{x}\|}\right) \Big|_{\mathbf{x}=\mathbf{u}} \cdot \mathbf{v}_j \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial \mathbf{v}_j} \left(\frac{x_1}{\|\mathbf{x}\|} \right) (\mathbf{u}) \\ \frac{\partial}{\partial \mathbf{v}_j} \left(\frac{x_2}{\|\mathbf{x}\|} \right) (\mathbf{u}) \\ \vdots \\ \frac{\partial}{\partial \mathbf{v}_j} \left(\frac{x_n}{\|\mathbf{x}\|} \right) (\mathbf{u}) \end{pmatrix}.$$

Moreover

$$\begin{aligned} \frac{\partial}{\partial \mathbf{v}_j} \left(\frac{x_i}{\|\mathbf{x}\|} \right) (\mathbf{u}) - \frac{v_{j,i}}{\|\mathbf{u}\|} &= \lim_{h \rightarrow 0} \frac{1}{h} \left(\frac{u_i + hv_{j,i}}{\|\mathbf{u} + h\mathbf{v}_j\|} - \frac{u_i}{\|\mathbf{u}\|} - \frac{hv_{j,i}}{\|\mathbf{u}\|} \right) \\ &= \lim_{h \rightarrow 0} \frac{(u_i + hv_{j,i})(\|\mathbf{u}\| - \|\mathbf{u} + h\mathbf{v}_j\|)}{h\|\mathbf{u}\| \|\mathbf{u} + h\mathbf{v}_j\|} = \lim_{h \rightarrow 0} \frac{(u_i + hv_{j,i})(\|\mathbf{u}\|^2 - \|\mathbf{u} + h\mathbf{v}_j\|^2)}{h\|\mathbf{u}\| \|\mathbf{u} + h\mathbf{v}_j\|(\|\mathbf{u}\| + \|\mathbf{u} + h\mathbf{v}_j\|)} \\ &= \lim_{h \rightarrow 0} \frac{(u_i + hv_{j,i})(-2h\mathbf{u} \cdot \mathbf{v}_j - h^2\|\mathbf{v}_j\|^2)}{h\|\mathbf{u}\| \|\mathbf{u} + h\mathbf{v}_j\|(\|\mathbf{u}\| + \|\mathbf{u} + h\mathbf{v}_j\|)} = \lim_{h \rightarrow 0} \frac{-h(u_i + hv_{j,i})\|\mathbf{v}_j\|^2}{\|\mathbf{u}\| \|\mathbf{u} + h\mathbf{v}_j\|(\|\mathbf{u}\| + \|\mathbf{u} + h\mathbf{v}_j\|)} = 0 \end{aligned}$$

because $\mathbf{u}$ is orthogonal to $\mathbf{v}_j$ for $1 \leq j \leq n-1$. Therefore $(D\rho(\mathbf{u}))\mathbf{v}_j = \|\mathbf{u}\|^{-1}\mathbf{v}_j$ for $1 \leq j \leq n-1$. Hence, for any permutation $\sigma \in S_{n-1}$, we have

$$\begin{aligned} \rho^*(\nu)(\mathbf{u})((\mathbf{u}, \mathbf{v}_{\sigma(1)}), \dots, (\mathbf{u}, \mathbf{v}_{\sigma(n-1)})) &= \nu(\rho(\mathbf{u}))(\rho_*(\mathbf{u}, \mathbf{v}_{\sigma(1)}), \dots, \rho_*(\mathbf{u}, \mathbf{v}_{\sigma(n-1)})) \\ &= \nu(\|\mathbf{u}\|^{-1}\mathbf{u})((\|\mathbf{u}\|^{-1}\mathbf{u}, \|\mathbf{u}\|^{-1}\mathbf{v}_{\sigma(1)}), \dots, (\|\mathbf{u}\|^{-1}\mathbf{u}, \|\mathbf{u}\|^{-1}\mathbf{v}_{\sigma(n-1)})) \\ &= \det(\|\mathbf{u}\|^{-1}\mathbf{u} \quad \|\mathbf{u}\|^{-1}\mathbf{v}_{\sigma(1)} \quad \dots \quad \|\mathbf{u}\|^{-1}\mathbf{v}_{\sigma(n-1)}) \\ &= \|\mathbf{u}\|^{-n} \det(\mathbf{u} \quad \mathbf{v}_{\sigma(1)} \quad \dots \quad \mathbf{v}_{\sigma(n-1)}) = \|\mathbf{u}\|^{-n} \tilde{\nu}(\mathbf{u})((\mathbf{u}, \mathbf{v}_{\sigma(1)}), \dots, (\mathbf{u}, \mathbf{v}_{\sigma(n-1)})). \end{aligned}$$

This complete the proof of (5.7.11). ■

The use of polar and spherical coordinates can be generalized to $\mathbb{R}^n$ with $n > 3$.

Proposition 5.7.15

Suppose that $f : B_1(\mathbf{0}) \rightarrow \mathbb{R}$ is a continuous function. Then

$$\int_{B_1(\mathbf{0})} f \, dx_1 \wedge dx_2 \wedge \dots \wedge dx_n = \int_{S^{n-1}} F \, \nu.$$

where $F(\mathbf{u}) = \int_0^1 r^{n-1} f(r\mathbf{u}) \, dr$ for $\mathbf{u} \in S^{n-1}$.

Proof.

To prove the proposition, we use the projections $\pi_S : S^{n-1} \times \mathbb{R} \rightarrow S^{n-1}$ and $\pi_I : S^{n-1} \times \mathbb{R} \rightarrow \mathbb{R}$. We consider the differential n -form on $S^{n-1} \times \mathbb{R}$ defined by $\pi_S^*(\nu) \wedge \pi_I^*(dr)$, where dr is the differential 1-form on $\mathbb{R}$ defined by $dr(r)(r, y) = y$ for all $r \in \mathbb{R}$ and $(r, y) \in T_r\mathbb{R} \cong \mathbb{R}$.

i) Let $g : S^{n-1} \times [0, 1] \rightarrow \mathbb{R}$ be the function defined by $g(\mathbf{u}, r) = r^{n-1} f(r\mathbf{u})$ for $\mathbf{u} \in S^{n-1}$ and $r \in [0, 1]$.

Since S^{n-1} is a compact manifold, we may choose a finite atlas $\mathcal{A} = \{(W_i, U_i, \phi_i)\}_{1 \leq i \leq I}$ of S^{n-1} and a finite partition of unity $\{\psi_i\}_{1 \leq i \leq I}$ subordinate to $\mathcal{A}$ such that $\text{supp } \psi_i \subset U_i$ for all i . We have that $\{(W_i \times]0, 1[, U_i \times]0, 1[, \phi_i \times \text{Id})\}_{1 \leq i \leq I}$ is an atlas of $S^{n-1} \times]0, 1[$. We assume that $\phi_i^*(\nu) = q_i dw_1 \wedge dw_2 \wedge \dots \wedge dw_{n-1}$ for all i . Hence

$$\begin{aligned} \int_{S^{n-1}} F \, \nu &= \int_{S^{n-1}} \left(\int_{]0, 1[} g \, dr \right) \nu = \sum_{i=1}^I \int_{S^{n-1}} \psi_i \left(\int_{]0, 1[} g \, dr \right) \nu \\ &= \sum_{i=1}^I \int_{W_i} \psi_i(\phi_i(\mathbf{w})) \left(\int_{]0, 1[} g(\phi_i(\mathbf{w}), r) \, dr \right) \phi_i^*(\nu)(\mathbf{w}) \\ &= \sum_{i=1}^I \int_{W_i} \psi_i(\phi_i(\mathbf{w})) \left(\int_{]0, 1[} g(\phi_i(\mathbf{w}), r) \, dr \right) q_i(\mathbf{w}) \, dw_1 \wedge dw_2 \wedge \dots \wedge dw_{n-1} \\ &= \sum_{i=1}^I \int_{W_i} \psi_i(\phi_i(\mathbf{w})) \left(\int_{]0, 1[} g(\phi_i(\mathbf{w}), r) \, dr \right) q_i(\mathbf{w}) \, dw_1 \, dw_2 \dots \, dw_{n-1} \\ &= \sum_{i=1}^I \int_{W_i} \int_{]0, 1[} \psi_i(\phi_i(\mathbf{w})) g(\phi_i(\mathbf{w}), r) q_i(\mathbf{w}) \, dr \, dw_1 \, dw_2 \dots \, dw_{n-1} \\ &= \sum_{i=1}^I \int_{W_i \times]0, 1[} \psi_i(\phi_i(\mathbf{w})) g(\phi_i(\mathbf{w}), r) q_i(\mathbf{w}) \, dr \wedge dw_1 \wedge dw_2 \wedge \dots \wedge dw_{n-1} \\ &= (-1)^{n-1} \sum_{i=1}^I \int_{W_i \times]0, 1[} \psi_i(\phi_i(\mathbf{w})) g(\phi_i(\mathbf{w}), r) q_i(\mathbf{w}) \, dw_1 \wedge dw_2 \wedge \dots \wedge dw_{n-1} \wedge dr \\ &= (-1)^{n-1} \sum_{i=1}^I \int_{W_i \times]0, 1[} (\phi \times \text{Id})^* (\psi_i(\mathbf{u}) g(\mathbf{u}, r) \pi_S^*(\nu) \wedge \pi_I^*(dr)) \end{aligned}$$

$$\begin{aligned}
&= (-1)^{n-1} \sum_{i=1}^I \int_{S^{n-1} \times]0,1[} \psi_i g \pi_S^*(\nu) \wedge \pi_I^*(dr) \\
&= (-1)^{n-1} \int_{S^{n-1} \times]0,1[} g \pi_S^*(\nu) \wedge \pi_I^*(dr) .
\end{aligned} \tag{5.7.12}$$

ii) Let $h : B_1(\mathbf{0}) \setminus \{\mathbf{0}\} \rightarrow S^{n-1} \times]0,1[$ be the map defined by $h(\mathbf{x}) = (\rho(\mathbf{x}), N(\mathbf{x}))$ for all $\mathbf{x} \in B_1(\mathbf{0}) \setminus \{\mathbf{0}\}$, where $\rho : \mathbb{R}^n \setminus \{\mathbf{0}\} \rightarrow S^{n-1}$ is the retraction previously defined and $N : \mathbb{R}^n \setminus \{\mathbf{0}\} \rightarrow]0, \infty[$ is defined by $N(\mathbf{x}) = \|\mathbf{x}\|$ for $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Using the previous proposition, we have

$$\begin{aligned}
h^*(\pi_S^*(\nu) \wedge \pi_I^*(dr))(\mathbf{x}) &= (\pi_S \circ h)^*(\nu)(\mathbf{x}) \wedge (\pi_I \circ h)^*(dr)(\mathbf{x}) = \rho^*(\nu)(\mathbf{x}) \wedge N^*(dr)(\mathbf{x}) \\
&= \left(\|\mathbf{x}\|^{-n} \sum_{j=1}^n (-1)^{j-1} x_j dx_1 \wedge \dots \wedge \widehat{dx_j} \wedge \dots \wedge dx_n \right) \wedge \left(\sum_{i=1}^n \frac{\partial}{\partial x_i} \|\mathbf{x}\| dx_i \right) \\
&= \left(\|\mathbf{x}\|^{-n} \sum_{j=1}^n (-1)^{j-1} x_j dx_1 \wedge \dots \wedge \widehat{dx_j} \wedge \dots \wedge dx_n \right) \wedge \left(\sum_{i=1}^n \frac{x_i}{\|\mathbf{x}\|} dx_i \right) \\
&= (-1)^{n-1} \|\mathbf{x}\|^{-n+1} dx_1 \wedge \dots \wedge dx_j \wedge \dots \wedge dx_n
\end{aligned}$$

for all $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. It follows that

$$\begin{aligned}
h^*(g \pi_S^*(\nu) \wedge \pi_I^*(dr))(\mathbf{x}) &= g(h(\mathbf{x})) h^*(\pi_S^*(\nu) \wedge \pi_I^*(dr))(\mathbf{x}) \\
&= \|\mathbf{x}\|^{n-1} f(\mathbf{x}) (-1)^{n-1} \|\mathbf{x}\|^{-n+1} dx_1 \wedge \dots \wedge dx_j \wedge \dots \wedge dx_n \\
&= (-1)^{n-1} f(\mathbf{x}) dx_1 \wedge \dots \wedge dx_j \wedge \dots \wedge dx_n
\end{aligned}$$

for all $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Using Proposition 5.6.8 and the fact that $\{\mathbf{0}\}$ is a set of measure zero where f is continuous, we finally get from (5.7.12) that

$$\begin{aligned}
\int_{B_1(\mathbf{0})} f dx_1 \wedge dx_2 \wedge \dots \wedge dx_n &= \int_{B_1(\mathbf{0}) \setminus \{\mathbf{0}\}} f dx_1 \wedge dx_2 \wedge \dots \wedge dx_n \\
&= (-1)^{n-1} \int_{B_1(\mathbf{0}) \setminus \{\mathbf{0}\}} h^*(g \pi_S^*(\nu) \wedge \pi_I^*(dr)) \\
&= (-1)^{n-1} \int_{S^{n-1} \times]0,1[} g \pi_S^*(\nu) \wedge \pi_I^*(dr) = \int_{S^{n-1}} F \nu . \quad \blacksquare
\end{aligned}$$

Chapter 6

Classical Vector Calculus

In this chapter, we briefly review the classical results of differential geometry commonly used in vector calculus. We show along the way how they are related to the theory of differential forms and integration on manifolds.

As we have seen, Stokes' theorem given in Theorem 5.6.11 can be expanded to manifolds with singular points as given in Theorem 5.6.23 or equivalently to manifolds with a piecewise smooth boundary as given in Theorem 5.6.31. Hence, the Green's theorem, the classical Stokes' theorem and the divergence theorem (i.e. Theorems 6.4.1, 6.4.2 and 6.4.6 respectively) that we will introduce in this chapter can also be used with these types of manifolds as it will often be the case in the examples and applications.

6.1 Gradient, Curl and Divergence

Let Ω be an open subset of $\mathbb{R}^3$ and $F : \Omega \rightarrow \mathbb{R}^3$ be a differentiable vector field. The **curl** of $F = F_1\mathbf{e}_1 + F_2\mathbf{e}_2 + F_3\mathbf{e}_3$ is defined (symbolically) by

$$\begin{aligned}\operatorname{curl} F &= \nabla \times F = \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ F_1 & F_2 & F_3 \end{pmatrix} \\ &= \left(\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3} \right) \mathbf{e}_1 + \left(\frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1} \right) \mathbf{e}_2 + \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) \mathbf{e}_3\end{aligned}$$

if the determinant is developed along the first row only. If $f : \Omega \rightarrow \mathbb{R}$ is a twice continuously differentiable function, then the reader can verify by a simple computation that $\operatorname{curl}(\nabla f) = \mathbf{0}$.

The **divergence** of the vector field F is defined (symbolically) by

$$\operatorname{div} F = \nabla \cdot F = \frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} + \frac{\partial F_3}{\partial x_3}.$$

If $F : \Omega \rightarrow \mathbb{R}^3$ is a twice continuously differentiable vector field, then the reader can again verify by a simple computation that $\operatorname{div}(\operatorname{curl} F) = 0$.

Everything that we just said about the curl and the divergence can be rephrased in term of differential forms. We first match functions and vector fields to differential forms.

Let $f : \Omega \rightarrow \mathbb{R}$ and $F : \Omega \rightarrow \mathbb{R}^3$ be sufficiently continuously differentiable. We define the following associations.

$$\begin{aligned} f &\longleftrightarrow f \\ f &\longleftrightarrow f \, dx_1 \wedge dx_2 \wedge dx_3 \\ F_1 \mathbf{e}_1 + F_2 \mathbf{e}_2 + F_3 \mathbf{e}_3 &\longleftrightarrow F_1 \, dx_1 + F_2 \, dx_2 + F_3 \, dx_3 \\ F_1 \mathbf{e}_1 + F_2 \mathbf{e}_2 + F_3 \mathbf{e}_3 &\longleftrightarrow F_1 \, dx_2 \wedge dx_3 + F_2 \, dx_3 \wedge dx_1 + F_3 \, dx_1 \wedge dx_2 \end{aligned}$$

Hence, the gradient of f is associated to the differential 1-form df .

$$\nabla f = \frac{\partial f}{\partial x_1} \mathbf{e}_1 + \frac{\partial f}{\partial x_2} \mathbf{e}_2 + \frac{\partial f}{\partial x_3} \mathbf{e}_3 \longleftrightarrow df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \frac{\partial f}{\partial x_3} dx_3 .$$

The curl of the vector field F is associated to the differential 2-form $d\omega$ with $\omega = F_1 \, dx_1 + F_2 \, dx_2 + F_3 \, dx_3$.

$$\begin{aligned} \operatorname{curl} F &= \left(\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3} \right) \mathbf{e}_1 + \left(\frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1} \right) \mathbf{e}_2 + \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) \mathbf{e}_3 \longleftrightarrow \\ d\omega &= \left(\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3} \right) dx_2 \wedge dx_3 + \left(\frac{\partial F_3}{\partial x_1} - \frac{\partial F_1}{\partial x_3} \right) dx_1 \wedge dx_3 + \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) dx_1 \wedge dx_2 \end{aligned}$$

The divergence of the vector field F is associated to the differential 3-form $d\omega$ with $\omega = F_1 \, dx_2 \wedge dx_3 + F_2 \, dx_3 \wedge dx_1 + F_3 \, dx_1 \wedge dx_2$.

$$\operatorname{div} F = \frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} + \frac{\partial F_3}{\partial x_3} \longleftrightarrow d\omega = \left(\frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} + \frac{\partial F_3}{\partial x_3} \right) dx_1 \wedge dx_2 \wedge dx_3 .$$

Hence, $\operatorname{curl}(\nabla f) = \mathbf{0}$ and $\operatorname{div}(\operatorname{curl} F) = 0$ are nothing else then $d(df) = 0$ and $d(d\omega) = 0$ respectively as given by (3) of Theorem 4.5.7.

6.2 Line Integral

We review the classical definition of line integrals and show how it is related to the integral of differential forms over manifolds that we have introduced in the previous chapter.

We will use the following definition from now on.

Definition 6.2.1

A **curve** or **path** is a singular 1-cube represented by a continuously differentiable function $\sigma : I_1 = [0, 1] \rightarrow \mathbb{R}^n$. If $\sigma(0) = \mathbf{x}_0$ and $\sigma(1) = \mathbf{x}_1$, then we say that σ is a **curve or path from $\mathbf{x}_0$ to $\mathbf{x}_1$** .

Without any consequence, we may replace the closed interval $[0, 1]$ by any closed interval of the form $[a, b]$. Recall that continuously differentiable means that $\sigma = \sum_{j=1}^n \sigma_j \mathbf{e}_j$ with $\sigma_j \in C^1([a, b])$ for $1 \leq j \leq n$.

We will interchangeably use the term “curve” to designate the function σ or its image $C = \sigma([a, b])$. In classical differential geometry, the function σ is called a **parametric representation** of the curve C . The end points of the curve C are $\mathbf{a} = \sigma(a)$ and $\mathbf{b} = \sigma(b)$ (Figure 6.1).

In some rare situation, we will tolerate the existence of a finite subset N of $[a, b]$ such that σ is one-to-one on $[a, b] \setminus N$ but not on the full interval $[a, b]$. This means that the curve C may intersect itself a finite number of times only. A curve that intersect itself does not define a 1-dimensional manifold because no local chart can be defined in the neighbourhood of an intersection point. Such curves have their use in differential geometry and differential topology but cannot be used with Stokes’ theorem without careful handling.

We can recast our definition of orientation on 1-dimensional manifolds to the context of curves. Suppose that C has a continuously differentiable parametric representation $\sigma : [a, b] \rightarrow \mathbb{R}^n$ and assume that $\sigma'(w) \neq \mathbf{0}$ for all $w \in [a, b]$. We may continuously assign to each point $\mathbf{u} \in C$ a unit tangent vector $\mathbf{t}(\mathbf{u})$. The **orientation** or **direction** on C is given by continuously selecting a unit tangent vector $\mathbf{t}$ at each point of C . The curve C has therefore two possible orientations. The parametric representation $\sigma : [a, b] \rightarrow \mathbb{R}^n$ of C is consistent with the orientation on C if

$$\mathbf{t}(\sigma(w)) = \|\sigma'(w)\|^{-1} \sigma'(w) \quad (6.2.1)$$

for $w \in [a, b]$. Intuitively, the tangent vector points in the direction of the motion along C as w increases (Figure 6.1). For this reason, it is common to give the orientation along a curve as the direction of the motion along the curve. For instance, we may say that the direction along a curve C between two points $\mathbf{a}$ and $\mathbf{b}$ is from $\mathbf{a}$ to $\mathbf{b}$. This means that a parametric representation $\sigma : [a, b] \rightarrow \mathbb{R}^n$ of C must satisfy $\sigma(a) = \mathbf{a}$ and $\sigma(b) = \mathbf{b}$. Thus, we travel from $\mathbf{a}$ to $\mathbf{b}$ along C as t increases from a to b . The direction along the curve C is therefore given by the unit tangent vector defined by (6.2.1).

The **line integral** of a function $f : C \rightarrow \mathbb{R}$ along the curve C is defined by

$$\int_C f \, ds \equiv \int_a^b f(\sigma(w)) \|\sigma'(w)\| \, dw \quad (6.2.2)$$

if this integral exists, where as usual $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^n$. Using the change of variable formula, we may show that the line integral of f along C does not depend on the parametric representation used. We write

$$ds = \|\sigma'\| \, dw. \quad (6.2.3)$$

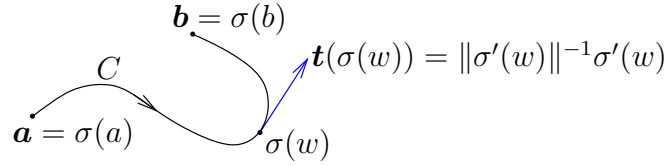


Figure 6.1: $\sigma : [a, b] \rightarrow \mathbb{R}^n$ is a parametric representation of the curve C consistent with its orientation.

Remark 6.2.2

The expression in (6.2.3) is associated to the local representation of the volume element $\nu = ds$ ¹ for a 1-dimensional oriented manifold given in Example 5.7.6 where (W, U, ϕ) is replaced by $(]a, b[, C \setminus \{\mathbf{a}, \mathbf{b}\}, \sigma)$. See Remark 5.2.7.

The formula in (6.2.2) is nothing else than $\int_S f \nu$. We have

$$\int_{C \setminus \{\mathbf{a}, \mathbf{b}\}} f \nu = \int_{]a, b[} \sigma^*(f \nu) = \int_{]a, b[} (f \circ \sigma) \sigma^*(\nu) = \int_{]a, b[} f(\sigma(w)) \|\sigma'(w)\| dw = \int_C f ds$$

where the third equality comes from (5.7.3) with $\nu = ds$. ♠

Let $F : V \rightarrow \mathbb{R}^n$ be a continuous **vector fields**; namely, a continuous vector valued function from V to $\mathbb{R}^n$. The **line integral** of a vector field $F : V \rightarrow \mathbb{R}^n$ along a curve C is defined by

$$\int_C F \cdot ds \equiv \int_C F \cdot \mathbf{t} ds = \int_a^b F(\sigma(w)) \cdot \sigma'(w) dw \quad (6.2.4)$$

if this integral exists, where $\mathbf{t}(\mathbf{u})$ is a tangent unit vector to the curve C at the point $\mathbf{u} \in C$ and $\sigma : [a, b] \rightarrow \mathbb{R}^n$ is a parametric representation of C , both of them consistent with the selected orientation on C . The integral is independent of the parametric representation of C used as long as this parametric representation is consistent with the orientation on C . We write

$$ds = \mathbf{t} ds = \sigma'(w) dw. \quad (6.2.5)$$

Remark 6.2.3

The definition of the line integral of a vector field in (6.2.4) is nothing else than the integral of a differential 1-form along a 1-dimensional manifold.

The integral $\int_C F \cdot \mathbf{t} ds$ in (6.2.4) is the integral $\int_{C \setminus \{\mathbf{a}, \mathbf{b}\}} \omega$ defined in Definition 5.6.6 where

ω is the differential 1-form $\omega = \sum_{j=1}^n F_j dx_j$. ♠

¹This is ds as a differential 1-form, not ds defined in (6.2.3) though they end up to have the same action in line integrals.

Theorem 6.2.4 (Fundamental Theorem for Line Integral)

Let V be an open subset of $\mathbb{R}^n$. Suppose that $C \subset V$ is a curve with a parametric representation $\sigma : [a, b] \rightarrow V$ consistent with its orientation. If $f \in C^1(V)$, then

$$\int_C \nabla f \cdot \mathbf{t} \, ds = f(\sigma(b)) - f(\sigma(a)) ,$$

where $\mathbf{t}(\mathbf{u})$ is a tangent unit vector to the curve C at the point $\mathbf{u} \in C$ that is consistent with the selected orientation on C .

Proof.

The Fundamental Theorem for Line Integral is just a special case of Theorem 4.7.5, Stokes' theorem, with the differential 0-form ω replaced by f and the singular 1-cube σ replaced by $\tilde{\sigma} : [0, 1] \rightarrow \mathbb{R}^n$ defined by $\tilde{\sigma}(w) = \sigma(a + w(b - a))$ for $0 \leq w \leq 1$. We get that $\int_{\tilde{\sigma}} df = \int_{\partial \tilde{\sigma}} f$.

Since $df = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i$, we have

$$\begin{aligned} \int_C \nabla f \cdot \mathbf{t} \, ds &= \int_{[a,b]} (\nabla f(\sigma(v))) \cdot \sigma'(v) \, dv = \int_{[0,1]} (\nabla f(\tilde{\sigma}(w))) \cdot \tilde{\sigma}'(w) \, dw \\ &= \int_{[0,1]} \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\tilde{\sigma}(w)) \tilde{\sigma}'_i(w) \, dw = \int_{[0,1]} \tilde{\sigma}^*(df) = \int_{\tilde{\sigma}} df . \end{aligned}$$

Moreover, since $\partial \tilde{\sigma} = \tilde{\sigma}_{1,1} - \tilde{\sigma}_{1,0}$ where $\tilde{\sigma}_{1,1}(0) = \sigma(b)$ and $\tilde{\sigma}_{1,0}(0) = \sigma(a)$, we have

$$\int_{\partial \tilde{\sigma}} f = \int_{\tilde{\sigma}_{1,1}} f - \int_{\tilde{\sigma}_{1,0}} f = f(\sigma(b)) - f(\sigma(a)) . \quad \blacksquare$$

Remark 6.2.5

The formulae in (6.2.2) and (6.2.4) can be used to define the line integral along a curve C that has only piecewise continuously differentiable parametric representations. In that situation, the integral is the sum of the integrals along each section of the curve C represented by a continuously differentiable parametric representation.

It is easy to prove that the fundamental theorem of line integral is also valid for a curve C that has only piecewise continuously differentiable parametric representations. It suffices to apply the fundamental theorem of line integral to each section of the curve C represented by a continuously differentiable parametric representation, and then sum up the results. ♠

For the rest of this section, we consider curves with a continuously differentiable or a piecewise continuously differentiable parametric representations.

Definition 6.2.6

A curve C with a parametric representation $\sigma : [a, b] \rightarrow \mathbb{R}^n$ is **closed** if $\sigma(a) = \sigma(b)$.

There are a couple of interesting consequences to the fundamental theorem of line integral.

Proposition 6.2.7

Let V be an open subset of $\mathbb{R}^n$ and $F : V \rightarrow \mathbb{R}^n$ be a continuous vector field. The integral $\int_C F \cdot \mathbf{t} \, ds$ is independent of the curve $C \subset V$ joining two points $\mathbf{a}$ and $\mathbf{b}$ in V if and only if $\oint_C F \cdot \mathbf{t} \, ds = 0$ for all closed curves C in V . As usual, $\mathbf{t}(\mathbf{u})$ is a tangent unit vector to the curve C at the point $\mathbf{u} \in C$ consistent with the orientation on C . The symbol $\oint$ is used to denote an integral over a closed curve.

Proof.

$\Rightarrow$) Suppose that C is a closed curve with a parametric representation $\sigma : [a, b] \rightarrow V$ consistent with its orientation. Let C_1 be the curve with parametric representation $\sigma_1 : [0, 1] \rightarrow V$ given by $\sigma_1(w) = \sigma(a + v(b - a)/2)$ for $0 \leq v \leq 1$ and let C_2 be the curve with parametric representation $\sigma_2 : [0, 1] \rightarrow V$ given by $\sigma_2(w) = \sigma(b + v(a - b)/2)$ for $0 \leq v \leq 1$. Those are two curves from $\sigma(a) = \sigma(b)$ to $\sigma((a + b)/2)$. So $\int_{C_1} F \cdot \mathbf{t} \, ds = \int_{C_2} F \cdot \mathbf{t} \, ds$ by hypothesis. Hence

$$\begin{aligned} 0 &= \int_{C_1} F \cdot \mathbf{t} \, ds - \int_{C_2} F \cdot \mathbf{t} \, ds = \int_{[0,1]} F(\sigma_1(v)) \cdot \sigma'_1(v) \, dv - \int_{[0,1]} F(\sigma_2(v)) \cdot \sigma'_2(v) \, dv \\ &= \int_{[0,1]} F(\sigma(a + v(b - a)/2)) \cdot \sigma'(a + v(b - a)/2) (b - a)/2 \, dv \\ &\quad - \int_{[0,1]} F(\sigma(b + v(a - b)/2)) \cdot \sigma'(b + v(a - b)/2) (a - b)/2 \, dv \\ &= \int_a^{(a+b)/2} F(\sigma(w)) \cdot \sigma'(w) \, dw - \int_b^{(a+b)/2} F(\sigma(w)) \cdot \sigma'(w) \, dw \\ &= \int_a^b F(\sigma(w)) \cdot \sigma'(w) \, dw = \int_C F \cdot \mathbf{t} \, ds \end{aligned}$$

$\Leftarrow$) Suppose that C_1 and C_2 are two curves from $\mathbf{a}$ to $\mathbf{b}$, and that $\sigma_i : [a, b] \rightarrow V$ is a parametric representation of C_i consistent with its orientation for $i = 1, 2$. Let C be the closed curve with the parametric representation $\sigma : [0, 1] \rightarrow V$ defined by

$$\sigma(v) = \begin{cases} \sigma_1(a + 2v(b - a)) & \text{if } 0 \leq v \leq 1/2 \\ \sigma_2((2b - a) + 2v(a - b)) & \text{if } 1/2 \leq v \leq 1 \end{cases}$$

By hypothesis, we have that $\int_C F \cdot \mathbf{t} \, ds = 0$. It follows that

$$\begin{aligned} 0 &= \int_C F \cdot \mathbf{t} \, ds = \int_0^{1/2} F(\sigma(v)) \cdot \sigma'(v) \, dv + \int_{1/2}^1 F(\sigma(v)) \cdot \sigma'(v) \, dv \\ &= \int_0^{1/2} F(\sigma_1(a + 2v(b - a))) \cdot \sigma'_1(a + 2v(b - a)) 2(b - a) \, dv \end{aligned}$$

$$\begin{aligned}
& + \int_{1/2}^1 F(\sigma_2((2b-a) + 2v(a-b))) \cdot \sigma_2'((2b-a) + 2v(a-b)) 2(a-b) dv \\
& = \int_a^b F(\sigma_1(w)) \cdot \sigma_1'(w) dw + \int_b^a F(\sigma_2(w)) \cdot \sigma_2'(w) dw \\
& = \int_a^b F(\sigma_1(w)) \cdot \sigma_1'(w) dw - \int_a^b F(\sigma_2(w)) \cdot \sigma_2'(w) dw = \int_{C_1} F \cdot \mathbf{t} ds - \int_{C_2} F \cdot \mathbf{t} ds .
\end{aligned}$$

Hence $\int_{C_1} F \cdot \mathbf{t} ds = \int_{C_2} F \cdot \mathbf{t} ds$. ■

Remark 6.2.8

The previous proof contains all the arguments that could be used to prove the following properties of line integral. We leave it to the reader to prove these properties.

If C_1 is a curve where the direction is from $\mathbf{a}$ to $\mathbf{b}$ and C_2 is the same curve where the direction is from $\mathbf{b}$ to $\mathbf{a}$, then $\int_{C_1} F \cdot \mathbf{t} ds = - \int_{C_2} F \cdot \mathbf{t} ds$.

If C_1 is a curve from $\mathbf{a}$ to $\mathbf{b}$ and C_2 is a curve from $\mathbf{b}$ to $\mathbf{c}$, then the union of these two curves forms a curve C_3 from $\mathbf{a}$ to $\mathbf{c}$ and $\int_{C_3} F \cdot \mathbf{t} ds = \int_{C_1} F \cdot \mathbf{t} ds + \int_{C_2} F \cdot \mathbf{t} ds$. ♠

Definition 6.2.9

Let $F : V \rightarrow \mathbb{R}^n$ be a continuous vector field. If $\int_C F \cdot \mathbf{t} ds$ is independent of the path C in V joining two points $\mathbf{a}$ and $\mathbf{b}$, then F is called a **conservative** vector field.

Proposition 6.2.10

Let V be an open subset $\mathbb{R}^n$ and $F : V \rightarrow \mathbb{R}^n$ be a continuous vector field. If F is conservative, then $F = \nabla f$ for some real valued function f defined on V .

Proof.

We have that V is the union of open connected components ². Let U be an open connected component of V .

Choose any point $\mathbf{x}_0 \in U$. We define a function $f : U \rightarrow \mathbb{R}$ as it follows. For $\mathbf{x} \in U$, let $f(\mathbf{x}) = \int_C F \cdot \mathbf{t} ds$ where $C \subset U$ is a curve from $\mathbf{x}_0$ to $\mathbf{x}$, and $\mathbf{t}(\mathbf{u})$ is a tangent unit vector to the curve C at the point $\mathbf{u} \in C$ that is consistent with the orientation on C from $\mathbf{x}_0$ to $\mathbf{x}$. Since F is conservative, this definition is independent of the curve $C \subset U$ from $\mathbf{x}_0$ to $\mathbf{x}$ used.

²The readers may consult an introductory textbook in topology (e.g. [20]) if they are not familiar with the concept of connected and path-connected subsets. In $\mathbb{R}^n$, connected and path-connected open subsets are the same. It is not true in all general topological spaces and it is not even true for close subsets of $\mathbb{R}^n$. For the present context, it suffices to know that $W \subset \mathbb{R}^n$ is connected if for each $\mathbf{a}, \mathbf{b} \in W$ there exists a curve from $\mathbf{a}$ to $\mathbf{b}$.

We prove that $\frac{\partial f}{\partial x_i}(\mathbf{x}) = F_i(\mathbf{x})$ for all $\mathbf{x} \in W$ and $1 \leq i \leq n$. Suppose that $\sigma : [0, 1] \rightarrow U$ is the parametric representation of the curve C consistent with the orientation from $\mathbf{x}_0$ to $\mathbf{x}$. Since U is open, there exists $\delta > 0$ such that $B_\delta(\mathbf{x}) \subset U$. Hence $\mathbf{x} + h\mathbf{e}_i \in U$ for $|h| < \delta$. Consider the curve $C_{i,h}$ from $\mathbf{x}_0$ to $\mathbf{x} + h\mathbf{e}_i$ with the parametric representation $\sigma_{i,h} : [0, 1+h] \rightarrow U$ defined by

$$\sigma_{i,h}(w) = \begin{cases} \sigma(w) & \text{if } 0 \leq w \leq 1 \\ \mathbf{x} + (w-1)\mathbf{e}_i & \text{if } 1 < w \leq 1+h \end{cases}$$

Hence

$$\begin{aligned} \frac{f(\mathbf{x} + h\mathbf{e}_i) - f(\mathbf{x})}{h} &= \frac{1}{h} \left(\int_{C_{i,h}} F \cdot \mathbf{t} \, ds - \int_C F \cdot \mathbf{t} \, ds \right) \\ &= \frac{1}{h} \int_1^{1+h} F(\mathbf{x} + (w-1)\mathbf{e}_i) \cdot (w-1)\mathbf{e}_i \, dw = \frac{1}{h} \int_0^h F(\mathbf{x} + w\mathbf{e}_i) \cdot \mathbf{e}_i \, dw \\ &= \frac{1}{h} \int_0^h F_i(\mathbf{x} + w\mathbf{e}_i) \, dw \rightarrow F_i(\mathbf{x}) \end{aligned}$$

as $h \rightarrow 0$ because F is continuous on V . ■

The converse to the previous proposition is true. If $F = \nabla f$ in V , then it follows from the fundamental theorem for line integral that $\int_C F \cdot \mathbf{t} \, ds = \int_C \nabla f \cdot \mathbf{t} \, ds = f(\sigma(b)) - f(\sigma(a)) = 0$ for all closed curves C in V , where $\sigma : [a, b] \rightarrow W$ is a parametric representation of C consistent with its orientation. Therefore, F is conservative according to Proposition 6.2.7.

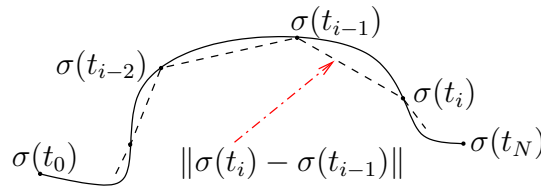
6.2.1 A Little Note on Arc Length

Suppose that C is a curve in $\mathbb{R}^n$ with the parametric representation $\sigma : [a, b] \rightarrow \mathbb{R}^n$. Recall that we may tolerate that there be a finite subset N of $[a, b]$ such that σ is one-to-one on $[a, b] \setminus N$ but not on the full interval $[a, b]$.

Given a partition $P = \{t_0, t_1, \dots, t_N\}$ of $[a, b]$, let

$$L_P(\sigma) = \sum_{i=1}^N \|\sigma(t_i) - \sigma(t_{i-1})\|.$$

This situation is represented in the figure below.



As the partition becomes finer, we expect to get better approximations of what we intuitively define as the “length” of C .

Definition 6.2.11

Let C be a curve in $\mathbb{R}^n$ with the parametric representation $\sigma : [a, b] \rightarrow \mathbb{R}^n$. The **length** L of the curve C is defined by

$$L = \sup\{L_P(\sigma) : P \text{ is a partition of } [a, b]\} . \quad (6.2.6)$$

The value of L is independent of the chosen parametric representation of C because only the points on C are used to define L in (6.2.6). If $L < \infty$, then we say that the curve C is **rectifiable**.

Since we assume that our parametric representations are continuously differentiable, we have the following result.

Theorem 6.2.12

Suppose that C is a curve in $\mathbb{R}^n$ with the parametric representation $\sigma : [a, b] \rightarrow \mathbb{R}^n$, then the length L of the curve C is given by

$$L = \int_C ds = \int_a^b \|\sigma'(t)\| dt .$$

To prove this theorem, we need the following lemma. We have already stated in Proposition 2.1.12 a version of this lemma for integrable real valued functions.

Lemma 6.2.13

If $f : [a, b] \rightarrow \mathbb{R}^n$ is continuous, then

$$\left\| \int_a^b f(t) dt \right\| \leq \int_a^b \|f(t)\| dt .$$

Proof.

Recall that

$$\int_a^b f(t) dt = \left(\int_a^b f_1(t) dt \quad \int_a^b f_2(t) dt \quad \dots \quad \int_a^b f_n(t) dt \right)^\top .$$

Given any unit vector $\mathbf{v} \in \mathbb{R}^n$, we get

$$\begin{aligned} \left| \left(\int_a^b f(t) dt \right) \cdot \mathbf{v} \right| &= \left| \int_a^b (f(t) \cdot \mathbf{v}) dt \right| \leq \int_a^b |f(t) \cdot \mathbf{v}| dt \\ &\leq \int_a^b \|f(t)\| \|\mathbf{v}\| dt = \int_a^b \|f(t)\| dt , \end{aligned} \quad (6.2.7)$$

where we have used Schwarz inequality in $\mathbb{R}^n$ for the second inequality. We may assume that $\int_a^b f(t) dt \neq \mathbf{0}$ since the result is trivial in that case. We get the conclusion of the lemma

by using the unit vector $\mathbf{v} = \left\| \int_a^b f(t) dt \right\|^{-1} \int_a^b f(t) dt$ in (6.2.7). ■

Proof (of Theorem 6.2.12).

i) Let $P = \{t_0, t_1, t_2, \dots, t_N\}$ be a partition of $[a, b]$. From the previous lemma, we have

$$\|\sigma(t_i) - \sigma(t_{i-1})\| = \left\| \int_{t_{i-1}}^{t_i} \sigma'(t) dt \right\| \leq \int_{t_{i-1}}^{t_i} \|\sigma'(t)\| dt$$

for $1 \leq i \leq N$. Hence

$$L_P(\sigma) = \sum_{i=1}^N \|\sigma(t_i) - \sigma(t_{i-1})\| \leq \sum_{i=1}^N \int_{t_{i-1}}^{t_i} \|\sigma'(t)\| dt = \int_a^b \|\sigma'(t)\| dt.$$

Since this is true for any partition P of $[a, b]$, we get that $L \leq \int_a^b \|\sigma'(t)\| dt$.

ii) Given $\epsilon > 0$, choose $\delta > 0$ such that $\|\sigma'(t) - \sigma'(s)\| < \epsilon$ for $|t - s| < \delta$. This is possible because $\sigma' : [a, b] \rightarrow \mathbb{R}^n$ is continuous on the compact set $[a, b]$. Therefore, it is uniformly continuous on $[a, b]$.

Let $P = \{t_0, t_1, t_2, \dots, t_N\}$ be a partition of $[a, b]$ such that $t_i - t_{i-1} < \delta$ for $1 \leq i \leq N$. Therefore $\|\sigma'(t)\| - \|\sigma'(t_i)\| \leq \|\sigma'(t) - \sigma'(t_i)\| < \epsilon$ for $t_{i-1} \leq t \leq t_i$. Hence

$$\begin{aligned} \int_{t_{i-1}}^{t_i} \|\sigma'(t)\| dt &\leq \int_{t_{i-1}}^{t_i} (\|\sigma'(t_i)\| + \epsilon) dt = \int_{t_{i-1}}^{t_i} \|\sigma'(t_i)\| dt + \epsilon(t_i - t_{i-1}) \\ &= \left\| \int_{t_{i-1}}^{t_i} \sigma'(t_i) dt \right\| + \epsilon(t_i - t_{i-1}) = \left\| \int_{t_{i-1}}^{t_i} (\sigma'(t) + (\sigma'(t_i) - \sigma'(t))) dt \right\| + \epsilon(t_i - t_{i-1}) \\ &\leq \left\| \int_{t_{i-1}}^{t_i} \sigma'(t) dt \right\| + \left\| \int_{t_{i-1}}^{t_i} (\sigma'(t_i) - \sigma'(t)) dt \right\| + \epsilon(t_i - t_{i-1}) \\ &\leq \underbrace{\left\| \int_{t_{i-1}}^{t_i} \sigma'(t) dt \right\|}_{=\|\sigma(t_i) - \sigma(t_{i-1})\|} + \underbrace{\int_{t_{i-1}}^{t_i} \|\sigma'(t_i) - \sigma'(t)\| dt}_{\leq \epsilon(t_i - t_{i-1})} + \epsilon(t_i - t_{i-1}) \\ &\leq \|\sigma(t_i) - \sigma(t_{i-1})\| + 2\epsilon(t_i - t_{i-1}) \end{aligned}$$

where the second equality comes from $\int_{t_{i-1}}^{t_i} \|\sigma'(t_i)\| dt = \left\| \int_{t_{i-1}}^{t_i} \sigma'(t_i) dt \right\|$ because $\sigma'(t_i)$ does not depend on t . We conclude that

$$\begin{aligned} \int_a^b \|\sigma'(t)\| dt &= \sum_{i=1}^N \int_{t_{i-1}}^{t_i} \|\sigma'(t)\| dt \leq \sum_{i=1}^N (\|\sigma(t_i) - \sigma(t_{i-1})\| + 2\epsilon(t_i - t_{i-1})) \\ &= L_P(\sigma) + 2\epsilon(b - a) \leq L + 2\epsilon(b - a). \end{aligned}$$

Since this is true for any $\epsilon > 0$, we get the inequality $\int_a^b \|\sigma'(t)\| dt \leq L$. ■

The reader will read in Section 6.3.1 a totally different story for the area of a surface.

Definition 6.2.14

Let C be a curve in $\mathbb{R}^n$ with the parametric representation $\sigma : [a, b] \rightarrow \mathbb{R}^n$. The **arc length function** along the curve C is defined by

$$s(t) = \int_a^t \|\sigma'(r)\| \, dr$$

for $a \leq t \leq b$.

Definition 6.2.15

A curve C in $\mathbb{R}^n$ is said to be **parameterized by arc length** if the parametric representation $\sigma : [a, b] \rightarrow \mathbb{R}^n$ of C satisfies $\|\sigma'(t)\| = 1$ for $a \leq t \leq b$.

If C is parameterized by arc length with the parametric representation $\sigma : [a, b] \rightarrow \mathbb{R}^n$, then the arc length function satisfies $s(t) = t - a$. In particular, the length L of C is $L = b - a$.

If $\sigma : [a, b] \rightarrow \mathbb{R}^n$ is a parametric representation of a curve C such that $\|\sigma'(t)\| \neq 0$ for $a \leq t \leq b$, then the arc length function $s : [a, b] \rightarrow [0, L]$ is a continuously differentiable and strictly increasing function. Thus $s^{-1} : [0, L] \rightarrow [a, b]$ exists and is a continuously differentiable function. If we set $\check{\sigma}(t) = \sigma(s^{-1}(t))$ for $0 \leq t \leq L$, then the curve C is parameterized by arc length with the parametric representation $\check{\sigma}$ because

$$\check{\sigma}'(t) = \sigma'(s^{-1}(t)) \frac{d}{dt} s^{-1}(t) = \frac{1}{\|\sigma'(s^{-1}(t))\|} \sigma'(s^{-1}(t))$$

and thus $\|\check{\sigma}'(t)\| = 1$ for $0 \leq t \leq L$. Note that, according to the Inverse Function Theorem,

$$\frac{d}{dt} s^{-1}(t) = \frac{1}{s'(s^{-1}(t))} = \frac{1}{\|\sigma'(s^{-1}(t))\|}$$

for $0 \leq t \leq L$.

6.3 Surface Integrals

In this section, a **surface** is a continuously differentiable function $\rho : W \rightarrow \mathbb{R}^n$ where W is an open subset of $\mathbb{R}^2$ and $D\rho(\mathbf{w})$ is of rank 2 for all $\mathbf{w} \in W$. Thus, a surface is a 2-dimensional manifold without boundary of class C^1 of $\mathbb{R}^n$ which can be represented with only one local chart.

As we did for the term “curve”, we will also interchangeably use the term “surface” to designate the function ρ or its image $S = \rho(W)$. In classical differential geometry, the function ρ is called a **parametric representation** of the surface S .

There may be occasions where we will tolerate that $\rho : W \rightarrow \mathbb{R}^n$ be one-to-one except on a set of measure zero.

Suppose that $f : S \rightarrow \mathbb{R}$ is a function. The **surface integral** of f over S is defined by

$$\iint_S f \, dS \equiv \iint_W f(\rho(\mathbf{w})) \left\| \left(\frac{\partial \rho}{\partial w_1} \times \frac{\partial \rho}{\partial w_2} \right) (\mathbf{w}) \right\| dw_1 dw_2 \quad (6.3.1)$$

if this integral exists, where (symbolically)

$$(\rho_{w_1} \times \rho_{w_2})(\mathbf{w}) = \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial \rho_1}{\partial w_1}(\mathbf{w}) & \frac{\partial \rho_2}{\partial w_1}(\mathbf{w}) & \frac{\partial \rho_3}{\partial w_1}(\mathbf{w}) \\ \frac{\partial \rho_1}{\partial w_2}(\mathbf{w}) & \frac{\partial \rho_2}{\partial w_2}(\mathbf{w}) & \frac{\partial \rho_3}{\partial w_2}(\mathbf{w}) \end{pmatrix}$$

if the determinant is developed along the top row only. We write

$$dS = \left\| \frac{\partial \rho}{\partial w_1} \times \frac{\partial \rho}{\partial w_2} \right\| dw_1 dw_2. \quad (6.3.2)$$

Remark 6.3.1

The expression in (6.3.2) is associated to the local representation of the volume element $\nu = dA$ for a 2-dimensional oriented manifold given in Example 5.7.10 where (W, U, ϕ) is replaced by (W, S, ρ) . See Remark 5.2.7.

As we did in Remark 6.2.2, we can show that (6.3.1) is nothing else than $\int_S f \, \nu$. We have

$$\begin{aligned} \int_S f \, \nu &= \int_W \rho^*(f \, \nu) = \int_W f(\rho(\mathbf{w}))(\rho^* \nu) = \int_W f(\rho(\mathbf{w})) \left\| \frac{\partial \rho}{\partial w_1}(\mathbf{w}) \times \frac{\partial \rho}{\partial w_2}(\mathbf{w}) \right\| dw_1 \wedge dw_2 \\ &= \int_W f(\rho(\mathbf{w})) \left\| \frac{\partial \rho}{\partial w_1}(\mathbf{w}) \times \frac{\partial \rho}{\partial w_2}(\mathbf{w}) \right\| dw_1 dw_2 \end{aligned}$$

where the third equality comes from (5.7.7) in Example 5.7.10 with $\nu = dA$. ♠

As we did for curves, we can recast our definition of orientation on a 2-dimensional manifold to the context of surfaces. Since we assume that the parametric representation ρ of a surface S is continuously differentiable and $D\rho(\mathbf{w})$ has rank 2 for all $\mathbf{w} \in W$, we may continuously assign to each point $\mathbf{u}$ of C a unit normal vector $\mathbf{n}(\mathbf{u})$.

If S is a surface in $\mathbb{R}^3$, then there is a close link between the selected orientation on S and the outward unit normal to S at each point of S . As we have seen in Remark 5.7.9, if $\mu_{\mathbf{u}} = [(\mathbf{u}, \mathbf{v}_{\mathbf{u},1}), (\mathbf{u}, \mathbf{v}_{\mathbf{u},2})]$ is the selected orientation on $T_{\mathbf{u}}S$, where $\{(\mathbf{u}, \mathbf{v}_{\mathbf{u},1}), (\mathbf{u}, \mathbf{v}_{\mathbf{u},2})\}$ is an orthonormal basis of $T_{\mathbf{u}}S$, then the outward unit normal is $\mathbf{n}_{\mathbf{u}} = (\mathbf{u}, \mathbf{v}_{\mathbf{u},1} \times \mathbf{v}_{\mathbf{u},2})$. Suppose that ρ is a parametric representation consistent with the orientation on S , then $\mu_{\rho(\mathbf{w})} = \left[\left(\rho(\mathbf{w}), \frac{\partial \rho}{\partial w_1}(\mathbf{w}) \right), \left(\rho(\mathbf{w}), \frac{\partial \rho}{\partial w_2}(\mathbf{w}) \right) \right]$ for all $\mathbf{w} \in W$ and

$$\mathbf{n}(\rho(\mathbf{w})) = \left\| \frac{\partial \rho}{\partial w_1}(\mathbf{w}) \times \frac{\partial \rho}{\partial w_2}(\mathbf{w}) \right\|^{-1} \left(\frac{\partial \rho}{\partial w_1}(\mathbf{w}) \times \frac{\partial \rho}{\partial w_2}(\mathbf{w}) \right) \quad (6.3.3)$$

for $w \in W$. Hence, the orientation determines the choice of the outward unit normal and vice-versa if the outward unit normal is consistent with the orientation on S . We may define

on orientation on S by continuously selecting an outward unit normal vector $\mathbf{n}$ at each point of S . This is another confirmation that the surface S has two possible orientations.

Suppose that $F : \Omega \rightarrow \mathbb{R}^3$ is a vector field. The **surface integral of F over the oriented surface S** is defined by

$$\iint_S F \cdot d\mathbf{S} \equiv \iint_S F \cdot \mathbf{n} dS = \iint_W F(\rho(\mathbf{w})) \cdot \left(\frac{\partial \rho}{\partial w_1} \times \frac{\partial \rho}{\partial w_2} \right) (\mathbf{u}) dw_1 dw_2, \quad (6.3.4)$$

where we assume that the parametric representation is consistent with the orientation on S and therefore $\mathbf{n}(\rho(\mathbf{w}))$ is given by (6.3.3) for all $\mathbf{w} \in W$. We write

$$d\mathbf{S} = \mathbf{n} dS = \left(\frac{\partial \rho}{\partial w_1} \times \frac{\partial \rho}{\partial w_2} \right) dw_1 dw_2. \quad (6.3.5)$$

Remark 6.3.2

The definition of the surface integral of a vector field in (6.3.4) is nothing else than the integral of a differential 2-form over a 2-dimensional manifold. namely, the integral $\int_S F \cdot \mathbf{n} dS$ in (6.3.4) is nothing else than the integral $\int_S \omega$ defined in (5.6.6) where ω is the differential 2-form $\omega = F_1 dx_2 \wedge dx_3 + F_2 dx_3 \wedge dx_1 + F_3 dx_1 \wedge dx_2$. This can easily be seen if we assume that the parametric representation ρ is consistent with the selected orientation on S . Then (W, S, ρ) is a local chart which is consistent with the orientation on S . As usual, we assume that $W \subset H_2$. Then

$$\begin{aligned} \int_S \omega &= \int_S F_1 dx_2 \wedge dx_3 + F_2 dx_3 \wedge dx_1 + F_3 dx_1 \wedge dx_2 \\ &= \int_W \rho^* (F_1 dx_2 \wedge dx_3 + F_2 dx_3 \wedge dx_1 + F_3 dx_1 \wedge dx_2) \\ &= \int_W \left(F_1(\rho(\mathbf{w})) \left(\frac{\partial \rho_2}{\partial w_1}(\mathbf{w}) \frac{\partial \rho_3}{\partial w_2}(\mathbf{w}) - \frac{\partial \rho_2}{\partial w_2}(\mathbf{w}) \frac{\partial \rho_3}{\partial w_1}(\mathbf{w}) \right) \right. \\ &\quad + F_2(\rho(\mathbf{w})) \left(\frac{\partial \rho_3}{\partial w_1}(\mathbf{w}) \frac{\partial \rho_1}{\partial w_2}(\mathbf{w}) - \frac{\partial \rho_3}{\partial w_2}(\mathbf{w}) \frac{\partial \rho_1}{\partial w_1}(\mathbf{w}) \right) \\ &\quad \left. + F_3(\rho(\mathbf{w})) \left(\frac{\partial \rho_1}{\partial w_1}(\mathbf{w}) \frac{\partial \rho_2}{\partial w_2}(\mathbf{w}) - \frac{\partial \rho_1}{\partial w_2}(\mathbf{w}) \frac{\partial \rho_2}{\partial w_1}(\mathbf{w}) \right) \right) dw_1 \wedge dw_2 \\ &= \int_W F(\rho(\mathbf{w})) \cdot \left(\frac{\partial \rho}{\partial w_1}(\mathbf{w}) \times \frac{\partial \rho}{\partial w_2}(\mathbf{w}) \right) dw_1 dw_2. \end{aligned}$$



Remark 6.3.3

If the surface S is given by $w_3 = g(w_1, w_2)$ for $(w_1, w_2) \in W \subset \mathbb{R}^2$, then

$$\begin{aligned} \iint_S f dS &= \iint_W f(w_1, w_2, g(w_1, w_2)) \left(\left(\frac{\partial g}{\partial w_1}(w_1, w_2) \right)^2 + \left(\frac{\partial g}{\partial w_2}(w_1, w_2) \right)^2 + 1 \right)^{1/2} dw_1 dw_2. \end{aligned}$$

To prove this claim, we note that in this case a parametric representation for S is given by $\rho(w_1, w_2) = w_1 \mathbf{e}_1 + w_2 \mathbf{e}_2 + g(w_1, w_2) \mathbf{e}_3$ for (w_1, w_2) in W . We have

$$\frac{\partial \rho}{\partial w_1}(\mathbf{w}) \times \frac{\partial \rho}{\partial w_2}(\mathbf{w}) = \left(-\frac{\partial g}{\partial w_1}(\mathbf{w}), -\frac{\partial g}{\partial w_2}(\mathbf{w}), 1 \right)$$

and

$$dS = \left\| \left(\frac{\partial \rho}{\partial w_1} \times \frac{\partial \rho}{\partial w_2} \right) (\mathbf{w}) \right\| dw_1 dw_2 = \left(\left(\frac{\partial g}{\partial w_1}(\mathbf{w}) \right)^2 + \left(\frac{\partial g}{\partial w_2}(\mathbf{w}) \right)^2 + 1 \right)^{1/2} dw_1 dw_2 .$$

We also have

$$\begin{aligned} \iint_S F \cdot \mathbf{n} dS &= \epsilon \iint_U \left(-F_1(w_1, w_2, g(w_1, w_2)) \frac{\partial g}{\partial w_1}(w_1, w_2) \right. \\ &\quad \left. -F_2(w_1, w_2, g(w_1, w_2)) \frac{\partial g}{\partial w_2}(w_1, w_2) + F_3(w_1, w_2, g(w_1, w_2)) \right) dw_1 dw_2 , \end{aligned}$$

where $\epsilon = 1$ if $\left(-\frac{\partial g}{\partial w_1}(\mathbf{w}), -\frac{\partial g}{\partial w_2}(\mathbf{w}), 1 \right)$ points in the same direction than the normal $\mathbf{n}(w_1, w_2, g(w_1, w_2))$ associated to the orientation on S and $\epsilon = -1$ otherwise. This is a special case of (6.3.4) with the parametric representation for S given by $\rho(w_1, w_2) = w_1 \mathbf{e}_1 + w_2 \mathbf{e}_2 + g(w_1, w_2) \mathbf{e}_3$ for (w_1, w_2) in W . In particular, we have

$$\mathbf{n}(\rho(\mathbf{w})) = \epsilon \left\| \left(-\frac{\partial g}{\partial w_1}(\mathbf{w}), -\frac{\partial g}{\partial w_2}(\mathbf{w}), 1 \right) \right\|^{-1} \left(-\frac{\partial g}{\partial w_1}(\mathbf{w}), -\frac{\partial g}{\partial w_2}(\mathbf{w}), 1 \right)$$

and

$$\mathbf{n} dS = \epsilon \left(-\frac{\partial g}{\partial w_1}, -\frac{\partial g}{\partial w_2}, 1 \right) dw_1 dw_2 .$$



Remark 6.3.4

The reader may have noted that we have only considered surfaces given by parametric representations defined on open sets in this section while we considered curves given by parametric representations defined on closed interval in the previous section. One of the reasons why we did that was to introduce the fundamental theorem of line integral. It also greatly simplified the issue about local charts on surfaces.

We will address the issue of the boundary of a surface in Section 6.4.



6.3.1 A Little Note on Surface Area

We would like to use an approach similar to the approach that we have used in Section 6.2.1 to define the area of a surface. For instance, we would like to use polygonal surfaces (like triangles or rectangles) to approximate the area of a surface. Unfortunately, this approach does not work as expected.

Suppose that $\rho : R \rightarrow \mathbb{R}^3$ is a parametric representation of a surface S where $R \subset \mathbb{R}^2$ is an open rectangle. Suppose that we cut R into small triangles R_i . For each triangle R_i , let A_i be the area of the triangles in $\mathbb{R}^3$ having vertices at $\rho(\mathbf{a}_i)$, $\rho(\mathbf{b}_i)$ and $\rho(\mathbf{c}_i)$ where $\mathbf{a}_i$, $\mathbf{b}_i$ and $\mathbf{c}_i$ are the vertices of R_i . We approximate the area of $\rho(R_i) \subset \mathbb{R}^3$ by A_i . If the triangles R_i are very small, then the area of $\rho(R)$ should be approximately equal to the sum of the A_i over all the triangles R_i that compose R , should not it? Unfortunately, this is not always the case. H. A. Schwarz (see [22] and [26]) gave an example for a simple cylinder where we can choose the triangles R_i such that the sum of the A_i over all the triangles R_i converges to infinity as the maximum size of the triangles R_i goes to zero. The approximation of the surface using little polygonal surface require a stronger notion of convergence than pointwise convergence. We will not cover this vast subject in the present document. The reader should consult [22] for a more general view of the subject.

Thus, there is no chance that we could use a formula like (6.2.6) to define the area of a surface. If S is a surface as defined at the beginning of this section, then the tradition is to define the **area of the surface** S as

$$A = \iint_S dS = \iint_W \left\| \frac{\partial \rho}{\partial w_1} \times \frac{\partial \rho}{\partial w_2} \right\| dw_1 dw_2 .$$

6.4 The Classical trio

The trio that we have in mind is composed of the Green's theorem, the classical Stokes' theorem and the divergence theorem. We present all of them below.

Theorem 6.4.1 (Green's Theorem)

Let $S \subset \mathbb{R}^2$ be a compact 2-dimensional manifold of class C^1 with boundary. Suppose that $F_i : \mathbb{R}^2 \rightarrow \mathbb{R}$ is of class C^1 for $1 \leq i \leq 2$. Then

$$\oint_{\partial S} F_1 dx_1 + F_2 dx_2 = \int_S \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) dx_1 \wedge dx_2 \quad (6.4.1)$$

The orientation on the boundary ∂S is such that the interior of S is to the left when travelling along ∂S . In such case, we say that the orientation or direction of ∂S is **positive** with respect to S .

Since S is a 2-dimensional manifold in $\mathbb{R}^2$, the integral on the right hand side of (6.4.1) is simply $\int_S \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) dx_1 dx_2$. Note that $(S^\circ, S^\circ, \text{Id})$ is a local chart for S consistent with the orientation on S . As we said in the last section, the value of the integral does not change if we include or not the boundary of S .

Proof.

This is a special case of Stokes' theorem, Theorem 5.6.11, with $\omega = F_1 dx_1 + F_2 dx_2$ because

$$d\omega = \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) dx_1 \wedge dx_2 .$$

The orientation on $T_u S$ for $\mathbf{u} \in S$ is the standard orientation given by $\mu_u = [(\mathbf{u}, \mathbf{e}_1), (\mathbf{u}, \mathbf{e}_2)]$. According to Definition 5.3.15, the induced orientation on ∂S from S is given by the tangent vector $(\mathbf{u}, \mathbf{t}(\mathbf{u}))$ for $\mathbf{u} \in \partial S$ such that $\mu_u = [(\mathbf{u}, \mathbf{n}(\mathbf{u})), (\mathbf{u}, \mathbf{t}(\mathbf{u}))]$ where $(\mathbf{u}, \mathbf{n}(\mathbf{u}))$ is the outward unit normal to ∂S at $\mathbf{u} \in \partial S$ as defined in Definition 5.3.14 (Figure 6.2). ■

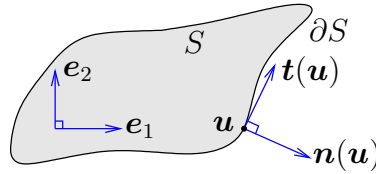


Figure 6.2: Illustration associated to the statement of Green's Theorem, Theorem 6.4.1

Theorem 6.4.2 (Classical Stokes' Theorem)

Let $S \subset \mathbb{R}^3$ be a compact and oriented 2-dimension manifold of class C^1 with boundary. Suppose that $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is a vector field of class C^1 . Then

$$\oint_{\partial S} F \cdot \mathbf{t} \, ds = \int_S \text{curl } F \cdot \mathbf{n} \, dA , \quad (6.4.2)$$

where $(\mathbf{u}, \mathbf{n}(\mathbf{u}))$ is the outward unit normal to S at $\mathbf{u} \in S$ as defined in Definition 5.3.18 and $(\mathbf{u}, \mathbf{t}(\mathbf{u}))$ is a unit tangent vector to ∂S at $\mathbf{u} \in \partial S$ such that $[(\mathbf{u}, \mathbf{n}(\mathbf{u})), (\mathbf{u}, \tilde{\mathbf{n}}(\mathbf{u})), (\mathbf{u}, \mathbf{t}(\mathbf{u}))] = [(\mathbf{u}, \mathbf{e}_1), (\mathbf{u}, \mathbf{e}_2), (\mathbf{u}, \mathbf{e}_3)]$ if $(\mathbf{u}, \tilde{\mathbf{n}}(\mathbf{u}))$ is the outward unit normal to ∂S at $\mathbf{u} \in \partial S$ as defined in Definition 5.3.14 (Figure 6.3).

The expressions dA and ds in the statement of the previous theorem refer to the volume elements on a 2-dimensional manifold and a 1-dimensional manifold respectively. If the 2-dimensional manifold S is a surface given by a parametric representation, then dA can be replaced by the expression in (6.3.2) when computing the integral. Similarly, if the boundary ∂S is a curve given by a parametric representation, then ds can be replaced by the expression in (6.2.3) when computing the integral.

Proof.

This is a special case of Stokes' theorem, Theorem 5.6.11, with $\omega = F_1 dx_1 + F_2 dx_2 + F_3 dx_3$. We get

$$\int_{\partial S} \omega = \int_S d\omega . \quad (6.4.3)$$

We have from (5.7.9) of Theorem 5.7.11 that $n_1 dA = dx_2 \wedge dx_3$, $n_2 dA = dx_3 \wedge dx_1$ and $n_3 dA = dx_1 \wedge dx_2$ on S . Thus

$$\begin{aligned} (\operatorname{curl} F \cdot \mathbf{n}) dA &= \left(\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3} \right) n_1 dA + \left(\frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1} \right) n_2 dA + \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) n_3 dA \\ &= \left(\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3} \right) dx_2 \wedge dx_3 + \left(\frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1} \right) dx_3 \wedge dx_1 + \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) dx_1 \wedge dx_2 \\ &= d\omega \end{aligned} \quad (6.4.4)$$

We have from (5.7.5) of Theorem 5.7.7 that $t_i ds = dx_i$ for $1 \leq i \leq 3$ on ∂S . Hence

$$(F \cdot \mathbf{t}) ds = (F_1 t_1) ds + (F_2 t_2) ds + (F_3 t_3) ds = F_1 dx_1 + F_2 dx_2 + F_3 dx_3 = \omega. \quad (6.4.5)$$

We get the conclusion of the theorem by substituting (6.4.4) and (6.4.5) into (6.4.3). $\blacksquare$

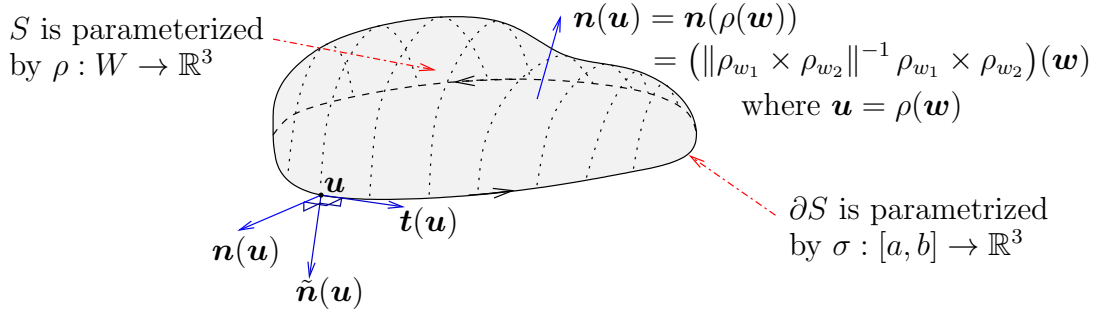


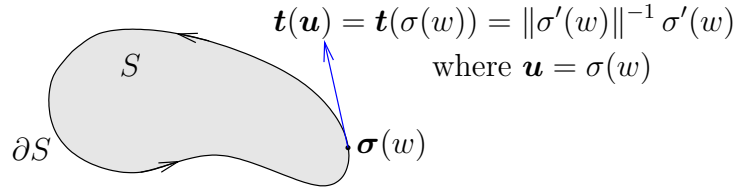
Figure 6.3: Illustration associated to the statement of Stokes' theorem, Theorem 6.4.2.

In the classical Stokes' theorem, we have that the outward unit normal $\tilde{\mathbf{n}}(\mathbf{u})$ to S at $\mathbf{u} \in \partial S$ is given by $\tilde{\mathbf{n}}(\mathbf{u}) = \mathbf{t}(\mathbf{u}) \times \mathbf{n}(\mathbf{u})$ according to Definition 5.3.18 because $[(\mathbf{u}, \mathbf{n}(\mathbf{u})), (\mathbf{u}, \tilde{\mathbf{n}}(\mathbf{u})), (\mathbf{u}, \mathbf{t}(\mathbf{u}))] = [(\mathbf{u}, \mathbf{e}_1), (\mathbf{u}, \mathbf{e}_2), (\mathbf{u}, \mathbf{e}_3)]$.

In the statement of the classical Stokes' theorem, the **positive orientation** or **direction** on ∂S with respect to S is the direction for which the surface S is to our left when walking along the curve C in that direction assuming that our standing position is in the direction of the outward unit normal $\mathbf{n}(\mathbf{u})$ to S at $\mathbf{u} \in \partial S$ (Figure 6.3).

Green's Theorem is a special case of the classical Stokes' theorem. Suppose that S is a surface in the x_1, x_2 plane and the orientation on S is given by $[(\mathbf{u}, \mathbf{e}_1), (\mathbf{u}, \mathbf{e}_2)]$. Then the outward unit normal to S at $\mathbf{u} \in S$ is given by $\mathbf{n}(\mathbf{u}) = \mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3$. The classical Stokes' theorem becomes

$$\oint_{\partial S} F \cdot \mathbf{t} ds = \int_S \operatorname{curl} F \cdot \mathbf{e}_3 dA. \quad (6.4.6)$$



If we develop both sides of (6.4.6) and assume that $F_3 = 0$ on S , then we get

$$\oint_C F_1 dx_1 + F_2 dx_2 = \int_S \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) dx_1 \wedge dx_2, \quad (6.4.7)$$

where we have used (6.4.4) and (6.4.5) with $F_3 = 0$ and $\mathbf{n} = \mathbf{e}_3$. We may also note that $dx_3 = 0$ because S is in the x_1, x_2 plane. Therefore, all local charts (W, U, ϕ) of S must satisfy $\phi_3(\mathbf{w}) = 0$ for all $\mathbf{w} \in W$.

Remark 6.4.3

After having stated Green's and Stokes' theorems for compact manifolds of class C^1 , we could restate them based on Theorem 5.6.23 or Theorem 5.6.31. For instance, if we used Theorem 5.6.23, then we assume that S is a bounded manifold without boundary of class C^1 . In that situation, ∂S should then be replaced by the set of regular points S_r of S in the statement of Green's and Stokes' theorems.

If the differential form ω that is integrated has a compact support, then the manifold S does not have to be bounded as in the statement of Theorem 5.6.23. If the differential form ω does not have a compact support, then the manifold S must be bounded. We then use Theorem 5.6.23 with ω replaced by $\psi\omega$ where ψ is a function of class C^∞ with compact support and such that $\psi(\mathbf{u}) = 1$ for all $\mathbf{u} \in S$.

From now on, S_r will be denoted ∂S even if the points of S_r are not in S . This is closer to the notation usually adopted in the literature. This should not cause, we hope, any ambiguity in the examples, applications and questions.

There are two issues to address when studying Lines and surfaces as defined in this chapter.

1. The parametric representations of the surfaces that we use are defined on open sets. We are basically ignoring the boundary of the surface. The integral of a differential form on surface S does not change if we include or not the boundary because the boundary represent a set of measure zero and, very often in applications, of zero content. Therefore, we may ignore any local charts that cover the boundary as long as we have local charts that cover the interior of S . This is the situation that we have with our parametric representations of surfaces. In fact, the parametric representation yields a single local chart that cover the entire interior of S .
2. The parametric representations of Lines or surfaces that we use are not always providing local charts for all points of the line or the surface. For instance, suppose that we want to integrate a continuous differential 1-form ω over the unit circle S^1 . We may use the parametric representation $\sigma(\theta) = (\cos(\theta), \sin(\theta))$ for $0 \leq \theta < 2\pi$. But this does not

represent a local chart. This yields the local chart $C_1 = ([0, 2\pi[, S^1 \setminus \{e_1\}, \sigma)$ but we are missing a local chart to cover e_1 . Rigorously, we must add another local chart like $C_2 = ([-\epsilon, \epsilon[, \sigma[-\epsilon, \epsilon], \sigma)$ to completely cover S^1 . Then we use a partition of unity $\{\psi_1, \psi_2\}$ subordinate to $\{C_1, C_2\}$. The integral is therefore

$$\int_{S^1} \omega = \sum_{j=1}^2 \int_{S^1} \psi_j \omega = \int_{]0, 2\pi[} \sigma^*(\psi_1 \omega) + \int_{]-\epsilon, \epsilon[} \sigma^*(\psi_2 \omega) .$$

However, since ω is continuous of the compact set S^1 , we can show using one of the convergence theorems for integrals (i.e. Theorem 2.4.2 or Lebesgue dominated convergence theorem) that $\int_{]0, 2\pi[} \sigma^*(\psi_1 \omega) \rightarrow \int_{]0, 2\pi[} \sigma^*(\omega)$ and $\int_{]-\epsilon, \epsilon[} \sigma^*(\psi_2 \omega) \rightarrow 0$ as $\epsilon \rightarrow 0$.

We have a similar issue when using cylindrical or spherical coordinates. For instance, to integrate over the unit ball $\{\mathbf{x} \in \mathbb{R}^3 : \|\mathbf{x}\| = 1\}$, we may use the parametric representation $\rho(\theta, \phi) = (\cos(\theta) \sin(\phi), \sin(\theta) \sin(\phi), \cos(\phi))$ for $0 \leq \theta < 2\pi$ and $0 \leq \phi < \pi$. First of all, this parametric representation is not properly defined at the poles because ρ is not one-to-one at the poles. This parametric representation yields a local chart $C_1 = (\{(\theta, \phi); 0 < \theta < 2\pi, 0 < \phi < \pi\}, B_1(\mathbf{0}) \setminus \{\sin(\phi), 0, \cos(\phi) : 0 \leq \phi < \pi\}, \rho)$. To rigorously compute the integral over D , we must add local charts to cover the region $\{\sin(\phi), 0, \cos(\phi) : 0 \leq \phi < \pi\}$. For instance, $C_2 = (\{(\theta, \phi); -\epsilon < \theta < \epsilon, 0 < \phi < \pi\}, \rho(\{(\theta, \phi); -\epsilon < \theta < \epsilon, 0 < \phi < \pi\}), \rho)$, $C_3 = (B_\epsilon(\mathbf{0}) \subset \mathbb{R}^2; \rho_3(B_\epsilon(\mathbf{0})); \rho_3(x, y) = (x, y, \sqrt{1 - x^2 - y^2}))$ and $C_4 = (B_\epsilon(\mathbf{0}) \subset \mathbb{R}^2; \rho_4(B_\epsilon(\mathbf{0})); \rho_4(x, y) = (x, y, -\sqrt{1 - x^2 - y^2}))$. Then we use a partition of unity $\{\psi_j\}_{1 \leq j \leq 4}$ subordinate to $\{C_j\}_{1 \leq j \leq 4}$. The integral is therefore

$$\begin{aligned} \int_{S^2} \omega = \sum_{j=1}^4 \int_{S^1} \psi_j \omega = & \int_{\{(\theta, \phi); 0 < \theta < 2\pi, 0 < \phi < \pi\}} \rho^*(\psi_1 \omega) + \int_{\{(\theta, \phi); -\epsilon < \theta < \epsilon, 0 < \phi < \pi\}} \rho^*(\psi_2 \omega) \\ & + \int_{B_\epsilon(\mathbf{0})} \rho_3^*(\psi_3 \omega) + \int_{B_\epsilon(\mathbf{0})} \rho_4^*(\psi_4 \omega) . \end{aligned}$$

However, since ω is continuous of the compact set S^2 , we can show as we mentioned for the integral over S^1 that $\int_{\{(\theta, \phi); 0 < \theta < 2\pi, 0 < \phi < \pi\}} \rho^*(\psi_1 \omega) \rightarrow \int_{\{(\theta, \phi); 0 < \theta < 2\pi, 0 < \phi < \pi\}} \rho^*(\omega)$ and the other three integrals converge to 0 as $\epsilon \rightarrow 0$.

Those are the two situations frequently encountered in the examples, applications and problems presented later. We will not refer to this remark in similar situations but the reader should always have in mind that such a manipulation is done to justify the use of Green's and Stokes' theorems. Our use of Green's and Stokes' theorems will be more heuristic. ♠

Example 6.4.4

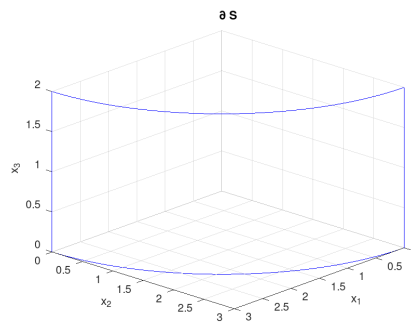
Consider the vector field

$$F(\mathbf{x}) = (x_1 x_2 \quad x_2 x_3 \quad x_3)^T . \quad (6.4.8)$$

Let S be the surface defined by the portion of the cylinder of radius 3 and axis x_3 which is inside the region defined by $x_1 > 0$, $x_2 > 0$ and $0 < x_3 < 2$. We compute the line integral of F along ∂S .

The selected orientation on the surface S is such that the outward unit normal to S is pointing away from the x_3 axis. The positive direction of the line integral on ∂S is the direction for which the interior of S is to the left when traveling along ∂S with the upward position in the direction of the outward unit normal to S .

We split ∂S in four curves: C_1 is the straight line from $(3, 0, 2)$ to $(3, 0, 0)$, C_2 is the arc of radius 3 from $(3, 0, 0)$ to $(0, 3, 0)$, C_3 is the straight line from $(0, 3, 0)$ to $(0, 3, 2)$, and C_4 is the arc of radius 3 from $(0, 3, 2)$ to $(3, 0, 2)$. The positive directions on C_1 , C_2 , C_3 and C_4 is obtained from the positive direction on ∂S . We plot ∂S in the following figure.



In theory, the corners should not be included in the parametric representation of the four curves above since they are singular points and do not belong to S_r but, as we have mentioned in Remark 5.6.24, they may be included without any consequences.

For each curve above, we choose a parametric representation of this curve which is consistent with the orientation on ∂S , and express F along this curve.

Curve	Parametric Representation	F along the curve	range of w
C_1	$\sigma_1(w) = \begin{pmatrix} 3 \\ 0 \\ 2-w \end{pmatrix}$	$F(\sigma_1(w)) = \begin{pmatrix} 0 \\ 0 \\ 2-w \end{pmatrix}$	$0 \leq w \leq 2$
C_2	$\sigma_2(w) = \begin{pmatrix} 3 \cos(w) \\ 3 \sin(w) \\ 0 \end{pmatrix}$	$F(\sigma_2(w)) = \begin{pmatrix} 9 \sin(w) \cos(w) \\ 0 \\ 0 \end{pmatrix}$	$0 \leq w \leq \pi/2$
C_3	$\sigma_3(w) = \begin{pmatrix} 0 \\ 3 \\ w \end{pmatrix}$	$F(\sigma_3(w)) = \begin{pmatrix} 0 \\ 3w \\ w \end{pmatrix}$	$0 \leq w \leq 2$
C_4	$\sigma_4(w) = \begin{pmatrix} 3 \sin(w) \\ 3 \cos(w) \\ 2 \end{pmatrix}$	$F(\sigma_4(w)) = \begin{pmatrix} 9 \sin(w) \cos(w) \\ 6 \cos(w) \\ 2 \end{pmatrix}$	$0 \leq w \leq \pi/2$

Hence,

$$\begin{aligned} \int_C F \cdot \mathbf{t} \, ds &= \sum_{i=1}^4 \int_{C_i} F \cdot \mathbf{t} \, ds = \int_0^2 (0, 0, 2-w) \cdot (0, 0, -1) \, dw \\ &+ \int_0^{\pi/2} (9 \sin(w) \cos(w), 0, 0) \cdot (-3 \sin(w), 3 \cos(w), 0) \, dw + \int_0^2 (0, 3w, w) \cdot (0, 0, 1) \, dw \\ &+ \int_0^{\pi/2} (9 \sin(w) \cos(w), 6 \cos(w), 2) \cdot (3 \cos(w), -3 \sin(w), 0) \, dw = -2 - 9 + 2 + 0 = -9. \end{aligned}$$

We now use the classical Stokes' theorem to compute the integral of the vector field (6.4.8) along ∂S in the positive direction of ∂S . The classical Stokes' theorem states that $\int_{\partial S} F \cdot \mathbf{t} \, ds = \int_S \operatorname{curl} F \cdot \mathbf{n} \, dS$. The curl of F is

$$\operatorname{curl} F = \left(\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3} \right) \mathbf{e}_1 - \left(\frac{\partial F_3}{\partial x_1} - \frac{\partial F_1}{\partial x_3} \right) \mathbf{e}_2 + \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) \mathbf{e}_3 = (-x_2 \quad 0 \quad -x_1)^\top.$$

To compute the surface integral on S , we use cylindrical coordinates to obtain a parametric representation for S . Namely, $\rho(w_1, w_2) = (3 \cos(w_1) \quad 3 \sin(w_1) \quad w_2)^\top$ for $0 < w_1 < \pi/2$ and $0 < w_2 < 2$. Hence $\operatorname{curl} F(\rho(\mathbf{w})) = (-3 \sin(w_1) \quad 0 \quad -3 \cos(w_1))^\top$ and

$$\frac{\partial \rho}{\partial w_1} \times \frac{\partial \rho}{\partial w_2} = (3 \cos(w_1) \quad 3 \sin(w_1) \quad 0)^\top.$$

This form of $\frac{\partial \rho}{\partial w_1} \times \frac{\partial \rho}{\partial w_2}$ is not surprising because the outward unit normal to S at $\rho(\mathbf{w}) \in S$ is given by $\mathbf{n}(\rho(\mathbf{w})) = (\cos(w_1) \quad \sin(w_1) \quad 0)^\top$. This shows that the parametric representation is consistent with the orientation on S . Hence

$$dA = \left\| \frac{\partial \rho}{\partial w_1} \times \frac{\partial \rho}{\partial w_2} \right\| dw_1 dw_2 = 3 dw_1 dw_2.$$

We finally have

$$\int_S \operatorname{curl} F \cdot \mathbf{n} \, dA = 3 \int_0^{\pi/2} \int_0^2 (-3 \sin(w_1), 0, -3 \cos(w_1)) \cdot (\cos(w_1), \sin(w_1), 0) dw_2 dw_1 = -9$$

as expected. ♣

Proposition 6.4.5

Let $F : V \rightarrow \mathbb{R}^3$ be a continuously differentiable vector field on a convex open subset V of $\mathbb{R}^3$. If F satisfies $\operatorname{curl} F = \mathbf{0}$ in V , then F is conservative.

Proof.

This is a special case of Theorem 4.5.12 for differential 1-forms $\omega = F_1 dx_1 + F_2 dx_2 + F_3 dx_3$

because a convex subset of $\mathbb{R}^3$ is star-shaped. Since $d\omega = 0$ (i.e. $\text{curl } F = \mathbf{0}$), it follows from Theorem 4.5.12 that there exists a 0-form f such that $df = \omega$ (i.e. $\nabla f = F$). The conclusion of the proposition follows from the fundamental theorem of line integral.

We provide another proof of this proposition. This proof is interesting in itself because it provides a glance at how Stokes' theorem could be used to generalize the proposition to surfaces.

Given any closed curve C in V , choose a surface $S \subset V$ with boundary C . A surface S exists because V is convex. We get from Stokes's Theorem that

$$\oint_C F \cdot d\mathbf{s} = \iint_S \text{curl } F \cdot d\mathbf{S} = 0 .$$

Since C is arbitrary, we get from Proposition 6.2.7 that $\oint_{\tilde{C}} F \cdot d\mathbf{s}$ is independent of the path $\tilde{C}$ in V connecting two points $\mathbf{a}$ and $\mathbf{b}$. Thus, F is conservative. ■

The set V in the previous proposition must be convex or, more generally, every curve in V must be continuously contractible in V to a point in V so that we may claim that it is the boundary of a surface in V (such curve is said to be “continuously homotopic” to a point). For instance, consider the non-convex set $V = \mathbb{R}^3 \setminus \{(0, 0, x_3) : x_3 \in \mathbb{R}\}$. There is no surface in V that has the circle $C = \{(x_1, x_2, 0) : x_1^2 + x_2^2 = 1\}$ as its boundary.

Theorem 6.4.6 (Divergence Theorem)

Let $S \subset \mathbb{R}^3$ be a compact 3-dimension manifold of class C^1 with a boundary. Suppose that $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is a vector field of class C^1 . Then

$$\int_S \text{div } F \, dV = \int_{\partial S} F \cdot \mathbf{n} \, dA , \quad (6.4.9)$$

where $\mathbf{n}(\mathbf{u})$ is the outward unit normal to ∂S at $\mathbf{u} \in \partial S$ as given in Definition 5.3.14 (Figure 6.4).

Recall that if $[(\mathbf{u}, \mathbf{v}_1(\mathbf{u})), (\mathbf{u}, \mathbf{v}_2(\mathbf{u}))]$ is the orientation on $T_{\mathbf{u}}\partial S$ for $\mathbf{u} \in \partial S$, then $[(\mathbf{u}, \mathbf{n}(\mathbf{u})), (\mathbf{u}, \mathbf{v}_1(\mathbf{u})), (\mathbf{u}, \mathbf{v}_2(\mathbf{u}))] = [(\mathbf{u}, \mathbf{e}_1), (\mathbf{u}, \mathbf{e}_2), (\mathbf{u}, \mathbf{e}_3)]$, the standard orientation on $T_{\mathbf{u}}S \cong \mathbb{R}^3$ for $\mathbf{u} \in \partial S$.

The expressions dA and dV in the statement of the previous theorem refer to the volume elements on a 2-dimensional manifold and a 3-dimensional manifold respectively. If the 2-dimensional manifold S is a surface given by a parametric representation, then dS can be replaced by the expression in (6.3.2) when computing the integral. As for $dV = dx_1 \wedge dx_2 \wedge dx_3$, it can simply be replaced by $dx_1 \, dx_2 \, dx_3$.

Proof.

This is a special case of Stokes' theorem, Theorem 5.6.11, with $\omega = F_1 \, dx_2 \wedge dx_3 + F_2 \, dx_3 \wedge dx_1 + F_3 \, dx_1 \wedge dx_2$. We have

$$\int_{\partial S} \omega = \int_S d\omega ,$$

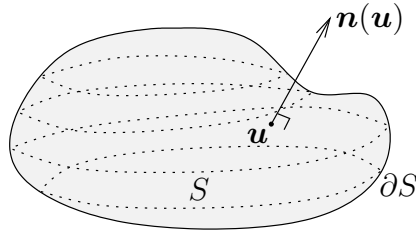


Figure 6.4: Illustration associated to the statement of the divergence theorem, Theorem 6.4.6.

where $d\omega = (F_1 + F_2 + F_3) dx_1 \wedge dx_2 \wedge dx_3 = (\operatorname{div} F) dx_1 \wedge dx_2 \wedge dx_3$ and

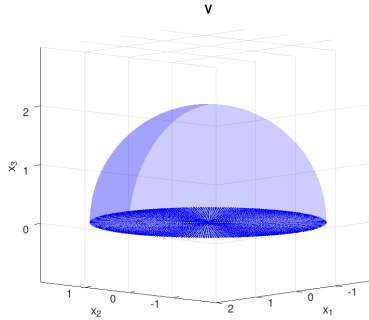
$$\begin{aligned} (F \cdot \mathbf{n}) dA &= F_1 n_1 dA + F_2 n_2 dA + F_3 n_3 dA \\ &= F_1 dx_2 \wedge dx_3 + F_2 dx_3 \wedge dx_1 + F_3 dx_1 \wedge dx_2 = \omega \end{aligned}$$

because $n_1 dA = dx_2 \wedge dx_3$, $n_2 dA = dx_3 \wedge dx_1$ and $n_3 dA = dx_1 \wedge dx_2$ on ∂S according to Theorem 5.7.11. ■

Remark 6.4.3 is also valid for the divergence theorem.

Example 6.4.7

We again consider the vector field F given in (6.4.8). Let V be the open half-sphere $x_1^2 + x_2^2 + x_3^2 < 4$ with $x_3 > 0$. The region V is drawn in following figure.



We compute the surface integral of F over $S = \partial V$. The orientation on S is associated to the normal pointing outside of V .

We split S in two: S_1 is the upper part of the sphere of radius 2 centred at the origin, and S_2 is the disk of radius 2 centred at the origin in the x_1, x_2 plane.

A parametric representation of S_1 is given by

$$\rho_1(\mathbf{w}) = (2 \cos(w_2) \sin(w_1) \quad 2 \sin(w_2) \sin(w_1) \quad 2 \cos(w_1))^T$$

for $0 < w_1 < \pi/2$ and $0 < w_2 < 2\pi$. Hence

$$\frac{\partial \rho_1}{\partial w_1}(\mathbf{w}) \times \frac{\partial \rho_1}{\partial w_2}(\mathbf{w}) = (4 \cos(w_2) \sin^2(w_1) \quad 4 \sin(w_2) \sin^2(w_1) \quad 4 \sin(w_1) \cos(w_1))^T.$$

Note that $\rho_1\left(\frac{\pi}{4}, 0\right) = (\sqrt{2} \ 0 \ \sqrt{2})^\top$ and $\left(\frac{\partial \rho_1}{\partial w_1} \times \frac{\partial \rho_1}{\partial w_2}\right)\left(\frac{\pi}{4}, 0\right) = (2 \ 0 \ 2)^\top$. This is a vector pointing outside V as required by the orientation on S_1 . Thus, our choice of orientation is consistent with the orientation on S_1 . If it has not been the case, then we could have changed the parametric representation by inverting the coordinate names w_1 and w_2 for instance to obtain a representation consistent with the orientation on S_1 . We have that $\left\|\frac{\partial \rho_1}{\partial w_1}(\mathbf{w}) \times \frac{\partial \rho_1}{\partial w_2}(\mathbf{w})\right\| = 4 \sin(w_1)$ and $dA = 4 \sin(w_1) dw_1 dw_2$.

A parametric representation of S_2 is given by

$$\rho_2(\mathbf{w}) = (w_2 \cos(w_1) \ w_2 \sin(w_1) \ 0)^\top$$

for $0 < w_1 < 2\pi$ and $0 < w_2 < 2$. Again, $\frac{\partial \rho_2}{\partial w_1}(\mathbf{w}) \times \frac{\partial \rho_2}{\partial w_2}(\mathbf{w}) = (0 \ 0 \ -w_2)^\top$ is a vector pointing outside V as required by the orientation on S_2 . The attentive reader may have realized that there is a problem at the origin with the parametric representation ρ_2 . The function ϕ_2 is not one-to-one because $\rho_2(0, w_1) = \mathbf{0}$ for $0 < w_1 < 2\pi$. Moreover, it cannot be used to define the outward unit vector since $\rho_2(\mathbf{0}) = \mathbf{0}$. Nevertheless, we may still use this parametric representation because $\int_{S_2} F \cdot \mathbf{n} dA = \int_{S_2 \setminus \{\mathbf{0}\}} F \cdot \mathbf{n} dA$ since a set with a single point represent a set of measure zero. We have that $\left\|\frac{\partial \rho_2}{\partial w_1}(\mathbf{w}) \times \frac{\partial \rho_2}{\partial w_2}(\mathbf{w})\right\| = w_2$ and $dA = w_2 dw_1 dw_2$ for $w_2 > 0$.

The integral on S is given by

$$\begin{aligned} \int_S F \cdot \mathbf{n} dA &= \sum_{i=1}^2 \int_{S_i} F \cdot \mathbf{n} dA \\ &= \int_0^{2\pi} \int_0^{\pi/2} (4 \sin(w_2) \cos(w_2) \sin^2(w_1), 4 \sin(w_2) \sin(w_1) \cos(w_1), 2 \cos(w_1)) \\ &\quad \cdot (4 \cos(w_2) \sin^2(w_1), 4 \sin(w_2) \sin^2(w_1), 4 \sin(w_1) \cos(w_1)) dw_1 dw_2 \\ &\quad + \int_0^2 \int_0^{2\pi} (w_2^2 \sin(w_1) \cos(w_1), 0, 0) \cdot (0, 0, -w_2) dw_1 dw_2 = \frac{28}{3}\pi + 0 = \frac{28}{3}\pi. \end{aligned}$$

We could have obtained the value of the integral of F over S_2 without doing any computation because at each point of S_2 the outward unit vector to S_2 , namely $\mathbf{e}_3$, is also perpendicular to the vector field F . In other words, the vector field F is tangent to S_2 and thus nothing goes in or out of V through S_2 .

We use the divergence theorem to compute the surface integral of the vector field (6.4.8) over S . The divergence theorem states that

$$\int_S F \cdot \mathbf{n} dA = \int_V \operatorname{div} F dV.$$

To compute the volume integral, we use the spherical coordinates provided by

$$\rho(\mathbf{w}) = (w_3 \cos(w_1) \sin(w_2) \ w_3 \sin(w_1) \sin(w_2) \ w_3 \cos(w_2))^\top$$

for $0 < w_3 < 2$, $0 < w_1 < 2\pi$ and $0 < w_2 < \pi/2$. The volume element on V is $dV = dx_1 \wedge dx_2 \wedge dx_3$. Its local representation using the spherical coordinates is $\rho^* dV = w_3^2 \sin(w_2) dw_1 \wedge dw_2 \wedge dw_3$. In the classical notation, this is $dV = w_3^2 \sin(w_2) dw_1 dw_2 dw_3$.

Since

$$\operatorname{div} F(\mathbf{x}) = \frac{\partial F_1}{\partial x_1}(\mathbf{x}) + \frac{\partial F_2}{\partial x_2}(\mathbf{x}) + \frac{\partial F_3}{\partial x_3}(\mathbf{x}) = x_2 + x_3 + 1 ,$$

we get

$$\begin{aligned} \int_V \operatorname{div} F \, dx_1 \, dx_2 \, dx_3 \\ = \int_0^2 \int_0^{2\pi} \int_0^{\pi/2} (w_3 \sin(w_1) \sin(w_2) + w_3 \cos(w_2) + 1) w_3^2 \sin(w_2) \, dw_2 \, dw_1 \, dw_3 = \frac{28}{3} \pi \end{aligned}$$

as expected. 

Proposition 6.4.8

Let $F : V \rightarrow \mathbb{R}^3$ be a continuously differentiable vector field on a convex open subset V of $\mathbb{R}^3$. If F satisfies $\operatorname{div} F = 0$ in V , then $F = \operatorname{curl} G$ for some continuously differentiable vector field $G : V \rightarrow \mathbb{R}^3$.

Proof.

This is in fact a special case of Theorem 4.5.12 for differential 2-forms $\omega = F_1 dx_2 \wedge dx_3 + F_2 dx_3 \wedge dx_1 + F_3 dx_1 \wedge dx_2$ because a convex subset of $\mathbb{R}^3$ is star-shaped. Since $d\omega = 0$ (i.e. $\operatorname{div} F = 0$), it follows from Theorem 4.5.12 that there exists a 1-form $\tau = G_1 dx_1 + G_2 dx_2 + G_3 dx_3$ such that $d\tau = \omega$ (i.e. $F = \operatorname{curl} G$). 

Remark 6.4.9

In the previous proposition, if $\operatorname{curl} G = F$, then $\operatorname{curl}(G + \nabla g) = F$ for all twice continuously differentiable functions $g : V \rightarrow \mathbb{R}$. We may therefore assume when we are trying to find G that the one of the components of G is null on V . For instance, we may look for $G = (G_1, G_2, 0)$ such that

$$\operatorname{curl} G = \left(-\frac{\partial G_2}{\partial z}, \frac{\partial G_1}{\partial z}, \frac{\partial G_2}{\partial x} - \frac{\partial G_1}{\partial y} \right) = (F_1, F_2, F_3) ,$$

for $G = (G_1, 0, G_3)$ such that

$$\operatorname{curl} G = \left(\frac{\partial G_3}{\partial y}, \frac{\partial G_1}{\partial z} - \frac{\partial G_3}{\partial x}, -\frac{\partial G_1}{\partial y} \right) = (F_1, F_2, F_3)$$

or for $G = (0, G_2, G_3)$ such that

$$\operatorname{curl} G = \left(\frac{\partial G_3}{\partial y} - \frac{\partial G_2}{\partial z}, -\frac{\partial G_3}{\partial x}, \frac{\partial G_2}{\partial x} \right) = (F_1, F_2, F_3) .$$

This is quite useful in applications. 

6.5 Applications

One of our goal in these lecture notes is to present the theory behind what is usually called vector calculus. However, it would not be acceptable to not include some geometrical and physical interpretation of the theory. This was the historical motivation for the development of differential geometry and it is still the main reason to use differential geometry. Our presentation will be elementary and heuristic.

In this section, the reader must be aware that we will sometime switch from the algebraic to the geometric notations for vectors and vice-versa when one notation is more convenient than the other. This is something that we have done occasionally in the past but will be done more frequently in this section. In particular, if $\mathbf{v}$ and $\mathbf{w}$ are two vectors in $\mathbb{R}^n$, then $\mathbf{v} \cdot \mathbf{w} = \mathbf{v}^\top \mathbf{w} = \mathbf{w}^\top \mathbf{v}$.

6.5.1 Tangent Vector, Curvature and Torsion

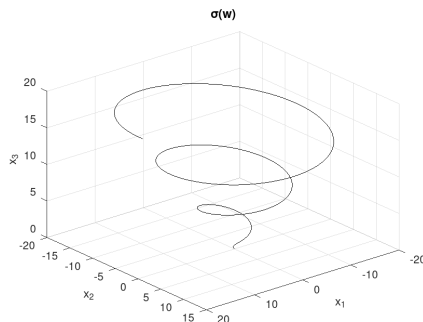
We start with the basic concept of tangent line to a curve. Let C is a curve in $\mathbb{R}^3$ with a continuously differentiable parametric representation $\sigma :]a, b[\rightarrow \mathbb{R}^3$. The basic definition of derivative as a limit of secants shows that $\sigma'(w)$ is a tangent vector to the curve C at $\sigma(w)$ for $w \in]a, b[$, and $\sigma'(w)$ points in the direction of increasing w .

If the parametric representation is consistent with the orientation on C , then $[(\sigma(w), \sigma'(w))]$ is the orientation on $T_{\sigma(w)}C$ for $w \in]a, b[$. In this section, we will always assume that the parametric representation is consistent with the orientation on C . Recall that the tangent unit vector to C at $\mathbf{u} = \sigma(w) \in C$ is given by

$$\mathbf{t}(\mathbf{u}) = \mathbf{t}(\sigma(w)) = \|\sigma'(w)\|^{-1} \sigma'(w) .$$

Example 6.5.1

Consider the curve C in $\mathbb{R}^3$ given by the parametric representation $\sigma(w) = (w \cos(w) \ w \sin(w) \ w)^\top$ for $w \in \mathbb{R}$. We have drawn a section of this curve in the following figure.

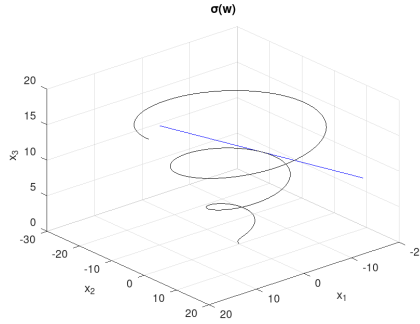


The vector $\sigma'(w)$ is tangent to the curve C at $\mathbf{u} = \sigma(w)$ and points in the direction of increasing w . From a mechanical point of view, σ' is the (directional) velocity of a particle traveling along the curve.

We have

$$\mathbf{t}(\mathbf{u}) = \mathbf{t}(\sigma(w)) = \|\sigma'(w)\|^{-1} \sigma'(w) = \frac{1}{\sqrt{2+w^2}} \begin{pmatrix} \cos(w) - w \sin(w) \\ \sin(w) + w \cos(w) \\ 1 \end{pmatrix}$$

for $\mathbf{u} = \sigma(w) \in C$. A parametric representation of the tangent line to the curve C at $\sigma(10)$ is given by $s \mapsto \sigma(10) + s \mathbf{t}(10)$ for $s \in \mathbb{R}$. We have drawn the curve C and its tangent line at $\sigma(10)$ (in blue) in the following figure.



As before, let $C \subset \mathbb{R}^3$ be a curve and $\sigma :]a, b[\rightarrow \mathbb{R}^3$ be a sufficiently differentiable parametric representation of C consistent with the orientation on C .

Lemma 6.5.2

Suppose that $g :]a, b[\rightarrow \mathbb{R}^n$ is a differentiable vector valued function such that $\|g(w)\| = c$, a constant, for all $w \in]a, b[$. Then $g'(w)$ is orthogonal to $g(w)$ for all $w \in]a, b[$.

Proof.

If we derive both sides of $g(w) \cdot g(w) = c^2$ with respect to w , then we get $2g'(w) \cdot g(w) = 0$ for all $w \in]a, b[$. Therefore $g'(w)$ is orthogonal to $g(w)$ for all $w \in]a, b[$. ■

Since $\|\mathbf{t}(\sigma(w))\| = 1$ for all $w \in]a, b[$, we may apply the previous lemma with $g(w) = \mathbf{t}(\sigma(w))$ to find that

$$\frac{d}{dw}(\mathbf{t}(\sigma(w))) = \frac{d}{dw}(\|\sigma'(w)\|^{-1} \sigma'(w)) = \|\sigma'(w)\|^{-3} (\sigma'(w) \times (\sigma'(w) \times \sigma''(w))) . \quad (6.5.1)$$

is orthogonal to $\mathbf{t}(\sigma(w))$ for all $w \in]a, b[$. It follows from the definition of $\mathbf{t}(\sigma(w))$ that $\frac{d}{dw}(\mathbf{t}(\sigma(w)))$ and $\sigma'(w)$ are orthogonal. This is also obvious from (6.5.1) and the properties of the cross product.

The **curvature vector** of C at $\mathbf{u} = \sigma(w) \in C$ is defined by

$$\mathbf{k}(\mathbf{u}) = \|\sigma'(w)\|^{-4} (\sigma'(w) \times (\sigma'(w) \times \sigma''(w))) . \quad (6.5.2)$$

For each $w \in]a, b[$, this vector points in the direction of the “centre of the arc” represented “locally” by the curve C near $\mathbf{u} = \sigma(w)$. The vector $\mathbf{k}(\mathbf{u})$ is perpendicular to the tangent vector $\mathbf{t}(\mathbf{u})$ to the curve C at $\mathbf{u} \in C$. The **curvature** $\kappa(\mathbf{u})$ of C at $\mathbf{u}$ is the Euclidean norm of $\mathbf{k}(\mathbf{u})$. For $\kappa(\mathbf{u}) \neq 0$, the ratio $1/\kappa(\mathbf{u})$ is called the **radius of curvature**. If $\kappa(\mathbf{u}) = 0$ for all $\mathbf{u} \in C$, then we have that C is a straight line.

Since $\sigma'(w)$ and $\sigma'(w) \times \sigma''(w)$ are orthogonal, we have that $\|\sigma'(w) \times (\sigma'(w) \times \sigma''(w))\| = \|\sigma'(w)\| \|\sigma'(w) \times \sigma''(w)\| \sin(\pi/2) = \|\sigma'(w)\| \|\sigma'(w) \times \sigma''(w)\|$. Hence, the curvature of C at $\mathbf{u} = \sigma(w) \in C$ is given by

$$\kappa(\mathbf{u}) = \frac{\|\sigma''(w) \times \sigma'(w)\|}{\|\sigma'(w)\|^3}. \quad (6.5.3)$$

We will generalize the definition of curvature vector and curvature in Sections 9.4 and 9.5 later.

Remark 6.5.3

Why did we introduce this extra factor $\|\sigma'(w)\|^{-1}$ in (6.5.2)? It does not even transform $\mathbf{k}(\mathbf{u})$ to a unit vector.

It is based on the fact the **tangent vector** and **curvature vector** to a curve C at a point $\mathbf{u} = \sigma(w)$ of C are defined by $\mathbf{t}(\mathbf{u}) = \sigma'(w)$ and $\mathbf{k}(\mathbf{u}) = \sigma''(w)$ if C is parameterized by arc length as we have introduced in Subsection 6.2.1.

Suppose that $\sigma : [a, b] \rightarrow \mathbb{R}^3$ is a curve of length L such that $\sigma'(w) \neq \mathbf{0}$ for all $w \in [a, b]$. Then we can use the arc length function $s : [a, b] \rightarrow [0, L]$ given in Definition 6.2.14 to obtain a parametric representation $\check{\sigma} = \sigma \circ s^{-1} : [0, L] \rightarrow \mathbb{R}^3$ such that C is parameterized by arc length with $\check{\sigma}$. Recall that $s'(t) = \|\sigma(t)\|$ for all $t \in]a, b[$. This parametric representation is often called the **natural parametric representation** of the curve C . With this representation, we have that the tangent unit vector to C at $\mathbf{u} = \check{\sigma}(t)$ is $\mathbf{t}(\mathbf{u}) = \check{\sigma}'(t)$ because $\|\check{\sigma}'(t)\| = 1$ for all $t \in]0, L[$. The curvature vector of C at $\mathbf{u} = \check{\sigma}(t) \in C$ is then simply $\mathbf{k}(\mathbf{u}) = \check{\sigma}''(t)$.

Using the chain rule, we find for $\mathbf{u} = \sigma(w) = \check{\sigma}(s(w))$ that

$$\begin{aligned} \mathbf{t}(\mathbf{u}) &= \check{\sigma}'(t)|_{t=s(w)} = \frac{d}{dw}(\sigma(w)) \Big|_{w=s^{-1}(s(w))} \frac{d}{dt}(s^{-1}(t)) \Big|_{t=s(w)} \\ &= (s'(w))^{-1} \sigma'(w) = \|\sigma'(w)\|^{-1} \sigma'(w) \end{aligned}$$

and

$$\begin{aligned} \mathbf{k}(\mathbf{u}) &= \check{\sigma}''(t)|_{t=s(w)} = \frac{d}{dw}(\check{\sigma}'(t)) \Big|_{w=s^{-1}(s(w))} \frac{d}{dt}(s^{-1}(t)) \Big|_{t=s(w)} \\ &= \|\sigma'(w)\|^{-1} \frac{d}{dw}(\|\sigma'(w)\|^{-1} \sigma'(w)) \end{aligned}$$

as we have defined above. This justify the extra factor $\|\sigma'(w)\|^{-1}$ in (6.5.2).

Though the formulae for the tangent unit vector and the curvature vector are much simpler when the natural parametric representation of the curve C is used, they are not very useful in practice because the natural parametric representation has to be found first. The natural parametric representation may often not be the simplest parametric representation for computations. ♠

Example 6.5.4 (Example 6.5.1 (Continued))

We compute the curvature of the curve C defined by the parametric representation σ given in Example 6.5.1.

Since the second derivative of σ is

$$\sigma''(w) = \begin{pmatrix} -2\sin(w) - w\cos(w) & 2\cos(w) - w\sin(w) & 0 \end{pmatrix}^\top$$

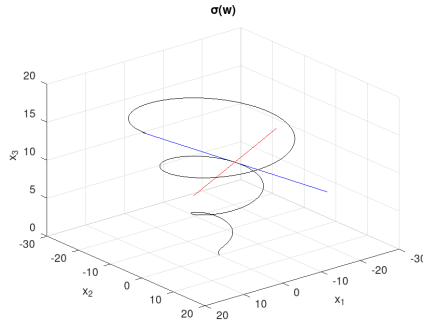
for all $w \in \mathbb{R}$. We get from (6.5.2) that the curvature vector of C at $\mathbf{u} = \sigma(w) \in C$ is

$$\mathbf{k}(u) = \frac{1}{(2+w^2)^2} \begin{pmatrix} -3\cos(w) - w^2\sin(w) - 3w\cos(w) - 4\sin(w) \\ -w^3\sin(w) + w^2\cos(w) - 3w\sin(w) + 4\cos(w) \\ -w \end{pmatrix}$$

and we get from (6.5.3) that the curvature at $\sigma(10)$ is $\kappa(10) \approx 0.099508407$.

From a mechanical point of view, $\sigma''(\mathbf{u})$ is the directional acceleration of a particle travelling along the curve.

A parametric representation of the line parallel to the curvature vector to C at $\sigma(10)$ is given by $s \mapsto \sigma(10) + s\mathbf{k}(10)$ for $s \in \mathbb{R}$. We have drawn the curve C , its tangent line (in blue) and a line parallel to its curvature vector (in red) at $\sigma(10)$ in the following figure.



Let $\mathbf{N}(\mathbf{u}) = \|\mathbf{k}(\mathbf{u})\|^{-1}\mathbf{k}(\mathbf{u}) = (\kappa(\mathbf{u}))^{-1}\mathbf{k}(\mathbf{u})$ and $\mathbf{B}(\mathbf{u}) = \mathbf{t}(\mathbf{u}) \times \mathbf{N}(\mathbf{u})$ for $\mathbf{u} \in C$. The vector $\mathbf{N}(\mathbf{u})$ is called the **principal normal** and $\mathbf{B}(\mathbf{u})$ is called the **binormal vector** to the curve C at $\mathbf{u}$. Since $\mathbf{B}(\mathbf{u})$ is a unit vector orthogonal to the plane generated by $\mathbf{t}(\mathbf{u})$ and $\mathbf{N}(\mathbf{u})$, we have that $\mathbf{t}(\mathbf{u})$, $\mathbf{N}(\mathbf{u})$ and $\mathbf{B}(\mathbf{u})$ are three unit vectors that form a right handed system of mutually orthogonal vectors moving along the curve C .

Since $\mathbf{B}(\sigma(w))$ is of norm one for all $w \in]a, b[$, we have from Lemma 6.5.2 that $\frac{d}{dw}\mathbf{B}(\sigma(w))$ is orthogonal to $\mathbf{B}(\sigma(w))$ for all $w \in]a, b[$. We also have that $\frac{d}{dw}\mathbf{B}(\sigma(w))$ is orthogonal to $\mathbf{t}(\sigma(w))$ for all $w \in]a, b[$ because

$$\frac{d}{dw}\mathbf{B}(\sigma(w)) = \left(\frac{d}{dw}\mathbf{t}(\sigma(w)) \right) \times \mathbf{N}(\sigma(w)) + \mathbf{t}(\sigma(w)) \times \frac{d}{dw}\mathbf{N}(\sigma(w)) = \mathbf{t}(\sigma(w)) \times \frac{d}{dw}\mathbf{N}(\sigma(w))$$

for all $w \in]a, b[$ since $\left(\frac{d}{dw}\mathbf{t}(\sigma(w))\right)$ and $\mathbf{N}(\sigma(w))$ are both parallel to $\mathbf{k}(\sigma(w))$. Hence $\frac{d}{dw}\mathbf{B}(\sigma(w)) \cdot \mathbf{t}(\sigma(w)) = 0$ for all $w \in]a, b[$ because $\mathbf{t}(\sigma(w)) \times \frac{d}{dw}\mathbf{N}(\sigma(w))$ is orthogonal to $\mathbf{t}(\sigma(w))$. Since $\frac{d}{dw}\mathbf{B}(\sigma(w))$ is orthogonal to both $\mathbf{t}(\sigma(w))$ and $\mathbf{B}(\sigma(w))$ for all $w \in]a, b[$, we get

$$\frac{d}{dw}\mathbf{B}(\sigma(w)) = -\tau(\sigma(w)) \|\sigma'(w)\| \mathbf{N}(\sigma(w))$$

for some number $\tau(\sigma(w))$. The **torsion** of the curve C at $\mathbf{u} = \sigma(w) \in C$ is $\tau(\mathbf{u})$.

If the torsion is zero for all $\mathbf{u} \in C$, then the curve C is in a fixed plane. With a little bit of patience, the reader may verify that

$$\tau(\sigma(w)) = \frac{(\sigma'(w) \times \sigma''(w)) \cdot \sigma'''(w)}{\|\sigma'(w) \times \sigma''(w)\|^2} \quad (6.5.4)$$

for all $w \in [a, b]$. The triple product $(\sigma'(w) \times \sigma''(w)) \cdot \sigma'''(w)$ is the determinant of the matrix $(\sigma'(w) \ \sigma''(w) \ \sigma'''(w))$.

Example 6.5.5 (Example 6.5.4 (Continued))

We compute the torsion of the curve C of Example 6.5.1 at $\sigma(10)$.

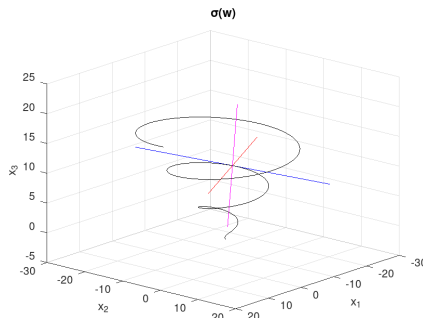
We have $\sigma'''(w) = (-3\cos(w) + w\sin(w) \ -3\sin(w) - w\cos(w) \ 0)^\top$ for all $w \in \mathbb{R}$.

Hence, we get from (6.5.4) that the torsion is $\tau(10) \approx 0.010087552$. We also have

$$\mathbf{B}(\sigma(w)) = \frac{1}{\kappa(\sigma(w))(2+w^2)^{5/2}} \begin{pmatrix} w^3 \sin(w) - 2w^2 \cos(w) + 2w \sin(w) - 4 \cos(w) \\ -w^3 \cos(w) - 2w^2 \sin(w) - 2w \cos(w) - 4 \sin(w) \\ w^4 + 4w^2 + 4 \end{pmatrix}$$

for $w \in \mathbb{R}$.

A parametric representation of the line parallel to the direction of the binormal vector $\mathbf{B}(\sigma(10))$ of C at $\sigma(10)$ is given by $s \mapsto \sigma(10) + s\mathbf{B}(10)$ for $s \in \mathbb{R}$. We have drawn the curve C , its tangent line (in blue), the line parallel to its principal normal (in red), and the line parallel to its binormal vector (in magenta) at $\sigma(10)$ in the following figure.



6.5.2 Physical Interpretations

As it is typical in physical applications of mathematics, we will allow ourselves a more liberal use of some mathematical concepts in this section. In particular, limits will be interpreted intuitively and with a more visual interpretation.

Flux of a Vector Field

The surface integral of a vector field $F = F_1\mathbf{e}_1 + F_2\mathbf{e}_2 + F_3\mathbf{e}_3$ over a oriented surface S gives the flux of the vector field F through the surface S in the direction of the unit normal to S associated to the orientation on S . To motivate this statement, we need to go back to the definition of the Riemann integral.

Let $I_2 = [0, 1] \times [0, 1]$. Suppose that $\rho : I_2 \rightarrow \mathbb{R}^3$ is a parametric representation of class C^1 consistent with the orientation on the surface S . This parametric representation is such that $\frac{\partial \rho}{\partial w_1} \times \frac{\partial \rho}{\partial w_2}$ points in the direction of the outward unit normal $\mathbf{n}$ to S associated with the orientation on S .

Suppose that $P = \{P_1, P_2\}$ is a partition of I_2 given by the partitions $P_1 = \{w_{1,i} : 0 \leq i \leq N_1\}$ and $P_2 = \{w_{2,j} : 0 \leq j \leq N_2\}$ of the w_1 and w_2 axes respectively. Moreover, suppose that the points in the two partitions P_1 and P_2 are equally spaced; namely, there are two constants Δw_1 and Δw_2 such that $\Delta w_1 = w_{1,i+1} - w_{1,i}$ for $0 < i \leq N_1$ and $\Delta w_2 = w_{2,j+1} - w_{2,j}$ for $0 < j \leq N_2$. We have that

$$\begin{aligned} & F(\rho(w_{1,i}, w_{2,j})) \cdot \left(\frac{\partial \rho}{\partial w_1}(w_{1,i}, w_{2,j}) \times \frac{\partial \rho}{\partial w_2}(w_{1,i}, w_{2,j}) \right) \Delta w_{1,i} \Delta w_{2,j} \\ &= \det \begin{pmatrix} F_1(\rho(w_{1,i}, w_{2,j})) & F_2(\rho(w_{1,i}, w_{2,j})) & F_3(\rho(w_{1,i}, w_{2,j})) \\ \frac{\partial \rho_1}{\partial w_1}(w_{1,i}, w_{2,j}) \Delta w_1 & \frac{\partial \rho_2}{\partial w_1}(w_{1,i}, w_{2,j}) \Delta w_1 & \frac{\partial \rho_3}{\partial w_1}(w_{1,i}, w_{2,j}) \Delta w_1 \\ \frac{\partial \rho_1}{\partial w_2}(w_{1,i}, w_{2,j}) \Delta w_2 & \frac{\partial \rho_2}{\partial w_2}(w_{1,i}, w_{2,j}) \Delta w_2 & \frac{\partial \rho_3}{\partial w_2}(w_{1,i}, w_{2,j}) \Delta w_2 \end{pmatrix} \end{aligned}$$

is the volume of the box defined by the vectors $F(\rho(w_{1,i}, w_{2,j}))$, $\frac{\partial \rho}{\partial w_1}(w_{1,i}, w_{2,j}) \Delta w_1$ and $\frac{\partial \rho}{\partial w_2}(w_{1,i}, w_{2,j}) \Delta w_2$ that we have drawn in Figure 6.5. This is an approximation of the volume of fluid that pass through the little region $\rho([w_{1,i}, w_{1,i+1}] \times [w_{2,j}, w_{2,j+1}])$ of S in the direction of the outward unit normal to S .

Since F is continuous and ρ are continuously differentiable, we have that $F \cdot \mathbf{n}$ is Riemann integrable. Hence, the total volume of fluid that goes through S in the direction of the outward unit normal to the surface S is given by the limit

$$\lim_{\max\{\Delta w_1, \Delta w_2\} \rightarrow 0} \sum_{\substack{1 \leq i \leq N_1 \\ 1 \leq j \leq N_2}} F(\rho(w_{1,i}, w_{2,j})) \cdot \left(\frac{\partial \rho}{\partial w_1}(w_{1,i}, w_{2,j}) \times \frac{\partial \rho}{\partial w_2}(w_{1,i}, w_{2,j}) \right) \Delta w_{1,i} \Delta w_{2,j}$$

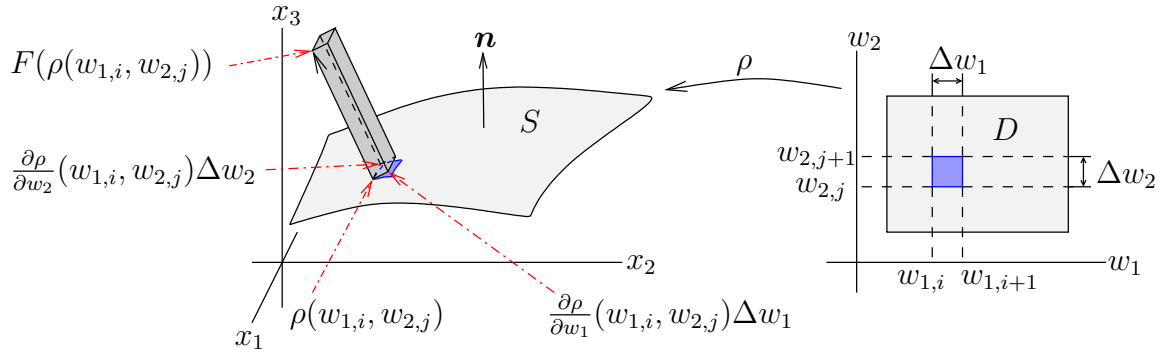


Figure 6.5: The representation of one of the terms of the Riemann sum for the integral of the vector field F on a surface. The volume of the box is an approximation of the flux through the little patch in the direction of the outward unit normal to the surface.

$$= \iint_{I_2} F(\rho(w_1, w_2)) \cdot \left(\frac{\partial \rho}{\partial w_1}(w_1, w_2) \times \frac{\partial \rho}{\partial w_2}(w_1, w_2) \right) dw_1 dw_2 = \iint_S F \cdot \mathbf{n} dA.$$

A Physical Interpretation of the Divergence

A vector field $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ can be associated to a system of differential equations $\sigma'(t) = F(\sigma(t))$. Students learn in their first course on ordinary differential equations that, given $\mathbf{y} \in \mathbb{R}^3$ and a vector field F of class C^1 on $\mathbb{R}^3$, there exist an open neighbourhood $I_{\mathbf{y}} \subset \mathbb{R}$ of 0 and a unique function $\sigma : I_{\mathbf{y}} \rightarrow \mathbb{R}^3$ such that $\sigma'(t) = F(\sigma(t))$ for $t \in I_{\mathbf{y}}$ and $\sigma(0) = \mathbf{y}$. The function $\sigma : I_{\mathbf{y}} \rightarrow \mathbb{R}^3$ is a parametric representation of a curve in $\mathbb{R}^3$ through $\mathbf{y}$ which is tangent to the vector field F at all points of the curve because $\sigma'(t) = F(\sigma(t))$ for all $t \in I_{\mathbf{y}}$.

Because F is of class C^1 , there exist an open neighbourhood $I \subset \mathbb{R}$ of 0 and an open subset U of $\mathbb{R}^3$ such that we may define a continuously differentiable map $\phi : I \times U \rightarrow \mathbb{R}^3$ by $\phi(t, \mathbf{x}) = \sigma(t)$ for $t \in I$ and $\mathbf{x} \in U$ where σ is the unique function such that $\sigma'(t) = F(\sigma(t))$ for $t \in I$ and $\sigma(0) = \mathbf{x} \in U$. The function ϕ defines what we call a local **flow**. For each $\mathbf{x} \in U$, we may view $\{\phi(t, \mathbf{x}) : t \in I\}$ as the trajectory of a particle starting at $\mathbf{x}$ and moving under the influence of the vector field F .

Let $P_0 = \{\mathbf{x} + s_1 \mathbf{e}_1 + s_2 \mathbf{e}_2 + s_3 \mathbf{e}_3 : 0 \leq s_i \leq \epsilon\}$ for really small ϵ . The set P_0 represents a small box that we have drawn in Figure 6.6. Using the Taylor's expansion theorem for functions from $\mathbb{R}^3$ to $\mathbb{R}^3$, we find that the image of a point $\mathbf{x} + s_1 \mathbf{e}_1 + s_2 \mathbf{e}_2 + s_3 \mathbf{e}_3$ of the box P_0 along the flow is given by

$$\begin{aligned} \phi(t, \mathbf{x} + s_1 \mathbf{e}_1 + s_2 \mathbf{e}_2 + s_3 \mathbf{e}_3) &= \phi(t, \mathbf{x}) + D_{\mathbf{x}} \phi(t, \mathbf{x}) (s_1 \mathbf{e}_1 + s_2 \mathbf{e}_2 + s_3 \mathbf{e}_3) \\ &\quad + O(s_1^2 + s_2^2 + s_3^2). \end{aligned}$$

If ϵ is really small, then the image of the box P_0 along the flow is

$$\phi(t, P_0) \approx P_t = \{\phi(t, \mathbf{x}) + D_{\mathbf{x}} \phi(t, \mathbf{x}) (s_1 \mathbf{e}_1 + s_2 \mathbf{e}_2 + s_3 \mathbf{e}_3) : 0 \leq s_i \leq \epsilon\}.$$

for t very small. The image of the box P_0 by the linear map $D_{\mathbf{x}} \phi(t, \mathbf{x})$ is approximately the box P_t with the sides given by the vectors

$$\mathbf{v}_i(t, \mathbf{x}) = \epsilon D_{\mathbf{x}} \phi(t, \mathbf{x}) \mathbf{e}_i \quad (6.5.5)$$

for $1 \leq i \leq 3$. Let $V(\mathbf{x}, t)$ be the volume of the box P_t that we have drawn in Figure 6.6. We now show that

$$\frac{1}{V(0, \mathbf{x})} \frac{dV}{dt}(0, \mathbf{x}) = \operatorname{div} F(\mathbf{x}) . \quad (6.5.6)$$

This relation can be interpreted as saying that the relative instantaneous rate of change of the volume at the point $\mathbf{x}$ is given by the divergence of the vector field F at the point $\mathbf{x}$.

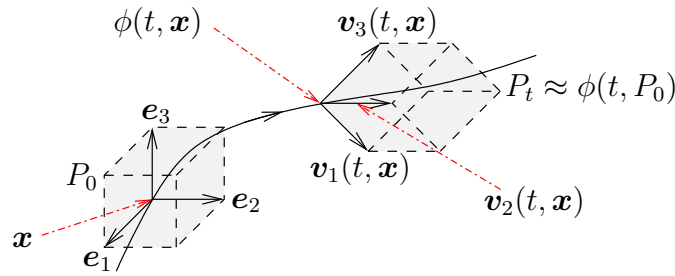


Figure 6.6: Physical interpretation of the divergence.

Since $\frac{\partial \phi}{\partial t}(t, \mathbf{x}) = F(\phi(t, \mathbf{x}))$, we have

$$\frac{\partial}{\partial t} D_{\mathbf{x}} \phi(t, \mathbf{x}) = D_{\mathbf{x}} \left(\frac{\partial \phi}{\partial t}(t, \mathbf{x}) \right) = D_{\mathbf{x}} (F(\phi(t, \mathbf{x}))) = D F(\phi(t, \mathbf{x})) D_{\mathbf{x}} \phi(t, \mathbf{x}) .$$

Thus

$$\frac{d\mathbf{v}_i}{dt}(t, \mathbf{x}) = \epsilon \frac{\partial}{\partial t} D_{\mathbf{x}} \phi(t, \mathbf{x}) \mathbf{e}_i = \epsilon D F(\phi(t, \mathbf{x})) D_{\mathbf{x}} \phi(t, \mathbf{x}) \mathbf{e}_i \quad (6.5.7)$$

for $1 \leq i \leq 3$. Since $\phi(0, \mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in U$, we get that $D_{\mathbf{x}} \phi(0, \mathbf{x}) = \operatorname{Id}$. Thus (6.5.5) yields $\mathbf{v}_i(0, \mathbf{x}) = \epsilon \mathbf{e}_i$ for $1 \leq i \leq 3$, as it should be for P_0 , and (6.5.7) becomes

$$\frac{d\mathbf{v}_i}{dt}(0, \mathbf{x}) = \epsilon D F(\mathbf{x}) \mathbf{e}_i \quad (6.5.8)$$

for $1 \leq i \leq 3$. The volume of P_t is given by $V(t, \mathbf{x}) = \mathbf{v}_1(t, \mathbf{x}) \cdot (\mathbf{v}_2(t, \mathbf{x}) \times \mathbf{v}_3(t, \mathbf{x}))$. Thus

$$\begin{aligned} \frac{dV}{dt}(t, \mathbf{x}) &= \frac{d\mathbf{v}_1}{dt}(t, \mathbf{x}) \cdot (\mathbf{v}_2(t, \mathbf{x}) \times \mathbf{v}_3(t, \mathbf{x})) \\ &\quad + \mathbf{v}_1(t, \mathbf{x}) \cdot \left(\frac{d\mathbf{v}_2}{dt}(t, \mathbf{x}) \times \mathbf{v}_3(t, \mathbf{x}) \right) + \mathbf{v}_1(t, \mathbf{x}) \cdot \left(\mathbf{v}_2(t, \mathbf{x}) \times \frac{d\mathbf{v}_3}{dt}(t, \mathbf{x}) \right) . \end{aligned}$$

If we evaluate this expression at $t = 0$ and use (6.5.8), then we get

$$\frac{dV}{dt}(0, \mathbf{x}) = (\epsilon D F(\mathbf{x}) \mathbf{e}_1) \cdot (\epsilon \mathbf{e}_2 \times \epsilon \mathbf{e}_3) + \epsilon \mathbf{e}_1 \cdot (\epsilon (D F(\mathbf{x}) \mathbf{e}_2) \times \epsilon \mathbf{e}_3)$$

$$\begin{aligned}
& + \epsilon \mathbf{e}_1 \cdot (\epsilon \mathbf{e}_2 \times (\epsilon D F(\mathbf{x}) \mathbf{e}_3)) \\
& = \epsilon^3 \left((D F(\mathbf{x}) \mathbf{e}_1) \cdot (\mathbf{e}_2 \times \mathbf{e}_3) + (D F(\mathbf{x}) \mathbf{e}_2) \cdot (\mathbf{e}_3 \times \mathbf{e}_1) \right. \\
& \quad \left. + (D F(\mathbf{x}) \mathbf{e}_3) \cdot (\mathbf{e}_1 \times \mathbf{e}_2) \right) \\
& = \epsilon^3 \left((D F(\mathbf{x}) \mathbf{e}_1) \cdot \mathbf{e}_1 + (D F(\mathbf{x}) \mathbf{e}_2) \cdot \mathbf{e}_2 + (D F(\mathbf{x}) \mathbf{e}_3) \cdot \mathbf{e}_3 \right) \\
& = \epsilon^3 \left(\frac{\partial F_1}{\partial x_1}(\mathbf{y}) + \frac{\partial F_2}{\partial x_2}(\mathbf{y}) + \frac{\partial F_3}{\partial x_3}(\mathbf{y}) \right) = \epsilon^3 \operatorname{div} F(\mathbf{y}) .
\end{aligned}$$

Since the volume $V(0, \mathbf{y})$ of P_0 is obviously ϵ^3 , we get (6.5.6).

If $\operatorname{div} F = 0$, then it follows from (6.5.6) that the volume of a region does not change as it moves along the flow.

A Physical Interpretation of the Curl

As in the previous section, we consider the flow ϕ associated to the vector field F . We consider a particle moving along the flow with $\mathbf{x}$ as the initial position of a particle. Let $\mathbf{v}_0$ and $\mathbf{w}_0$ be two vectors as drawn in Figure 6.7.

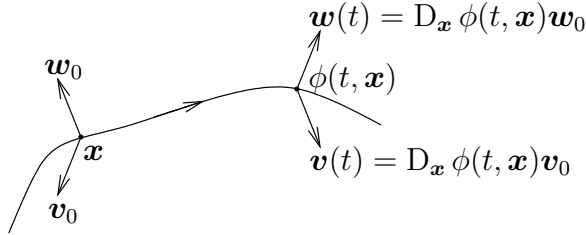


Figure 6.7: Physical interpretation of the curl for an object moving along a flow. The vectors $\mathbf{v}_0$ and $\mathbf{w}_0$ move along the flow.

We consider $\mathbf{w}(t) = D_x \phi(t, \mathbf{x}) \mathbf{w}_0$ and $\mathbf{v}(t) = D_x \phi(t, \mathbf{x}) \mathbf{v}_0$. Proceeding as we did to obtain (6.5.8), we get $\frac{d\mathbf{w}}{dt}(0) = D F(\mathbf{x}) \mathbf{w}_0$ and $\frac{d\mathbf{v}}{dt}(0) = D F(\mathbf{x}) \mathbf{v}_0$. The cosine of the angle between the vectors $\mathbf{w}(t)$ and $\mathbf{v}(t)$ is determined by the scalar product $\mathbf{v}(t) \cdot \mathbf{w}(t)$. Hence, if we want to measure any variation of this angle, it suffices to measure any variation of the scalar product $\mathbf{v}(t) \cdot \mathbf{w}(t)$ as t varies. The word variation suggests to compute the following derivative.

$$\begin{aligned}
\left. \frac{d}{dt} (\mathbf{v}(t) \cdot \mathbf{w}(t)) \right|_{t=0} &= \left(\frac{d\mathbf{v}}{dt}(t) \cdot \mathbf{w}(t) + \mathbf{v}(t) \cdot \frac{d\mathbf{w}}{dt}(t) \right) \Big|_{t=0} \\
&= (D F(\mathbf{x}) \mathbf{v}_0) \cdot \mathbf{w}_0 + \mathbf{v}_0 \cdot (D F(\mathbf{x}) \mathbf{w}_0) = \mathbf{w}_0^\top (D F(\mathbf{x}) + D F(\mathbf{x})^\top) \mathbf{v}_0 .
\end{aligned} \tag{6.5.9}$$

We may write $D F(\mathbf{x}) = S(\mathbf{x}) + W(\mathbf{x})$, where $S(\mathbf{x}) = (D F(\mathbf{x}) + D F(\mathbf{x})^\top) / 2$ and $W(\mathbf{x}) = (D F(\mathbf{x}) - D F(\mathbf{x})^\top) / 2$. The symmetric matrix S is called the **deformation matrix** while the antisymmetric matrix W is called the **matrix of rotation**. We conclude from (6.5.9)

that the matrix of rotation is responsible for the rotation of an object moving along the flow while the deformation matrix is responsible for the variation of the angle between particles that form the object as the object is moving along with the flow; it is responsible for the deformation of the object.

There is a connection between the matrix of rotation W and the curl of F . We first note that

$$\begin{aligned} W(\mathbf{x}) &\equiv \frac{1}{2} (\mathbf{D} F(\mathbf{x}) - \mathbf{D} F(\mathbf{x})^\top) \\ &= \frac{1}{2} \begin{pmatrix} 0 & \frac{\partial F_1}{\partial x_2}(\mathbf{x}) - \frac{\partial F_2}{\partial x_1}(\mathbf{x}) & \frac{\partial F_1}{\partial x_3}(\mathbf{x}) - \frac{\partial F_3}{\partial x_1}(\mathbf{x}) \\ \frac{\partial F_2}{\partial x_1}(\mathbf{x}) - \frac{\partial F_1}{\partial x_2}(\mathbf{x}) & 0 & \frac{\partial F_2}{\partial x_3}(\mathbf{x}) - \frac{\partial F_3}{\partial x_2}(\mathbf{x}) \\ \frac{\partial F_3}{\partial x_1}(\mathbf{x}) - \frac{\partial F_1}{\partial x_3}(\mathbf{x}) & \frac{\partial F_3}{\partial x_2}(\mathbf{x}) - \frac{\partial F_2}{\partial x_3}(\mathbf{x}) & 0 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 0 & -(\operatorname{curl} F(\mathbf{x}))_3 & (\operatorname{curl} F(\mathbf{x}))_2 \\ (\operatorname{curl} F(\mathbf{x}))_3 & 0 & -(\operatorname{curl} F(\mathbf{x}))_1 \\ -(\operatorname{curl} F(\mathbf{x}))_2 & (\operatorname{curl} F(\mathbf{x}))_1 & 0 \end{pmatrix}, \end{aligned}$$

where $(\operatorname{curl} F(\mathbf{x}))_i$ is the i^{th} coordinates of $\operatorname{curl} F(\mathbf{x})$.

When $\operatorname{curl} F = \mathbf{0}$, the flow is **irrotational**. Namely, an object will not rotate like a spinning top as it moves along with the flow because $W = 0$ but it will only be deformed because $S \neq 0$. We will give another justification of this result in the next section.

To better understand link between rotation and curl in the previous statement, we consider a solid object S that spin around an axis which runs through the object as illustrated in Figure 6.8. We may assume without lost of generality that the x_3 axis is the axis of rotation. Suppose that ω is the angular velocity and that $\mathbf{w}$ is a vector of length ω that points in the direction of the positive values of x_3 . Given $\mathbf{x} \in S$, we have that $\mathbf{x} = P(\mathbf{x}) + \mathbf{r}$ where P is the orthogonal projection on the x_3 axis and $\mathbf{r}$ is a vector perpendicular to the axis of rotation and from the axis of rotation to $\mathbf{x}$.

The vector $\mathbf{v} = \mathbf{w} \times \mathbf{r}$ is perpendicular to the plan containing the vectors $\mathbf{r}$ and $\mathbf{w}$, and points in the direction of the rotation (counterclockwise in Figure 6.8). Since $\mathbf{w} = (0, 0, \omega)$, we have

$$\mathbf{v} = \mathbf{w} \times \mathbf{r} = \mathbf{w} \times (\mathbf{x} - P(\mathbf{x})) = \mathbf{w} \times \mathbf{x} = (-\omega x_2, \omega x_1, 0)$$

because $P(\mathbf{x})$ is parallel to $\mathbf{w}$. Hence $\operatorname{curl} \mathbf{v} = 2\omega \mathbf{e}_3$.

In general, an object that spins like a spinning top with an angular velocity ω will satisfy $\operatorname{curl} \mathbf{v} = 2\omega \mathbf{u}$ where $\mathbf{u}$ is an unit vector in the direction of the axis of rotation.

Therefore, if $\operatorname{curl} \mathbf{v} = \mathbf{0}$, then $\omega = 0$. The object does not spin. If we think about an object moving along with the flow of a vector field F , then $\mathbf{v} = F(\mathbf{x})$. Therefore, if $\operatorname{curl} F = 0$, then the object does not spin like a spinning top.

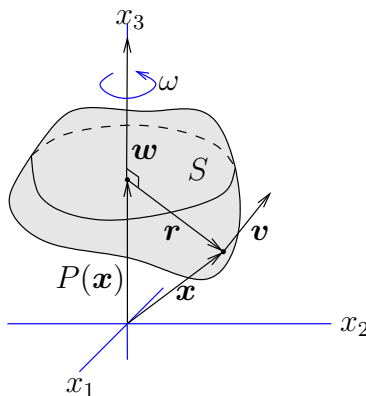


Figure 6.8: Physical interpretation of the curl for an solid object under the influence of a flow.

A Physical Interpretation of the Curl (Continued)

We consider a sufficiently smooth vector field $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ and a point $\mathbf{y} \in \mathbb{R}^3$. Let $\mathbf{n}$ be a unit vector not necessarily tangent to the flow near $\mathbf{y}$; namely, not necessarily parallel to $F(\mathbf{x})$ for some $\mathbf{x}$ near $\mathbf{y}$. Let S_r be a disk of radius r centred at $\mathbf{y}$ and perpendicular to the vector $\mathbf{n}$. We assume that $\mathbf{n}$ is the outward unit normal associated to the orientation on S_r . This situation is represented in Figure 6.9.

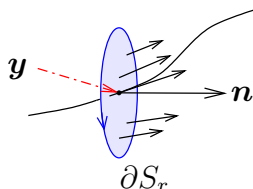


Figure 6.9: Physical interpretation of the curl using Stokes' theorem. We have drawn some of the vectors of the vector field F on the surface S_r . The vector $\mathbf{n}$ is not parallel to $F(\mathbf{x})$ for most $\mathbf{x} \in S_r$. However, $\mathbf{n}$ is perpendicular to the surface S_r .

From Stokes' theorem, we have

$$\int_{\partial S_r} F \cdot \mathbf{t} \, ds = \iint_{S_r} \text{curl } F \cdot \mathbf{n} \, dS ,$$

where ∂S_r is the boundary of the surface S_r and $\mathbf{t}(\mathbf{u})$ is the unit tangent vector to ∂S_r at $\mathbf{u} \in \partial S_r$ which is consistent with the orientation on ∂S_r that is induced from the orientation

on S_r . According to the Mean Value Theorem for Integrals (Question 3.6), we have

$$\frac{1}{\pi r^2} \iint_{S_r} \operatorname{curl} F \cdot \mathbf{n} \, dS = \operatorname{curl} F(\mathbf{x}_r) \cdot \mathbf{n}$$

for some $\mathbf{u}_r \in S_r$. Thus

$$\operatorname{curl} F(\mathbf{u}_r) \cdot \mathbf{n} = \frac{1}{\pi r^2} \int_{C_r} F \cdot \mathbf{t} \, ds.$$

Since $\lim_{r \rightarrow 0} \mathbf{u}_r = \mathbf{y}$, we get

$$\operatorname{curl} F(\mathbf{y}) \cdot \mathbf{n} = \lim_{r \rightarrow 0} \operatorname{curl} F(\mathbf{u}_r) \cdot \mathbf{n} = \lim_{r \rightarrow 0} \frac{1}{\pi r^2} \int_{\partial S_r} F \cdot \mathbf{t} \, ds.$$

Since $F(\mathbf{u}) \cdot \mathbf{t}(\mathbf{u})$ is the length of the component of the vector field F tangent to ∂S_r at $\mathbf{u} \in \partial S_r$, the integral $\int_{\partial S_r} F \cdot \mathbf{t} \, ds$ expresses the displacement along ∂S_r . Thus, for very small r , $\frac{1}{\pi r^2} \int_{\partial S_r} F \cdot \mathbf{t} \, ds$ can be used as a measure of the curl around the point $\mathbf{y}$ per unit of area in the plane perpendicular to $\mathbf{n}$. Hence, as $r \rightarrow 0$, we get that $\operatorname{curl} F(\mathbf{y}) \cdot \mathbf{n}$ measure the curl around the axis $\mathbf{n}$ at the point $\mathbf{y}$.

We have that $\operatorname{curl} F(\mathbf{y}) \cdot \mathbf{n}$ reaches its maximal value with $\mathbf{n} = \|\operatorname{curl} F(\mathbf{y})\|^{-1} \operatorname{curl} F(\mathbf{y})$. In this case, $\operatorname{curl} F(\mathbf{y}) \cdot \mathbf{n} = \|\operatorname{curl} F(\mathbf{y})\|$ and we may conclude that the angular velocity of the rotation around the axis given by $\operatorname{curl} F(\mathbf{y})$ at $\mathbf{y}$ is $\|\operatorname{curl} F(\mathbf{y})\|$.

6.6 Exercises

As we did for the chapter on integration of functions of several variables, we are including questions that are traditional questions that can be found in advanced calculus textbooks. We include them for the benefit of the readers who may want to strengthen their knowledge of classical vector calculus. They are problems that any mathematics students should be able to solve. This section could be used as a source of problems by instructors of advanced calculus.

The reader should keep in mind the comments in Remark 6.4.3. To avoid cumbersome details when using Green's, Stokes' and divergence theorems, we will liberally use them without rigorously verifying and explaining why they can be used. In particular, we will use them without mentioning that, in a lot of the cases, we are using Theorem 5.6.23 or Theorem 5.6.31.

6.6.1 Line and Surface Integrals

Question 6.1

Compute the arc length of the following curves.

- a) The curve C with the parametric representation $g : [0, 2\pi] \rightarrow \mathbb{R}^3$ given by $g(t) = (a \cos(t), bt, a \sin(t))$.
- b) The curve C with the parametric representation $h : [0, 2] \rightarrow \mathbb{R}^2$ given by $h(t) = (t^3/3 - t, t^2)$.
- c) The curve C with the parametric representation $f : [1, e] \rightarrow \mathbb{R}^3$ given by $f(t) = (\ln(t), 2t, t^2)$.

Solution

- a) The arc length is

$$\int_C \|g'(t)\| dt = \int_0^{2\pi} \sqrt{(-a \sin(t))^2 + b^2 + (a \cos(t))^2} dt = \int_0^{2\pi} \sqrt{a^2 + b^2} dt = 2\pi \sqrt{a^2 + b^2}.$$

- b) The arc length is

$$\int_C \|h'(t)\| dt = \int_0^2 \sqrt{(t^2 - 1)^2 + (2t)^2} dt = \int_0^2 (t^2 + 1) dt = \left(\frac{t^3}{3} + t \right) \Big|_{t=0}^2 = \frac{14}{3}.$$

- c) The arc length is

$$\int_C \|f'(t)\| dt = \int_1^e \sqrt{(1/t)^2 + 2^2 + (2t)^2} dt = \int_1^e \frac{1 + 2t^2}{t} dt = (\ln(t) + t^2) \Big|_{t=1}^e = e^2.$$

♡

Question 6.2

Integrals of the form $E(k) = \int_0^{\pi/2} \sqrt{1 - k^2 \sin^2(t)} dt$ with $0 < k < 1$ are called elliptic integrals.

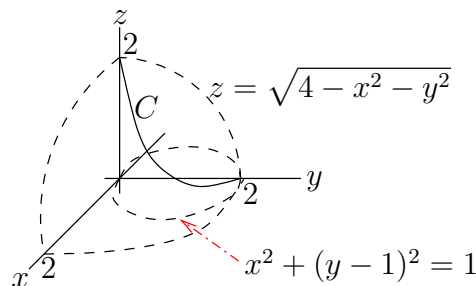
- a) Write the arc length of the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ in terms of an elliptic integral.
- b) Write the arc length of the curve C given by intersection of the sphere $x^2 + y^2 + z^2 = 4$ and the cylinder $x^2 + y^2 - 2xy = 0$ for $z \geq 0$.

Solution

- a) Let C be the curve represented by the ellipse. We assume that $a > b$. We get a similar result if $b > a$. A parametric representation for the ellipse is given by $g : [0, 2\pi] \rightarrow \mathbb{R}^2$ defined by $g(t) = (a \sin(t), b \cos(t))$. The arc length of C is

$$\begin{aligned} \int_0^{2\pi} \|g'(t)\| dt &= \int_0^{2\pi} \sqrt{(a \cos(t))^2 + (-b \sin(t))^2} dt = \int_0^{2\pi} \sqrt{a^2 \cos^2(t) + b^2 \sin^2(t)} dt \\ &= \int_0^{2\pi} \sqrt{a^2 + (b^2 - a^2) \sin^2(t)} dt = a \int_0^{2\pi} \sqrt{1 - \left(\frac{a^2 - b^2}{a^2} \right) \sin^2(t)} dt \\ &= 4a \int_0^{\pi/2} \sqrt{1 - \left(\frac{a^2 - b^2}{a^2} \right) \sin^2(t)} dt = 4aE \left(\sqrt{\frac{a^2 - b^2}{a^2}} \right). \end{aligned}$$

- b) The curve C is shown in the following figure.



We have

$$x^2 + y^2 - 2y = 0 \Rightarrow x^2 + (y - 1)^2 = 1.$$

This suggests that $g : [0, \pi] \rightarrow \mathbb{R}^3$ defined by

$$g(t) = \left(\sin(t), 1 - \cos(t), \sqrt{4 - \sin^2(t) - (1 + \cos(t))^2} \right) = \left(\sin(t), 1 - \cos(t), \sqrt{2 - 2\cos(t)} \right)$$

should be a good parametric representation for the curve C . The length of the curve C is

$$\begin{aligned} \int_0^\pi \|g'(t)\| dt &= \int_0^\pi \sqrt{\cos^2(t) + \sin^2(t) + \left(\frac{\sin(t)}{\sqrt{2 - 2\cos(t)}} \right)^2} dt \\ &= \int_0^\pi \sqrt{1 + \frac{\sin^2(t)}{2(1 - \cos(t))}} dt = \int_0^\pi \sqrt{1 + \frac{(1 - \cos(t))(1 + \cos(t))}{2(1 - \cos(t))}} dt \\ &= \int_0^\pi \sqrt{1 + \frac{1 + \cos(t)}{2}} dt = \int_0^\pi \sqrt{1 + \cos^2(t/2)} dt = \int_0^\pi \sqrt{2 - \sin^2(t/2)} dt \\ &= 2\sqrt{2} \int_0^{\pi/2} \sqrt{1 - \frac{1}{2} \sin^2(u)} du = 2\sqrt{2} E\left(\frac{1}{\sqrt{2}}\right) \end{aligned}$$

where we have used the substitution $u = t/2$ for the second to last equality. ♡

Question 6.3

Let C be a curve with the parametric representation $g : [0, 2\pi] \rightarrow \mathbb{R}^3$ defined by $g(t) = (2\cos(t), 2\sin(t), t^2)$. Compute $\int_C \sqrt{z} ds$.

Solution

Since $g'(t) = (-2\sin(t), 2\cos(t), 2t)$, we have that $\|g'(t)\| = 2\sqrt{1 + t^2}$. Thus

$$\int_C \sqrt{z} ds = 2 \int_0^{2\pi} t \sqrt{1 + t^2} dt = \frac{2}{3} (1 + t^2)^{3/2} \Big|_{t=0}^{2\pi} = \frac{2}{3} ((1 + 4\pi^2)^{3/2} - 1).$$

♡

Question 6.4

For each of the following vector fields F and curves C , compute $\int_C F \cdot ds$.

- a) C is the straight line from the origin to $(1, 1, 1)$ and $F(x, y, z) = (yz, x^2, xz)$.
 b) C is the curve $\{(x, x^2, x^3) : 0 \leq x \leq 1\}$ from $(0, 0, 0)$ to $(1, 1, 1)$ and $F(x, y, z) = (yz, x^2, xz)$.
 c) C is the circle $x^2 + y^2 = 1$ oriented clockwise and $F(x, y, z) = (x - y, x + y)$.

Solution

a) We select the parametric representation of C given by $g(t) = (t, t, t)$ for $0 \leq t \leq 1$. We have

$$\int_C F \cdot d\mathbf{s} = \int_C F \cdot g'(t) dt = \int_0^1 (t^2, t^2, t^2) \cdot (1, 1, 1) dt = \int_0^1 3t^2 dt = t^3 \Big|_{t=0}^1 = 1.$$

b) We use the parametric representation of C given by $g(t) = (t, t^2, t^3)$ for $0 \leq t \leq 1$. We have

$$\begin{aligned} \int_C F \cdot d\mathbf{s} &= \int_C F \cdot g'(t) dt = \int_0^1 (t^5, t^2, t^4) \cdot (1, 2t, 3t^2) dt \\ &= \int_0^1 (t^5 + 2t^3 + 3t^6) dt = \left(\frac{t^6}{6} + \frac{t^4}{2} + \frac{3t^7}{7} \right) \Big|_{t=0}^1 = \frac{23}{21}. \end{aligned}$$

c) We use the parametric representation of C given by

$$g(t) = (\cos(2\pi - t), \sin(2\pi - t)) = (\cos(t), -\sin(t))$$

for $0 \leq t \leq 2\pi$. The parametric representation is clockwise as required by the orientation on the curve C . We have

$$\begin{aligned} \int_C F \cdot d\mathbf{s} &= \int_C F \cdot g'(t) dt = \int_0^{2\pi} (\cos(t) + \sin(t), \cos(t) - \sin(t)) \cdot (-\sin(t), -\cos(t)) dt \\ &= \int_0^{2\pi} (-1) dt = -2\pi. \end{aligned}$$

♡

Question 6.5

For each of the following differential 1-form ω and curves C , compute $\int_C \omega$.

- a) C is the curve $\{(x, x^2) : 0 \leq x \leq 1\}$ from $(0, 0)$ to $(1, 1)$ and $\omega = xe^{-y} dx + \sin(\pi x) dy$.
 b) C is the curve $\{(\cos(t), \sin(t), t) : 0 \leq t \leq 2\pi\}$ from $(1, 0, 0)$ to $(1, 0, 2\pi)$ and $\omega = y dx + z dy + xy dz$.
 c) C is the triangle with vertices at $(0, 0)$, $(1, 0)$ and $(1, 1)$ oriented counterclockwise and $\omega = y^2 dx - 2x dy$.

Solution

a) The curve C is the singular 1-cube given by $\sigma(t) = (t, t^2)$ for $0 \leq t \leq 1$. We have

$$\int_\sigma \omega = \int_{[0,1]} \sigma^*(\omega) = \int_0^1 (te^{-t^2} + 2t \sin(\pi t)) dt$$

$$= \left(-\frac{e^{-t^2}}{2} - \frac{2t \cos(\pi t)}{\pi} + \frac{2 \sin(\pi t)}{\pi^2} \right) \Big|_{t=0}^1 = \frac{1}{2} - \frac{1}{2e} + \frac{2}{\pi}.$$

b) The curve C is the singular 1-cube given by $\sigma(t) = (\cos(2\pi t), \sin(2\pi t), 2\pi t)$ for $0 \leq t \leq 1$. We have

$$\begin{aligned} \int_{\sigma} \omega &= \int_{[0,1]} \sigma^*(\omega) = 2\pi \int_0^1 (-\sin^2(2\pi t) + t \cos(2\pi t) + \cos(2\pi t) \sin(2\pi t)) dt \\ &= \int_0^{2\pi} \left(-\frac{1}{2}(1 - \cos(2t)) + t \cos(t) + \cos(t) \sin(t) \right) dt \\ &= \left(-\frac{t}{2} + \frac{\sin(2t)}{4} + t \sin(t) + \cos(t) + \frac{\sin^2(t)}{2} \right) \Big|_{t=0}^{2\pi} = -\pi. \end{aligned}$$

c) We have that $C = C_1 \cup C_2 \cup C_3$, where the straight lines C_i for $1 \leq i \leq 3$ are the following singular 1-cube: C_1 from $(0, 0)$ to $(1, 0)$ is given by $\sigma_1(t) = (t, 0)$ for $0 \leq t \leq 1$, C_2 from $(1, 0)$ to $(1, 1)$ is given by $\sigma_2(t) = (1, t)$ for $0 \leq t \leq 1$, and C_3 from $(1, 1)$ to $(0, 0)$ is given by $\sigma_3(t) = (1 - t, 1 - t)$ for $0 \leq t \leq 1$. Since $\int_{\sigma_1} \omega = \int_{[0,1]} \sigma_1^*(\omega) = \int_0^1 0 dt = 0$, $\int_{\sigma_2} \omega = \int_{[0,1]} \sigma_2^*(\omega) = \int_0^1 (-2) dt = -2$ and $\int_{\sigma_3} \omega = \int_{[0,1]} \sigma_3^*(\omega) = \int_0^1 (-(1-t)^2 + 2(1-t)) dt = \frac{2}{3}$, we find that

$$\int_C \omega = \int_{\sigma_1} \omega + \int_{\sigma_2} \omega + \int_{\sigma_3} \omega = -2 + \frac{2}{3} = -\frac{4}{3}.$$

♡

Question 6.6

Let $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a continuous vector fields and $C \subset \mathbb{R}^n$ be a curve. Prove that $\left| \int_C F \cdot ds \right| \leq \int_C \|F\| ds$.

Solution

Let $\sigma : [0, 1] \rightarrow \mathbb{R}^n$ be a parametric representation of C . Then

$$\begin{aligned} \left| \int_C F \cdot ds \right| &= \left| \int_0^1 F(\sigma(t)) \cdot \sigma'(t) dt \right| \leq \int_0^1 |F(\sigma(t)) \cdot \sigma'(t)| dt \\ &\leq \int_0^1 \|F(\sigma(t))\| \|\sigma'(t)\| dt = \int_0^1 \|F\| ds, \end{aligned}$$

where the first inequality is standard for the Riemann integral of real valued functions and the second inequality comes from the Schwarz inequality. ♡

Question 6.7

Suppose that a curve C is the union of two curves C_1 and C_2 . Use definition 6.2.11 to prove that the length of C is the sum of the lengths of C_1 and C_2 .

Solution

Since C is a curve, there exists a parametric representation $g : [a, b] \rightarrow \mathbb{R}^n$ of C such that

g is one-to-one except possibly at a finite number of points. Since C is the union of two curves C_1 and C_2 , there exists c between a and b such that $g : [a, c] \rightarrow \mathbb{R}^n$ is a parametric representation of C_1 and $g : [c, b] \rightarrow \mathbb{R}^n$ is a parametric representation of C_2 .

Let L be the length of C and L_i be the length of C_i for $1 \leq i \leq 2$.

i) Let P be a partition of $[a, b]$ and consider the partition $Q = P \cup \{c\}$. The partition Q is a refinement of P . Moreover $Q_1 = Q \cap [a, c]$ is a partition of $[a, c]$ and $Q_2 = Q \cap [c, b]$ is a partition of $[c, b]$. We have that $L_P(g) \leq L_Q(g) = L_{Q_1}(g) + L_{Q_2}(g) \leq L_1 + L_2$. Since this is true for all partitions P of $[a, b]$, we get that $L \leq L_1 + L_2$.

ii) Choose $\epsilon > 0$. Let P_1 and P_2 be partitions of $[a, c]$ and $[c, b]$ respectively such that $L_1 - L_{P_1}(g) < \epsilon/2$ and $L_2 - L_{P_2}(g) < \epsilon/2$. We have that $P = P_1 \cup P_2$ is a partition of $[a, b]$ such that

$$0 \leq L_1 + L_2 - L_P(g) = (L_1 - L_{P_1}(g)) + (L_2 - L_{P_2}(g)) < \epsilon.$$

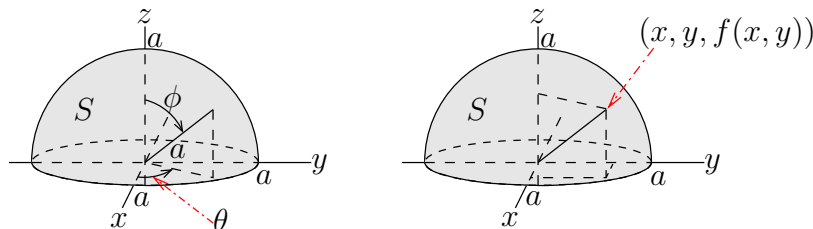
Thus $L_1 + L_2 < L_P(g) + \epsilon \leq L + \epsilon$. Since ϵ is arbitrary, we get the second inequality $L_1 + L_2 \leq L$. $\heartsuit$

Question 6.8

Compute the surface integral $\iint_S z \, dS$ over the upper hemisphere S of radius a using two different parametric representations.

Solution

The surface S is shown in the figure below.



First representation: The parametric representation is given by the spherical coordinates

$$g(\theta, \phi) = (a \cos(\theta) \sin(\phi), a \sin(\theta) \sin(\phi), a \cos(\phi))$$

for $0 \leq \theta < 2\pi$ and $0 \leq \phi \leq \pi/2$. We have that $\left\| \frac{\partial g}{\partial \theta} \times \frac{\partial g}{\partial \phi} \right\| = a^2 \sin(\phi)$. Thus

$$\begin{aligned} \iint_S z \, dS &= \int_0^{2\pi} \int_0^{\pi/2} a \cos(\phi) \left\| \frac{\partial g}{\partial \theta} \times \frac{\partial g}{\partial \phi} \right\| \, d\phi \, d\theta = \int_0^{2\pi} \int_0^{\pi/2} a^3 \cos(\phi) \sin(\phi) \, d\phi \, d\theta \\ &= \int_0^{2\pi} \left(a^3 \frac{\sin^2(\phi)}{2} \Big|_0^{\pi/2} \right) \, d\theta = \int_0^{2\pi} \frac{a^3}{2} \, d\theta = \pi a^3. \end{aligned}$$

Second representation: The parametric representation is given by the Cartesian coordinates $g(x, y) = (x, y, \sqrt{a^2 - x^2 - y^2})$ for $x^2 + y^2 \leq a^2$. We have

$$dS = \sqrt{\left(\frac{\partial z}{\partial x} \right)^2 + \left(\frac{\partial z}{\partial y} \right)^2 + 1} \, dy \, dx$$

$$= \sqrt{\left(\frac{-x}{\sqrt{a^2 - x^2 - y^2}}\right)^2 + \left(\frac{-y}{\sqrt{a^2 - x^2 - y^2}}\right)^2 + 1} \, dy \, dx = \frac{a}{\sqrt{a^2 - x^2 - y^2}} \, dy \, dx .$$

On the disk $x^2 + y^2 \leq a$, we have that $-a \leq x \leq a$ and $-\sqrt{a^2 - x^2} \leq y \leq \sqrt{a^2 - x^2}$. Hence

$$\begin{aligned} \iint_S z \, dS &= \int_{-a}^a \int_{-\sqrt{a^2 - x^2}}^{\sqrt{a^2 - x^2}} \sqrt{a^2 - x^2 - y^2} \frac{a}{\sqrt{a^2 - x^2 - y^2}} \, dy \, dx = \int_{-a}^a \int_{-\sqrt{a^2 - x^2}}^{\sqrt{a^2 - x^2}} a \, dy \, dx \\ &= \int_{-a}^a a y \Big|_{y=-\sqrt{a^2 - x^2}}^{\sqrt{a^2 - x^2}} \, dx = \int_{-a}^a 2a\sqrt{a^2 - x^2} \, dx = \pi a^3 , \end{aligned}$$

where we have used the trigonometric substitution $x = a \sin(t)$ with $-\pi/2 \leq t \leq \pi/2$ to compute the last integral.

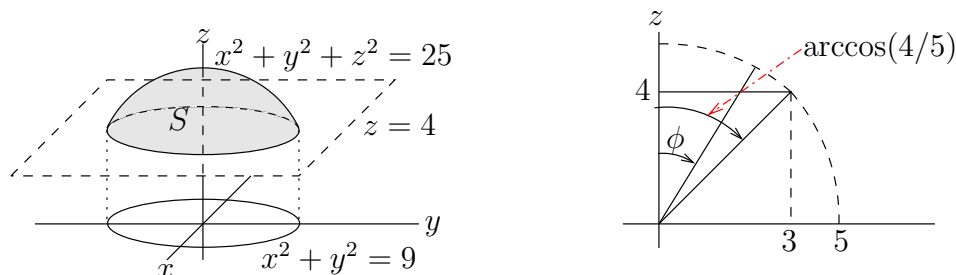
This should convince the reader that a good choice of parametric representation may seriously simplify the computations. $\heartsuit$

Question 6.9

Compute the surface integral $\iint_S (x^2 + y^2) \, dS$ over the portion S of the sphere $x^2 + y^2 + z^2 = 25$ above $z = 4$; namely, $z \geq 4$.

Solution

The surface S is shown in the figure below.



There are two possible parametric representations that immediately come to mind:

$h(r, \theta) = (r \cos(\theta), r \sin(\theta), \sqrt{25 - r^2})$ for $0 \leq r \leq 3$ and $0 \leq \theta \leq 2\pi$, and

$g(\theta, \phi) = (5 \cos(\theta) \sin(\phi), 5 \sin(\theta) \sin(\phi), 5 \cos(\phi))$ for $0 \leq \phi \leq \arccos(4/5)$ and $0 \leq \theta \leq 2\pi$. We will use the second parametric representation to compute the area of S . Since

$\left\| \frac{\partial g}{\partial \theta} \times \frac{\partial g}{\partial \phi} \right\| = 25 \sin(\phi)$, the area is

$$\begin{aligned} \iint_S (x^2 + y^2) \, dS &= \int_0^{2\pi} \int_0^{\arccos(4/5)} 625 \sin^3(\phi) \, d\phi \, d\theta \\ &= 1250\pi \int_0^{\arccos(4/5)} (1 - \cos^2(\phi)) \sin(\phi) \, d\phi \\ &= 1250\pi \int_{4/5}^1 (1 - u^2) \, du = 1250\pi \left(u - \frac{u^3}{3} \right) \Big|_{u=4/5}^1 = \frac{140\pi}{3} \end{aligned}$$

where we have used the substitution $u = \cos(\phi)$ for the third equality. ♡

Question 6.10

Compute $\int_S (x^2 + y^2 - 2z^2) dS$ where S is the unit sphere centred at the origin.

Solution

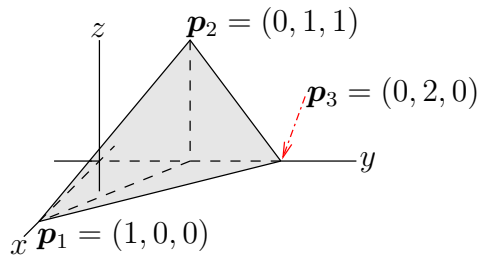
We have

$$\int_S (x^2 + y^2 - 2z^2) dS = \int_S x^2 dS + \int_S y^2 dS - 2 \int_S z^2 dS.$$

A little observation may save us of a long computation. Because of the symmetry, we have that $\int_S x^2 dS = \int_S y^2 dS = \int_S z^2 dS$. Thus $\int_S (x^2 + y^2 - 2z^2) dS = 0$. ♡

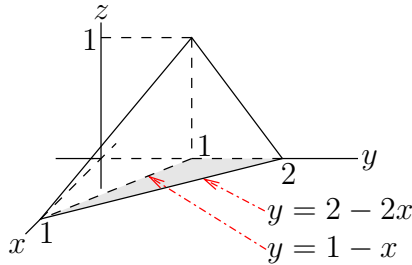
Question 6.11

Compute the integral $\iint_S xyz dS$ where S is the triangular surface represented in the following figure.



Solution

We first find the equation of the plane containing the triangular surface. The vectors $\mathbf{v}_1 = \mathbf{p}_1 - \mathbf{p}_2 = (1, -1, -1)$ and $\mathbf{v}_2 = \mathbf{p}_3 - \mathbf{p}_2 = (0, 1, -1)$ are parallel to that plane. A normal vector to the plane is given by $\mathbf{n} = \mathbf{v}_1 \times \mathbf{v}_2 = (2, 1, 1)$ and the equation of the plane is $\langle ((x, y, z) - \mathbf{p}_1), \mathbf{n} \rangle = 2x - 2 + y + z = 0$. A parametric representation of the triangular surface is given by $\phi(x, y) = (x, y, 2 - 2x - y)$ for $0 \leq x \leq 1$ and $1 - x \leq y \leq 2 - 2x$.



We have that $dS = \left\| \frac{\partial \phi}{\partial x} \times \frac{\partial \phi}{\partial y} \right\| dy dx = \|\mathbf{n}\| dy dx = \sqrt{6} dy dx$. Hence

$$\iint_S xyz dS = \int_0^1 \int_{1-x}^{2-2x} xy(2 - 2x - y) \sqrt{6} dy dx = \int_0^1 \left(xy^2 - x^2 y^2 - \frac{xy^3}{3} \right) \Big|_{y=1-x}^{2-2x} \sqrt{6} dx$$

$$\begin{aligned}
&= \frac{2\sqrt{6}}{3} \int_0^1 x(1-x)^3 dx = \frac{2\sqrt{6}}{3} \int_0^1 (-x^4 + 3x^3 - 3x^2 + x) dx \\
&= \frac{2\sqrt{6}}{3} \left(-\frac{x^5}{5} + \frac{3x^4}{4} - x^3 + \frac{x^2}{2} \right) \Big|_{x=0}^1 = \frac{\sqrt{6}}{30} .
\end{aligned}$$

Note that we could also have used the following formula to compute dS .

$$dS = \sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 + 1} dy dx = \sqrt{(-2)^2 + (-1)^2 + 1} dy dx = \sqrt{6} dy dx .$$

♡

Question 6.12

Find the area of the surface $z = xy$ inside the cylinder $x^2 + y^2 = a^2$.

Solution

Let S be the surface defined in the statement of the question. A parametric representation of this surface is given by $g(r, \theta) = (r \cos(\theta), r \sin(\theta), r^2 \cos(\theta) \sin(\theta))$ for $0 \leq r \leq a$ and

$0 \leq \theta \leq 2\pi$. Since $\frac{\partial g}{\partial r} = (\cos(\theta), \sin(\theta), 2r \cos(\theta) \sin(\theta))$ and

$\frac{\partial g}{\partial \theta} = (-r \sin(\theta), r \cos(\theta), r^2(\cos^2(\theta) - \sin^2(\theta)))$, we get that

$\left\| \frac{\partial g}{\partial r} \times \frac{\partial g}{\partial \theta} \right\| = \|(-r^2 \sin(\theta), -r^2 \cos(\theta), r)\| = r\sqrt{r^2 + 1}$. Hence, the area is

$$\iint_S dA = \int_0^{2\pi} \int_0^a r\sqrt{r^2 + 1} dr d\theta = \frac{2\pi}{3} (r^2 + 1)^{3/2} \Big|_{r=0}^a = \frac{2\pi}{3} ((a^2 + 1)^{3/2} - 1) .$$

♡

Question 6.13

Find the area of the surface $z = x^2 + y^2$ inside the cylinder $x^2 + y^2 = a^2$.

Solution

Let S be the surface defined in the statement of the question. A parametric representation of this surface is given by $g(r, \theta) = (r \cos(\theta), r \sin(\theta), r^2)$ for $0 \leq r \leq a$ and $0 \leq \theta \leq 2\pi$.

Since $\frac{\partial g}{\partial r} = (\cos(\theta), \sin(\theta), 2r)$ and $\frac{\partial g}{\partial \theta} = (-r \sin(\theta), r \cos(\theta), 0)$, we get that $\left\| \frac{\partial g}{\partial r} \times \frac{\partial g}{\partial \theta} \right\| =$

$\|(-2r^2 \cos(\theta), -2r^2 \sin(\theta), r)\| = r\sqrt{4r^2 + 1}$. Hence, the area is

$$\iint_S dA = \int_0^{2\pi} \int_0^a r\sqrt{4r^2 + 1} dr d\theta = \frac{\pi}{6} (4r^2 + 1)^{3/2} \Big|_{r=0}^a = \frac{\pi}{6} ((4a^2 + 1)^{3/2} - 1) .$$

♡

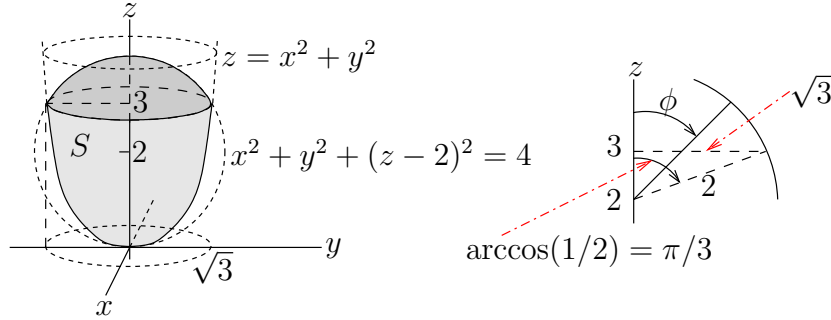
Question 6.14

Let S be the solid bounded below by the paraboloid $z = x^2 + y^2$ and above by the sphere $x^2 + y^2 + z^2 = 4z$ for $z \geq 0$. Find the area of the surface of S .

Solution

The value of z at the intersection of the paraboloid with the sphere is given by $z + z^2 = 4z$; namely, $z^2 - 3z = z(z - 3) = 0$. For $z = 0$, the sphere touches the paraboloid at its lowest point. For $z = 3$, we get the circle $x^2 + y^2 = 3$.

The surface S is shown in the following figure.



A parametric representation of the part of the sphere is given by

$$\phi_1(\theta, \phi) = (2 \cos(\theta) \sin(\phi), 2 \sin(\theta) \sin(\phi), (2 + 2 \cos(\phi)))$$

for $0 \leq \theta \leq 2\pi$ and $0 \leq \phi \leq \pi/3 = \arccos(1/2)$. A parametric representation of the part of the paraboloid is given by

$$\phi_2(\theta, r) = (r \cos(\theta), r \sin(\theta), r^2)$$

for $0 \leq \theta \leq 2\pi$ and $0 \leq r \leq \sqrt{3}$. We have

$$\left\| \frac{\partial \phi_1}{\partial \theta} \times \frac{\partial \phi_1}{\partial \phi} \right\| = \left\| (-4 \sin(\theta) \sin^2(\phi), -4 \cos(\theta) \sin^2(\phi), -4 \sin(\phi) \cos(\phi)) \right\| = 4 \sin(\phi)$$

and

$$\left\| \frac{\partial \phi_2}{\partial \theta} \times \frac{\partial \phi_2}{\partial r} \right\| = \left\| (2r^2 \cos(\theta), 2r^2 \sin(\theta), -r) \right\| = r\sqrt{4r^2 + 1}.$$

Hence, the area is

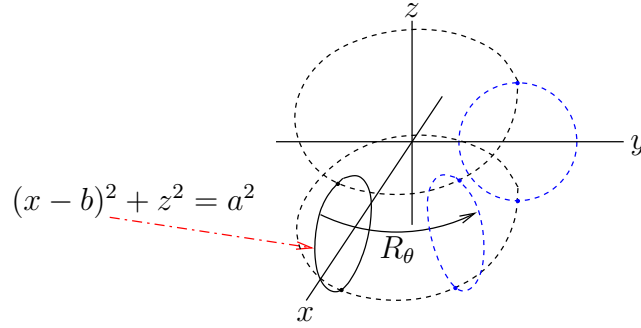
$$\begin{aligned} \int_{\partial S} dA &= \int_0^{2\pi} \int_0^{\pi/3} 4 \sin(\phi) d\phi d\theta + \int_0^{2\pi} \int_0^{\sqrt{3}} r\sqrt{1 + 4r^2} dr d\theta \\ &= -8\pi \cos(\phi) \Big|_{\phi=0}^{\pi/3} + \frac{\pi}{6} (1 + 4r^2)^{3/2} \Big|_{r=0}^{\sqrt{3}} = 4\pi + \frac{\pi}{6} (13^{3/2} - 1). \end{aligned}$$

♡

Question 6.15

Given $0 < a < b$, compute the area of the torus obtained from the rotation of the circle $(x - b)^2 + z^2 = a^2$ about the z axis.

Solution



A parametric representation of the circle $(x - b)^2 + z^2 = a^2$ is given by $h(\phi) = (b + a \cos(\phi) \ 0 \ a \sin(\phi))^T$ for $0 \leq \phi \leq 2\pi$. To obtain the torus, it suffices to rotate this circle about the z axis. This is done by multiplying $h(\phi)$ by the matrix of rotation by θ radians counterclockwise about the z axis given by

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

We get the following parametric representation for the torus.

$$g(\phi, \theta) = R_\theta \begin{pmatrix} b + a \cos(\phi) \\ 0 \\ a \sin(\phi) \end{pmatrix} = \begin{pmatrix} (b + a \cos(\phi)) \cos(\theta) \\ (b + a \cos(\phi)) \sin(\theta) \\ a \sin(\phi) \end{pmatrix}$$

for $0 \leq \phi \leq 2\pi$ and $0 \leq \theta \leq 2\pi$. Since $\frac{\partial g}{\partial \phi} = \begin{pmatrix} -a \sin(\phi) \cos(\theta) \\ -a \sin(\phi) \sin(\theta) \\ a \cos(\phi) \end{pmatrix}$ and

$\frac{\partial g}{\partial \theta} = \begin{pmatrix} -(b + a \cos(\phi)) \sin(\theta) \\ (b + a \cos(\phi)) \cos(\theta) \\ 0 \end{pmatrix}$, we get

$$\left\| \frac{\partial g}{\partial \phi} \times \frac{\partial g}{\partial \theta} \right\| = \left\| \begin{pmatrix} -a(b + a \cos(\phi)) \cos(\theta) \cos(\phi) \\ -a(b + a \cos(\phi)) \sin(\theta) \cos(\phi) \\ -a(b + a \cos(\phi)) \sin(\phi) \end{pmatrix} \right\| = a(b + a \cos(\phi)).$$

Hence, the area is

$$\iint_S dA = \int_0^{2\pi} \int_0^{2\pi} a(b + a \cos(\phi)) d\phi d\theta = 2\pi \left(ab\phi + a^2 \sin(\phi) \right) \Big|_{\phi=0}^{2\pi} = 4\pi^2 ab.$$



Question 6.16

A Mobius strip can be parameterized by

$$g(u, v) = ((2 - v \sin(u/2)) \cos(u), (2 - v \sin(u/2)) \sin(u), v \cos(u/2))$$

for $0 \leq u < 2\pi$ and $|v| < 1$. Set up the integral to compute the area of this Möbius strip. What is the surface of a Möbius strip?

Question 6.17

Find the surface area of the portion of the paraboloid $z = x^2 + y^2$ which is inside the region bounded by the cylinders $x^2 + y^2 = 2$ and $x^2 + y^2 = 6$.

Question 6.18

Find the area of the portion of the sphere of radius R centred at the origin which is inside the cylinder $x^2 + y^2 - Ry \leq 0$ and satisfies $x \geq 0$.

Question 6.19

Suppose that $f : [a, b] \rightarrow \mathbb{R}$ is a continuously differentiable function. We learn in calculus that the area A of the surface of revolution S obtained by revolving the graph of f (in the x, y plane) around the x axis is $A = 2\pi \int_a^b |f(x)| \sqrt{1 + (f'(x))^2} dx$. Prove this formula using a surface integral.

Solution

A parametric representation of the surface of revolution is given by $\phi(x, \theta) = (x, f(x) \cos(\theta), f(x) \sin(\theta))$ for $a \leq x \leq b$ and $0 \leq \theta \leq 2\pi$. Hence

$$\begin{aligned} \left\| \frac{\partial \phi}{\partial \theta} \times \frac{\partial \phi}{\partial x} \right\| &= \left\| \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ 0 & -f(x) \sin(\theta) & f(x) \cos(\theta) \\ 1 & f'(x) \cos(\theta) & f'(x) \sin(\theta) \end{pmatrix} \right\| \\ &= \|(-f(x)f'(x), f(x) \cos(\theta), f(x) \sin(\theta))\| = |f(x)| \sqrt{1 + (f'(x))^2}. \end{aligned}$$

Thus, the area is

$$\int_0^{2\pi} \int_a^b |f(x)| \sqrt{1 + (f'(x))^2} dx d\theta = 2\pi \int_a^b |f(x)| \sqrt{1 + (f'(x))^2} dx.$$

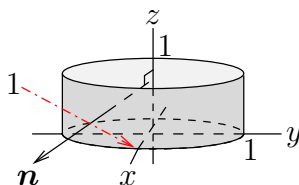
♡

Question 6.20

Evaluate directly the surface integral $\iint_S \mathbf{F} \cdot d\mathbf{S}$ where $F(x, y, z) = (x, y, z)$ and S is the cylinder defined by $x^2 + y^2 = 1$ and $0 \leq z \leq 1$. The orientation on S is given by the unit normal to S pointing away from the z axis.

Solution

The surface S is shown in the following figure.



We choose the parametric representation $\phi(\theta, z) = (\cos(\theta), \sin(\theta), z)$ for $0 \leq \theta < 2\pi$ and $0 \leq z \leq 1$. Hence

$$\frac{\partial \phi}{\partial \theta} \times \frac{\partial \phi}{\partial z} = \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix} = (\cos(\theta), \sin(\theta), 0) .$$

This is a vector pointing away from the z axis as required by the orientation on S . Since it is a vector of norm one, we have that $\mathbf{n} = \frac{\partial \phi}{\partial \theta} \times \frac{\partial \phi}{\partial z}$ is the unit normal associated to the orientation on S . Thus

$$\begin{aligned} \iint_S \mathbf{F} \cdot d\mathbf{S} &= \int_S \mathbf{F} \cdot \mathbf{n} dS = \int_0^{2\pi} \int_0^1 (\cos(\theta), \sin(\theta), z) \cdot (\cos(\theta), \sin(\theta), 0) dz d\theta \\ &= \int_0^{2\pi} \int_0^1 dz d\theta = 2\pi . \end{aligned}$$

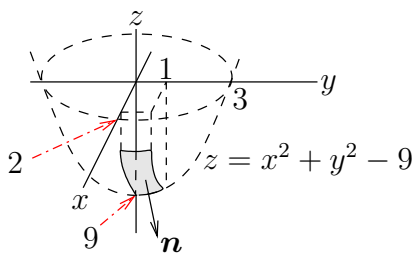
♥

Question 6.21

Evaluate directly the surface integral $\iint_S \mathbf{F} \cdot d\mathbf{S}$ where $F(x, y, z) = (x^2y, -3xy^2, 4y^3)$ and S is the portion of the surface $z = x^2 + y^2 + 9$ with $0 \leq x \leq 2$ and $0 \leq y \leq 1$. The orientation on S is given by the unit normal to S pointing in the direction of z negative.

Solution

The surface S is shown in the following figure.



We choose the parametric representation $\phi(x, y) = (x, y, (x^2 + y^2 + 9))$ for $0 \leq x < 2$ and $0 \leq y \leq 1$. Hence

$$\frac{\partial \phi}{\partial x} \times \frac{\partial \phi}{\partial y} = \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ 1 & 0 & 2x \\ 0 & 1 & 2y \end{pmatrix} = (-2x, -2y, 1) .$$

This is a vector pointing in the direction of z positive. Thus, in the opposite direction of the unit normal that gives the orientation on S . Therefore

$$\mathbf{n} = - \left\| \frac{\partial \phi}{\partial x} \times \frac{\partial \phi}{\partial y} \right\|^{-1} \left(\frac{\partial \phi}{\partial x} \times \frac{\partial \phi}{\partial y} \right)$$

$$= \left(\frac{2x}{\sqrt{4x^2 + 4y^2 + 1}}, \frac{2y}{\sqrt{4x^2 + 4y^2 + 1}}, -\frac{1}{\sqrt{4x^2 + 4y^2 + 1}} \right)$$

is a normal vector pointing in the direction of z negative as required and

$$dS = \left\| \frac{\partial \phi}{\partial x} \times \frac{\partial \phi}{\partial y} \right\| dy dx = \sqrt{4x^2 + 4y^2 + 1} dy dx. \text{ Hence}$$

$$\begin{aligned} \iint_S F \cdot d\mathbf{S} &= \int_S F \cdot \mathbf{n} dS = \int_0^2 \int_0^1 (x^2y, -3xy^2, 4y^3) \cdot (2x, 2y, -1) dy dx \\ &= \int_0^2 \int_0^1 (2x^3y - 6xy^3 - 4y^3) dy dx = \int_0^2 \left(x^3y^2 - \frac{3}{2}xy^4 - y^4 \right) \Big|_{y=0}^1 dx \\ &= \int_0^2 \left(x^3 - \frac{3x}{2} - 1 \right) dx = \left(\frac{x^4}{4} - \frac{3x^2}{4} - x \right) \Big|_{x=0}^2 = -1. \end{aligned}$$

♡

Question 6.22

Evaluate directly the surface integral $\iint_R F \cdot d\mathbf{S}$ where $F(x, y, z) = (x, 0, y)$ and R is one of the surfaces described below.

- a) R is the surface with the parametric representation $g(u, v) = (u + v, u^2 - v^2, v \cos(u/2))$ for $0 \leq u \leq 1$ and $0 \leq v \leq 1$. We assume that this parametric representation is consistent with the chosen orientation on R .
- b) R is the surface of the cylinder $y^2 + z^2 = 1$ with $0 \leq x \leq 1$. The orientation is given by the unit normal pointing away from the x axis.
- c) R is the boundary of the unit cube $\{(x, y, z) : 0 \leq x, y, z \leq 1\}$. The orientation is given by the unit normal pointing outward.

Question 6.23

Compute $\int_S F \cdot d\mathbf{S}$ for each of the following vector fields F and surfaces S .

- a) $F(x, y, z) = (xz, 0, -xy)$ and S is the portion of the surface $z = xy$ for $0 \leq x \leq 1$ and $0 \leq y \leq 2$. The orientation on S is give by a normal pointing in the direction of z positive.
- b) $F(x, y, z) = (x^2, z, -y)$ and S is the unit sphere centred at the origin. The orientation on S is give by a normal pointing outward.
- c) $F(x, y, z) = (xy, z, 0)$ and S is the triangular surface with vertices at $(2, 0, 0)$, $(0, 2, 0)$ and $(0, 0, 2)$. The orientation on S is give by a normal pointing in the direction of z positive.
- d) $F(x, y, z) = (0, 0, z^2)$ and S is the boundary of the region enclosed by the cylinder $x^2 + y^2 = 1$ and the planes $z = a, b$ with $a < b$. The orientation on S is given by a normal pointing outward.
- e) $F(x, y, z) = (x, y, z)$ and S is the boundary of the region enclosed by the surface $z = x^2 + y^2$ and the sphere $x^2 + y^2 + z^2 = 20$. The orientation on S is given by a normal pointing outward.

Solution

- a) we choose the parametric representation $g(u, v) = (u, v, uv)$ for $0 \leq u \leq 1$ and $0 \leq v \leq 2$. Since $\frac{\partial g}{\partial u} = (1, 0, v)$ and $\frac{\partial g}{\partial v} = (0, 1, u)$, we get that $\frac{\partial g}{\partial u} \times \frac{\partial g}{\partial v} = (-v, -u, 1)$. Since $(-v, -u, 1)$

is a vector that point in the direction of z positive, the parametric representation is consistent with the required orientation on S . Thus

$$\begin{aligned}\iint_S F \cdot d\mathbf{S} &= \int_0^1 \int_0^2 (u^2v, 0, -uv) \cdot (-v, -u, 1) dv du = \int_0^1 \int_0^2 (-u^2v^2 - uv) dv du \\ &= \int_0^1 \left(-\frac{u^2v^3}{3} - \frac{uv^2}{2} \right) \Big|_{v=0}^2 du = \int_0^1 \left(-\frac{8u^2}{3} - 2u \right) du \\ &= \left(-\frac{8u^3}{9} - u^2 \right) \Big|_{u=0}^1 = -\frac{17}{9}.\end{aligned}$$

b) We have that $\iint_S F \cdot d\mathbf{S} = \iint_S F \cdot \mathbf{n} dS$ where $\mathbf{n}$ is a unit vector pointing outside S at each point of S . Since S is a unit sphere centred at the origin, we have that $\mathbf{n} = (x, y, z)$ at every point (x, y, z) of S . Thus

$$\iint_S F \cdot d\mathbf{S} = \iint_S (x^2, z, -y) \cdot (x, y, z) dS = \iint_S x^3 dS.$$

Since $x \mapsto x^3$ is an odd function and S is symmetric with respect to the y, z plane, the last integral is null. Hence $\iint_S F \cdot d\mathbf{S} = 0$.

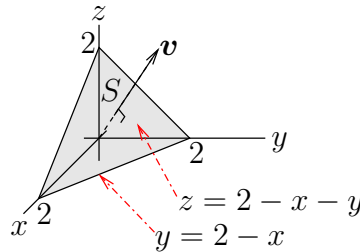
c) We first need to find the equation of the plane that contains the points $(2, 0, 0)$, $(0, 2, 0)$ and $(0, 0, 2)$. A normal to this plane is given by

$$\mathbf{v} = ((0, 2, 0) - (2, 0, 0)) \times ((0, 0, 2) - (2, 0, 0)) = (4, 4, 4).$$

An equation for the plane is given by

$$((x, y, z) - (2, 0, 0)) \cdot (4, 4, 4) = (x - 2, y, z) \cdot (4, 4, 4) = 4x - 8 + 4y + 4z = 0;$$

namely, $z = 2 - x - y$. We choose the parametric representation $g(u, v) = (u, v, 2 - u - v)$ for $0 \leq u \leq 2$ and $0 \leq v \leq 2 - u$.



Since $\frac{\partial g}{\partial u} = (1, 0, -1)$ and $\frac{\partial g}{\partial v} = (0, 1, -1)$, we get $\frac{\partial g}{\partial u} \times \frac{\partial g}{\partial v} = (1, 1, 1)$. Since $(1, 1, 1)$ is a vector that point in the direction of z positive, the parametric representation is consistent with the required orientation on S . Thus

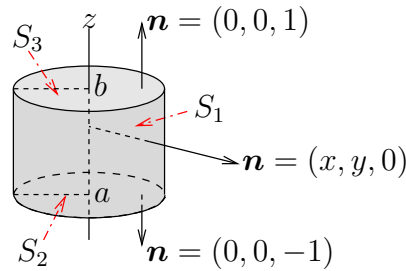
$$\iint_S F \cdot d\mathbf{S} = \int_0^2 \int_0^{2-u} (uv, 2 - u - v, 0) \cdot (1, 1, 1) dv du = \int_0^2 \int_0^{2-u} (2 - u - v + uv) dv du$$

$$\begin{aligned}
&= \int_0^2 \left(2v - uv - \frac{v^2}{2} + \frac{uv^2}{2} \right) \Big|_{v=0}^{2-u} du = \int_0^2 \left(2 - \frac{3u^2}{2} + \frac{u^3}{2} \right) du \\
&= \left(2u - \frac{u^3}{2} + \frac{u^4}{8} \right) \Big|_{u=0}^2 = 2 .
\end{aligned}$$

d) We have

$$\iint_S F \cdot d\mathbf{S} = \iint_S F \cdot \mathbf{n} dS = \iint_{S_1} F \cdot \mathbf{n} dS + \iint_{S_2} F \cdot \mathbf{n} dS + \iint_{S_3} F \cdot \mathbf{n} dS ,$$

where $S_1 = \{(x, y, z) : x^2 + y^2 \leq 1 \text{ and } z = a\}$ is the bottom of S , $S_2 = \{(x, y, z) : x^2 + y^2 = 1 \text{ and } a \leq z \leq b\}$ is the side of S , and $S_3 = \{(x, y, z) : x^2 + y^2 \leq 1 \text{ and } z = b\}$ is the top of S .



As usual, $\mathbf{n}$ is a unit vector pointing outside S . So $\mathbf{n} = (0, 0, -1)$ on S_1 , $\mathbf{n} = (x, y, 0)$ on S_2 and $\mathbf{n} = (0, 0, 1)$ on S_3 . Hence

$$\begin{aligned}
\iint_S F \cdot d\mathbf{S} &= \iint_{S_1} (0, 0, a^2) \cdot (0, 0, -1) dS + \iint_{S_2} (0, 0, z^2) \cdot (x, y, 0) dS \\
&\quad + \iint_{S_3} (0, 0, b^2) \cdot (0, 0, 1) dS = -a^2 \iint_{S_1} dS + b^2 \iint_{S_3} dS .
\end{aligned}$$

Since $\iint_{S_1} dS = \iint_{S_3} dS = \iint_{x^2+y^2 \leq 1} dS = \pi$, the area of a disk of radius 1, we get that

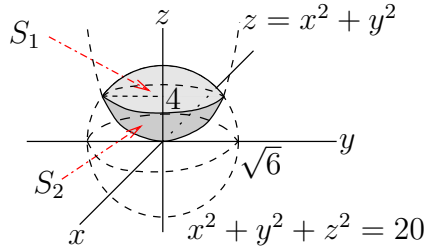
$$\iint_S F \cdot d\mathbf{S} = (b^2 - a^2)\pi .$$

e) We first find the intersection of the two surfaces. On the intersection, z must satisfy $z^2 + z - 20 = (z + 5)(z - 4) = 0$. Since $z = x^2 + y^2$ is non negative, we get $z = 4$. Thus the intersection is the circle $x^2 + y^2 = 4$ with $z = 4$.

We have

$$\iint_S F \cdot d\mathbf{S} = \iint_S F \cdot \mathbf{n} dS = \iint_{S_1} F \cdot \mathbf{n} dS + \iint_{S_2} F \cdot \mathbf{n} dS ,$$

where $S_1 = \{(x, y, \sqrt{20 - x^2 - y^2}) : x^2 + y^2 \leq 4\}$ is the top of S and $S_2 = \{(x, y, x^2 + y^2) : x^2 + y^2 \leq 1\}$ is the bottom of S .

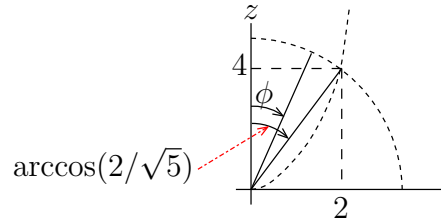


As usual, $\mathbf{n}$ is a unit vector pointing outside S .

On S_1 , we choose the parametric representation

$$g_1(\phi, \theta) = (2\sqrt{5} \cos(\theta) \sin(\phi), 2\sqrt{5} \sin(\theta) \sin(\phi), 2\sqrt{5} \cos(\phi))$$

for $0 \leq \theta \leq 2\pi$ and $0 \leq \phi \leq \arccos(2/\sqrt{5})$.



Since

$$\frac{\partial g_1}{\partial \phi} = (2\sqrt{5} \cos(\theta) \cos(\phi), 2\sqrt{5} \sin(\theta) \cos(\phi), -2\sqrt{5} \sin(\phi))$$

and

$$\frac{\partial g_1}{\partial \theta} = (-2\sqrt{5} \sin(\theta) \sin(\phi), 2\sqrt{5} \cos(\theta) \sin(\phi), 0) ,$$

we get

$$\frac{\partial g_1}{\partial \phi} \times \frac{\partial g_1}{\partial \theta} = (20 \cos(\theta) \sin^2(\phi), 20 \sin(\theta) \sin^2(\phi), 20 \cos(\phi) \sin(\phi)) .$$

This is a vector that point in the direction of z positive, thus the parametric representation is consistent with the required orientation on S_1 . We have

$$\begin{aligned} \iint_{S_1} F \cdot d\mathbf{S} &= \iint_{S_1} (2\sqrt{5} \cos(\theta) \sin(\phi), 2\sqrt{5} \sin(\theta) \sin(\phi), 2\sqrt{5} \cos(\phi)) \\ &\quad \cdot (20 \cos(\theta) \sin^2(\phi), 20 \sin(\theta) \sin^2(\phi), 20 \cos(\phi) \sin(\phi)) dS \\ &= \iint_{S_1} 40\sqrt{5} \sin(\phi) dS = \int_0^{2\pi} \int_0^{\arccos(2/\sqrt{5})} 40\sqrt{5} \sin(\phi) d\phi d\theta \\ &= -80\pi\sqrt{5} \cos(\phi) \Big|_{\phi=0}^{\arccos(2/\sqrt{5})} = 80\pi\sqrt{5} - 160\pi . \end{aligned}$$

On S_2 , we choose the parametric representation $g_2(\theta, r) = (r \cos(\theta), r \sin(\theta), r^2)$ for $0 \leq \theta \leq 2\pi$ and $0 \leq r \leq 2$. Since $\frac{\partial g_2}{\partial \theta} = (-r \sin(\theta), r \cos(\theta), 0)$ and $\frac{\partial g_2}{\partial r} = (\cos(\theta), \sin(\theta), 2r)$, we

get that $\frac{\partial g_2}{\partial \theta} \times \frac{\partial g_2}{\partial r} = (2r^2 \cos(\theta), 2r^2 \sin(\theta), -r)$. This is a vector that point in the direction of z negative, thus the parametric representation is consistent with the required orientation on S_2 . We have

$$\begin{aligned} \iint_{S_2} F \cdot d\mathbf{S} &= \iint_{S_2} (r \cos(\theta), r \sin(\theta), r^2) \cdot (2r^2 \cos(\theta), 2r^2 \sin(\theta), -r) dS \\ &= \iint_{S_2} r^3 dS = \int_0^{2\pi} \int_0^2 r^3 dr d\theta = 2\pi \left(\frac{r^4}{4} \right) \Big|_{r=0}^2 = 8\pi. \end{aligned}$$

We therefore have that $\iint_S F \cdot d\mathbf{S} = 80\pi\sqrt{5} - 160\pi + 8\pi = 8\pi(10\sqrt{5} - 19)$. ♡

6.6.2 Green's, Stokes' and Divergence Theorems

Question 6.24

Using two different methods, evaluate the line integral $\int_C F \cdot d\mathbf{s}$ where $F(x, y) = (y, x)$ and C is the closed curve defined by the parametric representation $\sigma(t) = (t^2 - 1, t^3 - t)$ for $-1 \leq t \leq 1$. We assume that the parametric representation is consistent with the orientation on C .

Solution

i) Since $\sigma'(t) = (2t, 3t^2 - 1)$, the line integral is

$$\begin{aligned} \int_C F \cdot d\mathbf{s} &= \int_{-1}^1 (t^3 - t, t^2 - 1) \cdot (2t, 3t^2 - 1) dt \\ &= \int_{-1}^1 (5t^4 - 6t^2 + 1) dt = (t^5 - 2t^3 + t) \Big|_{t=-1}^1 = 0. \end{aligned}$$

ii) Suppose that D is the region enclosed by C . Note that C is a closed curve like a circle. To justify this claim, consider $t^3 - t = t(t^2 - 1)$. We have that $x = t^2 - 1$ decreases from 0 to -1 with $y = t^3 - t > 0$ when t goes from -1 to 0, and $x = t^2 - 1$ increases from -1 to 0 with $y = t^3 - t < 0$ when t goes from 0 to 1.

We select the orientation on C such that D is to the left when travelling along C . Using Green's Theorem, we get

$$\epsilon \int_C F \cdot d\mathbf{s} = \iint_D \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dx dy = \iint_D (1 - 1) dx dy = 0$$

where $\epsilon = 1$ if the orientation on C used for the Green's theorem is the orientation on C in the statement of the question or -1 if it is the opposite orientation. For this question, the value of ϵ does not matter because the value of the integral is null. ♡

Question 6.25

Compute $\int_C F \cdot d\mathbf{s}$ using Green's Theorem in each of the following cases.

- a) $F(x, y) = (x - y, x + y)$ and C is the unit circle oriented clockwise.
b) $F(x, y) = (y^2, -2x)$ and C is the triangle with vertices $(0, 0)$, $(1, 0)$ and $(1, 1)$ oriented counterclockwise.
c) $F(x, y) = (x^2 + 10xy + y^2, 5x^2 + 5xy)$ and C is the square with vertices $(0, 0)$, $(2, 0)$, $(0, 2)$ and $(2, 2)$ oriented counterclockwise.
d) $F(x, y) = (3x^2 \sin(y^2), 2x^3 y \cos(y^2))$ and C is the boundary of a bounded region D with a continuously differentiable boundary where the orientation on C is the positive orientation with respect to D .

Solution

- a) Since clockwise is not the positive orientation on the boundary of the disk $D = \{(x, y) : x^2 + y^2 \leq 1\}$, we have

$$\int_C F \cdot d\mathbf{s} = - \int_{\partial D} F \cdot d\mathbf{s} = - \iint_D \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dA = -2 \iint_D dA = -2\pi .$$

- b) We have that $C = \partial D$ where $D = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq x\}$. Moreover C has the positive orientation with respect to D . Thus

$$\begin{aligned} \int_C F \cdot d\mathbf{s} &= \iint_D \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dA = \iint_D (-2 - 2y) dA = -2 \int_0^1 \int_0^x (1 + y) dy dx \\ &= -2 \int_0^1 \left(y + \frac{y^2}{2} \right) \Big|_{y=0}^x dx = -2 \int_0^1 \left(x + \frac{x^2}{2} \right) dx = -2 \left(\frac{x^2}{2} + \frac{x^3}{6} \right) \Big|_{x=0}^1 = -\frac{4}{3} . \end{aligned}$$

- c) We have that $C = \partial D$ where $D = \{(x, y) : 0 \leq x, y \leq 2\}$. Moreover C has the positive orientation with respect to D . Thus

$$\begin{aligned} \int_C F \cdot d\mathbf{s} &= \iint_D \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dA = \iint_D (10x + 5y - 10x - 2y) dA \\ &= 3 \int_0^2 \int_0^2 y dy dx = 3 \int_0^2 \frac{y^2}{2} \Big|_{y=0}^2 dx = 6 \int_0^2 dx = 12 . \end{aligned}$$

- d) We have that $C = \partial D$ with the positive orientation with respect to D . Thus

$$\int_C F \cdot d\mathbf{s} = \iint_D \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dA = \iint_D 0 dA = 0 .$$

♡

Question 6.26

Let D be the annulus $\{(x, y) : 1 \leq x^2 + y^2 \leq 4\}$. Compute $\int_{\partial D} xy^2 dy - x^2 y dx$ where ∂D has the positive orientation with respect to D .

Solution

We have $\int_{\partial D} F \cdot d\mathbf{s}$ where $F(x, y) = (-x^2 y, xy^2)$. Thus

$$\int_{\partial D} xy^2 dy - x^2 y dx = \int_{\partial D} F \cdot d\mathbf{s} = \iint_D \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dA = \iint_D (y^2 + x^2) dA .$$

To compute the double integral, we use the change of variables $x = r \cos(\theta)$ and $y = r \sin(\theta)$ for $1 \leq r \leq 2$ and $0 \leq \theta \leq 2\pi$. Thus $\left| \frac{\partial(x, y)}{\partial(r, \theta)} \right| = r$ and

$$\iint_D (y^2 + x^2) \, dA = \int_0^{2\pi} \int_1^2 r^3 \, dr \, d\theta = 2\pi \frac{r^4}{4} \Big|_{r=1}^2 = 8\pi - \frac{\pi}{2}.$$

♥

Question 6.27

Find the simple closed curve C of class C^1 that maximize the line integral $I_C = \int_C F \cdot d\mathbf{s}$ where $F(x, y) = (y^3, 3x - x^3)$.

Solution

Suppose that $C = \partial D$ is the boundary of a bounded region D with a boundary of class C^1 . Moreover, suppose that the orientation on C is the positive orientation with respect to D . We have from Green's Theorem that

$$I_C = \int_C F \cdot d\mathbf{s} = \iint_D \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) \, dA = 3 \iint_D (1 - x^2 - y^2) \, dA.$$

The maximum of the double integral is reached for $D = \{(x, y) : x^2 + y^2 \leq 1\}$, the largest region on which $1 - x^2 - y^2$ is non-negative. Thus C is the unit circle with the counterclockwise orientation, the positive orientation of ∂D with respect to D . ♥

Question 6.28

Let $f : [a, b] \rightarrow [0, \infty[$ be a continuously differentiable function. Show that the area under the graph of f for $a \leq x \leq b$ is given by $-\int_{\partial S} y \, dx$ where $S = \{(x, y) : a \leq x \leq b \text{ and } 0 \leq y \leq f(x)\}$ and the orientation on ∂S is counterclockwise.

Solution

We show that $\int_a^b f(x) \, dx = -\int_{\partial S} y \, dx$. We have that $\int_{\partial S} y \, dx = \sum_{i=1}^4 \int_{C_i} y \, dx$ where C_1 is the straight line with parametric representation $g_1(t) = (t, 0)$ for $a \leq t \leq b$, C_2 is the straight line with parametric representation $g_2(t) = (b, t)$ for $0 \leq t \leq f(b)$, C_3 is the curve with parametric representation $g_3(t) = (b + t(a - b), f(b + t(a - b)))$ for $0 \leq t \leq 1$, and C_4 is the straight line with parametric representation $g_4(t) = (a, f(a) - t)$ for $0 \leq t \leq f(a)$. All parametric representations are respecting the positive orientation on ∂S with respect to S .

We can easily see that $\int_{C_i} y \, dx = 0$ for $i = 1, 2$ and 4 because either x is constant or $y = 0$.

Moreover

$$\int_{C_3} y \, dx = \int_0^1 f(b + t(a - b))(a - b) \, dt = \int_b^a f(x) \, dx = -\int_a^b f(x) \, dx,$$

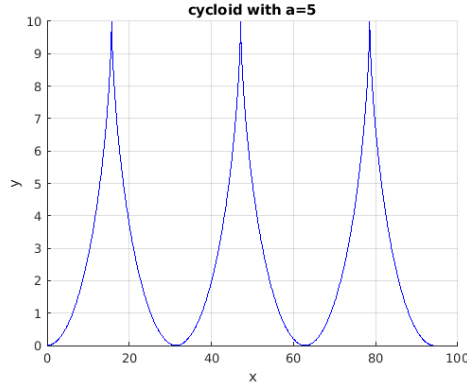
where we have used the substitution $x = b + t(a - b)$. Thus $\int_{\partial S} y \, dx = -\int_a^b f(x) \, dx$ as required. ♥

Question 6.29

Compute the area under one arch of the cycloid given by $x = a(u - \sin(u))$ and $y = a(1 - \cos(u))$.

Solution

Since $\frac{dx}{du} = a(1 - \cos(u)) > 0$ for $u \neq 2n\pi$ with $n \in \mathbb{Z}$, we have that x is strictly increasing. Moreover, $y \geq 0$ for all u and $y = 0$ only for $u = 2n\pi$ with $n \in \mathbb{Z}$. Therefore, there is an arch for each interval of the form $2n\pi \leq u \leq 2(n+1)\pi$ with $n \in \mathbb{Z}$.



The area under such an arch is the area of the region D enclosed by the horizontal line C_1 from $(2n\pi a, 0)$ to $(2(n+1)\pi a, 0)$ and the section C_2 of the cycloid given by $\{(a(u - \sin(u)), a(1 - \cos(u))) : 2n\pi \leq u \leq 2(n+1)\pi\}$.

On C_1 , we use the parametric representation $g_1(u) = (2\pi a u, 0)$ for $n \leq u \leq n+1$. On C_2 , we use the parametric representation $g_2(u) = (a(2\pi u - \sin(2\pi u)), a(1 - \cos(2\pi u)))$ for $n \leq u \leq n+1$. We note that this last parametric representation is not consistent with the positive orientation on ∂D .

Using Green's Theorem with $F(x, y) = (0, x)$, we find that the area of the arch is given by

$$\iint_D dA = \int_{\partial D} x dy = \int_{C_1} x dy - \int_{C_2} x dy .$$

However $\int_{C_1} x dy = 0$ and

$$\begin{aligned} \int_{C_2} x dy &= 2\pi a^2 \int_n^{n+1} (2\pi u - \sin(2\pi u)) \sin(2\pi u) du \\ &= (2\pi a)^2 \int_n^{n+1} u \sin(2\pi u) du - 2\pi a^2 \int_n^{n+1} \sin^2(2\pi u) du \\ &= (2\pi a)^2 \int_n^{n+1} u \sin(2\pi u) du - \pi a^2 \int_n^{n+1} (1 - \cos(4\pi u)) du \\ &= (2\pi a)^2 \left(\frac{\sin(2\pi u)}{(2\pi)^2} - \frac{u \cos(2\pi u)}{2\pi} \right) \Big|_{u=n}^{n+1} - \pi a^2 \left(u - \frac{\sin(4\pi u)}{4\pi} \right) \Big|_{u=n}^{n+1} \end{aligned}$$

$$= -2\pi a^2 - \pi a^2 = -3\pi a^2 .$$

Thus $\iint_D dA = 3\pi a^2$.

We could have used $n = 0$ in the computation above because each arch has the same shape and area due to the 2π periodicity of the cosine and sine functions. $\heartsuit$

Question 6.30

For each of the following vector fields F and closed surfaces S , compute the integral $\iint_S F \cdot d\mathbf{S}$ using the divergence theorem where the positive orientation on S is given by the unit normal pointing outside the region enclosed by S .

- a) $F(x, y, z) = (x^2, z, -y)$ and S is the sphere of radius 1 centred at the origin.
- b) $F(x, y, z) = (x, y, z)$ and S is the boundary of the region $R = \{x^2 + y^2 \leq z \leq \sqrt{20 - x^2 - y^2}\}$.
- c) $F(x, y, z) = (x^2, y^2, z^2)$ and S is the surface of the cube $R = \{(x, y, z) : a \leq x, y, z \leq b\}$.
- d) $F(x, y, z) = \left(\frac{x}{a^2}, \frac{y}{b^2}, \frac{z}{c^2}\right)$ and S is the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$.
- e) $F(x, y, z) = (x^2, -2xy, z^2)$ and S is the boundary of the cylinder $R = \{(x, y, z) : (x, y) \in D \text{ and } a \leq z \leq b\}$, where $D \subset \mathbb{R}^2$ is a surface with a boundary of class C^1 and an area that is equal to A .

Solution

a) According to the divergence theorem, we have

$$\iint_S F \cdot d\mathbf{S} = \iiint_R \operatorname{div} F \, dV = \iiint_R 2x \, dV$$

where $R = \{(x, y, z) : x^2 + y^2 + z^2 \leq 1\}$ is the ball of radius 1 centred at the origin and $S = \partial R$. To compute the triple integral, we note that $x \mapsto 2x$ is a symmetric function with respect to the y, z plane and that the integral is computed over a symmetric domain with respect to the y, z plane; namely, the ball R . Thus $\iiint_R 2x \, dV = 0$.

b) A sketch of the region R and additional details about this region can be found in the solution of Question 6.23 (e). We use the information in that question to answer the present question. According to the divergence theorem, we have

$$\iint_S F \cdot d\mathbf{S} = \iiint_R \operatorname{div} F \, dV = \iiint_R 3 \, dV ,$$

To compute this triple integral, we use the cylindrical coordinates $x = r \cos(\theta)$, $y = r \sin(\theta)$ and $z = z$ for $0 \leq r \leq 2$, $0 \leq \theta \leq 2\pi$ and $r^2 \leq z \leq \sqrt{20 - r^2}$. Since $\left| \frac{\partial(x, y, z)}{\partial(r, \theta, z)} \right| = r$, we get

$$\begin{aligned} \iiint_R 3 \, dV &= 3 \int_0^{2\pi} \int_0^2 \int_{r^2}^{\sqrt{20-r^2}} r \, dz \, dr \, d\theta = 6\pi \int_0^2 (rz) \Big|_{z=r^2}^{\sqrt{20-r^2}} dr \\ &= 6\pi \int_0^2 \left(r\sqrt{20-r^2} - r^3 \right) dr = 6\pi \left(-\frac{(20-r^2)^{3/2}}{3} - \frac{r^4}{4} \right) \Big|_{r=0}^2 = 8\pi(10\sqrt{5} - 19) . \end{aligned}$$

c) According to the divergence theorem, we have

$$\begin{aligned}\iint_S F \cdot d\mathbf{S} &= \iiint_R \operatorname{div} F \, dV = \iiint_R (2x + 2y + 2z) \, dV \\ &= \int_a^b \int_a^b \int_a^b (2x + 2y + 2z) \, dz \, dy \, dx = 3 \int_a^b \int_a^b \int_a^b 2x \, dz \, dy \, dx \\ &= 3(b-a)^2 \int_a^b 2x \, dx = 3(b-a)^2 x^2 \Big|_{x=a}^b = 3(b-a)^2(b^2 - a^2) .\end{aligned}$$

d) According to the divergence theorem, we have

$$\iint_S F \cdot d\mathbf{S} = \iiint_R \operatorname{div} F \, dV = \iiint_R \left(\frac{1}{a^2} + \frac{1}{b^2} + \frac{1}{c^2} \right) dV = \left(\frac{1}{a^2} + \frac{1}{b^2} + \frac{1}{c^2} \right) \iiint_R dV ,$$

where $R = \left\{ (x, y, z) : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1 \right\}$. Let us assume that we do not know the formula to compute the volume of an ellipsoid. To compute the integral $\iiint_R dV$, we first reduce it to an integral over a sphere of radius 1 centred at the origin. We use the change of variables $x = au$, $y = bv$ and $z = cw$ for $u^2 + v^2 + w^2 \leq 1$. Since $\left| \frac{\partial(x, y, z)}{\partial(u, v, w)} \right| = abc$, we get

$$\iiint_R dV = abc \iiint_{x^2+y^2+z^2 \leq 1} dV = \frac{4\pi}{3} abc ,$$

where we have used the well known formula $V = 4\pi r^3/3$ to compute the volume of a ball of radius r . Thus

$$\iint_S F \cdot d\mathbf{S} = \frac{4\pi}{3} abc \left(\frac{1}{a^2} + \frac{1}{b^2} + \frac{1}{c^2} \right) .$$

e) According to the divergence theorem, we have

$$\begin{aligned}\iint_S F \cdot d\mathbf{S} &= \iiint_R \operatorname{div} F \, dV = \iiint_R 2z \, dV = \iint_W \left(\int_a^b 2z \, dz \right) dA \\ &= \iint_W \left(z^2 \Big|_{z=a}^b \right) dA = (b^2 - a^2) \iint_W dA = (b^2 - a^2)A .\end{aligned}$$

♡

Question 6.31

Let F be the vector field $F(x, y, z) = (x^2 + y^2 + z^2)(x, y, z)$ and S be the sphere of radius a centred at the origin. Suppose that the orientation on S is given by the unit normal pointing away from the origin. Compute $\iint_S F \cdot d\mathbf{S}$ with and without the divergence theorem.

Solution

i) To compute the integral directly, we use the parametric representation of S given by

$g(\phi, \theta) = (a \cos(\theta) \sin(\phi), a \sin(\theta) \sin(\phi), a \cos(\phi))$ for $0 \leq \theta \leq 2\pi$ and $0 \leq \phi \leq \pi$. We have

$$\frac{\partial g}{\partial \phi} \times \frac{\partial g}{\partial \theta} = (a^2 \cos(\theta) \sin^2(\phi), a^2 \sin(\theta) \sin^2(\phi), a^2 \cos(\phi) \sin(\phi)) .$$

We note that this parametric representation is consistent with the selected orientation on S ; namely, $\frac{\partial g}{\partial \phi} \times \frac{\partial g}{\partial \theta}$ points outside the ball R of radius a centred at the origin. Thus

$$\begin{aligned} \iint_S F \cdot d\mathbf{S} &= \int_0^{2\pi} \int_0^\pi (a^3 \cos(\theta) \sin(\phi), a^3 \sin(\theta) \sin(\phi), a^3 \cos(\phi)) \\ &\quad \cdot (a^2 \cos(\theta) \sin^2(\phi), a^2 \sin(\theta) \sin^2(\phi), a^2 \cos(\phi) \sin(\phi)) d\phi d\theta \\ &= a^5 \int_0^{2\pi} \int_0^\pi \sin(\phi) d\phi d\theta = 4\pi a^5 . \end{aligned}$$

ii) According to the divergence theorem, we have

$$\iint_S F \cdot d\mathbf{S} = \iiint_R \operatorname{div} F dV = \iiint_R 5(x^2 + y^2 + z^2) dV$$

where R is the ball of radius a centred at the origin. To compute the triple integral, we use the spherical coordinates $x = r \cos(\theta) \sin(\phi)$, $y = r \sin(\theta) \sin(\phi)$ and $z = r \cos(\phi)$ for $0 \leq r \leq a$, $0 \leq \theta \leq 2\pi$ and $0 \leq \phi \leq \pi$. Since $\left| \frac{\partial(x, y, z)}{\partial(r, \theta, \phi)} \right| = r^2 \sin(\phi)$, we get

$$\begin{aligned} \iiint_R 5(x^2 + y^2 + z^2) dV &= \int_0^{2\pi} \int_0^\pi \int_0^a 5r^4 \sin(\phi) dr d\phi d\theta \\ &= 2\pi \int_0^\pi \left(r^5 \Big|_{r=0}^a \right) \sin(\phi) d\phi = 2\pi a^5 \int_0^\pi \sin(\phi) d\phi = 4\pi a^5 . \end{aligned}$$

♡

Question 6.32

Consider the vector field $F(x, y, z) = (x^3 + y \sin(z), y^3 + z \sin(x), 2z)$ and the solid $R = \{(x, y, z) : 1 \leq x^2 + y^2 + z^2 \leq 9 \text{ and } z \geq 0\}$. Compute $\iint_{\partial R} F \cdot d\mathbf{S}$ where the positive orientation on ∂R is given by the unit normal pointing outside R .

Solution

From the divergence theorem, we have that $\iint_{\partial R} F \cdot d\mathbf{S} = \iiint_R \operatorname{div} F dx dy dz$ where $\operatorname{div} F = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z} = 3x^2 + 3y^2 + 2$. We use the spherical coordinates to compute the triple integral. The region R is the image of $g(\theta, \phi, r) = (r \cos(\theta) \sin(\phi), r \sin(\theta) \sin(\phi), r \cos(\phi))$ for $1 \leq r \leq 3$, $0 \leq \theta \leq 2\pi$ and $0 \leq \phi \leq \pi/2$. Since $|\det Dg(\theta, \phi, r)| = r^2 \sin(\phi)$, we get

$$\iiint_R \operatorname{div} F dx dy dz$$

$$\begin{aligned}
&= \int_0^{2\pi} \int_0^{\pi/2} \int_1^3 (3r^2 \cos^2(\theta) \sin^2(\phi) + 3r^2 \sin^2(\theta) \sin^2(\phi) + 2) r^2 \sin(\phi) \, dr \, d\phi \, d\theta \\
&= \int_0^{2\pi} \int_0^{\pi/2} \int_1^3 (3r^2 \sin^2(\phi) + 2) r^2 \sin(\phi) \, dr \, d\phi \, d\theta \\
&= \int_0^{2\pi} \int_0^{\pi/2} \int_1^3 (3r^4(1 - \cos^2(\phi)) \sin(\phi) + 2r^2 \sin(\phi)) \, dr \, d\phi \, d\theta \\
&= 2\pi \left(\frac{3r^5}{5} \right) \Big|_{r=1}^3 \left(-\cos(\phi) + \frac{1}{3} \cos^3(\phi) \right) \Big|_{\phi=0}^{\pi/2} + 2\pi \left(\frac{2r^3}{3} \right) \Big|_{r=1}^3 (-\cos(\phi)) \Big|_{\phi=0}^{\pi/2} \\
&= \frac{3432\pi}{15} .
\end{aligned}$$

♡

Question 6.33

Consider the vector field $F(x, y, z) = (x^2y, xy^2, 2xyz)$. Let R be the tetrahedron with vertices at $(0, 0, 0)$, $(2, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 2)$. Compute $\iint_{\partial R} F \cdot d\mathbf{S}$ where the orientation on ∂R is given by the unit normal pointing outside R .

Question 6.34

Consider the vector Field $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ defined by $F(x, y, z) = (xz^2, z, x^2z)$. Compute the integral $\iint_D F \cdot d\mathbf{S}$ where D is the surface of the solid $R = \{(x, y, z) : 1 \leq x^2 + z^2 \leq 4 \text{ and } -1 \leq y \leq 1\}$ and the orientation on D is given by the unit normal pointing outside of R .

Solution

From the divergence theorem, we have that $\iint_D F \cdot d\mathbf{S} = \int_R \operatorname{div} F$ where $\operatorname{div} F = z^2 + x^2$. To compute the integral on R , we use the cylindrical coordinates; namely, $g(r, \theta, y) = (r \cos(\theta), y, r \sin(\theta))$ for $-1 \leq y \leq 1$, $0 \leq \theta \leq 2\pi$ and $1 \leq r \leq 2$. Since

$$|\det Dg| = \left| \det \begin{pmatrix} \cos(\theta) & -r \sin(\theta) & 0 \\ 0 & 0 & 1 \\ \sin(\theta) & r \cos(\theta) & 0 \end{pmatrix} \right| = |-r| = r ,$$

we get

$$\iiint_R \operatorname{div} F = \int_0^{2\pi} \int_{-1}^1 \int_1^2 r^3 \, dr \, dy \, d\theta = 4\pi \frac{r^4}{4} \Big|_{r=1}^2 = 15\pi .$$

♡

Question 6.35

Let R be a bounded region in $\mathbb{R}^3$ with a piecewise continuously differentiable boundary ∂R . Show that the volume of R is given by $\frac{1}{3} \iint_{\partial R} F \cdot d\mathbf{S}$ where $F(x, y, z) = (x, y, z)$ and the orientation on ∂R is given by the unit normal to ∂R pointing outside R .

Solution

Since $\operatorname{div} F = 3$, we get from the divergence theorem that

$$\iint_{\partial R} F \cdot d\mathbf{S} = \iiint_R \operatorname{div} F \, dV = 3 \iiint_R dV$$

where $\iiint_R dV$ is the volume of R .

♡

Question 6.36

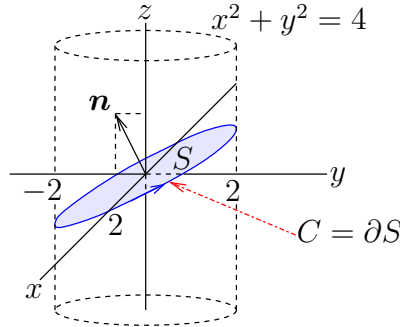
Use Stokes' theorem to evaluate $\int_C F \cdot d\mathbf{s}$ in each of the following cases.

a) $F(x, y, z) = (x - z, x + y, y + z)$ and C is the ellipse resulting from the intersection of the plane $z = my + b$ with the cylinder $x^2 + y^2 = 4$. The orientation on C is counterclockwise when viewed from the positive z axis.

b) $F(x, y, z) = (y, y^2, x + 2z)$ and C is the curve resulting from the intersection of the plane $z = a - y$ with the sphere $x^2 + y^2 + z^2 = a^2$. The orientation on C is counterclockwise when viewed from the positive z axis.

Solution

a) The ellipse C is the boundary of the region S of the plane $z = my + b$ that is enclosed by C . For C to be positively oriented with respect to S , the orientation on S must be given by the unit normal $\mathbf{n} = (0, -m/\sqrt{1+m^2}, 1/\sqrt{1+m^2})$. This is shown in the following figure for $m > 0$ and $b = 0$.



According to Stokes' theorem, we have that $\int_C F \cdot d\mathbf{s} = \iint_S \operatorname{curl} F \cdot d\mathbf{S}$ where

$$\operatorname{curl} F = \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ x - z & x + y & y + z \end{pmatrix} = (1, -1, 1).$$

The parametric representation for S that we use is given by

$$g(r, \theta) = (r \cos(\theta), r \sin(\theta), b + mr \sin(\theta))$$

for $0 \leq r \leq 2$ and $0 \leq \theta \leq 2\pi$. We have

$$\frac{\partial g}{\partial r} \times \frac{\partial g}{\partial \theta} = \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \cos(\theta) & \sin(\theta) & m \sin(\theta) \\ -r \sin(\theta) & r \cos(\theta) & mr \cos(\theta) \end{pmatrix} = -mr \mathbf{e}_2 + r \mathbf{e}_3.$$

Thus, this parametric representation is consistent with the orientation on S since $(0, -mr, r)$ points in the same direction as $\mathbf{n}$. Hence

$$\begin{aligned} \iint_S \operatorname{curl} F \cdot d\mathbf{S} &= \int_0^{2\pi} \int_0^2 (1, -1, 1) \cdot (0, -mr, r) dr d\theta = \int_0^{2\pi} \int_0^2 (1+m)r dr d\theta \\ &= 2\pi \left. \frac{(1+m)r^2}{2} \right|_0^2 = 4(1+m)\pi . \end{aligned}$$

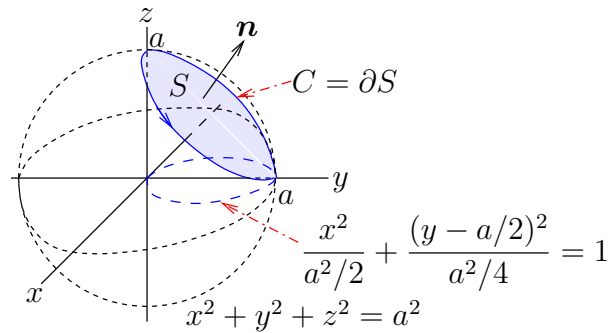
The answer does not depend on b .

b) To find the intersection of the sphere and the plane $z = a - y$, we solve the following equation.

$$\begin{aligned} x^2 + y^2 + (a - y)^2 &= a^2 \Rightarrow x^2 + 2y^2 - 2ay = 0 \Rightarrow x^2 + 2\left(y - \frac{a}{2}\right)^2 = \frac{a^2}{2} \\ &\Rightarrow \frac{x^2}{a^2/2} + \frac{(y - a/2)^2}{a^2/4} = 1 . \end{aligned}$$

This is an ellipse centred at $(0, a/2)$ with major axis parallel to the y axis and minor axis parallel to the x axis.

Let S be the region S of the plane $z = a - y$ that is enclosed by C . For C to be positively oriented with respect to S , the orientation on S must be given by the unit normal $\mathbf{n} = (0, 1/\sqrt{2}, 1/\sqrt{2})$. This is shown in the following figure.



According to Stokes' theorem, we have that $\int_C F \cdot d\mathbf{s} = \iint_S \operatorname{curl} F \cdot d\mathbf{S}$ where

$$\operatorname{curl} F = \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ y & y^2 & x + 2z \end{pmatrix} = (0, -1, -1) .$$

We use the parametric representation for S given by

$$g(r, \theta) = \left(\frac{ar}{\sqrt{2}} \cos(\theta), \frac{a}{2} + \frac{ar}{2} \sin(\theta), \frac{a}{2} - \frac{ar}{2} \sin(\theta) \right)$$

for $0 \leq r \leq 1$ and $0 \leq \theta \leq 2\pi$. We have

$$\frac{\partial g}{\partial r} \times \frac{\partial g}{\partial \theta} = \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ (a/\sqrt{2})\cos(\theta) & (a/2)\sin(\theta) & -(a/2)\sin(\theta) \\ -(a/\sqrt{2})\sin(\theta) & (a/2)\cos(\theta) & -(a/2)\cos(\theta) \end{pmatrix} = \left(0, \frac{a^2 r}{2\sqrt{2}}, \frac{a^2 r}{2\sqrt{2}}\right).$$

Therefore, the parametric representation is consistent with the orientation on S because $\left(0, \frac{a^2 r}{2\sqrt{2}}, \frac{a^2 r}{2\sqrt{2}}\right)$ points in the same direction as $\mathbf{n}$. Hence

$$\begin{aligned} \iint_S \operatorname{curl} F \cdot d\mathbf{S} &= \int_0^{2\pi} \int_0^1 (0, -1, -1) \cdot \left(0, \frac{a^2 r}{2\sqrt{2}}, \frac{a^2 r}{2\sqrt{2}}\right) dr d\theta = \int_0^{2\pi} \int_0^1 \frac{-a^2 r}{\sqrt{2}} dr d\theta \\ &= 2\pi \left. \frac{-a^2 r^2}{2\sqrt{2}} \right|_0^1 = -\frac{a^2 \pi}{\sqrt{2}}. \end{aligned}$$

♡

Question 6.37

Compute $\iint_S \operatorname{curl} F \cdot d\mathbf{S}$ where $F(x, y, z) = (y, x - 2x^3z, xy^3)$ and S is the upper half of the sphere of radius a centred at the origin. The orientation on S is given by a normal pointing away from the origin.

Solution

We could obviously compute directly the integral but it may require less work to use Stokes' theorem; namely, $\iint_S \operatorname{curl} F \cdot d\mathbf{S} = \int_{\partial S} F \cdot d\mathbf{s}$ where ∂S is the circle of radius a centred at the origin and contained in the plane $z = 0$. The positive orientation on ∂S with respect to S is counterclockwise.

A parametric representation of ∂S is given by $g(\theta) = (a \cos(\theta), a \sin(\theta), 0)$ for $0 \leq \theta < 2\pi$. It is consistent with the orientation on ∂S . Thus

$$\begin{aligned} \int_{\partial S} F \cdot d\mathbf{s} &= \int_0^{2\pi} (a \sin(\theta), a \cos(\theta), a^4 \cos(\theta) \sin^3(\theta)) \cdot (-a \sin(\theta), a \cos(\theta), 0) d\theta \\ &= a^2 \int_0^{2\pi} (-\sin^2(\theta) + \cos^2(\theta)) d\theta = a^2 \int_0^{2\pi} \cos(2\theta) d\theta = \left. \frac{a^2}{2} \sin(2\theta) \right|_{\theta=0}^{2\pi} = 0 \end{aligned}$$

Note: Instead of computing the line integral $\int_{\partial S} F \cdot d\mathbf{s}$, we can use Stokes' theorem again to replace this integral by a surface integral. We have that $\int_{\partial S} F \cdot d\mathbf{s} = \iint_S \operatorname{curl} F \cdot d\mathbf{S}$ where D is the disk $D = \{(x, y, 0) : x^2 + y^2 \leq a^2\}$ and the orientation on D is the orientation given by the unit vector $\mathbf{e}_3$. A parametric representation of D is given by $g(r, \theta) = (r \cos(\theta), r \sin(\theta), 0)$ for $0 \leq r \leq a$ and $0 \leq \theta \leq 2\pi$. We have

$$\frac{\partial g}{\partial r} \times \frac{\partial g}{\partial \theta} = \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \cos(\theta) & \sin(\theta) & 0 \\ -r \sin(\theta) & r \cos(\theta) & 0 \end{pmatrix} = (0, 0, r).$$

Therefore, the parametric representation is consistent with the orientation on D . Since

$$\operatorname{curl} F = \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ y & x - 2x^3z & xy^3 \end{pmatrix} = (3xy^2 + 2x^3, -y^3, -6x^2z),$$

we get

$$\begin{aligned} \iint_D \operatorname{curl} F \cdot d\mathbf{S} &= \int_0^{2\pi} \int_0^a (3a^3 \cos(\theta) \sin^2(\theta) + 2a^3 \cos^3(\theta), -a^3 \sin^3(\theta), 0) \cdot (0, 0, r) dr d\theta \\ &= \int_0^{2\pi} \int_0^a 0 dr d\theta = 0. \end{aligned}$$

♡

Question 6.38

Let $F(x, y, z) = (-x^2y, x^3, x^2 + yz)$ and S be the lower half of the ellipsoid $\frac{x^2}{4} + \frac{y^2}{9} + \frac{z^2}{25} = 1$; namely, for $z \leq 0$. Compute the flow of $\operatorname{curl} F$ across the surface S if the orientation on S is given by the unit vector pointing in the direction of z positive.

Solution

To compute the flow of $\operatorname{curl} F$ across the surface S , we need to compute $\iint_S \operatorname{curl} F \cdot d\mathbf{S}$. We could obviously compute directly this integral but it may require less work to use Stokes' theorem; namely $\iint_S \operatorname{curl} F \cdot d\mathbf{S} = \int_{\partial S} F \cdot d\mathbf{s}$ where ∂S is the ellipse $\frac{x^2}{4} + \frac{y^2}{9} = 1$ contained in the plane $z = 0$. The positive orientation on ∂S is counterclockwise when view from the positive side of the z axis.

A parametric representation of ∂S is given by $g(\theta) = (2 \cos(\theta), 3 \sin(\theta), 0)$ for $0 \leq \theta < 2\pi$. It is consistent with the positive orientation on ∂S . Thus

$$\begin{aligned} \int_{\partial S} F \cdot d\mathbf{s} &= \int_0^{2\pi} (-12 \cos^2(\theta) \sin(\theta), 8 \cos^3(\theta), 4 \cos^2(\theta)) \cdot (-2 \sin(\theta), 3 \cos(\theta), 0) d\theta \\ &= \int_0^{2\pi} (24 \cos^2(\theta) \sin^2(\theta) + 24 \cos^4(\theta)) d\theta \\ &= 24 \int_0^{2\pi} \cos^2(\theta) d\theta = 12 \int_0^{2\pi} (1 + \cos(2\theta)) d\theta = 24\pi. \end{aligned}$$

Note: Instead of computing the line integral $\int_{\partial S} F \cdot d\mathbf{s}$, we may use Stokes' theorem again to replace this integral by a surface integral; namely, $\int_{\partial S} F \cdot d\mathbf{s} = \iint_D \operatorname{curl} F \cdot d\mathbf{S}$ where D is the region $D = \left\{ (x, y, 0) : \frac{x^2}{4} + \frac{y^2}{9} \leq 1 \right\}$ with the orientation given by the unit vector $\mathbf{e}_3$.

A parametric representation of D is given by $g(r, \theta) = (2r \cos(\theta), 3r \sin(\theta), 0)$ for $0 \leq r \leq 1$ and $0 \leq \theta \leq 2\pi$. We have

$$\frac{\partial g}{\partial r} \times \frac{\partial g}{\partial \theta} = \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ 2 \cos(\theta) & 3 \sin(\theta) & 0 \\ -2r \sin(\theta) & 3r \cos(\theta) & 0 \end{pmatrix} = (0, 0, 6r) .$$

Thus, the parametric representation is consistent with the orientation on D . Since

$$\operatorname{curl} F = \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 2xy & 3x & x^2 + z^2 \end{pmatrix} = (z, -2x, 4x^2) ,$$

we get

$$\begin{aligned} \iint_D \operatorname{curl} F \cdot d\mathbf{S} &= \int_0^{2\pi} \int_0^1 (0, -4r \cos(\theta), 16r^2 \cos^2(\theta)) \cdot (0, 0, 6r) dr d\theta \\ &= 96 \int_0^{2\pi} \int_0^1 r^3 \cos^2(\theta) dr d\theta = 24r^4 \Big|_{r=0}^1 \int_0^{2\pi} \cos^2(\theta) d\theta = 24\pi . \end{aligned}$$

♡

Question 6.39

Let $F(x, y, z) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2}, 0 \right)$ for $(x, y) \neq (0, 0)$.

- Show that $\operatorname{curl} F = \mathbf{0}$.
- Show that $\int_C F \cdot d\mathbf{s} = 2\pi$ where C is the circle $C = \{(x, y, b) : x^2 + y^2 = a^2\}$ oriented counterclockwise for $a, b \in \mathbb{R}$ arbitrary but fixed.
- Explain why (a) and (b) do not contradict Stokes' theorem.

Solution

a) We have

$$\operatorname{curl} F = \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{-y}{x^2 + y^2} & \frac{x}{x^2 + y^2} & 0 \end{pmatrix} = \left(0, 0, \frac{\partial}{\partial x} \left(\frac{x}{x^2 + y^2} \right) - \frac{\partial}{\partial y} \left(\frac{-y}{x^2 + y^2} \right) \right) = \mathbf{0}$$

b) A parametric representation of C is given by $g(\theta) = (a \cos(\theta), a \sin(\theta), b)$ for $0 \leq \theta < 2\pi$. It is consistent with the orientation on C . Thus

$$\begin{aligned} \int_C F \cdot d\mathbf{s} &= \int_0^{2\pi} \left(\frac{-\sin(\theta)}{a}, \frac{\cos(\theta)}{a}, 0 \right) \cdot (-a \sin(\theta), a \cos(\theta), 0) d\theta \\ &= \int_0^{2\pi} (\sin^2(\theta) + \cos^2(\theta)) d\theta = \int_0^{2\pi} d\theta = 2\pi . \end{aligned}$$

c) We cannot use Stokes' theorem because there is no surface $S \subset \mathbb{R}^3$ such that $C = \partial S$ and F is continuously differentiable on an open set containing S . The vector field F is not defined along the z axis and it cannot be continuously extended to the z axis. ♡

Question 6.40

Consider the vector field $F(x, y, z) = (xyz^3, xyz, \sin(xyz))$. Let R be the portion of the sphere $x^2 + y^2 + z^2 = 9$ between the planes $z = 0$ and $z = 2$. The orientation on R is given by the unit normal pointing away from the z axis. Compute the integral $\iint_R \text{curl } F \cdot d\mathbf{S}$.

Solution

According to Stokes' theorem, we have that $\iint_R \text{curl } F \cdot d\mathbf{S} = \int_{\partial R} F \cdot d\mathbf{s}$ where ∂R is the union of two circles: C_1 with the parametric representation $\sigma_1(\theta) = (3 \cos(\theta), 3 \sin(\theta), 0)$ and C_2 with the parametric representation $\sigma_2(\theta) = (\sqrt{5} \sin(\theta), \sqrt{5} \cos(\theta), 2)$ for $0 \leq \theta \leq 2\pi$. The positive orientation on C_1 with respect to R is counterclockwise when view from the positive side of the z axis, and the positive orientation on C_2 with respect to R is clockwise when view from the positive side of the z axis. The parametric representations are consistent with these orientations. Since $\sigma'_1(\theta) = (-3 \sin(\theta), 3 \cos(\theta), 0)$ and $\sigma'_2(\theta) = (\sqrt{5} \cos(\theta), -\sqrt{5} \sin(\theta), 0)$, we get

$$\begin{aligned} \iint_R \text{curl } F \cdot d\mathbf{S} &= \int_{C_1} F \cdot d\mathbf{s} + \int_{C_2} F \cdot d\mathbf{s} \\ &= \int_0^{2\pi} 0 \, d\theta + \int_0^{2\pi} \left(40\sqrt{5} \cos^2(\theta) \sin(\theta) - 10\sqrt{5} \cos(\theta) \sin^2(\theta) \right) d\theta \\ &= \left(-40\sqrt{5} \frac{\cos^3(\theta)}{3} - 10\sqrt{5} \frac{\sin^3(\theta)}{3} \right) \Big|_0^{2\pi} = 0. \end{aligned}$$

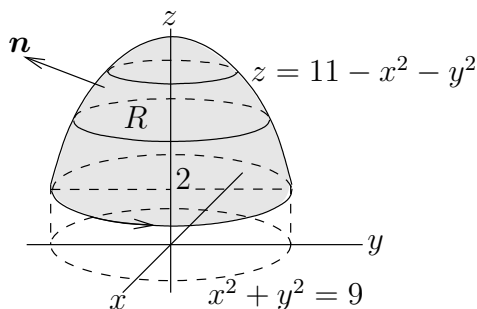
♡

Question 6.41

Consider the vector field $F(x, y, z) = (yz, xyz, \cos(xy)e^{xz})$. Let R be the portion of the paraboloid $z = 11 - x^2 - y^2$ above the plane $z = 2$. Compute $\int_{\partial R} F \cdot d\mathbf{s}$ where the orientation on ∂R is counterclockwise when view from the positive side of the z axis.

Solution

The region R is shown in the following figure.



According to Stokes' theorem, we have that $\int_{\partial R} F \cdot d\mathbf{s} = \int_R \text{curl } F \cdot d\mathbf{S}$ where we choose the orientation on R to be given by the unit normal pointing in the direction of z positive to ensure that the orientation on ∂R induced by R matches the orientation on ∂R stated in the question. Both integrals are terrible to compute. However, since the vector field $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is smooth, we may use any surface V whose boundary is ∂R . We take V to be the disk $x^2 + y^2 \leq 9$ with $z = 2$. A parametric representation of V is given by $g(r, \theta) = (r \cos(\theta), r \sin(\theta), 2)$ for $0 \leq \theta \leq 2\pi$ and $0 \leq r \leq 3$. We have that

$$\left(\frac{\partial g}{\partial r} \times \frac{\partial g}{\partial \theta} \right) = \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \cos(\theta) & \sin(\theta) & 0 \\ -r \sin(\theta) & r \cos(\theta) & 0 \end{pmatrix} = (0, 0, r)$$

points in the direction of z positive. This is the direction of the normal that we need on V to get that the positive orientation on ∂R with respect to V be counterclockwise when view from the positive side of the z axis. The parametric representation is therefore consistent with the orientation that we need on V . Moreover, because the unit normal to V is $\mathbf{e}_3$, we only need the third component of $\text{curl } F$, that is $(\text{curl } F)_3 = \left(\frac{\partial}{\partial x}(xyz) - \frac{\partial}{\partial y}(yz) \right) \mathbf{e}_3 = z(y - 1)\mathbf{e}_3$. Hence

$$\begin{aligned} \int_{\partial R} F \cdot d\mathbf{s} &= \int_V \text{curl } F \cdot d\mathbf{S} = \int_0^{2\pi} \int_0^3 ((\text{curl } F)_1, (\text{curl } F)_2, 2(r \sin(\theta) - 1)) \cdot (0, 0, r) dr d\theta \\ &= \int_0^{2\pi} \int_0^3 (2r^2 \sin(\theta) - 2r) dr d\theta = \left(\frac{2r^3}{3} \Big|_{r=0}^3 \right) \left(-\cos(\theta) \Big|_{\theta=0}^{2\pi} \right) - 2\pi r^2 \Big|_{r=0}^3 = -18\pi. \end{aligned}$$

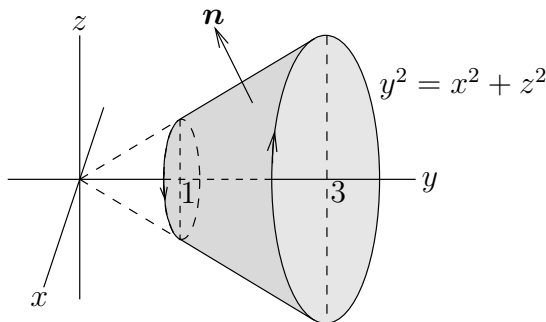
♥

Question 6.42

Consider the vector field $F(x, y, z) = (xy^3z, \sin(xyz), xyz)$. Let R be the portion of the cone $y^2 = x^2 + z^2$ between the planes $y = 1$ and $y = 3$. The orientation on R is given by the unit normal pointing away from the y axis. Compute the integral $\int_R \text{curl } F \cdot d\mathbf{S}$.

Solution

The region R is shown in the following figure.



We present two solutions.

i) According to Stokes' theorem, we have that $\int_R \text{curl } F \cdot d\mathbf{S} = \int_{\partial R} F \cdot d\mathbf{s}$ where ∂R is the union of two circles: C_1 with the parametric representation $\sigma_1(\theta) = (\cos(\theta), 1, -\sin(\theta))$ and C_3 with the parametric representation $\sigma_3(\theta) = (3\cos(\theta), 3, 3\sin(\theta))$ for $0 \leq \theta \leq 2\pi$. The positive orientation on the circle C_1 with respect to R is clockwise when viewed from the negative side of the y axis, and the positive orientation on the circle C_3 with respect to R is counterclockwise when viewed from the negative side of the y axis. The parametric representations are consistent with these orientations. Since $\sigma'_1(\theta) = (-\sin(\theta), 0, -\cos(\theta))$ and $\sigma'_3(\theta) = (-3\sin(\theta), 0, 3\cos(\theta))$, we get

$$\begin{aligned} \int_R \text{curl } F \cdot d\mathbf{S} &= \int_{C_3} F \cdot d\mathbf{s} + \int_{C_1} F \cdot d\mathbf{s} \\ &= \int_0^{2\pi} (-3^6 \cos(\theta) \sin^2(\theta) + 3^4 \cos^2(\theta) \sin(\theta)) d\theta + \int_0^{2\pi} (\cos(\theta) \sin^2(\theta) + \cos^2(\theta) \sin(\theta)) d\theta \\ &= (-3^5 \sin^3(\theta) - 3^3 \cos^3(\theta)) \Big|_0^{2\pi} + \left(\frac{\sin^3(\theta)}{3} - \frac{\cos^3(\theta)}{3} \right) \Big|_0^{2\pi} = 0. \end{aligned}$$

ii) For $\rho = 1$ and $\rho = 3$, let B_ρ be the disk in the plane $y = \rho$ of radius ρ centred at $(0, \rho, 0)$. If we keep the same orientation as before on $C_\rho = \partial B_\rho$ for $\rho = 1$ and 3 , and assume that the orientation on B_ρ is given by the unit normal $\mathbf{e}_2$, then

$$\int_{\partial R} F \cdot d\mathbf{s} = \int_{\partial B_1} F \cdot d\mathbf{s} + \int_{\partial B_3} F \cdot d\mathbf{s} = \int_{B_1} \text{curl } F \cdot d\mathbf{S} - \int_{B_3} \text{curl } F \cdot d\mathbf{S} \quad (6.6.1)$$

according to Stokes' theorem. The minus sign in front of the second integral comes from the fact that the orientation on C_3 is not the positive orientation on C_3 with respect to B_3 .

A parametric representation of B_ρ is given by $g_\rho(r, \theta) = (r \sin(\theta), \rho, r \cos(\theta))$ for $0 \leq \theta \leq 2\pi$ and $0 \leq r \leq \rho$. We have that

$$\left(\frac{\partial g_\rho}{\partial r} \times \frac{\partial g_\rho}{\partial \theta} \right) = \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \sin(\theta) & 0 & \cos(\theta) \\ r \cos(\theta) & 0 & -r \sin(\theta) \end{pmatrix} = (0, r, 0).$$

This vector is pointing in the direction of $\mathbf{e}_2$. Therefore, the parametric representation is consistent with the orientation on B_ρ . Since

$$\begin{aligned} \text{curl } F &= \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ xy^3z & \sin(xyz) & xyz \end{pmatrix} \\ &= (xz - xy \cos(xyz), -(yz - xy^3), yz \cos(xyz) - 3x^2y^2z), \end{aligned}$$

we get

$$\int_{B_\rho} \text{curl } F \cdot d\mathbf{S} = \int_0^{2\pi} \int_0^\rho (\rho r \cos(\theta) - \rho^3 r \sin(\theta)) r dr d\theta = 0$$

for $\rho = 1$ and $\rho = 3$. It follows from (6.6.1) that $\int_{\partial R} F \cdot d\mathbf{s} = 0$.

Note: A parametric representation of the cone is given by $g(r, \theta) = (r \cos(\theta), r, r \sin(\theta))$ for $0 \leq \theta \leq 2\pi$ and $1 \leq r \leq 3$. We have

$$\left(\frac{\partial g}{\partial r} \times \frac{\partial g}{\partial \theta} \right) = \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \cos(\theta) & 1 & \sin(\theta) \\ -r \sin(\theta) & 0 & r \cos(\theta) \end{pmatrix} = (r \cos(\theta), -r, r \sin(\theta)).$$

Since this vector points away from the y axis, the parametric representation is consistent with the orientation on R . Hence, $\int_R \text{curl } F \cdot d\mathbf{S} = \int_0^{2\pi} \int_1^3 \dots dr d\theta$ is an integral of a function involving expressions of the form $\cos(r^3 \cos(\theta) \sin(\theta))$ which cannot be integrated with standard methods. $\heartsuit$

Question 6.43

Let $F : \mathbb{R}^2 \setminus \{\mathbf{0}\} \rightarrow \mathbb{R}^2$ be a continuously differentiable vector fields such that $\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} = 0$ on $\mathbb{R}^2 \setminus \{\mathbf{0}\}$.

a) Prove that $\int_{C_r} F \cdot d\mathbf{s} = \alpha$, a constant, for all circles C_r of radius r centred at the origin if we assume that the orientation on C_r is counterclockwise.

b) Expand (a) by showing that $\int_C F \cdot d\mathbf{s} = \alpha$ for all closed curves C that surround the origin if we assume that the orientation on C is “globally” counterclockwise. We assume here that the closed curves C can be continuously deformed into a circle without intersecting the origin.

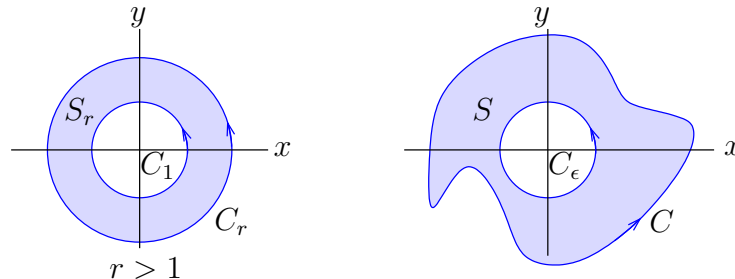
c) Let $F_0(x, y) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right)$. Prove that $G = F - \frac{\alpha}{2\pi} F_0$ is conservative on $\mathbb{R}^2 \setminus \{\mathbf{0}\}$ and thus is the gradient of a function. Such a result is fundamental in the studies of Green’s Functions.

Solution

Let $C_r = \{(x, y) : x^2 + y^2 = r^2\}$ and

$$S_r = \begin{cases} \{(x, y) : 1 \leq x^2 + y^2 \leq r^2\} & \text{if } r > 1 \\ \{(x, y) : r^2 \leq x^2 + y^2 \leq 1\} & \text{if } 0 < r < 1 \end{cases}$$

The orientation on ∂S_r is counterclockwise.



a) Using Green's Theorem, we get

$$\int_{C_r} F \cdot d\mathbf{s} - \int_{C_1} F \cdot d\mathbf{s} = \epsilon \int_{\partial S_r} F \cdot d\mathbf{s} = \epsilon \iint_{S_r} \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dA = 0$$

where $\epsilon = 1$ for $r > 1$ and $\epsilon = -1$ for $0 < r < 1$. Note that the orientation on C_1 is not consistent with the orientation on ∂S_r if $r < 1$ while it is the orientation on C_r that is not consistent with the orientation on ∂S_r if $r < 1$. Thus $\int_{C_r} F \cdot d\mathbf{s} = \int_{C_1} F \cdot d\mathbf{s}$ for all r . If we

let $\alpha = \int_{C_1} F \cdot d\mathbf{s}$, then $\int_{C_r} F \cdot d\mathbf{s} = \alpha$ for all $r > 0$.

b) Let C be any closed curve surrounding the origin as described in (b). Since the origin is an interior point of the region enclosed by the curve C , there exists $\epsilon > 0$ such that C_ϵ is included in the interior of the region enclosed by the curve C . See the figure above. If S is the region enclosed by C and the circle C_ϵ , then we get from Green's Theorem that

$$\int_C F \cdot d\mathbf{s} - \int_{C_\epsilon} F \cdot d\mathbf{s} = \int_{\partial S} F \cdot d\mathbf{s} = \iint_S \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dA = 0.$$

Note that the orientation on C_ϵ is not consistent with the orientation on ∂S . Thus $\int_C F \cdot d\mathbf{s} = \int_{C_\epsilon} F \cdot d\mathbf{s} = \alpha$.

c) If we use the parametric representation $\sigma(\theta) = (r \cos(\theta), r \sin(\theta))$ for $0 \leq \theta < 2\pi$, then we find that

$$\int_{C_r} F_0 \cdot d\mathbf{s} = \int_0^{2\pi} \left(\frac{-\sin(\theta)}{r}, \frac{\cos(\theta)}{r} \right) \cdot (-r \sin(\theta), r \cos(\theta)) d\theta = \int_0^{2\pi} d\theta = 2\pi$$

for all $r > 0$. As in (b), we get that $\int_C F_0 \cdot d\mathbf{s} = 2\pi$ for any closed curve surrounding the origin as described in (b). Hence

$$\int_C G \cdot d\mathbf{s} = \int_C \left(F - \frac{\alpha}{2\pi} F_0 \right) \cdot d\mathbf{s} = \int_C F \cdot d\mathbf{s} - \frac{\alpha}{2\pi} \int_C F_0 \cdot d\mathbf{s} = 0$$

for any closed curve surrounding the origin as described in (b). Moreover

$$\begin{aligned} \frac{\partial G_2}{\partial x} - \frac{\partial G_1}{\partial y} &= \left(\frac{\partial F_2}{\partial x} - \frac{\alpha}{2\pi} \frac{\partial}{\partial x} \left(\frac{x}{x^2 + y^2} \right) \right) - \left(\frac{\partial F_1}{\partial y} - \frac{\alpha}{2\pi} \frac{\partial}{\partial y} \left(\frac{-y}{x^2 + y^2} \right) \right) \\ &= \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) + \frac{\alpha}{2\pi} \left(\frac{\partial}{\partial x} \left(\frac{x}{x^2 + y^2} \right) - \frac{\partial}{\partial y} \left(\frac{-y}{x^2 + y^2} \right) \right) \\ &= \frac{\alpha}{2\pi} \left(\frac{1}{x^2 + y^2} - \frac{2x^2}{(x^2 + y^2)^2} + \frac{1}{x^2 + y^2} - \frac{2y^2}{(x^2 + y^2)^2} \right) = 0 \end{aligned}$$

on $\mathbb{R}^2 \setminus \{\mathbf{0}\}$. Hence, using Green's Theorem as we did in (a) and (b), we get that $\int_C G \cdot d\mathbf{s} = 0$ for any closed curve C that does not surround the origin.

It follows from Proposition 6.2.7 that G is conservative in $\mathbb{R}^2 \setminus \{0\}$. That G is the gradient of a function follows from Proposition 6.2.10. $\heartsuit$

Question 6.44

a) Suppose that $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is a vector field of class C^1 . Prove that

$$\int_{\partial S} F \cdot \mathbf{n} \, ds = \iint_S \left(\frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} \right) \, dA, \quad (6.6.2)$$

where S is a bounded region in $\mathbb{R}^2$ with a smooth boundary, $\mathbf{n}$ is the unit normal to ∂S pointing outside of S , and ∂S is positively oriented with respect to S .

b) Use the result in (a) to prove the **Green's formulas**:

$$\int_{\partial S} f \nabla g \cdot \mathbf{n} \, ds = \iint_S (\nabla f \cdot \nabla g + f \nabla^2 g) \, dA \quad (6.6.3)$$

and

$$\int_{\partial S} (f \nabla g - g \nabla f) \cdot \mathbf{n} \, ds = \iint_S (f \nabla^2 g - g \nabla^2 f) \, dA \quad (6.6.4)$$

where $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ and $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ are two twice continuously differentiable functions.

Note that $\nabla^2 f = \operatorname{div}(\nabla f)$ is the **Laplacian** of f and is more often denoted Δf .

c) Let $g(x, y) = \ln(x^2 + y^2)$. Verify that $\nabla g = \frac{2}{x^2 + y^2} (x, y)$ and $\nabla^2 g(x, y) = 0$ for $(x, y) \neq (0, 0)$.

d) Suppose that $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ is twice continuously differentiable and $\nabla^2 f = 0$. Use g defined in (c) and the results in (b) to prove that mean value of f on circles r_1 and r_2 centred at the origin are equal. Namely,

$$\frac{1}{2\pi r_2} \int_{C_{r_2}} f \, ds = \frac{1}{2\pi r_1} \int_{C_{r_1}} f \, ds \quad (6.6.5)$$

where C_r is the circle of radius r centred at the origin and the orientation on C_r is counterclockwise.

e) Use (d) to prove that the mean value of f on circles centred at the origin is equal to the value of f at the origin if we assume that the orientation on the circles is counterclockwise.

Solution

a) We apply the classical Stokes' theorem, Theorem 6.4.2, to the vector field $\tilde{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ defined by $\tilde{F}(x, y, z) = (-F_2(x, y), F_1(x, y), 0)$ and the region $\tilde{S} = \{(x, y, 0) : (x, y) \in S\}$. We get

$$\oint_{\partial \tilde{S}} \tilde{F} \cdot \mathbf{t} \, ds = \int_{\tilde{S}} \operatorname{curl} \tilde{F} \cdot \mathbf{e}_3 \, dS \quad (6.6.6)$$

where $\mathbf{t}(\mathbf{x}) = (t_1(\mathbf{x}), t_2(\mathbf{x}), 0)$ is a unit vector tangent to the curve $\partial \tilde{S}$ at $\mathbf{x} \in \partial \tilde{S}$. In the present case, $\mathbf{t}(\mathbf{x})$ points in the direction of the positive orientation on $\partial \tilde{S}$ with respect to $\tilde{S}$; namely, the orientation associated to the counterclockwise motion on $\partial \tilde{S}$.

However

$$\oint_{\partial \tilde{S}} \tilde{F} \cdot \mathbf{t} \, ds = \oint_{\partial \tilde{S}} (-F_2, F_1, 0) \cdot (t_1, t_2, 0) \, ds = \oint_{\partial \tilde{S}} (F_1, F_2, 0) \cdot (t_2, -t_1, 0) \, ds \quad (6.6.7)$$

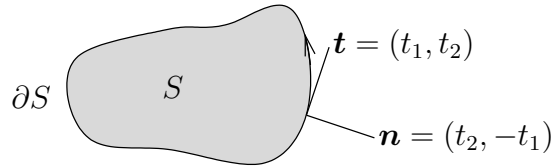
and

$$\int_{\tilde{S}} \operatorname{curl} \tilde{F} \cdot \mathbf{e}_3 \, dS = \iint_{\tilde{S}} \left(\frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} \right) \, dS. \quad (6.6.8)$$

Since everything in the previous discussion is independent of the third component in $\mathbb{R}^3$, we may substitute (6.6.7) and (6.6.8) in (6.6.6) to get the relation

$$\oint_{\partial S} F \cdot \mathbf{n} \, ds = \oint_{\partial S} (F_1, F_2) \cdot (t_2, -t_1) \, ds = \iint_S \left(\frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} \right) \, dS,$$

where we have used the fact that $\mathbf{n}(\mathbf{x}) = (t_2(\mathbf{x}), -t_1(\mathbf{x}))$ is a unit vector perpendicular to ∂S at $\mathbf{x} \in \partial S$ and $\mathbf{n}(\mathbf{x})$ points outside S . This is shown in the following figure.



b) To prove (6.6.3), we substitute $F = f \nabla g = \left(f \frac{\partial g}{\partial x}, f \frac{\partial g}{\partial y} \right)$ in (6.6.2) to get

$$\begin{aligned} \int_{\partial S} f \nabla g \cdot \mathbf{n} \, ds &= \iint_S \left(\frac{\partial}{\partial x} \left(f \frac{\partial g}{\partial x} \right) + \frac{\partial}{\partial y} \left(f \frac{\partial g}{\partial y} \right) \right) \, dA \\ &= \iint_S \left(\frac{\partial f}{\partial x} \frac{\partial g}{\partial x} + f \frac{\partial^2 g}{\partial x^2} + \frac{\partial f}{\partial y} \frac{\partial g}{\partial y} + f \frac{\partial^2 g}{\partial y^2} \right) \, dA = \iint_S (\nabla f \cdot \nabla g + f \nabla^2 g) \, dA. \end{aligned}$$

If we interchange f and g in (6.6.3) and subtract from the original (6.6.3), then we get (6.6.4).

c) We have

$$\nabla g(x, y) = \left(\frac{\partial g}{\partial x}(x, y), \frac{\partial g}{\partial y}(x, y) \right) = \left(\frac{2x}{x^2 + y^2}, \frac{2y}{x^2 + y^2} \right).$$

and

$$\begin{aligned} \nabla^2 g(x, y) &= \frac{\partial^2 g}{\partial x^2}(x, y) + \frac{\partial^2 g}{\partial y^2}(x, y) = \frac{\partial}{\partial x} \left(\frac{2x}{x^2 + y^2} \right) + \frac{\partial}{\partial y} \left(\frac{2y}{x^2 + y^2} \right) \\ &= \frac{2}{x^2 + y^2} - \frac{4x^2}{(x^2 + y^2)^2} + \frac{2}{x^2 + y^2} - \frac{4y^2}{(x^2 + y^2)^2} = \frac{4}{x^2 + y^2} - \frac{4(x^2 + y^2)}{(x^2 + y^2)^2} = 0. \end{aligned}$$

d) Since $\nabla^2 f = 0$ and $\nabla^2 g = 0$, we get from (6.6.4) that $\int_{\partial S} (f \nabla g - g \nabla f) \cdot \mathbf{n} \, ds = 0$ where S is the annulus bounded by the circle C_{r_1} and C_{r_2} . Thus

$$\int_{C_{r_1}} (f \nabla g - g \nabla f) \cdot \mathbf{n} \, ds = \int_{C_{r_2}} (f \nabla g - g \nabla f) \cdot \mathbf{n} \, ds \quad (6.6.9)$$

where both C_{r_1} and C_{r_2} have the counterclockwise orientation. Note that the orientation on C_{r_1} is not consistent with the orientation on ∂S if $r_1 < r_2$ while it is the orientation on C_{r_2} .

that is not consistent with the orientation on ∂S if $r_2 < r_1$. This justifies the equality above. However, for $(x, y) \in C(r_1)$, we have

$$\nabla g(x, y) = \frac{2}{x^2 + y^2}(x, y) = \frac{2}{\sqrt{x^2 + y^2}}\mathbf{n} = \frac{2}{r_1}\mathbf{n}$$

and $g(x, y) = \ln(x^2 + y^2) = \ln(r_1^2)$. Thus $\int_{C_{r_1}} f \nabla g \cdot \mathbf{n} \, ds = \frac{2}{r_1} \int_{C_{r_1}} f \, ds$ and

$$\int_{C_{r_1}} g \nabla f \cdot \mathbf{n} \, ds = \ln(r_1^2) \int_{C_{r_1}} \nabla f \cdot \mathbf{n} \, ds = \ln(r_1^2) \iint_{B_{r_1}(\mathbf{0})} \nabla^2 f \, dS = 0$$

because of (6.6.2). Similarly, $\int_{C_{r_2}} f \nabla g \cdot \mathbf{n} \, ds = \frac{2}{r_2} \int_{C_{r_2}} f \, ds$ and $\int_{C_{r_2}} g \nabla f \cdot \mathbf{n} \, ds = 0$. We get from (6.6.9) that $\frac{2}{r_2} \int_{C_{r_2}} f \, ds = \frac{2}{r_1} \int_{C_{r_1}} f \, ds$. Hence, after multiplying this equality by $1/(4\pi)$, we get (6.6.5).

e) From (d), we have that $\frac{1}{2\pi r} \int_{C_r} f \, ds = M$ for all r , where M is a constant. From the Mean Value Theorem for Integrals, we have for each r that $\frac{1}{2\pi r} \int_{C(r)} f \, ds = f(\mathbf{a}_r)$ for some $\mathbf{a}_r \in C_r$. So $f(\mathbf{a}_r) = M$ for all $\mathbf{a}_r$. Since $\lim_{r \rightarrow 0} \mathbf{a}_r = \mathbf{0}$ and f is continuous, we get that $M = \lim_{r \rightarrow 0^+} f(\mathbf{a}_r) = f(\mathbf{0})$. $\heartsuit$

Question 6.45

Let R be a bounded region in $\mathbb{R}^3$ with a piecewise smooth boundary ∂R . Suppose that $f, g : \mathbb{R}^3 \rightarrow \mathbb{R}$ are two continuously differentiable functions. Prove that

$$\iiint_R f \frac{\partial g}{\partial x_1} \, dV = - \iiint_R g \frac{\partial f}{\partial x_1} \, dV + \iint_{\partial R} f g n_1 \, dS \quad (6.6.10)$$

where $\mathbf{n} = (n_1, n_2, n_3)$ is the unit normal pointing outside R at every point of ∂R and the orientation on ∂R is given by this unit normal. This is the orientation induced on ∂R by the orientation on R as defined in Definition 5.3.15.

This is an higher dimensional version of integration by parts. One can get a similar result for the derivatives with respect to x_2 and x_3 .

Solution

Let $F(\mathbf{x}) = (f(\mathbf{x})g(\mathbf{x}), 0, 0)$. Since $\operatorname{div} F = \frac{\partial f}{\partial x_1}g + f \frac{\partial g}{\partial x_1}$, we get from the divergence theorem that

$$\begin{aligned} \iint_{\partial R} f g n_1 \, dS &= \iint_{\partial R} F \cdot \mathbf{n} \, dS = \iint_{\partial R} F \cdot d\mathbf{S} = \iiint_R \operatorname{div} F \, dV \\ &= \iiint_R \left(\frac{\partial f}{\partial x_1}g + f \frac{\partial g}{\partial x_1} \right) dV. \end{aligned}$$

We get (6.6.10) by subtracting $\iiint_R \frac{\partial f}{\partial x_1} g \, dV$ on both sides of the previous equality. $\heartsuit$

Question 6.46

Suppose that $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ is a function of class C^2 . Let R be a bounded region in $\mathbb{R}^3$ with a piecewise smooth boundary ∂R and $\mathbf{n}$ be the unit normal pointing outside R at every point of ∂R . The orientation on ∂R given by this unit normal is the orientation induced on ∂R by the orientation on R as defined in Definition 5.3.15.

a) Show that

$$\iint_{\partial R} \frac{\partial f}{\partial \mathbf{n}} \cdot d\mathbf{S} = \iiint_R \nabla^2 f \, dV .$$

b) If $\nabla^2 f = 0$, show that

$$\iint_{\partial R} f \frac{\partial f}{\partial \mathbf{n}} \, dS = \iiint_R \|\nabla f\|^2 \, dV .$$

Solution

a) From the divergence theorem, we have

$$\iint_{\partial R} \frac{\partial f}{\partial \mathbf{n}} \, dS = \iint_{\partial R} \nabla f \cdot \mathbf{n} \, dS = \iint_{\partial R} \nabla f \cdot d\mathbf{S} = \iiint_R \operatorname{div}(\nabla f) \, dV = \iiint_R \nabla^2 f \, dV .$$

b) We have

$$\begin{aligned} \iint_{\partial R} f \frac{\partial f}{\partial \mathbf{n}} \, dS &= \iint_{\partial R} f (\nabla f \cdot \mathbf{n}) \, dS = \iint_{\partial R} f \nabla f \cdot d\mathbf{S} \\ &= \iiint_R \operatorname{div}(f \nabla f) \, dV = \iiint_R (\nabla f \cdot \nabla f + f \nabla^2 f) \, dV = \iiint_R \|\nabla f\|^2 \, dV . \end{aligned}$$

because $\nabla^2 f = 0$ on R by assumption. $\heartsuit$

Question 6.47

Let $g(\mathbf{x}) = \frac{1}{\|\mathbf{x}\|}$ for $\mathbf{x} \in \mathbb{R}^3 \setminus \{\mathbf{0}\}$.

a) Compute ∇g and show that $\nabla^2 g(\mathbf{x}) = 0$ for all $\mathbf{x} \neq \mathbf{0}$.

b) Show that $\iint_{\partial B_r(\mathbf{0})} \frac{\partial g}{\partial \mathbf{n}} \, dS = -4\pi$ where $B_r(\mathbf{0})$ is the open ball of radius r centred at the origin, $\mathbf{n}$ is the unit normal to $\partial B_r(\mathbf{0})$ pointing outside $B_r(\mathbf{0})$, and the orientation on $\partial B_r(\mathbf{0})$ is given by the vector $\mathbf{n}$.

c) Explain why (a) and (b) do not contradict (a) of Question 6.46.

d) Show that $\iint_{\partial R} \frac{\partial g}{\partial \mathbf{n}} \, dS = -4\pi$ for any bounded region R with a piecewise smooth boundary such that $\mathbf{0} \in R^\circ$. The vector $\mathbf{n}$ is the unit normal to ∂R pointing outside R and the orientation on ∂R is given by $\mathbf{n}$.

Solution

a) We have

$$\nabla g(\mathbf{x}) = \sum_{i=1}^3 \frac{\partial}{\partial x_i} (x_1^2 + x_2^2 + x_3^2)^{-1/2} \mathbf{e}_i = \sum_{i=1}^3 (-x_i (x_1^2 + x_2^2 + x_3^2)^{-3/2}) \mathbf{e}_i = -\frac{1}{\|\mathbf{x}\|^3} \mathbf{x}$$

and

$$\begin{aligned}\nabla^2 g(\mathbf{x}) &= \sum_{i=1}^3 \frac{\partial^2}{\partial x_i^2} (x_1^2 + x_2^2 + x_3^2)^{-1/2} = \sum_{i=1}^3 \frac{\partial}{\partial x_i} (-x_i (x_1^2 + x_2^2 + x_3^2)^{-3/2}) \\ &= \left(\frac{2x_1^2 - x_2 - x_3}{(x_1^2 + x_2^2 + x_3^2)^{5/2}} \right) + \left(\frac{2x_2^2 - x_1 - x_3}{(x_1^2 + x_2^2 + x_3^2)^{5/2}} \right) + \left(\frac{2x_3^2 - x_1 - x_2}{(x_1^2 + x_2^2 + x_3^2)^{5/2}} \right) = 0\end{aligned}$$

for $\mathbf{x} \neq \mathbf{0}$.

b) We have

$$\begin{aligned}\iint_{\partial B_r(\mathbf{0})} \frac{\partial g}{\partial \mathbf{n}} dS &= \iint_{\partial B_r(\mathbf{0})} \nabla g \cdot \mathbf{n} dS = \iint_{\partial B_r(\mathbf{0})} \left(\frac{-1}{\|\mathbf{x}\|^3} \mathbf{x} \right) \cdot \left(\frac{1}{\|\mathbf{x}\|} \mathbf{x} \right) dS \\ &= - \iint_{\partial B_r(\mathbf{0})} \frac{1}{\|\mathbf{x}\|^2} dS = -\frac{1}{r^2} \iint_{\partial B_r(\mathbf{0})} dS\end{aligned}$$

where we have used the fact that $\mathbf{n}(\mathbf{x}) = \|\mathbf{x}\|^{-1} \mathbf{x}$ for $\mathbf{x} \in \partial B_r(\mathbf{0})$.

For those who do not know by heart the formula for the area of the sphere of radius r , they may use the parametric representation $g(\theta, \phi) = (r \cos(\theta) \sin(\phi), r \sin(\theta) \sin(\phi), r \cos(\phi))$ for $0 \leq \theta \leq 2\pi$ and $0 \leq \phi \leq \pi$ to find that the area is

$$\iint_{\partial B_r(\mathbf{0})} dS = \int_0^{2\pi} \int_0^\pi \left| \frac{\partial g}{\partial \theta} \times \frac{\partial g}{\partial \phi} \right| d\phi d\theta = \int_0^{2\pi} \int_0^\pi r^2 \sin(\phi) d\phi d\theta = 4\pi r^2.$$

Thus $\iint_{\partial B_r(\mathbf{0})} \frac{\partial g}{\partial \mathbf{n}} dS = -4\pi$.

c) We get from (a) and (b) that $\iint_{\partial B_r(\mathbf{0})} \frac{\partial g}{\partial \mathbf{n}} dS = -4\pi \neq 0 = \iiint_{B_r(\mathbf{0})} \nabla^2 g dV$. The formula in (a) of Question 6.46 cannot be used because g is not defined at the origin (and cannot be continuously extended to $\mathbb{R}^3$) and thus it is certainly not of class C^2 in any open set containing $B_r(\mathbf{0})$ as required.

d) Choose $\epsilon > 0$ such that $B_\epsilon(\mathbf{0}) \subset R$. This is possible because $\mathbf{0}$ is in the interior of R . Let $R_\epsilon = R \setminus B_\epsilon(\mathbf{0})$.

If we use (a) of Question 6.46 and (a) of the present question, then we get $\iint_{\partial R_\epsilon} \frac{\partial g}{\partial \mathbf{n}} dS = \iiint_{R_\epsilon} \nabla^2 g dV = 0$ where the orientation on ∂R is given by the unit normal that point outside of R at all points of ∂R . Thus

$$0 = \iint_{\partial R_\epsilon} \frac{\partial g}{\partial \mathbf{n}} dS = \iint_{\partial R} \frac{\partial g}{\partial \mathbf{n}} dS - \iint_{\partial B_\epsilon(\mathbf{0})} \frac{\partial g}{\partial \mathbf{n}} dS = \iint_{\partial R} \frac{\partial g}{\partial \mathbf{n}} dS + 4\pi$$

because of (b). The minus sign in front of the integral $\iint_{\partial B_\epsilon(\mathbf{0})} \frac{\partial g}{\partial \mathbf{n}} dS$ comes from the fact that the orientation on $\partial B_\epsilon(\mathbf{0})$ given by the unit normal that point outside $B_\epsilon(\mathbf{0})$ for all $\mathbf{x} \in B_\epsilon(\mathbf{0})$ is not consistent with the orientation on ∂R_ϵ . Hence $\iint_{\partial R} \frac{\partial g}{\partial \mathbf{n}} dS = -4\pi$. $\heartsuit$

Question 6.48

In this question, we generalize the results presented in Question 6.44; in particular, we generalize the Green's formulas presented in part (b) of that question.

a) Prove the **Green's formulas**:

$$\iint_{\partial R} f \nabla g \cdot \mathbf{n} \, dS = \iiint_R (\nabla f \cdot \nabla g + f \nabla^2 g) \, dV$$

and

$$\iint_{\partial R} (f \nabla g - g \nabla f) \cdot \mathbf{n} \, dS = \iiint_R (f \nabla^2 g - g \nabla^2 f) \, dV \quad (6.6.11)$$

where $f, g : \mathbb{R}^3 \rightarrow \mathbb{R}$ are two functions of class C^2 , the vector $\mathbf{n}(\mathbf{x})$ is the unit normal to ∂R pointing outside R at $\mathbf{x} \in \partial R$, and the orientation on ∂R is given by the vector $\mathbf{n}$. This is the orientation induced on ∂R by the orientation on R as defined in Definition 5.3.15.

b) Suppose that $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ is a function of class C^2 and $\nabla^2 f = 0$, prove that

$$\frac{1}{4\pi r_1^2} \iint_{\partial B_{r_1}(\mathbf{0})} f \, dS = \frac{1}{4\pi r_2^2} \iint_{\partial B_{r_2}(\mathbf{0})} f \, dS$$

where $B_r(\mathbf{0})$ is the ball of radius r centred at the origin, and the orientation on $\partial B_r(\mathbf{0})$ is given by the unit normal to $\partial B_r(\mathbf{0})$ pointing outside $B_r(\mathbf{0})$.

c) Use (b) to prove that $f(\mathbf{0}) = \frac{1}{4\pi r^2} \iint_{\partial B_r(\mathbf{0})} f \, dS$ for all $r > 0$. In general, prove that

$$f(\mathbf{a}) = \frac{1}{4\pi r^2} \iint_{\partial B_r(\mathbf{a})} f \, dS \text{ for all } \mathbf{a} \in \mathbb{R}^3 \text{ and } r > 0.$$

Solution

a) We apply the divergence theorem to the vector field $F = f \nabla g$.

$$\begin{aligned} \iint_{\partial R} f (\nabla g \cdot \mathbf{n}) \, dS &= \iint_{\partial R} f \nabla g \cdot d\mathbf{S} = \iiint_R \operatorname{div}(f \nabla g) \, dV \\ &= \iiint_R (\nabla f \cdot \nabla g + f \nabla^2 g) \, dV. \end{aligned}$$

If we interchange f and g in the previous equation and subtract the resulting equation from the previous equation, then we get (6.6.11).

b) Without loss of generality, we may assume that $r_1 > r_2$. Let $R = \overline{B_{r_1}(\mathbf{0})} \setminus \overline{B_{r_2}(\mathbf{0})}$. From the second Green's formula with f and $g(\mathbf{x}) = \frac{1}{\|\mathbf{x}\|}$ for $\mathbf{x} \in R$, we get

$$\iint_{\partial R} (f \nabla g - g \nabla f) \cdot d\mathbf{S} = \iiint_R (f \nabla^2 g - g \nabla^2 f) \, dV = 0$$

because $\nabla^2 f = 0$ by hypothesis and $\nabla^2 g = 0$ as we have shown in (a) of Question 6.47. The orientation on ∂R is given by the unit normal to ∂R that points outside R . Thus

$$\iint_{\partial B_{r_1}(\mathbf{0})} (f \nabla g - g \nabla f) \cdot d\mathbf{S} = \iint_{\partial B_{r_2}(\mathbf{0})} (f \nabla g - g \nabla f) \cdot d\mathbf{S}, \quad (6.6.12)$$

where the orientation on $\partial B_r(\mathbf{0})$ is given by the unit normal to $\partial B_r(\mathbf{0})$ that points outside $B_r(\mathbf{0})$. To get (6.6.12), we note that the orientation on $\partial B_{r_2}(\mathbf{0})$ is not consistent with the orientation on ∂R . Using (a) of Question 6.47, we get

$$\begin{aligned} \iint_{\partial B_{r_1}(\mathbf{0})} f \nabla g \cdot d\mathbf{S} &= \iint_{\partial B_{r_1}(\mathbf{0})} f(\mathbf{x}) \left(\frac{1}{\|\mathbf{x}\|^3} \mathbf{x} \right) \cdot \left(\frac{-1}{\|\mathbf{x}\|} \mathbf{x} \right) dS \\ &= - \iint_{\partial B_{r_1}(\mathbf{0})} \left(\frac{1}{\|\mathbf{x}\|^2} \right) f(\mathbf{x}) dS = -\frac{1}{r_1^2} \iint_{\partial B_{r_1}(\mathbf{0})} f dS. \end{aligned}$$

Using the divergence theorem, we get

$$\begin{aligned} \iint_{\partial B_{r_1}(\mathbf{0})} g \nabla f \cdot d\mathbf{S} &= \iint_{\partial B_{r_1}(\mathbf{0})} \left(\frac{1}{\|\mathbf{x}\|} \right) \nabla f \cdot d\mathbf{S} \\ &= \frac{1}{r_1} \iint_{\partial B_{r_1}(\mathbf{0})} \nabla f \cdot d\mathbf{S} = \frac{1}{r_1} \iiint_{B_{r_1}(\mathbf{0})} \nabla^2 f dV = 0 \end{aligned}$$

where we have used (a) of Question 6.46 to get that last equality. Thus

$$\iint_{\partial B_{r_1}(\mathbf{0})} (f \nabla g - g \nabla f) \cdot d\mathbf{S} = -\frac{1}{r_1^2} \iint_{\partial B_{r_1}(\mathbf{0})} f dS.$$

Similarly, we get $\iint_{\partial B_{r_2}(\mathbf{0})} f \nabla g \cdot d\mathbf{S} = -\frac{1}{r_2^2} \iint_{\partial B_{r_2}(\mathbf{0})} f dS$ and $\iint_{\partial B_{r_2}(\mathbf{0})} g \nabla f \cdot d\mathbf{S} = 0$.

Thus

$$\iint_{\partial B_{r_2}(\mathbf{0})} (f \nabla g - g \nabla f) \cdot d\mathbf{S} = -\frac{1}{r_2^2} \iint_{\partial B_{r_2}(\mathbf{0})} f dS.$$

Hence, we get the result in (b) from (6.6.12) after dividing by -4π .

c) We have from (b) that $\frac{1}{4\pi r^2} \iint_{\partial B_r(\mathbf{0})} f dS = C$, a constant, for all $r > 0$. Hence

$\lim_{r \rightarrow 0^+} \frac{1}{4\pi r^2} \iint_{\partial B_r(\mathbf{0})} f dS = C$. Moreover, from the Mean Value Theorem for Integrals, we have that $\frac{1}{4\pi r^2} \iint_{\partial B_r(\mathbf{0})} f dS = f(\mathbf{a}_r)$ for some $\mathbf{a}_r \in \partial B_r(\mathbf{0})$. Note that $\iint_{\partial B_r(\mathbf{0})} f dS = 4\pi r^2$ is the area of the sphere of radius r . Since f is continuous and $\mathbf{a}_r \rightarrow \mathbf{0}$ as $r \rightarrow 0^+$, we get that $C = \lim_{r \rightarrow 0^+} \frac{1}{4\pi r^2} \iint_{\partial B_r(\mathbf{0})} f dS = \lim_{r \rightarrow 0^+} f(\mathbf{a}_r) = f(\mathbf{0})$.

We get the formula at $\mathbf{x} = \mathbf{a} \in \mathbb{R}$ arbitrary from the formula at the origin with the change of variable $\mathbf{x} = \mathbf{y} + \mathbf{a}$. ♥

Question 6.49

Suppose that $S \subset \mathbb{R}^3$ is a surface with a piecewise continuously differentiable boundary ∂S . Moreover, assume that the orientation on ∂S with respect to S is positive; namely, the orientation on ∂S is the orientation induced from the orientation on S . If $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ is a function of class C^1 and $g : \mathbb{R}^3 \rightarrow \mathbb{R}$ is function of class C^2 , prove that

$$\int_{\partial S} f \nabla g \cdot d\mathbf{s} = \iint_S (\nabla f \times \nabla g) \cdot d\mathbf{S}. \quad (6.6.13)$$

Solution

We prove that (6.6.13) is given by Stokes' theorem and the relation $\text{curl}(f \nabla g) = \nabla f \cdot \nabla g$. We present two methods to prove this relation.

First method) We use brute force.

$$\begin{aligned} \text{curl}(f \nabla g) &= \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ f \frac{\partial g}{\partial x} & f \frac{\partial g}{\partial y} & f \frac{\partial g}{\partial z} \end{pmatrix} \\ &= \left(\frac{\partial}{\partial y} \left(f \frac{\partial g}{\partial z} \right) - \frac{\partial}{\partial z} \left(f \frac{\partial g}{\partial y} \right), - \left(\frac{\partial}{\partial x} \left(f \frac{\partial g}{\partial z} \right) - \frac{\partial}{\partial z} \left(f \frac{\partial g}{\partial x} \right) \right), \frac{\partial}{\partial x} \left(f \frac{\partial g}{\partial y} \right) - \frac{\partial}{\partial y} \left(f \frac{\partial g}{\partial x} \right) \right) \\ &= \left(\frac{\partial f}{\partial y} \frac{\partial g}{\partial z} + f \frac{\partial^2 g}{\partial y \partial z} - \frac{\partial f}{\partial z} \frac{\partial g}{\partial y} - f \frac{\partial^2 g}{\partial z \partial y}, - \left(\frac{\partial f}{\partial x} \frac{\partial g}{\partial z} + f \frac{\partial^2 g}{\partial x \partial z} - \frac{\partial f}{\partial z} \frac{\partial g}{\partial x} - f \frac{\partial^2 g}{\partial z \partial x} \right), \right. \\ &\quad \left. \frac{\partial f}{\partial x} \frac{\partial g}{\partial y} + f \frac{\partial^2 g}{\partial x \partial y} - \frac{\partial f}{\partial y} \frac{\partial g}{\partial x} - f \frac{\partial^2 g}{\partial y \partial x} \right) \\ &= \left(\frac{\partial f}{\partial y} \frac{\partial g}{\partial z} - \frac{\partial f}{\partial z} \frac{\partial g}{\partial y}, - \left(\frac{\partial f}{\partial x} \frac{\partial g}{\partial z} - \frac{\partial f}{\partial z} \frac{\partial g}{\partial x} \right), \frac{\partial f}{\partial x} \frac{\partial g}{\partial y} - \frac{\partial f}{\partial y} \frac{\partial g}{\partial x} \right) = \nabla f \times \nabla g \end{aligned}$$

because $\frac{\partial^2 g}{\partial z \partial y} = \frac{\partial^2 g}{\partial y \partial z}$, $\frac{\partial^2 g}{\partial z \partial x} = \frac{\partial^2 g}{\partial x \partial z}$ and $\frac{\partial^2 g}{\partial y \partial x} = \frac{\partial^2 g}{\partial x \partial y}$ for functions of class C^2 like g .

Second method) We use differential forms. We have that $\nabla g \longleftrightarrow dg$. Thus $f \nabla g \longleftrightarrow f dg$ and $\text{curl}(f \nabla g) \longleftrightarrow d(f dg) = df \wedge dg + f d(dg) = df \wedge dg$ where we have used item (3) of Theorem 4.5.7 to conclude that $d(dg) = 0$. Moreover, we have $\nabla f \times \nabla g \longleftrightarrow df \wedge dg$. Thus $\text{curl}(f \nabla g) = \nabla f \times \nabla g$. The second method is much shorter than the first one.

To answer the question, we use Stokes' theorem to get

$$\int_{\partial S} f \nabla g \cdot d\mathbf{s} = \iint_S \text{curl}(f \nabla g) \cdot d\mathbf{S} = \iint_S (\nabla f \times \nabla g) \cdot d\mathbf{S}.$$

♡

Question 6.50

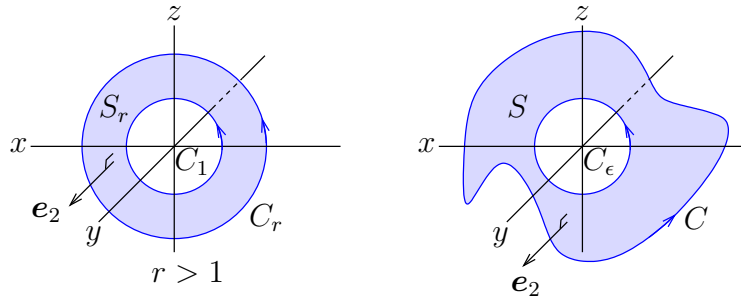
Suppose that $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is a continuously differentiable vector field such that $\text{curl } F \cdot \mathbf{e}_2 = 0$ everywhere in the x, z plane. Show that $\int_C F \cdot d\mathbf{s}$ yields the same value for every closed curve C entirely in the x, z plane that surrounds the origin if we assume that the orientation on C is “globally” counterclockwise when view from the positive side of the y axis. We assume here that the closed curves C can be continuously deformed in the x, z plane into a circle without intersecting the origin.

Solution

Let $C_r = \{(x, 0, z) : x^2 + z^2 = r^2\}$ and

$$S_r = \begin{cases} \{(x, 0, z) : 1 \leq x^2 + z^2 \leq r^2\} & \text{if } r > 1 \\ \{(x, 0, z) : r^2 \leq x^2 + z^2 \leq 1\} & \text{if } 0 < r < 1 \end{cases}$$

The orientation on S_r is given by the unit vector $\mathbf{e}_2$ and ∂S_r is given the positive orientation consistent with the orientation on S_r .



i) Using Stokes' theorem, we have

$$\int_{C_r} F \cdot d\mathbf{s} - \int_{C_1} F \cdot d\mathbf{s} = \epsilon \int_{\partial S_r} F \cdot d\mathbf{s} = \epsilon \iint_{S_r} \text{curl } F \cdot d\mathbf{S} = \epsilon \iint_{S_r} \text{curl } F \cdot \mathbf{e}_2 dS = 0,$$

where $\epsilon = 1$ for $r > 1$ and $\epsilon = -1$ for $0 < r < 1$. Note that the orientation on C_1 is not consistent with the orientation on ∂S_r if $r > 1$ while it is the orientation on C_r that is not consistent with the orientation on ∂S_r if $r < 1$. Thus $\int_{C_r} F \cdot d\mathbf{s} = \int_{C_1} F \cdot d\mathbf{s}$ for all r . If

$K = \int_{C_1} F \cdot d\mathbf{s}$, then $\int_{C_r} F \cdot d\mathbf{s} = K$ for all $r > 0$.

ii) Let C be any closed curves entirely in the x, z plane that surrounds the origin as described in the statement of the question. Since the origin is an interior point of the region enclosed by the curve C in the x, z plane, there exists $\epsilon > 0$ such that C_ϵ is included in the interior of the region enclosed by the curve C in the x, z plane. See the figure above. If S is the region in the x, z plane enclosed by C and the circle C_ϵ , then we get from Stokes' theorem that

$$\int_C F \cdot d\mathbf{s} - \int_{C_\epsilon} F \cdot d\mathbf{s} = \int_{\partial S} F \cdot d\mathbf{s} = \iint_S \text{curl } F \cdot d\mathbf{S} = \iint_S \text{curl } F \cdot \mathbf{e}_2 dS = 0.$$

Note that the orientation on C_ϵ is not consistent with the orientation on ∂S . Thus $\int_C F \cdot d\mathbf{s} = \int_{C_\epsilon} F \cdot d\mathbf{s} = K$. ♥

6.6.3 Vector Fields

Question 6.51

For each of the vector fields F below, determine if it is conservative and if there exists a function f such that $F = \nabla f$. If it is possible, find a function f .

a) $F(x, y, z) = (yz^2, xz^2, 2xyz)$.

b) $F(x, y, z) = (2xy + x^2, x^2 - y^2, 0)$.

- c) $F(x, y, z) = (3y^2 + 5x^4, x^5 + 6xy, 0)$.
d) $F(x, y, z) = (yz - y \sin(xy), xz - x \sin(xy) + z \cos(yz), xy + y \cos(yz))$.
e) $F(x, y, z) = (y - z, x - z, x - y)$.

Solution

a) Since

$$\operatorname{curl} F(x, y, z) = \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ yz^2 & xz^2 & 2xyz \end{pmatrix} = \mathbf{0}$$

for all $(x, y, z) \in \mathbb{R}^3$ and F is a continuously differentiable vector field on the convex set $\mathbb{R}^3$, we get from Proposition 6.4.5 that F is conservative and we get from Proposition 6.2.10 that there exists a function $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ such that $\nabla f = F$. The function f must satisfy the relations

$$\frac{\partial f}{\partial x} = yz^2, \quad \frac{\partial f}{\partial y} = xz^2 \quad \text{and} \quad \frac{\partial f}{\partial z} = 2xyz. \quad (6.6.14)$$

From the first equation in (6.6.14), we get that $f(x, y, z) = xyz^2 + \phi(y, z)$. If we substitute this expression for f in the second equation in (6.6.14), then we get that $xz^2 + \frac{\partial \phi}{\partial y} = xz^2$. Thus

$\frac{\partial \phi}{\partial y} = 0$ and we conclude that ϕ is only a function of z . If we substitute $f(x, y, z) = xyz^2 + \phi(z)$

in the last equation in (6.6.14), then we get that $2xyz + \frac{d\phi}{dz} = 2xyz$. Thus $\frac{d\phi}{dz} = 0$ and $\phi(z) = C$, a constant. We find that $f(x, y, z) = xyz^2 + C$.

b) Since

$$\operatorname{curl} F(x, y, z) = \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 2xy + x^2 & x^2 - y^2 & 0 \end{pmatrix} = \mathbf{0}$$

for all $(x, y, z) \in \mathbb{R}^3$ and F is a continuously differentiable vector field on the convex set $\mathbb{R}^3$, we get from Proposition 6.4.5 that F is conservative and we get from Proposition 6.2.10 that there exists a function $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ such that $\nabla f = F$. The function f must satisfy the relations

$$\frac{\partial f}{\partial x} = 2xy + x^2, \quad \frac{\partial f}{\partial y} = x^2 - y^2 \quad \text{and} \quad \frac{\partial f}{\partial z} = 0. \quad (6.6.15)$$

From the first equation in (6.6.15), we get that $f(x, y, z) = x^2y + \frac{x^3}{3} + \phi(y, z)$. If we substitute this expression for f in the second equation in (6.6.15), then we get that $x^2 + \frac{\partial \phi}{\partial y} = x^2 - y^2$.

Thus $\frac{\partial \phi}{\partial y} = -y^2$ and we conclude that $\phi(y, z) = -\frac{y^3}{3} + \phi(z)$. If we substitute $f(x, y, z) =$

$x^2y + \frac{x^3}{3} - \frac{y^3}{3} + \phi(z)$ in the last equation in (6.6.15), then we get that $\frac{d\phi}{dz} = 0$. Thus

$\phi(z) = C$, a constant. We find that $f(x, y, z) = x^2y + \frac{x^3}{3} - \frac{y^3}{3} + C$.

c) We have

$$\operatorname{curl} F(x, y, z) = \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 3y^2 + 5x^4 & x^5 + 6xy & 0 \end{pmatrix} = 5x^4 \mathbf{e}_3 \neq \mathbf{0}$$

except if $x = 0$. Since F is a continuously differentiable vector field on the convex set $\mathbb{R}^3$, we get from Proposition 6.2.10 that F cannot be conservative because we would then have that $F = \nabla f$ for a function $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ and hence $\operatorname{curl} F = \operatorname{curl}(\nabla f) = \mathbf{0}$ on $\mathbb{R}^3$.

d) Since

$$\operatorname{curl} F(x, y, z) = \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ yz - y \sin(xy) & xz - x \sin(xy) + z \cos(yz) & xy + y \cos(yz) \end{pmatrix} = \mathbf{0}$$

for all $(x, y, z) \in \mathbb{R}^3$ and F is a continuously differentiable vector field on the convex set $\mathbb{R}^3$, we get from Proposition 6.4.5 that F is conservative and we get from Proposition 6.2.10 that there exists a function $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ such that $\nabla f = F$. The function f must satisfy the relations

$$\frac{\partial f}{\partial x} = yz - y \sin(xy), \quad \frac{\partial f}{\partial y} = xz - x \sin(xy) + z \cos(yz) \quad \text{and} \quad \frac{\partial f}{\partial z} = xy + y \cos(yz) \quad (6.6.16)$$

From the first equation in (6.6.16), we get that $f(x, y, z) = xyz + \cos(xy) + \phi(y, z)$. If we substitute this expression for f in the second equation in (6.6.16), then we get that $xz - x \sin(xy) + \frac{\partial \phi}{\partial y} = xz - x \sin(xy) + z \cos(yz)$. Thus $\frac{\partial \phi}{\partial y} = z \cos(yz)$ and we conclude that $\phi(y, z) = \sin(yz) + \phi(z)$. If we substitute $f(x, y, z) = xyz + \cos(xy) + \sin(yz) + \phi(z)$ in the last equation in (6.6.16), then we get that $xy + y \cos(yz) + \frac{d\phi}{dz} = xy + y \cos(yz)$. Thus $\frac{d\phi}{dz} = 0$ and $\phi(z) = C$, a constant. We find that $f(x, y, z) = xyz + \cos(xy) + \sin(yz) + C$.

e) We have

$$\operatorname{curl} F(x, y, z) = \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ y - z & x - z & x - y \end{pmatrix} = -2\mathbf{e}_2 \neq \mathbf{0}.$$

Since F is a continuously differentiable vector field on the convex set $\mathbb{R}^3$, we get from Proposition 6.2.10 that F cannot be conservative because we would then have that $F = \nabla f$ for a function $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ and hence $\operatorname{curl} F = \operatorname{curl}(\nabla f) = \mathbf{0}$ on $\mathbb{R}^3$. $\heartsuit$

Question 6.52

For each of the vector fields F defined on $\mathbb{R}^3$ below, determine if it is conservative and if there exists a vector field G such that $F = \operatorname{curl} G$. If it is possible, find a vector field G .

a) $F(x, y, z) = (3y^2z - 3xz^2, x^2y, z^3 - x^2z)$.

- b) $F(x, y, z) = (0, 2yz + z, z)$.
 c) $F(x, y, z) = (x^3 + yz, y - 3x^2y, 4y^2 - z)$.
 d) $F(x, y, z) = (xy + z, xz, -yz + x)$.
 e) $F(x, y, z) = (xe^{-x^2z^2} - 6x, 5y + 2z, z - ze^{-x^2z^2})$.

Solution

a) Since

$$\operatorname{div} F(x, y, z) = \frac{\partial}{\partial x}(3y^2z - 3xz^2) + \frac{\partial}{\partial y}(x^2y) + \frac{\partial}{\partial z}(z^3 - x^2z) = -3z^2 + x^2 + 3z^2 - x^2 = 0$$

for all $(x, y, z) \in \mathbb{R}^3$ and $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is continuously differentiable on the convex set $\mathbb{R}^3$, it follows from Proposition 6.4.8 that there exists $G : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ such that $\operatorname{curl} G = F$. As explained in Remark 6.4.9, we may seek $G : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ such that

$$\begin{aligned} G_1(x, y, z) &= \int_0^z F_2(x, y, s) \, ds + \psi(x, y) = \int_0^z x^2y \, ds + \psi(x, y) = x^2yz + \psi(x, y), \\ G_2(x, y, z) &= -\int_0^z F_1(x, y, s) \, ds + \phi(x, y) = -\int_0^z (3y^2s - 3xs^2) \, ds + \phi(x, y) \\ &= -\frac{3}{2}y^2z^2 + xz^3 + \phi(x, y) \quad \text{and} \\ G_3(x, y, z) &= 0 \end{aligned}$$

with

$$\frac{\partial G_2}{\partial x} - \frac{\partial G_1}{\partial y} = F_3 \Rightarrow z^3 + \frac{\partial \phi}{\partial x} - x^2z - \frac{\partial \psi}{\partial y} = z^3 - x^2z \Rightarrow \frac{\partial \phi}{\partial x}(x, y) - \frac{\partial \psi}{\partial y}(x, y) = 0.$$

If we choose $\phi(x, y) = 0$ and $\psi(x, y) = 0$ for all (x, y) , then we get that

$$G(x, y, z) = \left(x^2yz, -\frac{3}{2}y^2z^2 + xz^3, 0\right).$$

b) Since

$$\operatorname{div} F(x, y, z) = \frac{\partial}{\partial y}(2yz + z) + \frac{\partial}{\partial z}(z) = 2z + 1 \neq 0$$

for $z \neq -1/2$, we cannot write F as $F = \operatorname{curl} G$ for some vector field $G : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ because we would then have that $\operatorname{div} F = \operatorname{div}(\operatorname{curl} F) = 0$ on $\mathbb{R}^3$.

c) Since

$$\operatorname{div} F(x, y, z) = \frac{\partial}{\partial x}(x^3 + yz) + \frac{\partial}{\partial y}(y - 3x^2y) + \frac{\partial}{\partial z}(4y^2 - z) = 3x^2 + 1 - 3x^2 - 1 = 0$$

for all $(x, y, z) \in \mathbb{R}^3$ and $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is continuously differentiable on the convex set $\mathbb{R}^3$, it follows from Proposition 6.4.8 that there exists $G : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ such that $\operatorname{curl} G = F$. As explained in Remark 6.4.9, we may seek $G : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ such that

$$G_1(x, y, z) = \int_0^z F_2(x, y, s) \, ds + \psi(x, y) = \int_0^z (y - 3x^2y) \, ds + \psi(x, y)$$

$$\begin{aligned}
&= yz - 3x^2yz + \psi(x, y) , \\
G_2(x, y, z) &= - \int_0^z F_1(x, y, s) \, ds + \phi(x, y) = - \int_0^z (x^3 + ys) \, ds + \phi(x, y) \\
&= -x^3z - \frac{yz^2}{2} + \phi(x, y) \quad \text{and} \\
G_3(x, y, z) &= 0
\end{aligned}$$

with

$$\begin{aligned}
\frac{\partial G_2}{\partial x} - \frac{\partial G_1}{\partial y} &= F_3 \Rightarrow -3x^2z + \frac{\partial \phi}{\partial x} - z + 3x^2z - \frac{\partial \psi}{\partial y} \\
&= 4y^2 - z \Rightarrow \frac{\partial \phi}{\partial x}(x, y) - \frac{\partial \psi}{\partial y}(x, y) = 4y^2 .
\end{aligned}$$

If we choose $\phi(x, y) = 4xy^2$ and $\psi(x, y) = 0$ for all (x, y) , then we get that $G(x, y, z) = \left(yz - 3x^2yz, -x^3z - \frac{yz^2}{2} + 4xy^2, 0 \right)$.

d) Since

$$\operatorname{div} F(x, y, z) = \frac{\partial}{\partial x}(xy + z) + \frac{\partial}{\partial y}(xz) + \frac{\partial}{\partial z}(-yz + x) = y - y = 0$$

for all $(x, y, z) \in \mathbb{R}^3$ and $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is continuously differentiable on the convex set $\mathbb{R}^3$, it follows from Proposition 6.4.8 that there exists $G : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ such that $\operatorname{curl} G = F$. As explained in Remark 6.4.9, we may seek $G : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ such that

$$\begin{aligned}
G_1(x, y, z) &= \int_0^z F_2(x, y, s) \, ds + \psi(x, y) = \int_0^z xs \, ds + \psi(x, y) = \frac{xz^2}{2} + \psi(x, y) , \\
G_2(x, y, z) &= - \int_0^z F_1(x, y, s) \, ds + \phi(x, y) = - \int_0^z (xy + s) \, ds + \phi(x, y) \\
&= -xyz - \frac{z^2}{2} + \phi(x, y) \quad \text{and} \\
G_3(x, y, z) &= 0
\end{aligned}$$

with

$$\frac{\partial G_2}{\partial x} - \frac{\partial G_1}{\partial y} = F_3 \Rightarrow -yz + \frac{\partial \phi}{\partial x} - \frac{\partial \psi}{\partial y} = -yz + x \Rightarrow \frac{\partial \phi}{\partial x}(x, y) - \frac{\partial \psi}{\partial y}(x, y) = x .$$

If we choose $\phi(x, y) = \frac{x^2}{2}$ and $\psi(x, y) = 0$ for all (x, y) , then we get that $G(x, y, z) = \left(\frac{xz^2}{2}, -xyz - \frac{z^2}{2} + \frac{x^2}{2}, 0 \right)$.

e) Since

$$\operatorname{div} F(x, y, z) = \frac{\partial}{\partial x}(xe^{-x^2z^2} - 6x) + \frac{\partial}{\partial y}(5y + 2z) + \frac{\partial}{\partial z}(z - ze^{-x^2z^2})$$

$$= (e^{-x^2z^2} - 2x^2z^2e^{-x^2z^2} - 6) + 5 + (1 - e^{-x^2z^2} + 2x^2z^2e^{-x^2z^2}) = 0$$

for all $(x, y, z) \in \mathbb{R}^3$ and $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is continuously differentiable on the convex set $\mathbb{R}^3$, it follows from Proposition 6.4.8 that there exists $G : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ such that $\text{curl } G = F$. As explained in Remark 6.4.9, we may seek $G : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ such that

$$\begin{aligned} G_1(x, y, z) &= - \int_0^y F_3(x, s, z) \, ds + \psi(x, z) = - \int_0^y (z - ze^{-x^2z^2}) \, ds + \psi(x, z) \\ &= -yz + yze^{-x^2z^2} + \psi(x, z), \\ G_2(x, y, z) &= 0 \quad \text{and} \\ G_3(x, y, z) &= \int_0^y F_1(x, s, z) \, ds + \phi(x, z) = \int_0^y (xe^{-x^2z^2} - 6x) \, ds + \phi(x, z) \\ &= xye^{-x^2z^2} - 6xy + \phi(x, z) \end{aligned}$$

with

$$\begin{aligned} \frac{\partial G_1}{\partial z} - \frac{\partial G_3}{\partial x} &= F_2 \Rightarrow -y + ye^{-x^2z^2} - 2x^2yz^2e^{-x^2z^2} + \frac{\partial \psi}{\partial z} \\ &\quad - ye^{-x^2z^2} + 2x^2yz^2e^{-x^2z^2} + 6y - \frac{\partial \phi}{\partial x} = 5y + 2z \\ &\Rightarrow \frac{\partial \psi}{\partial z}(x, z) - \frac{\partial \phi}{\partial x}(x, z) = 2z. \end{aligned}$$

If we choose $\psi(x, z) = z^2$ and $\phi(x, z) = 0$ for all (x, z) , then we get that $G(x, y, z) = (-yz + yze^{-x^2z^2} + z^2, 0, xye^{-x^2z^2} - 6xy)$. $\heartsuit$

Question 6.53

Consider the vector field $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ defined by $F(x, y, z) = (e^x \sin(y), e^x \cos(y), z^2)$. Evaluate the integral $\int_C F \cdot d\mathbf{s}$ where C is the curve with the parametric representation $\sigma(t) = (t^{1/3}, t^5, e^{t^{1/2}})$ for $0 \leq t \leq 1$. We assume that the parametric representation of C is consistent with the orientation on C .

Solution

Since

$$\text{curl } F = \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ e^x \sin(y) & e^x \cos(y) & z^2 \end{pmatrix} = \mathbf{0}$$

for all $(x, y, z) \in \mathbb{R}^3$ and F is a continuously differentiable vector field on the convex set $\mathbb{R}^3$, we get from Proposition 6.4.5 that F is conservative and we get from Proposition 6.2.10 that there exists a function $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ such that $\nabla f = F$. Hence

$$\begin{aligned} \int_C F \cdot d\mathbf{s} &= \int_C \nabla f \cdot d\mathbf{s} = \int_0^1 (\nabla f)(\sigma(t)) \sigma'(t) \, dt \\ &= \int_0^1 \frac{d}{dt} f(\sigma(t)) \, dt = f(\sigma(1)) - f(\sigma(0)) = f(1.1.e) - f(0, 0, 1). \end{aligned}$$

We only have to find f . The function f must satisfy the relations

$$\frac{\partial f}{\partial x} = e^x \sin(y) \quad , \quad \frac{\partial f}{\partial y} = e^x \cos(y) \quad \text{and} \quad \frac{\partial f}{\partial z} = z^2 . \quad (6.6.17)$$

From the first equation in (6.6.17), we get that $f(x, y, z) = e^x \sin(y) + \phi(y, z)$. If we substitute this expression for f in the second equation in (6.6.17), then we get that $e^x \cos(y) + \frac{\partial \phi}{\partial y} = e^x \cos(y)$. Thus $\frac{\partial \phi}{\partial y} = 0$ and we conclude that ϕ is only a function of z . If we substitute $f(x, y, z) = e^x \sin(y) + \phi(z)$ in the last equation in (6.6.17), then we get that $\frac{d\phi}{dz} = z^2$. Thus $\phi(z) = \frac{z^3}{3}$. We may ignore the constant of integration because we are going to evaluate f between two points later and the constant will cancel out. We have found that $f(x, y, z) = e^x \sin(y) + \frac{z^3}{3}$. Hence $\int_C F \cdot d\mathbf{s} = f(1, 1, e) - f(0, 0, 1) = e \sin(1) + (e^3 - 1)/3$. $\heartsuit$

Question 6.54

Let C be the path in $\mathbb{R}^3$ defined by the segments from $(0, 0, 0)$ to $(1, 1, 1)$, from $(1, 1, 1)$ to $(1, 1, 0)$, from $(1, 1, 0)$ to $(1, 0, 0)$, from $(1, 0, 0)$ to $(2, 2, 2)$, from $(2, 2, 2)$ to $(2, 2, 0)$, from $(2, 2, 0)$ to $(2, 0, 0)$, ..., from $(100, 100, 0)$ to $(100, 0, 0)$ in the given order. Compute $\int_C 2xyz \, dx + x^2z \, dy + x^2y \, dz$.

Question 6.55

Compute the integral $\int_C (2y - 3z) \, dx + (2x + z) \, dy + (y - 3x) \, dz$ where C is the curve with the parametric representation $\sigma(t) = (t^{1/3}, \sqrt{t}e^{-t}, e^{\cos(t\pi)})$ for $0 \leq t \leq 1$. We assume that the parametric representation of C is consistent with the orientation on C .

Question 6.56

Let $F(x, y, z) = -\frac{GMm}{(x^2 + y^2 + z^2)^{3/2}}(x, y, z)$. This is the gravitational force field of a object of mass M on a particle of mass m . Show that the work done by F on the particle as it moves from (x_0, y_0, z_0) to (x_1, y_1, z_1) is a function of $\|(x_0, y_0, z_0)\|$ and $\|(x_1, y_1, z_1)\|$ only. Note that the work to move a particle from (x_0, y_0, z_0) to (x_1, y_1, z_1) is given by the integral of F along the path from (x_0, y_0, z_0) to (x_1, y_1, z_1) .

Part II

Differential Geometry

Chapter 7

Cohomology

We have seen in Theorem 4.5.12 that closed differential k -forms defined on a star-shaped domain are exact. In this chapter, we show that the property that closed differential k -forms are exact is determined by the domain of the differential k -form only.

Depending of the researcher to whom we talk, the topic of cohomology is either part of algebraic topology, differential topology or differential geometry. This is another illustration that the boundaries between different fields of mathematics are often fluid and blur.

The presentation in this section is inspired by the presentation of cohomology in [10, 27]. The first section introduce some of the important tools that we will use in our study of cohomology.

To simplify the presentation in this chapter, we will only consider smooth manifolds; namely, manifolds of class C^∞ . Similarly, all the differential forms will be of class C^∞ . These restrictions may often be reduced to C^j with j larger than the dimension of the manifold.

7.1 Preliminaries

Definition 7.1.1

A k -dimensional smooth manifold S is **(smoothly) contractible to a point** if there exist a map $H : S \times [0, 1] \rightarrow S$ of class C^∞ ¹ and a point $\tilde{\mathbf{u}} \in S$ such that $H(\mathbf{u}, 1) = \mathbf{u}$ and $H(\mathbf{u}, 0) = \tilde{\mathbf{u}}$ for all $\mathbf{u} \in S$.

It is easy to show that $\mathbb{R}^n$ is contractible to the origin. It suffices to use the map $H : \mathbb{R}^n \times [0, 1] \rightarrow \mathbb{R}^n$ defined by $H(\mathbf{x}, t) = t\mathbf{x}$.

¹To follow our convention about function of class C^∞ on a closed interval, we assume that there exists an open set $V \supset [0, 1]$ such that H can be extended to a C^∞ function from $S \times V$ to S .

Remark 7.1.2

Recall that the **direct sum** of two linear spaces M and N over an integral domain R (e.g. $\mathbb{R}$), denoted $M \oplus N$, is the linear space $M \times N$ with the addition defined by $(m_1, n_1) + (m_2, n_2) = (m_1 + m_2, n_1 + n_2)$ for $m_1, m_2 \in M$ and $n_1, n_2 \in N$, and the product by a scalar defined by $c(m, n) = (cm, cn)$ for $m \in M$, $n \in N$ and $c \in R$. In particular, if M and N are two subspaces of a vector space E such that $\dim(M) + \dim(N) = \dim(E)$ and $M \cap N = \{\mathbf{0}\}$, then we may write $E = M \oplus N$ by identifying (m, n) to $m + n$ because every element of E has a unique representation of the form $m + n$ with $m \in M$ and $n \in N$. If $M, N, \tilde{M}$ and $\tilde{N}$ are four linear spaces, and $f : M \rightarrow \tilde{M}$ and $g : N \rightarrow \tilde{N}$ are two linear maps, then $(f, g) : M \oplus N \rightarrow \tilde{M} \oplus \tilde{N}$, defined by $(f, g)(m, n) = (f(m), g(n))$ for all $(m, n) \in M \times N$, is a linear map. ♠

Let S be a k -dimensional smooth manifold and $J \subset \mathbb{R}$ be an open interval. Then $S \times J$ with is a $(k + 1)$ -dimensional smooth manifold. In particular, if $\{(W_\alpha, U_\alpha, \phi_\alpha)\}_{\alpha \in A}$ is an atlas for S , then $\{(W_\alpha \times J, U_\alpha \times J, (\phi_\alpha, \text{Id}))\}_{\alpha \in A}$ is an atlas for $S \times J$ ². We also have that $T_{(\mathbf{u}, x)}(S \times J) = T_{\mathbf{u}}S \oplus T_xJ$ for all $(\mathbf{u}, x) \in S \times J$, where $T_{\mathbf{u}}S \cong \{((\mathbf{u}, x), (\mathbf{y}, 0)) : (\mathbf{u}, \mathbf{y}) \in T_{\mathbf{u}}S\}$ and $T_xJ \cong \{((\mathbf{u}, x), (\mathbf{0}, z)) : (x, z) \in T_xJ\}$.

Proposition 7.1.3

Suppose that S is a k -dimensional smooth manifold and $J \subset \mathbb{R}$ is an open interval. Any differential q -form ω on $S \times J$ can be expressed as $\omega = \omega_1 + dx \wedge \omega_2$ where

1. ω_1 is a differential q -form on $S \times J$ such that $\omega_1(\mathbf{w}, x)((\mathbf{u}, x), (\mathbf{v}_1, y_1), \dots, ((\mathbf{u}, x), (\mathbf{v}_q, y_q))) = 0$ if $\mathbf{v}_i = 0$ for some $1 \leq i \leq q$,
2. ω_2 is a differential $(q - 1)$ -form on $S \times J$ such that $\omega_2(\mathbf{w}, x)((\mathbf{u}, x), (\mathbf{v}_1, y_1), \dots, ((\mathbf{u}, x), (\mathbf{v}_{q-1}, y_{q-1}))) = 0$ if $\mathbf{v}_i = 0$ for some $1 \leq i \leq q - 1$, and
3. dx is the differential 1-form on $S \times J$ defined by $dx(\mathbf{u}, x)((\mathbf{u}, x), (\mathbf{v}, y)) = y$ for all $(\mathbf{u}, x) \in S \times J$ and $((\mathbf{u}, x), (\mathbf{v}, y)) \in T_{(\mathbf{u}, x)}(S \times J)$.

Proof.

We consider a local chart $(W \times J, U \times J, (\phi, \text{Id}))$ of $S \times J$.

Let $\{\psi_j(\mathbf{u})\}_{1 \leq j \leq k} \subset (T_{\mathbf{u}}S)^*$ be the dual basis associated to the basis $\{(\mathbf{u}, (D\phi)(\phi^{-1}(\mathbf{u}))e_j)\}_{1 \leq j \leq k}$ of $T_{\mathbf{u}}S$. Consider

$$\begin{aligned} \tilde{\psi}_j(\mathbf{u}, x) : T_{(\mathbf{u}, x)}(S \times J) &\rightarrow \mathbb{R} \\ ((\mathbf{u}, x), (\mathbf{v}, y)) &\mapsto \psi_j(\mathbf{u})(\mathbf{u}, \mathbf{v}) \end{aligned}$$

for $1 \leq j \leq k$ and

$$\tilde{\psi}_{k+1}(\mathbf{u}, x) : T_{(\mathbf{u}, x)}(S \times J) \rightarrow \mathbb{R}$$

²The map $(\phi_\alpha, \text{Id}) : (W_i \times J) \rightarrow (U_i \times J)$ is obviously defined as $(\phi_i, \text{Id})(\mathbf{w}, x) = (\phi_i(\mathbf{w}), x)$ for all $(\mathbf{w}, x) \in W_i \times J$.

$$((\mathbf{u}, x), (\mathbf{v}, y)) \mapsto dx(x)(x, y) = y$$

We have that $\{\tilde{\psi}_j(\mathbf{u}, x)\}_{1 \leq j \leq k+1} \subset (T_{(\mathbf{u}, x)}(S \times J))^*$ is the dual basis associated to the basis $\{((\mathbf{u}, x), (D\phi(\phi^{-1}(\mathbf{u}))\mathbf{e}_j, 0))\}_{1 \leq j \leq k} \cup \{((\mathbf{u}, x), (\mathbf{0}, 1))\}$ of $T_{(\mathbf{u}, x)}(S \times J) = T_{\mathbf{u}}S \oplus T_xJ$.

It follows from Theorem 4.1.15 that

$$\begin{aligned} \omega &= \sum_{1 \leq i_1 < i_2 < \dots < i_q \leq k+1} \omega_{i_1, i_2, \dots, i_q} \tilde{\psi}_{i_1} \wedge \tilde{\psi}_{i_2} \wedge \dots \wedge \tilde{\psi}_{i_q} \\ &= \underbrace{\sum_{1 \leq i_1 < i_2 < \dots < i_q \leq k} \omega_{i_1, i_2, \dots, i_q} \tilde{\psi}_{i_1} \wedge \tilde{\psi}_{i_2} \wedge \dots \wedge \tilde{\psi}_{i_q}}_{=\omega_1} \\ &\quad + \tilde{\psi}_{k+1} \wedge \underbrace{\left((-1)^{q-1} \sum_{1 \leq i_1 < i_2 < \dots < i_{q-1} \leq k} \omega_{i_1, i_2, \dots, i_{q-1}, k+1} \tilde{\psi}_{i_1} \wedge \tilde{\psi}_{i_2} \wedge \dots \wedge \tilde{\psi}_{i_{q-1}} \right)}_{=\omega_2} \end{aligned}$$

on $U \times J$, where the functions $\omega_{i_1, i_2, \dots, i_q} : U \times J \rightarrow J$ and $\omega_{i_1, i_2, \dots, i_{q-1}, k+1} : U \times J \rightarrow J$ are of class C^∞ . From the definition of $\tilde{\psi}_j$ for $1 \leq j \leq k$, it is clear that ω_1 satisfies the condition (1) in the statement of the proposition and ω_2 satisfies the condition (2). $\tilde{\psi}_{k+1}$ is the differential 1-form dx defined in the statement of the proposition.

Note that the local representation of ω_1 and ω_2 may change for a different local chart but their action will not. As for dx , it is invariant with respect to the local charts $(W \times J, U \times J, (\phi, \text{Id}))$ chosen. ■

The inclusion $\iota_x : S \rightarrow S \times J$ is defined by $\iota_x(\mathbf{u}) = (\mathbf{u}, x)$ for all $\mathbf{u} \in S$ and $x \in J$. If ω is a differential q -form on $S \times J$, then $\iota_x^*(\omega)$ is the differential q -form on S defined by

$$\begin{aligned} (\iota_x^*(\omega))(\mathbf{u})((\mathbf{u}, \mathbf{v}_1), \dots, (\mathbf{u}, \mathbf{v}_q)) &= \iota_x^*(\omega(\mathbf{u}, x))((\mathbf{u}, \mathbf{v}_1), \dots, (\mathbf{u}, \mathbf{v}_q)) \\ &= \omega(\mathbf{u}, x)((\mathbf{u}, x), (\mathbf{v}_1, 0), \dots, (\mathbf{u}, x), (\mathbf{v}_q, 0)) \end{aligned}$$

for all $x \in J$, $\mathbf{u} \in S$ and $(\mathbf{u}, \mathbf{v}_i) \in T_{\mathbf{u}}S$ with $1 \leq i \leq q$.

Remark 7.1.4

Consider a local chart $(W \times J, U \times J, (\phi, \text{Id}))$ of $S \times J$ where (W, U, ϕ) is a local chart of S . We have that $\iota_x : S \rightarrow S \times J$ for $x \in J$ and its local representation $\tilde{\iota}_x : W \rightarrow W \times J$ satisfy the relation $(\phi, \text{Id}) \circ \tilde{\iota}_x = \iota_x \circ \phi$; namely, we have the following commutative diagram.

$$\begin{array}{ccc} S & \xrightarrow{\iota_x} & S \times J \\ \phi \uparrow & & \uparrow (\phi, \text{Id}) \\ W & \xrightarrow{\tilde{\iota}_x} & W \times J \end{array}$$

Since $(\iota_x \circ \phi)(\mathbf{w}) = (\phi(\mathbf{w}), x)$, we get that $\tilde{\iota}_x(\mathbf{w}) = (\mathbf{w}, x)$ for all $\mathbf{w} \in W$. Since $D\tilde{\iota}_x(\mathbf{w})(\mathbf{y}) = (\mathbf{y}, 0)$ and

$$(\tilde{\iota}_x)_*(\mathbf{w}, \mathbf{y}) = (\tilde{\iota}_x(\mathbf{w}), D\tilde{\iota}_x(\mathbf{w})(\mathbf{y})) = ((\mathbf{w}, x), (\mathbf{y}, 0)) \in T_{\mathbf{w}, x}(W \times J),$$

we get from

$$(\iota_x)_* \circ \phi_* = (\iota_x \circ \phi)_* = ((\phi, \text{Id}) \circ \tilde{\iota}_x)_* = (\phi, \text{Id})_* \circ (\tilde{\iota}_x)_*$$

that

$$\begin{aligned} (\iota_x)_*(\phi_*(\mathbf{w}, \mathbf{y})) &= (\phi, \text{Id})_*((\tilde{\iota}_x)_*(\mathbf{w}, \mathbf{y})) = (\phi, \text{Id})_*((\mathbf{w}, x), (\mathbf{y}, 0)) \\ &= ((\phi(\mathbf{w}), x), (D\phi(\mathbf{w})\mathbf{y}, 0)) \end{aligned}$$

for all $\mathbf{w} \in W$ and $\mathbf{y} \in \mathbb{R}^k$. This translates into $(\iota_x)_*(\mathbf{u}, \mathbf{v}) = ((\mathbf{u}, x), (\mathbf{v}, 0)) \in T_{\mathbf{u},x}(S \times J)$ for $\mathbf{u} \in U$ and $(\mathbf{u}, \mathbf{v}) \in T_{\mathbf{u}}S$. ♠

From now on, we assume that $J \subset \mathbb{R}$ is an open interval containing $[0, 1]$.

We define an operator F that maps differential q -forms on $S \times J$ to differential $(q-1)$ -forms on S as it follows. Suppose that ω is a differential q -form on $S \times J$. From Proposition 7.1.3, we may express ω as $\omega = \omega_1 + dx \wedge \omega_2$. We define $F(\omega)$ as

$$\begin{aligned} F(\omega)(\mathbf{u})((\mathbf{u}, \mathbf{v}_1), \dots, (\mathbf{u}, \mathbf{v}_{q-1})) &= \int_0^1 (\iota_x)^*(\omega_2)(\mathbf{u})((\mathbf{u}, \mathbf{v}_1), \dots, (\mathbf{u}, \mathbf{v}_{q-1})) dx \\ &= \int_0^1 \omega_2(\mathbf{u}, x)((\iota_x)_*(\mathbf{u}, \mathbf{v}_1), \dots, (\iota_x)_*(\mathbf{u}, \mathbf{v}_{q-1})) dx \\ &= \int_0^1 \omega_2(\mathbf{u}, x)((\mathbf{u}, x), (\mathbf{v}_1, 0), \dots, (\mathbf{u}, x), (\mathbf{v}_{q-1}, 0)) dx \end{aligned} \quad (7.1.1)$$

for all $\mathbf{u} \in S$ and $(\mathbf{u}, \mathbf{v}_1) \in T_{\mathbf{u}}S$.

Proposition 7.1.5

Suppose that S is a k -dimensional smooth manifold and that $J \subset \mathbb{R}$ is an open interval containing $[0, 1]$. If ω is a differential q -form on $S \times \mathbb{R}$, then $\iota_1^*(\omega) - \iota_0^*(\omega) = d(F(\omega)) + F(d\omega)$.

Proof.

From Proposition 7.1.3, we may express ω as $\omega = \omega_1 + dx \wedge \omega_2$, where ω_1 satisfies the condition (1) in the statement of the proposition and ω_2 satisfies the condition (2).

Let $(W \times J, U \times J, (\phi, \text{Id}))$ be a local chart of $S \times J$. We have

$$\begin{aligned} (\phi, \text{Id})^*(\omega) &= (\phi, \text{Id})^*(\omega_1) + (\phi, \text{Id})^*(dx) \wedge (\phi, \text{Id})^*(\omega_2) \\ &= \underbrace{\sum_{1 \leq i_1 < i_2 < \dots < i_q \leq k} g_{i_1, i_2, \dots, i_q} d\tilde{w}_{i_1} \wedge d\tilde{w}_{i_2} \wedge \dots \wedge d\tilde{w}_{i_q}}_{=\tilde{\omega}_1} \\ &\quad + dx \wedge \underbrace{\sum_{1 \leq i_1 < i_2 < \dots < i_{q-1} \leq k} h_{i_1, i_2, \dots, i_{q-1}} d\tilde{w}_{i_1} \wedge d\tilde{w}_{i_2} \wedge \dots \wedge d\tilde{w}_{i_{q-1}}}_{=\tilde{\omega}_2} \end{aligned} \quad (7.1.2)$$

for some continuously differentiable functions $g_{i_1, i_2, \dots, i_q} : W \rightarrow \mathbb{R}$ and $h_{i_1, i_2, \dots, i_{q-1}} : W \rightarrow \mathbb{R}$, where $d\tilde{w}_j(\mathbf{w}, x)((\mathbf{w}, x), (\mathbf{v}, y)) = dw_j(\mathbf{w})(\mathbf{w}, \mathbf{v}) = v_j$ and $dx(\mathbf{w}, x)((\mathbf{w}, x), (\mathbf{v}, y)) = y$ for $1 \leq j \leq k$ and $((\mathbf{w}, x), (\mathbf{v}, y)) \in T_{(\mathbf{w}, x)}(W \times J) \cong T_{\mathbf{w}}W \oplus T_xJ$.

Moreover, we have

$$\begin{aligned} & (\phi^*(F(\omega)))(\mathbf{w})((\mathbf{w}, \mathbf{v}_1), \dots, (\mathbf{w}, \mathbf{v}_{q-1})) \\ &= F(\omega)(\phi(\mathbf{w}))((\phi(\mathbf{w}), D\phi(\mathbf{w})\mathbf{v}_1), \dots, (\phi(\mathbf{w}), D\phi(\mathbf{w})\mathbf{v}_{q-1})) \\ &= \int_0^1 \omega_2(\phi(\mathbf{w}), x) \left(((\phi(\mathbf{w}), x), (D\phi(\mathbf{w})\mathbf{v}_1, 0)), \dots, ((\phi(\mathbf{w}), x), (D\phi(\mathbf{w})\mathbf{v}_{q-1}, 0)) \right) dx \\ &= \int_0^1 (\phi, \text{Id})^*(\omega_2)(\mathbf{w}, x) \left(((\mathbf{w}, x), (\mathbf{v}_1, 0)), \dots, ((\mathbf{w}, x), (\mathbf{v}_{q-1}, 0)) \right) dx \\ &= F((\phi, \text{Id})^*(\omega))(\mathbf{w})((\mathbf{w}, \mathbf{v}_1), \dots, (\mathbf{w}, \mathbf{v}_{q-1})) \end{aligned}$$

for all $\mathbf{w} \in W$ and $(\mathbf{w}, \mathbf{v}_i) \in T_{\mathbf{w}}W$ with $1 \leq i \leq q-1$ ³. Similarly.

$$\begin{aligned} & (\phi^*(\iota_x^*(\omega)))(\mathbf{w})((\mathbf{w}, \mathbf{v}_1), \dots, (\mathbf{w}, \mathbf{v}_{q-1})) \\ &= \iota_x^*(\omega)(\phi(\mathbf{w}))((\phi(\mathbf{w}), D\phi(\mathbf{w})\mathbf{v}_1), \dots, (\phi(\mathbf{w}), D\phi(\mathbf{w})\mathbf{v}_{q-1})) \\ &= \omega(\phi(\mathbf{w}), x) \left(((\phi(\mathbf{w}), x), (D\phi(\mathbf{w})\mathbf{v}_1, 0)), \dots, ((\phi(\mathbf{w}), x), (D\phi(\mathbf{w})\mathbf{v}_{q-1}, 0)) \right) \\ &= (\phi, \text{Id})^*(\omega)(\mathbf{w}, x) \left(((\mathbf{w}, x), (\mathbf{v}_1, 0)), \dots, ((\mathbf{w}, x), (\mathbf{v}_{q-1}, 0)) \right) \\ &= \tilde{\iota}_x^*((\phi, \text{Id})^*(\omega))(\mathbf{w})((\mathbf{w}, \mathbf{v}_1), \dots, (\mathbf{w}, \mathbf{v}_{q-1})) \end{aligned}$$

for all $\mathbf{w} \in W$ and $(\mathbf{w}, \mathbf{v}_i) \in T_{\mathbf{w}}W$ with $1 \leq i \leq q-1$, where $\tilde{\iota}_x(\mathbf{w}) = (\mathbf{w}, x)$ for all $\mathbf{w} \in W$ and $x \in J$ as we have seen in Remark 7.1.4. Therefore, it suffices to prove that

$$\tilde{\iota}_1^*((\phi, \text{Id})^*(\omega)) - \tilde{\iota}_0^*((\phi, \text{Id})^*(\omega)) = d(F((\phi, \text{Id})^*(\omega))) + F(d((\phi, \text{Id})^*(\omega))) .$$

to get

$$\begin{aligned} \phi^*(\iota_1^*(\omega)) - \phi^*(\iota_0^*(\omega)) &= \tilde{\iota}_1^*((\phi, \text{Id})^*(\omega)) - \tilde{\iota}_0^*((\phi, \text{Id})^*(\omega)) \\ &= d(F((\phi, \text{Id})^*(\omega))) + F(d((\phi, \text{Id})^*(\omega))) \\ &= d(\phi^*(F(\omega))) + F((\phi, \text{Id})^*(d\omega)) = \phi^*(dF(\omega)) + \phi^*(F(d\omega)) . \end{aligned}$$

Using (7.1.2) and the linearity of the maps F and ι_x , we have to prove that

$$\tilde{\iota}_1^*(\tilde{\omega}_1) - \tilde{\iota}_0^*(\tilde{\omega}_1) = d(F(\tilde{\omega}_1)) + F(d(\tilde{\omega}_1)) \quad (7.1.3)$$

and

$$\tilde{\iota}_1^*(dx \wedge \tilde{\omega}_2) - \tilde{\iota}_0^*(dx \wedge \tilde{\omega}_2) = d(F(dx \wedge \tilde{\omega}_2)) + F(d(dx \wedge \tilde{\omega}_2)) . \quad (7.1.4)$$

³We should probably have used $\tilde{F}$ to denote the local representation of F . Hopefully, the context should be enough to determine when we are referring to the local representation of F .

i) We first prove (7.1.3). Since $F(\tilde{\omega}_1) = 0$, we have that $d(F(\tilde{\omega}_1)) = 0$. Since

$$\begin{aligned} d(\tilde{\omega}_1) = & \sum_{1 \leq i_1 < i_2 < \dots < i_q \leq k} \left(\sum_{j=1}^k \frac{\partial g_{i_1, i_2, \dots, i_q}}{\partial w_j} d\tilde{w}_j \right) d\tilde{w}_{i_1} \wedge d\tilde{w}_{i_2} \wedge \dots \wedge d\tilde{w}_{i_q} \\ & + dx \wedge \sum_{1 \leq i_1 < i_2 < \dots < i_q \leq k} \frac{\partial g_{i_1, i_2, \dots, i_q}}{\partial x} d\tilde{w}_{i_1} \wedge d\tilde{w}_{i_2} \wedge \dots \wedge d\tilde{w}_{i_q}, \end{aligned}$$

we get

$$\begin{aligned} & F(d(\tilde{\omega}_1))(\mathbf{w})((\mathbf{w}, \mathbf{v}_1), \dots, (\mathbf{w}, \mathbf{v}_q)) \\ &= \sum_{1 \leq i_1 < i_2 < \dots < i_q \leq k} \left(\int_0^1 \frac{\partial g_{i_1, i_2, \dots, i_q}}{\partial x}(\mathbf{w}, x) (d\tilde{w}_{i_1} \wedge d\tilde{w}_{i_2} \wedge \dots \wedge d\tilde{w}_{i_q})(\mathbf{w}, x) \right. \\ & \quad \left. \left(((\mathbf{w}, x), (\mathbf{v}_1, 0)), \dots, ((\mathbf{w}, x), (\mathbf{v}_q, 0)) \right) dx \right) \\ &= \sum_{1 \leq i_1 < i_2 < \dots < i_q \leq k} \left(\int_0^1 \frac{\partial g_{i_1, i_2, \dots, i_q}}{\partial x}(\mathbf{w}, x) (dw_{i_1} \wedge dw_{i_2} \wedge \dots \wedge dw_{i_q})(\mathbf{w}) \right. \\ & \quad \left. ((\mathbf{w}, \mathbf{v}_1), \dots, (\mathbf{w}, \mathbf{v}_q)) dx \right) \\ &= \sum_{1 \leq i_1 < i_2 < \dots < i_q \leq k} \left(\int_0^1 \frac{\partial g_{i_1, i_2, \dots, i_q}}{\partial x}(\mathbf{w}, x) dx \right) (dw_{i_1} \wedge dw_{i_2} \wedge \dots \wedge dw_{i_q})(\mathbf{w}) \\ & \quad ((\mathbf{w}, \mathbf{v}_1), \dots, (\mathbf{w}, \mathbf{v}_q)) \\ &= \sum_{1 \leq i_1 < i_2 < \dots < i_q \leq k} (g_{i_1, i_2, \dots, i_q}(\mathbf{w}, 1) - g_{i_1, i_2, \dots, i_q}(\mathbf{w}, 0)) (dw_{i_1} \wedge dw_{i_2} \wedge \dots \wedge dw_{i_q})(\mathbf{w}) \\ & \quad ((\mathbf{w}, \mathbf{v}_1), \dots, (\mathbf{w}, \mathbf{v}_q)) \\ &= \sum_{1 \leq i_1 < i_2 < \dots < i_q \leq k} (g_{i_1, i_2, \dots, i_q}(\mathbf{w}, 1) - g_{i_1, i_2, \dots, i_q}(\mathbf{w}, 0)) (d\tilde{w}_{i_1} \wedge d\tilde{w}_{i_2} \wedge \dots \wedge d\tilde{w}_{i_q})(\mathbf{w}, x) \\ & \quad \left(((\mathbf{w}, x), (\mathbf{v}_1, 0)), \dots, ((\mathbf{w}, x), (\mathbf{v}_q, 0)) \right) \\ &= (\tilde{l}_1^*(\tilde{\omega}_1))(\mathbf{w})((\mathbf{w}, \mathbf{v}_1), \dots, (\mathbf{w}, \mathbf{v}_q)) - (\tilde{l}_0^*(\tilde{\omega}_1))(\mathbf{w})((\mathbf{w}, \mathbf{v}_1), \dots, (\mathbf{w}, \mathbf{v}_q)). \end{aligned}$$

Thus (7.1.3) is satisfied.

ii) We now prove (7.1.4). We have that $\tilde{l}_1^*(dx \wedge \tilde{\omega}_2) = \tilde{l}_0^*(dx \wedge \tilde{\omega}_2) = 0$ because $dx((\mathbf{w}, x), (\mathbf{v}, 0)) = 0$ for all $((\mathbf{w}, x), (\mathbf{v}, 0)) \in T_{(\mathbf{w}, x)}(W \times J)$. Since

$$\begin{aligned} & F(dx \wedge \tilde{\omega}_2)(\mathbf{w})((\mathbf{w}, \mathbf{v}_1), \dots, (\mathbf{w}, \mathbf{v}_q)) \\ &= \sum_{1 \leq i_1 < i_2 < \dots < i_{q-1} \leq k} \left(\int_0^1 h_{i_1, i_2, \dots, i_{q-1}}(\mathbf{w}, x) (d\tilde{w}_{i_1} \wedge d\tilde{w}_{i_2} \wedge \dots \wedge d\tilde{w}_{i_{q-1}})(\mathbf{w}, x) \right. \\ & \quad \left. \left(((\mathbf{w}, x), (\mathbf{v}_1, 0)), \dots, ((\mathbf{w}, x), (\mathbf{v}_{q-1}, 0)) \right) dx \right) \end{aligned}$$

$$= \sum_{1 \leq i_1 < i_2 < \dots < i_{q-1} \leq k} \left(\int_0^1 h_{i_1, i_2, \dots, i_{q-1}}(\mathbf{w}, x) dx \right) (dw_{i_1} \wedge dw_{i_2} \wedge \dots \wedge dw_{i_{q-1}})(\mathbf{w}) \\ ((\mathbf{w}, \mathbf{v}_1), \dots, (\mathbf{w}, \mathbf{v}_{q-1}))$$

for all $\mathbf{w} \in W$ and $(\mathbf{w}, \mathbf{v}_i) \in T_{\mathbf{w}}W$ with $1 \leq i \leq k$, we get

$$F(dx \wedge \tilde{\omega}_2) = \sum_{1 \leq i_1 < i_2 < \dots < i_{q-1} \leq k} \left(\int_0^1 h_{i_1, i_2, \dots, i_{q-1}}(\cdot, x) dx \right) dw_{i_1} \wedge dw_{i_2} \wedge \dots \wedge dw_{i_{q-1}}.$$

It follows that

$$d(F(dx \wedge \tilde{\omega}_2)) = \sum_{1 \leq i_1 < i_2 < \dots < i_{q-1} \leq k} \sum_{j=1}^k \left(\int_0^1 \frac{\partial h_{i_1, i_2, \dots, i_{q-1}}(\cdot, x)}{\partial w_j} dx \right) \\ dw_j \wedge dw_{i_1} \wedge dw_{i_2} \wedge \dots \wedge dw_{i_{q-1}}. \quad (7.1.5)$$

Since

$$d(dx \wedge \omega_2) = -dx \wedge \left(\sum_{1 \leq i_1 < i_2 < \dots < i_{q-1} \leq k} \sum_{j=1}^k \left(\frac{\partial h_{i_1, i_2, \dots, i_{q-1}}}{\partial w_j} \right) d\tilde{w}_j \wedge d\tilde{w}_{i_1} \wedge d\tilde{w}_{i_2} \wedge \dots \wedge d\tilde{w}_{i_{q-1}} \right),$$

we get

$$F(d(dx \wedge \omega_2))(\mathbf{w})((\mathbf{w}, \mathbf{v}_1), \dots, (\mathbf{w}, \mathbf{v}_q)) \\ = - \sum_{1 \leq i_1 < i_2 < \dots < i_{q-1} \leq k} \sum_{j=1}^k \left(\int_0^1 \frac{\partial h_{i_1, i_2, \dots, i_{q-1}}(\mathbf{w}, x)}{\partial w_j} (d\tilde{w}_j \wedge d\tilde{w}_{i_1} \wedge d\tilde{w}_{i_2} \wedge \dots \wedge d\tilde{w}_{i_{q-1}})(\mathbf{w}, x) \right. \\ \left. \left(((\mathbf{w}, x), (\mathbf{v}_1, 0)), \dots, ((\mathbf{w}, x), (\mathbf{v}_q, 0)) \right) dx \right) \\ = - \sum_{1 \leq i_1 < i_2 < \dots < i_{q-1} \leq k} \sum_{j=1}^k \left(\int_0^1 \frac{\partial h_{i_1, i_2, \dots, i_{q-1}}(\mathbf{w}, x)}{\partial w_j} dx \right) \\ (dw_j \wedge dw_{i_1} \wedge dw_{i_2} \wedge \dots \wedge dw_{i_{q-1}})(\mathbf{w})((\mathbf{w}, \mathbf{v}_1), \dots, (\mathbf{w}, \mathbf{v}_q))$$

for all $\mathbf{w} \in W$ and $(\mathbf{w}, \mathbf{v}_i) \in T_{\mathbf{w}}W$ with $1 \leq i \leq k$. Thus

$$F(d(dx \wedge \omega_2)) = - \sum_{1 \leq i_1 < i_2 < \dots < i_{q-1} \leq k} \sum_{j=1}^k \left(\int_0^1 \frac{\partial h_{i_1, i_2, \dots, i_{q-1}}(\cdot, x)}{\partial w_j} dx \right) \\ dw_j \wedge dw_{i_1} \wedge dw_{i_2} \wedge \dots \wedge dw_{i_{q-1}}. \quad (7.1.6)$$

It follows from (7.1.5) and (7.1.6) that $d(F(dx \wedge \tilde{\omega}_2)) + F(d(dx \wedge \omega_2)) = 0$. This completes the proof of (7.1.4). ■

We may now generalize Theorem 4.5.12.

Theorem 7.1.6 (Poincaré Lemma)

Let S be a k -dimensional smooth manifold which is smoothly contractible to the point $\tilde{\mathbf{u}} \in S$. Suppose that ω is a closed differential q -form on S with $q > 0$. Then there exists a differential $(q-1)$ -form η on S such that $d\eta = \omega$.

Proof.

Suppose that $H : S \times [0, 1] \rightarrow S$ is a map of class C^∞ such that $H(\mathbf{u}, 1) = \mathbf{u}$ and $H(\mathbf{u}, 0) = \tilde{\mathbf{u}}$ for all $\mathbf{u} \in S$.

Since $\text{Id} = H \circ \iota_1 : S \rightarrow S$ and $H \circ \iota_0 : S \rightarrow \{\tilde{\mathbf{u}}\}$, we get that $\iota_1^*(H^*(\omega)) = (H \circ \iota_1)^*(\omega) = \omega$ and $\iota_0^*(H^*(\omega)) = (H \circ \iota_0)^*(\omega) = 0$. Since $d\omega = 0$, we also get that $d(H^*(\omega)) = H^*(d\omega) = 0$. It follows from Proposition 7.1.5 that

$$\omega = \iota_1^*(H^*(\omega)) - \iota_0^*(H^*(\omega)) = d(F(H^*(\omega))) + F(d(H^*(\omega))) = d(F(H^*(\omega))) .$$

We may take $\eta = F(H^*(\omega))$. ■

The following result coming out of Proposition 7.1.5 will be quite useful for our presentation of cohomology in the next section.

Proposition 7.1.7

Suppose that $\omega = g dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_q}$ with $1 \leq i_1 < i_2 < \dots < i_q \leq k$ and $q > 0$ is a differential q -form on $\mathbb{R}^k$, and that η is the differential $(q-1)$ -form on $\mathbb{R}^k$ defined by

$$\eta(\mathbf{x}) = \left(\sum_{j=1}^q (-1)^{j-1} \left(\int_0^1 r^{q-1} g(r\mathbf{x}) dr \right) x_{i_j} \right) (dx_{i_1} \wedge \dots \wedge \widehat{dx_{i_j}} \wedge \dots \wedge dx_{i_q})(\mathbf{x})$$

for $\mathbf{x} \in \mathbb{R}^k$. Then $d\eta = \omega$.

Proof.

The reader should review Remark 4.5.10 before studying this proof.

Let $H : \mathbb{R}^k \times \mathbb{R} \rightarrow \mathbb{R}^k$ be the map defined by $H(\mathbf{x}, r) = r\mathbf{x}$ for $(\mathbf{x}, r) \in \mathbb{R}^k \times \mathbb{R}$. If we apply Proposition 7.1.5 to $H^*(\omega)$, then we get

$$\iota_1^*(H^*(\omega)) - \iota_0^*(H^*(\omega)) = d(F(H^*(\omega))) + F(d(H^*(\omega))) . \quad (7.1.7)$$

Let $Z : \mathbb{R}^k \rightarrow \mathbb{R}^k$ be the function defined by $Z(\mathbf{x}) = \mathbf{0}$ for all $\mathbf{x} \in \mathbb{R}^k$. We have

$$\iota_1^*(H^*(\omega)) = (H \circ \iota_1)^*(\omega) = \text{Id}^*(\omega) = \omega \quad (7.1.8)$$

and

$$\iota_0^*(H^*(\omega)) = (H \circ \iota_0)^*(\omega) = Z^*(\omega) = 0 \quad (7.1.9)$$

on $\mathbb{R}^k$.

Since

$$\begin{aligned} H^*(dx_j)(\mathbf{x}, r)((\mathbf{x}, r), (\mathbf{v}, z)) &= dx_j(H(\mathbf{x}, r))(H_*((\mathbf{x}, r), (\mathbf{v}, z))) \\ &= dx_j(r\mathbf{x})(H(\mathbf{x}, r), D H(\mathbf{x}, r)(\mathbf{v}, z)) = dx_j(r\mathbf{x})(r\mathbf{x}, r\mathbf{v} + z\mathbf{x}) = rv_j + zx_j \\ &= r dx_j(\mathbf{x})(\mathbf{x}, \mathbf{v}) + x_j dr(r)(r, z) \end{aligned}$$

for all $((\mathbf{x}, r), (\mathbf{v}, z)) \in T_{(\mathbf{x}, r)}(\mathbb{R}^k \times \mathbb{R})$ and $(\mathbf{x}, r) \in \mathbb{R}^k \times \mathbb{R}$ with $1 \leq j \leq k$, we may write that

$$H^*(dx_j)(\mathbf{x}, r) = r d\tilde{x}_j(\mathbf{x}, r) + x_j d\tilde{r}(\mathbf{x}, r) ,$$

where $d\tilde{x}_j(\mathbf{x}, r)((\mathbf{x}, r), (\mathbf{v}, z)) = dx_j(\mathbf{x})(\mathbf{x}, \mathbf{v}) = v_j$ and $d\tilde{r}(\mathbf{x}, r)((\mathbf{x}, r), (\mathbf{v}, z)) = dr(r)(r, z) = z$ for all $((\mathbf{x}, r), (\mathbf{v}, z)) \in T_{(\mathbf{x}, r)}(\mathbb{R}^k \times \mathbb{R})$ and $(\mathbf{x}, r) \in \mathbb{R}^k \times \mathbb{R}$ with $1 \leq j \leq k$. Hence

$$\begin{aligned} H^*(\omega) &= (g \circ H)(r d\tilde{x}_{i_1} + x_{i_1} d\tilde{r}) \wedge \dots \wedge (r d\tilde{x}_{i_q} + x_{i_q} d\tilde{r}) \\ &= r^q g(r\mathbf{x}) d\tilde{x}_{i_1} \wedge \dots \wedge d\tilde{x}_{i_q} + r^{q-1} g(r\mathbf{x}) \left(\sum_{j=1}^q -1^{j-1} x_{i_j} d\tilde{r} \wedge d\tilde{x}_{i_1} \wedge \dots \wedge \widehat{d\tilde{x}_{i_j}} \wedge \dots \wedge d\tilde{x}_{i_q} \right) \\ &= r^q g(r\mathbf{x}) d\tilde{x}_{i_1} \wedge \dots \wedge d\tilde{x}_{i_q} + d\tilde{r} \wedge \left(\sum_{j=1}^q -1^{j-1} r^{q-1} g(r\mathbf{x}) x_{i_j} d\tilde{x}_{i_1} \wedge \dots \wedge \widehat{d\tilde{x}_{i_j}} \wedge \dots \wedge d\tilde{x}_{i_q} \right) \end{aligned}$$

on $\mathbb{R}^k \times \mathbb{R}$. It follows from (7.1.1) that

$$F(H^*(\omega)) = \sum_{j=1}^q -1^{j-1} \left(\int_0^1 r^{q-1} g(r\mathbf{x}) dr \right) x_{i_j} dx_{i_1} \wedge \dots \wedge \widehat{dx_{i_j}} \wedge \dots \wedge dx_{i_q} = \eta \quad (7.1.10)$$

on $\mathbb{R}^k$. Moreover $d(H^*(\omega)) = H^*(d\omega) = H^*(0) = 0$ because ω is a differential k -form on $\mathbb{R}^k$. Thus

$$F(d(H^*(\omega))) = 0 . \quad (7.1.11)$$

If we substitute (7.1.8), (7.1.9), (7.1.10) and (7.1.11) into (7.1.7), then we get $d\eta = \omega$. ■

7.2 Cohomology Modules

We will start with a more restrictive type of cohomology before moving on to the more general type. This will prepare us for the more general type of cohomology by giving us some useful results for the study of the more general type of cohomology. This more restrictive type of cohomology is really important in its own right.

7.2.1 Cohomology Modules with Compact Supports

Definition 7.2.1

Let S be a k -dimensional smooth manifold. Two closed differential q -forms with compact support, ω_1 and ω_2 , on S are **cohomologous** if there exists a differential $(q-1)$ -form with compact support η on S such that $\omega_1 - \omega_2 = d\eta$. We write $\omega_1 \sim \omega_2$.

In the previous definition, the condition that the support of the differential $(q-1)$ -form η be compact is stronger than the condition that the support of $d\eta$ be compact. The differential q -form $d\eta$ may have a compact support without the differential $(q-1)$ -form η having a compact support.

The relation defined above is an equivalence relation.

1. If ω is a closed differential q -form with compact support, then obviously $\omega \sim \omega$.
2. If ω_1 and ω_2 are two closed differential q -forms with compact support, then again it is obvious that $\omega_1 \sim \omega_2$ if and only if $\omega_2 \sim \omega_1$.
3. If ω_i for $1 \leq i \leq 3$ are three closed differential q -forms with compact support such that $\omega_1 \sim \omega_2$ and $\omega_2 \sim \omega_3$, then $\omega_1 \sim \omega_3$. To verify this claim, let η_1 be a differential $(q-1)$ -form with compact support such that $\omega_2 - \omega_1 = d\eta_1$ and η_2 be a differential $(q-1)$ -form with compact support such that $\omega_3 - \omega_2 = d\eta_2$. Then $\omega_3 - \omega_1 = (\omega_3 - \omega_2) + (\omega_2 - \omega_1) = d\eta_2 + d\eta_1 = d(\eta_2 + \eta_1)$ where $\eta_1 + \eta_2$ is a differential $(q-1)$ -form with compact support because the union of two compact sets is a compact set.

Definition 7.2.2

Let S be a k -dimensional smooth manifold. The **q^{th} (de Rham) Cohomology Module with compact support** of S , denoted $H_c^q(S)$, is the set of all equivalent classes of closed differential q -forms on S with compact support with respect to the equivalence relation given in Definition 7.2.1. The equivalence class for a closed differential q -form with compact support ω on S is called a **cohomology class** and is denoted $[\omega]$.

Example 7.2.3

We have that the dimension of $H_c^k(\mathbb{R}^k)$ is greater than 0.

To prove this statement, let $\omega = f dx_1 \wedge dx_2 \wedge \dots \wedge dx_k$ where $f : \mathbb{R}^k \rightarrow [0, \infty[$ is a continuous function with compact support such that $f(\mathbf{x}) > 0$ for some $\mathbf{x} \in \mathbb{R}^k$. Then $\int_{\mathbb{R}^k} \omega = \int_{\mathbb{R}^k} f(\mathbf{x}) d\mathbf{x} > 0$. Suppose that there is a differential $(k-1)$ -form η on S with compact support such that $d\eta = \omega$. Given $r > 0$ such that $\text{supp } \eta \subset B_r(\mathbf{0})$, we then have from Stokes' theorem that $\int_{\mathbb{R}^k} \omega = \int_{B_r(\mathbf{0})} \omega = \int_{B_r(\mathbf{0})} d\eta = \int_{\partial B_r(\mathbf{0})} \eta = 0$. This is a contradiction. ♣

Example 7.2.4

If S is an oriented k -dimensional smooth manifold, then the dimension of $H_c^k(S)$ is greater than 0.

To prove this statement, let (W, U, ϕ) be an orientation preserving local chart of S such that $U \cap \partial S = \emptyset$. Let ω be a differential k -form on S with compact support in U such that its local representation is $\phi^*(\omega) = f dw_1 \wedge dw_2 \wedge \dots \wedge dw_k$ where $f : W \rightarrow [0, \infty[$ is a continuous function with compact support and $f(\mathbf{w}) > 0$ for some $\mathbf{w} \in W$. Then $\int_S \omega = \int_W f(\mathbf{w}) d\mathbf{w} > 0$. Suppose that there is a differential $(k-1)$ -form η on S with compact support such that $d\eta = \omega$. We must have that $\text{supp } \eta \subset \text{supp } \omega \subset U$ because $d\eta(\mathbf{u}) = \omega(\mathbf{u}) = 0$ for $\mathbf{u} \in S \setminus \text{supp } \omega$. Thus $\eta(\mathbf{u})$ is constant for $\mathbf{u} \in S \setminus \text{supp } \omega$. Since η has a compact support, we must have that $\eta(\mathbf{u}) = 0$ for $\mathbf{u} \in S \setminus \text{supp } \omega$. In particular, we have that $\text{supp } \eta \cap \partial S \subset U \cap \partial S = \emptyset$. It follows from Stokes' theorem that $\int_S \omega = \int_S d\eta = \int_{\partial S} \eta = 0$. This is a contradiction. ♣

Proposition 7.2.5

Suppose that S is a k -dimensional smooth manifold. Then $H_c^q(S)$ is a vector space.

Proof.

i) We prove that $[\omega_1 + \omega_2] = [\omega_1] + [\omega_2]$ is well defined for all closed differential q -forms with compact support ω_1 and ω_2 on S .

We first note that $\omega_1 + \omega_2$ is closed because $d(\omega_1 + \omega_2) = d\omega_1 + d\omega_2 = 0$. Moreover, $\omega_1 + \omega_2$ has a compact support because the union of two compact sets is a compact set.

Suppose that $\tilde{\omega}_i \sim \omega_i$ for $i = 1, 2$. Then there exist two differential $(q-1)$ -forms with compact support, η_1 and η_2 , on S such that $\tilde{\omega}_i - \omega_i = d\eta_i$ for $i = 1, 2$. It follows that $(\tilde{\omega}_1 + \tilde{\omega}_2) - (\omega_1 + \omega_2) = (\tilde{\omega}_2 - \omega_2) + (\tilde{\omega}_1 - \omega_1) = d\eta_2 + d\eta_1 = d(\eta_2 + \eta_1)$. Moreover, $\eta_1 + \eta_2$ has a compact support because the union of two compact sets is a compact set. Thus $(\tilde{\omega}_1 + \tilde{\omega}_2) \sim (\omega_1 + \omega_2)$ and $[\tilde{\omega}_1 + \tilde{\omega}_2] = [\omega_1 + \omega_2]$.

ii) We prove that $a[\omega] = [a\omega]$ is well defined for all differential q -form with compact support ω on S and $a \in \mathbb{R}$.

We first note that $a\omega$ is closed because $d(a\omega) = a d\omega = 0$. Moreover, $a\omega$ has a compact support because $\text{supp}(a\omega) = \text{supp } \omega$ if $a \neq 0$.

Suppose that $\tilde{\omega} \sim \omega$. Then there exists a differential $(q-1)$ -form with compact support η on S such that $\tilde{\omega} - \omega = d\eta$. It follows that $a\tilde{\omega} + a\omega = a(\tilde{\omega} - \omega) = a d\eta = d(a\eta)$ where $a\eta$ is a differential $(q-1)$ -form with compact support. Thus $a\tilde{\omega} \sim a\omega$ and $[a\tilde{\omega}] = [a\omega]$. ■

Remark 7.2.6

Let S be a k -dimensional smooth manifold. The null element $[0]$ in $H_c^q(S)$ is the collection of all differential q -forms with compact support ω on S such that $\omega = d\eta$ for some differential $(q-1)$ -form with compact support η on S . This follows from the fact that ω in the equivalence class $[0]$ means that $\omega = \omega - 0 = d\eta$ for some differential $(q-1)$ -form with compact support on S . ♠

We want to compute the q^{th} cohomology modules with compact support for the connected and oriented smooth manifolds. To do that will however require some work.

Proposition 7.2.7

Suppose that S is a connected and oriented k -dimensional smooth manifold without boundary. Then $H_c^k(S) \cong \mathbb{R}$ if and only if the map $M : H_c^k(S) \rightarrow \mathbb{R}$ defined by $M([\omega]) = \int_S \omega$ is an isomorphism.

Proof.

We first have to prove that the map M is well defined. Suppose that $[\omega_1] = [\omega_2]$ for two differential k -forms ω_1 and ω_2 on S with compact support. Then there exists a differential $(k-1)$ -form on S with compact support such that $\omega_1 - \omega_2 = d\eta$. It follows from Stokes' theorem that $\int_S \omega_1 - \int_S \omega_2 = \int_S (\omega_1 - \omega_2) = \int_S d\eta = \int_{\partial S} \eta = 0$. So $\int_S \omega_1 = \int_S \omega_2 = 0$ and thus $M([\omega_1]) = M([\omega_2])$.

Suppose that M is an isomorphism between $H_c^k(S)$ and $\mathbb{R}$, and that ω_1 and ω_2 are two differential k -forms on S with compact support. Choose $a \in \mathbb{R}$ such that $M([\omega_1]) = aM([\omega_2])$. Then $M([\omega_1] - a[\omega_2]) = M([\omega_1]) - aM([\omega_2]) = 0$. Since M is an isomorphism, we have that $[\omega_1] - a[\omega_2] = [0]$; namely, $[\omega_1] = a[\omega_2]$.

Conversely, suppose that $H_c^k(S) \cong \mathbb{R}$ and that $M([\omega_1]) = 0$ with $[\omega_1] \neq [0]$, where ω_1 is a differential k -form on S with compact support. For every differential k -form ω_2 on S with compact support, we have that $[\omega_2] = a[\omega_1]$ for some $a \in \mathbb{R}$. Thus $M([\omega_2]) = aM([\omega_1]) = 0$. It follows that M is null on $H_c^k(S)$ which is not the case. An example of a differential k -form ω with $M([\omega]) > 0$ is given in Example 7.2.4. ■

Lemma 7.2.8

Suppose that S is a connected set and that $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$ is an open cover of S with $U_\alpha \subset S$ for all $\alpha \in I$. Given any two sets $U_a, U_b \in \mathcal{U}$, there exists a finite collection $\{U_{\alpha_j}\}_{0 \leq j \leq J}$ such that $U_{\alpha_0} = U_a$, $U_{\alpha_J} = U_b$ and $U_{\alpha_j} \cap U_{\alpha_{j-1}} \neq \emptyset$ for $1 \leq j \leq J$.

Proof.

Let $\mathcal{B}$ be the set of all $U_c \in \mathcal{U}$ such that there exists a finite collection $\{U_{\alpha_j}\}_{0 \leq j \leq J_c}$ satisfying $U_{\alpha_0} = U_a$, $U_{\alpha_{J_c}} = U_c$ and $U_{\alpha_j} \cap U_{\alpha_{j-1}} \neq \emptyset$ for $1 \leq j \leq J_c$.

Let $C = \bigcup_{U_c \in \mathcal{B}} U_c$. We have that C is an open set because it is the union of open sets. We now show that $S \setminus C = \emptyset$ and thus that $U_b \in \mathcal{B}$.

Suppose that $S \setminus C \neq \emptyset$. Given any $\mathbf{x} \in S \setminus C$, there exists $U_d \in \mathcal{U}$ such that $\mathbf{x} \in U_d$ because $\mathcal{U}$ is an open cover of S . We have that $U_d \cap C = \emptyset$ otherwise U_d would be in $\mathcal{B}$ and $\mathbf{x}$ in C . To be more precise, $U_d \cap C \neq \emptyset$ implies that $U_d \cap U_c \neq \emptyset$ for some $U_c \in \mathcal{B}$. Thus U_d may be added to the finite collection $\{U_{\alpha_j}\}_{0 \leq j \leq J_c}$ joining U_a to U_c with $U_{\alpha_j} \cap U_{\alpha_{j-1}} \neq \emptyset$ for $1 \leq j \leq J_c$ to create a finite collection $\{U_{\alpha_j}\}_{0 \leq j \leq J_c+1}$ joining U_a to U_d with $U_{\alpha_{J_c+1}} = U_d$ and $U_{\alpha_j} \cap U_{\alpha_{j-1}} \neq \emptyset$ for $1 \leq j \leq J_c+1$. Thus $S \setminus C$ is a non-empty open set. Since S is connected, this contradicts the fact that S cannot be the union of two disjoint non-empty open sets. ■

Theorem 7.2.9

If S is an connected and oriented k -dimensional smooth manifold without boundary, then $H_c^k(S) \cong \mathbb{R}$.

Proof.

i) We first prove that $H_c^1(\mathbb{R}) \cong \mathbb{R}$. To do so, we prove that the map $M : H_c^1(\mathbb{R}) \rightarrow \mathbb{R}$ defined in Proposition 7.2.7 is an isomorphism. Suppose that $\omega = g dx$ is a differential 1-form on S with compact support such that $\int_{\mathbb{R}} \omega = 0$. In particular, $g : \mathbb{R} \rightarrow \mathbb{R}$ is a continuous function with compact support.

There exists a function $f : \mathbb{R} \rightarrow \mathbb{R}$ such that $\omega = df$. In fact, it suffices to take $f(x) = \int_0^x g(y) dy$ for $x \in \mathbb{R}$.

Since $\omega = df = f' dx$ has a compact support, there exists an interval $[a, b] \subset \mathbb{R}$ such that $f'(x) = 0$ for $x \in \mathbb{R} \setminus [a, b]$. Thus f is constant on $\mathbb{R} \setminus [a, b]$. Suppose that $f(x) = c$ for $x < a$ and $f(x) = d$ for $x > b$. Then $c = d$ because $0 = \int_{\mathbb{R}} \omega = \int_{\mathbb{R}} f'(x) dx = \int_a^b f'(x) dx = f(b) - f(a) = d - c$.

We get that $\omega = dh$ for $h = f - c$ with $\text{supp } h \subset [a, b]$. Thus $[\omega] = [0]$ in $H^1(\mathbb{R})$.

ii) We prove that $H_c^{k-1}(S) \cong \mathbb{R}$ for all connected and oriented $(k-1)$ -dimensional smooth manifolds without boundary implies that $H_c^k(\mathbb{R}^k) \cong \mathbb{R}$. Again, we do so by proving that $M : H_c^k(\mathbb{R}^k) \rightarrow \mathbb{R}$ defined in Proposition 7.2.7 is an isomorphism.

Suppose that ω is a differential k -form on $\mathbb{R}^k$ with compact support and that $\int_{\mathbb{R}^k} \omega = 0$. Since $\text{supp } \omega$ is compact, there exists $b > 0$ such that $\text{supp } \omega \subset B_b(\mathbf{0})$.

Let $h : \mathbb{R}^k \rightarrow \mathbb{R}^k$ be the diffeomorphism defined by $h(\mathbf{x}) = b\mathbf{x}$ for $\mathbf{x} \in \mathbb{R}^k$. We have that $h^*(\omega)$ is a differential k -form on $\mathbb{R}^k$ with $\text{supp } h^*(\omega) \subset B_1(\mathbf{0})$. If we prove that there exists a differential $(k-1)$ -form η on $\mathbb{R}^k$ with compact support such that $h^*(\omega) = d\eta$, then $\omega = (h^{-1})^*(d\eta) = d((h^{-1})^*(\eta))$ where $(h^{-1})^*(\eta)$ is a differential $(k-1)$ -form η on $\mathbb{R}^k$ with compact support. Thus $[\omega] = [0]$ in $H_c^k(\mathbb{R}^k)$.

Therefore, we may assume without loss of generality that $\omega = g dx_1 \wedge dx_2 \wedge \dots \wedge dx_k$ with $\text{supp } g \subset B_1(\mathbf{0})$.

Let η be the differential $(k-1)$ -form on $\mathbb{R}^k$ defined by

$$\eta(\mathbf{x}) = \sum_{j=1}^k (-1)^{j-1} \left(\int_0^1 r^{k-1} g(r\mathbf{x}) dr \right) x_j \left(dx_1 \wedge \dots \wedge \widehat{dx_j} \wedge \dots \wedge dx_k \right)(\mathbf{x}) \quad (7.2.1)$$

for $\mathbf{x} \in \mathbb{R}^k$. We have from Proposition 7.1.7 that $d\eta = \omega$. However, η may not have a compact support. We have to slightly modify η to get a differential $(k-1)$ -form on S with compact support whose derivative is ω .

If we use the substitution $t = \|\mathbf{x}\|r$ in the integral in (7.2.1) and use Proposition 5.7.14, then we get

$$\begin{aligned} \eta(\mathbf{x}) &= \sum_{j=1}^k (-1)^{j-1} \left(\int_0^{\|\mathbf{x}\|} t^{k-1} g\left(\frac{t}{\|\mathbf{x}\|} \mathbf{x}\right) dt \right) \|\mathbf{x}\|^{-k} x_j \left(dx_1 \wedge \dots \wedge \widehat{dx_j} \wedge \dots \wedge dx_k \right)(\mathbf{x}) \\ &= \left(\int_0^1 t^{k-1} g\left(\frac{t}{\|\mathbf{x}\|} \mathbf{x}\right) dt \right) \left(\sum_{j=1}^k (-1)^{j-1} \|\mathbf{x}\|^{-k} x_j \left(dx_1 \wedge \dots \wedge \widehat{dx_j} \wedge \dots \wedge dx_k \right)(\mathbf{x}) \right) \\ &= \left(\int_0^1 t^{k-1} g\left(\frac{t}{\|\mathbf{x}\|} \mathbf{x}\right) dt \right) \rho^*(\nu)(\mathbf{x}) \end{aligned}$$

for $\mathbf{x} \in \mathbb{R}^k \setminus \{\mathbf{0}\}$ where $\rho : \mathbb{R}^k \setminus \{\mathbf{0}\} \rightarrow S^{k-1}$ is the retraction defined by $\rho(\mathbf{x}) = \|\mathbf{x}\|^{-1} \mathbf{x}$ and ν is the volume element on S^{k-1} . We have used the fact that $g(\mathbf{x}) = 0$ for $\|\mathbf{x}\| > 1$ to modify the upper limit of integration.

Let $G : \mathbb{R}^k \rightarrow \mathbb{R}$ be the function defined by $G(\mathbf{x}) = \int_0^1 t^{k-1} g(t\mathbf{x}) dt$ for $\mathbf{x} \in \mathbb{R}^k$. We have that $\eta(\mathbf{x}) = G(\rho(\mathbf{x})) \rho^*(\nu)(\mathbf{x}) = \rho^*(G\nu)(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^k \setminus \{\mathbf{0}\}$. Moreover, we get from Proposition 5.7.15 that

$$\int_{S^{k-1}} G\nu = \int_{B_1(\mathbf{0})} g dx_1 \wedge dx_2 \wedge \dots \wedge dx_n = \int_{\mathbb{R}^k} \omega = 0.$$

Using the induction hypothesis with S^{k-1} , we have that $M : H_c^{k-1}(S^{k-1}) \rightarrow \mathbb{R}$ defined by $M([\omega]) = \int_{S^{k-1}} \omega$ is an isomorphism. So $[G\nu] = [0]$ in $H_c^{k-1}(S^{k-1})$; namely, there exists a differential $(k-2)$ -form ξ on S^{k-1} with compact support such that $G\nu = d\xi$ on S^{k-1} . It follows that $\eta = \rho^*(G\nu) = \rho^*(d\xi) = d(\rho^*(\xi))$ on $\mathbb{R}^k \setminus \{\mathbf{0}\}$.

Choose $0 < \epsilon < \delta < 1$ and let $h : \mathbb{R}^k \rightarrow \mathbb{R}$ be a function of class C^∞ such that $h(\mathbf{x}) = 1$ for $\|\mathbf{x}\| \geq \delta$ and $h(\mathbf{x}) = 0$ for $\|\mathbf{x}\| < \epsilon$. Then $h\rho^*(\xi)$ is a differential $(k-2)$ -form on $\mathbb{R}^k \setminus \{\mathbf{0}\}$. Since $h\rho^*(\xi)$ is null in a neighbourhood of the origin, we may assume that $h\rho^*(\xi)$ is defined on $\mathbb{R}^k$ by setting $h\rho^*(\xi)(\mathbf{0}) = 0$.

We have that $\eta - d(h\rho^*(\xi))$ is a differential $(k-2)$ -form on $\mathbb{R}^k$ such that $d(\eta - d(h\rho^*(\xi))) = d\eta = \omega$ and

$$\begin{aligned} (\eta - d(h\rho^*(\xi)))(\mathbf{x}) &= \eta(\mathbf{x}) - dh(\mathbf{x}) \wedge (h\rho^*(\xi))(\mathbf{x}) - h(\mathbf{x}) d(\rho^*(\xi))(\mathbf{x}) \\ &= (1 - h(\mathbf{x}))\eta(\mathbf{x}) - dh(\mathbf{x}) \wedge (h\rho^*(\xi))(\mathbf{x}) = 0 \end{aligned}$$

for $\|\mathbf{x}\| > \delta$ because $h(\mathbf{x}) = 1$ for $\|\mathbf{x}\| > \delta$. So $\eta - d(h\rho^*(\xi))$ has a compact support in $B_1(\mathbf{0})$.

This shows that $[\omega] = [0]$ in $H_c^k(\mathbb{R}^k)$.

iii) We prove that $H_c^k(\mathbb{R}^k) \cong \mathbb{R}$ implies $H_c^k(S) \cong \mathbb{R}$ for all connected and oriented k -dimensional smooth manifolds without boundary.

Let $\mathcal{A} = \{(W_\alpha, U_\alpha, \phi_\alpha)\}_{\alpha \in A}$ be an atlas of orientation preserving local charts on S and $\{\psi_j\}_{j \in \mathbb{N}^+}$ be a partition of unity subordinate to $\mathcal{A}$. Without any loss of generality, we may assume that the W_i are open ball in $\mathbb{R}^k$ centred at the origin ⁴.

Choose a local chart $(W_{\tau_0}, U_{\tau_0}, \phi_{\tau_0})$ with $\tau_0 \in A$ and a differential k -form ω on S with compact support in U_{τ_0} such that $[\omega] \neq [0]$ in $H_c^k(S)$; namely, ω is not exact. Given another differential k -form $\tilde{\omega}$ on S with compact support, we want to show that there exists $a \in \mathbb{R}$ such that $[\tilde{\omega}] = a[\omega]$; namely, there exist $a \in \mathbb{R}$ and a differential $(k-1)$ form η on S with compact support such that $\tilde{\omega} - a\omega = d\eta$.

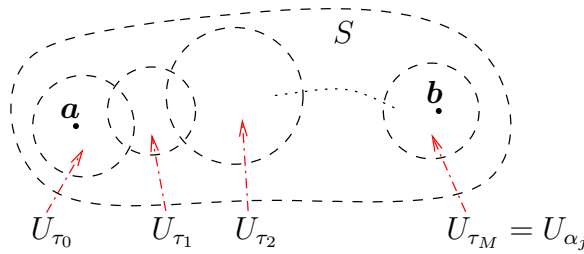
Since $\tilde{\omega}$ has a compact support, it follows from Proposition 5.6.5 that there exists a finite subset J of $\mathbb{N}^+$ such that $\tilde{\omega} = \sum_{j \in J} \psi_j \tilde{\omega}$. If we prove that for each $j \in J$ we can find $a_j \in \mathbb{R}$ and a differential $(k-1)$ -form η_j on S with compact support such that $\psi_j \tilde{\omega} - a_j \omega = d\eta_j$, then we will have

$$\tilde{\omega} = \sum_{j \in J} \psi_j \tilde{\omega} = \underbrace{\left(\sum_{j \in J} a_j \right)}_{=a} \omega + d \underbrace{\left(\sum_{j \in J} \eta_j \right)}_{=\eta},$$

where η is a differential $(k-1)$ -form on S with compact support because each η_j has a compact support and the finite union of compact sets is a compact set.

Therefore, we only have to prove that for each $j \in J$ we can find $a_j \in \mathbb{R}$ and a differential $(k-1)$ -form η_j on S with compact support such that $\psi_j \tilde{\omega} - a_j \omega = d\eta_j$. If $[\psi_j \tilde{\omega}] = [0]$ in $H_c^k(S)$, then $\psi_j \tilde{\omega} = d\eta_j$ for some differential $(k-1)$ -form η_j on S with compact support. Thus, there is nothing to prove. We may take $a_j = 0$. We assume that $[\psi_j \tilde{\omega}] \neq [0]$ in $H_c^k(S)$.

Using Lemma 7.2.8, we can find a finite sequence $\{U_{\tau_m}\}_{0 \leq m \leq M}$ with $\tau_m \in A$ and $U_{\tau_{m-1}} \cap U_{\tau_m} \neq \emptyset$ for $1 \leq m \leq M$ where U_{τ_0} was defined above and $U_{\tau_M} = U_{\alpha_j} \supset \text{supp } \psi_j \supset \text{supp } \psi_j \tilde{\omega}$. Such an index $\alpha_j \in A$ exists by definition of a partition of unity.



Let $\omega_0 = \omega$. For $0 \leq m < M$, choose a differential k -form ω_{m+1} on S with compact support in $U_{\tau_m} \cap U_{\tau_{m+1}}$ and such that $M([\omega_{m+1}]) = \int_S \omega_{m+1} > 0$ (see Example 7.2.4). This implies that $[\phi_{\tau_m}^*(\omega_{m+1})] \neq [0]$ in $H_c^k(\mathbb{R}^k)$ ⁵. Since we assume that $H_c^k(\mathbb{R}^k) \cong \mathbb{R}$, there

⁴For each $\mathbf{x}$, there exists an orientation preserving local chart (W, U, ϕ) on S with $\mathbf{x} \in U$. Choose a small ball $W_{\mathbf{x}} = B_r(\phi^{-1}(\mathbf{x})) \subset W$. Let $U_{\mathbf{x}} = \phi(B_{\mathbf{x}})$ and $\phi_{\mathbf{x}} = \phi$. We get a new local chart $(W_{\mathbf{x}}, U_{\mathbf{x}}, \phi_{\mathbf{x}})$ on S about $\mathbf{x} \in S$. The collection $\{(W_{\mathbf{x}}, U_{\mathbf{x}}, \phi_{\mathbf{x}})\}_{\mathbf{x} \in S}$ is an atlas on S which orientation preserving local charts such that the $W_{\mathbf{x}}$ are open ball in $\mathbb{R}^k$.

⁵Since $\text{supp } \phi_{\tau_m}^*(\omega_{m+1}) \subset W_{\tau_m}$, we may assume that $\phi_{\tau_m}^*(\omega_{m+1})$ is defined on $\mathbb{R}^k$ by setting it to 0 outside W_{τ_m} .

exists $a_{m+1} \in \mathbb{R}$ and a differential $(k-1)$ -form $\tilde{\xi}_m$ on $\mathbb{R}^k$ with compact support such that $\phi_{\tau_m}^*(\omega_{m+1}) - a_{m+1}\phi_{\tau_m}^*(\omega_m) = d\tilde{\xi}_m$. Proceeding as in the prove of (ii), since $\text{supp } \phi_{i_m}^*(\omega_{m+1})$ is a subset of the open ball W_{τ_m} centred at the origin, we may modify $\tilde{\xi}_m$ to get that $\text{supp } \tilde{\xi}_m \subset W_{\tau_m}$. We may therefore define a differential $(k-1)$ -form ξ_m on S with compact support in W_{τ_m} by $\xi_m = 0$ on $S \setminus U_{\tau_m}$ and $\phi_{\tau_m}^*(\xi_m) = \tilde{\xi}_m$ on U_{τ_m} . Hence, $d\xi_m = d(\phi_{\tau_m}^*(\xi_m)) = \phi_{\tau_m}^*(d\tilde{\xi}_m)$ and we get

$$\omega_{m+1} - a_{m+1}\omega_m = d\xi_m \quad (7.2.2)$$

for $0 \leq m < M$.

For $m = M$, we consider the differential k -form $\omega_{M+1} = \psi_j \tilde{\omega}$ which has a compact support in $U_{\tau_M} = U_{\alpha_j}$. We proceed exactly as in the previous paragraph to find $a_{M+1} \in \mathbb{R}$ and a differential $(k-1)$ -form ξ_M on S with compact support in U_{τ_M} such that

$$\omega_{M+1} - a_{M+1}\omega_M = d\xi_M. \quad (7.2.3)$$

Using induction, it follows from (7.2.2) and (7.2.3) that

$$\begin{aligned} \psi_j \tilde{\omega} &= \omega_{M+1} = a_{M+1}\omega_M + d\xi_M = a_{M+1}(a_M\omega_{M-1} + d\xi_{M-1}) + d\xi_M \\ &= a_{M+1}a_M\omega_{M-1} + (a_{M+1}d\xi_{M-1} + d\xi_M) \\ &= a_{M+1}a_M(a_{M-1}\omega_{M-2} + d\xi_{M-2}) + (a_{M+1}d\xi_{M-1} + d\xi_M) \\ &= a_{M+1}a_Ma_{M-1}\omega_{M-2} + (a_{M+1}a_Md\xi_{M-2} + a_{M+1}d\xi_{M-1} + d\xi_M) \\ &= \dots \\ &= \left(\prod_{i=1}^{M+1} a_i \right) \omega_0 + d \left(\xi_M + \sum_{i=0}^{M-1} \left(\prod_{m=i+2}^{M+1} a_m \right) \xi_i \right) \\ &= \underbrace{\left(\prod_{i=1}^{M+1} a_i \right)}_{=a_j} \omega + d \underbrace{\left(\xi_M + \sum_{i=0}^{M-1} \left(\prod_{m=i+2}^{M+1} a_m \right) \xi_i \right)}_{=\eta_j}, \end{aligned}$$

where η_j a differential $(k-1)$ -form on S with compact support because each ξ_j has a compact support and the finite union of compact sets is a compact set.

iv) We finally have that (i) and (iii) implies that $H_c^1(S) \cong \mathbb{R}$ for all connected and oriented 1-dimensional smooth manifolds without boundary. This result combined with (ii) implies that $H_c^2(\mathbb{R}^2) \cong \mathbb{R}$. Then, this result combined with (iii) implies that $H_c^2(S) \cong \mathbb{R}$ for all connected and oriented 2-dimensional smooth manifolds without boundary. And so on. ■

Example 7.2.10

It is interesting to note that $H_c^1(S^1) \cong \mathbb{R}$ could be proved directly.

We first note that all differential 1-form on S^1 are closed because there are no non-null differential 2-form on S^1 . They also all have compact support because S^1 is a compact set. Moreover, every exact differential 1-form ω on S^1 is of the form $\omega = df$ for some $f : S^1 \rightarrow \mathbb{R}$ of class C^∞ .

We have seen in Example 4.7.4 that

$$\omega = \frac{-x_2}{x_1^2 + x_2^2} dx_1 + \frac{x_1}{x_1^2 + x_2^2} dx_2$$

is a differential 1-form which is not exact. In fact, ω is the volume (or length) element on S^1 .

Suppose that μ is a differential 1-form. Let $\phi : \mathbb{R} \rightarrow \mathbb{R}^2$ be the function defined by $\phi(\theta) = (\cos(\theta), \sin(\theta))$ for all $\theta \in \mathbb{R}$. We have that $\phi^*(\mu) = f d\theta$ for some 2π -periodic function $f : \mathbb{R} \rightarrow \mathbb{R}$.

There are two cases to consider. If $\int_{[0, 2\pi]} \phi^*(\mu) = \int_0^{2\pi} f(\theta) d\theta = 0$, then $\tilde{f} : \mathbb{R} \rightarrow \mathbb{R}$ defined by $\tilde{f}(\theta) = \int_0^\theta f(t) dt$ is a 2π -periodic function of class C^∞ such that $d\tilde{f} = f d\theta$. Consider $F : S^1 \rightarrow \mathbb{R}$ defined by $F = \tilde{f} \circ (\phi|_J)^{-1}$ on $\phi(J)$ for all open interval of length less than 2π to ensure that $\phi|_J$ be one-to-one. This function is of class C^∞ on S^1 and $dF = \mu$ because $\phi^*(dF) = d(\phi^*(F)) = d\tilde{f} = f d\theta = \phi^*(\mu)$. Thus μ is exact. Hence $[\mu] = [0]$ in $H_c^1(S^1)$.

If $c = \int_{[0, 2\pi]} \phi^*(f) = \int_0^{2\pi} f(\theta) d\theta \neq 0$, then $\mu - c/(2\pi)\omega$ is exact. To prove this claim, we first note that $\phi^*(\omega) = d\theta$. Hence $\int_{[0, 2\pi]} \phi^*(\mu - c/(2\pi)\omega) = \int_0^{2\pi} (f(\theta) - c/(2\pi)) d\theta = 0$. Let $\tilde{f} : \mathbb{R} \rightarrow \mathbb{R}$ be the function of period 2π defined by $\tilde{f}(\theta) = \int_0^\theta f(t) dt - c\theta/(2\pi)$ for $0 \leq \theta \leq 2\pi$. This is a 2π -periodic function of class C^∞ such that $d\tilde{f} = f d\theta - c/(2\pi)$. Hence F defined as above is a function of class C^∞ on S^1 such that $dF = \mu - c/(2\pi)\omega$ because $\phi^*(dF) = d(\phi^*(F)) = d\tilde{f} - c\theta/(2\pi) = f d\theta - c/(2\pi) d\theta = \phi^*(\mu - c/(2\pi)\omega)$. Thus $[\mu] = [c/(2\pi)\omega]$ in $H_c^1(S^1)$.

We have an isomorphism between $H_c^1(S^1)$ and $\mathbb{R}$ given by $[a\omega] \mapsto a$ for all $a \in \mathbb{R}$. ♣

The next step should normally be to compute $H_c^q(S)$ for the connected and oriented k -dimensional smooth manifolds S and $0 \leq q < k$. We will prove the following result at the end of the next section.

Theorem 7.2.11

$$H_c^q(\mathbb{R}^k) = \{[0]\} \text{ for } 0 \leq q < k.$$

Note 7.2.12

Suppose that S is an oriented k -dimensional smooth manifold and $0 \leq q \leq k$. It is the tradition to write $H_c^q(S) = 0$ when $H_c^q(S) = \{[0]\}$. We will adopt this tradition from now on.

7.2.2 Cohomology Modules (General Case)

We now drop the condition that the differential q -forms have a compact support.

Definition 7.2.13

Let S be a k -dimensional smooth manifold. Two closed differential q -forms, ω_1 and ω_2 , on S are **cohomologous** if there exists a differential $(q-1)$ -form η on S such that $\omega_1 - \omega_2 = d\eta$. We write $\omega_1 \sim \omega_2$ ⁶.

As we had for the relation defined in Definition 7.2.1, the relation defined in Definition 7.2.13 is an equivalence relation. The justification is basically identical. We have two definitions of cohomologous forms. The context will determine which definition is to be used.

Definition 7.2.14

Let S be a k -dimensional smooth manifold. The **q^{th} (de Rham) Cohomology Module** of S , denoted $H^q(S)$, is the set of all equivalent classes of closed differential q -forms on S with respect to the equivalence relation given in Definition 7.2.13. An equivalence class for a closed differential q -form ω on S is called a **cohomology class** and is denoted $[\omega]$.

Remark 7.2.15

$H^q(S)$ is also called the **q^{th} cohomology group** because the space of differential q -form on S can be view as a free abelian group generated by a finite basis of differential q -forms. Moreover, since we are using $\mathbb{R}$ for the scalars, we may also have talked about vector spaces instead of module. However, the use of the word module is justified by the fact that the theory can be expanded to modules. ♠

A proof almost identical (and simpler) to the proof of Proposition 7.2.5 yields the following result.

Proposition 7.2.16

Suppose that S is a k -dimensional smooth manifold. Then $H^q(S)$ is a vector space.

Remark 7.2.17

Let S be a k -dimensional smooth manifold. The null element $[0]$ in $H^q(S)$ is the collection of all exact differential q -forms on S because ω in the class $[0]$ means that $\omega = \omega - 0 = d\eta$ for some differential $(q-1)$ -form on S . ♠

Example 7.2.18

⁶We use the same symbol to denote cohomologous differential forms with or without compact supports. The context should be enough to determine which cohomologous definition should be used. The readers may wonder why we did not use a different symbol. They will understand when reading the next chapter why we choose to not introduce a new symbol at this time. There will be enough different similarity symbols in the next chapter to drive someone crazy.

If S is a compact and oriented k -dimensional smooth manifold without boundary, then the dimension of $H^k(S)$ is larger than 0. To prove this claim, we only have to show that there exists a closed and non-exact differential k -form on S . This is what we did in Example 5.7.12.

We have a more precise result. Since S is compact, we get from Theorem 7.2.9 that $H^k(S) = H_c^k(S) \cong \mathbb{R}$. ♣

Example 7.2.19

We have from Example 5.7.13 that $H^{n-1}(\mathbb{R}^n \setminus \{0\})$ is of dimension greater than 0 for $n \geq 2$. ♣

Proposition 7.2.20

Suppose that S is a k -dimensional smooth manifold. The dimension of $H^0(S)$ is equal to the number of (connected) components of S .

Proof.

To justify this claim, we first note that a differential 0-form on S is a function. Let $f : S \rightarrow \mathbb{R}$ be smooth function.

If (W, U, ϕ) is a local chart of S , then $df = 0$ implies that $0 = \phi^*(df) = d(\phi^*f) = d(f \circ \phi) = \sum_{j=1}^k \frac{\partial}{\partial w_j}(f \circ \phi) dw_j = 0$ on W and so $\frac{\partial}{\partial w_j}(f \circ \phi) = 0$ on W for $1 \leq j \leq k$. Since ϕ is one-to-one, this implies that f is constant on U . Thus f is constant on open subsets of S and therefore on the components of S . Suppose that S is the union of the disjoint components S_i for $i \in I$. Let

$$\mathcal{X}_i(\mathbf{u}) = \begin{cases} 1 & \text{if } \mathbf{u} \in S_i \\ 0 & \text{if } \mathbf{u} \in S \setminus S_i \end{cases}$$

Then $\{[\mathcal{X}_i]\}_{i \in I}$ is a basis of $H^0(S)$.

It is important to note that each equivalent class $[f]$ of $H^0(S)$ consists in one element only, the function f . If g is an element of the class $[f]$, then we must have that $g - f = 0$ because there are no such thing as a differential -1 -forms η such that $d\eta = g - f$. ■

The following result is a consequence of the previous proposition.

Corollary 7.2.21

$$H^0(\mathbb{R}^k) \cong \mathbb{R}.$$

It follows from Poincaré Lemma, Theorem 7.1.6, that if S is a k -dimensional smooth manifold which is smoothly contractible to the point of S , then $H^q(S) = 0$ for $0 < q \leq k$. In particular, we get the following result.

Corollary 7.2.22

$$H^q(\mathbb{R}^k) = 0 \text{ for } 0 < q \leq k.$$

Now that we have computed the q^{th} cohomology module of $\mathbb{R}^k$ for $0 \leq q \leq k$, we would like to do the same for other smooth manifolds than $\mathbb{R}^k$. In fact, we can easily obtain the q^{th} cohomology module of the smooth manifolds that are diffeomorphic to $\mathbb{R}^k$.

Suppose that S_1 and S_2 are two k -dimensional smooth manifolds. Let $f : S_1 \rightarrow S_2$ be a function of class C^∞ . We have that f^* maps differential q -forms on S_2 to differential q -forms on S_1 . The following map is well defined.

$$\begin{aligned} f^\# : H^q(S_2) &\rightarrow H^q(S_1) \\ [\omega] &\mapsto [f^*(\omega)] \end{aligned}$$

Suppose that ω_1 and ω_2 are two differential q -forms on S_2 such that $\omega_1 \sim \omega_2$. Then there exists a differential $(q-1)$ -form η on S_2 such that $\omega_2 - \omega_1 = d\eta$. Using Proposition 5.5.6, we get that $f^*(\omega_2) - f^*(\omega_1) = f^*(\omega_2 - \omega_1) = f^*(d\eta) = d(f^*(\eta))$. Since $f^*(\eta)$ is a differential $(q-1)$ -form on S_1 , we have that $f^*(\omega_2) \sim f^*(\omega_1)$ and $[f^*(\omega_2)] = [f^*(\omega_1)]$.

Proposition 7.2.23

1. Suppose that S_1 and S_2 are two k -dimensional smooth manifolds and that $f : S_1 \rightarrow S_2$ is a function of class C^∞ . Then $f^\# : H^q(S_2) \rightarrow H^q(S_1)$ is a linear map.
2. Suppose that S_i for $1 \leq i \leq 3$ are three k -dimensional smooth manifolds and that $f_i : S_i \rightarrow S_{i+1}$ for $1 \leq i \leq 2$ are two function of class C^∞ . Then $(f_2 \circ f_1)^\# = f_1^\# \circ f_2^\#$.

We leave the proof of this proposition to the reader.

If $f : S_1 \rightarrow S_2$ is a diffeomorphism of class C^∞ between two k -dimensional smooth manifolds S_1 and S_2 , then $f^\# : H^q(S_2) \rightarrow H^q(S_1)$ is a linear isomorphism. Therefore $H^q(S_2) \cong H^q(S_1)$.

Our next goal is to compute the q^{th} cohomology modules of $\mathbb{R}^k \setminus \{0\}$ for $0 \leq q \leq k$. This will also give us the q^{th} cohomology module of S^{k-1} for $0 \leq q \leq k-1$ as we will show shortly.

Definition 7.2.24

Let S_1 and S_2 be two k -dimensional smooth manifolds and $f, g : S_1 \rightarrow S_2$ be two functions of class C^∞ . We say that f and g are **(smoothly) homotopic** if there exists a function $H : S_1 \times [0, 1] \rightarrow S_2$ of class C^∞ such that $H(\mathbf{u}, 0) = f(\mathbf{u})$ and $H(\mathbf{u}, 1) = g(\mathbf{u})$ for all $\mathbf{u} \in S_1$. The function H is called an **homotopy** between f and g .

We can reformulate Definition 7.1.1 by saying that a smooth manifold S is contractible to a point $\tilde{\mathbf{u}} \in S$ if the identity map $\text{Id} : S \rightarrow S$ and the map $g : S \rightarrow S$ defined by $g(\mathbf{u}) = \tilde{\mathbf{u}}$ for all $\mathbf{u} \in S$ are homotopic.

Proposition 7.2.25

Suppose that S_1 and S_2 are two k -dimensional smooth manifolds and that $f, g : S_1 \rightarrow S_2$ are two homotopic functions of class C^∞ . Then $f^\# = g^\#$ as maps from $H^q(S_2)$ to $H^q(S_1)$.

Proof.

Let F be the map defined in (7.1.1) that maps differential q -forms on $S_1 \times I$ to differential $(q-1)$ -forms on S_1 . Moreover, let $\iota_s : S_1 \rightarrow S_1 \times [0, 1]$ for $s \in [0, 1]$ be the inclusion defined by $\iota_s(\mathbf{u}) = (\mathbf{u}, s)$.

Since f and g are homotopic, there exists a smooth function $H : S_1 \times [0, 1] \rightarrow S_2$ such that $f = H \circ \iota_0$ and $g = H \circ \iota_1$.

Given ω , a closed differential q -form on S_2 , we have from Proposition 7.1.5 that

$$\begin{aligned} g^*(\omega) - f^*(\omega) &= (H \circ \iota_1)^*(\omega) - (H \circ \iota_0)^*(\omega) = \iota_1^*(H^*(\omega)) - \iota_0^*(H^*(\omega)) \\ &= d(F(H^*(\omega))) + F(d(H^*(\omega))) = d(F(H^*(\omega))) + F(H^*(d\omega)) = d(F(H^*(\omega))) . \end{aligned}$$

Thus $g^*(\omega) \sim f^*(\omega)$. it follows that $f^\#([\omega]) = [f^*(\omega)] = [g^*(\omega)] = g^\#([\omega])$. ■

Proposition 7.2.26

$H^q(S^{k-1}) \cong H^q(\mathbb{R}^k \setminus \{\mathbf{0}\})$ for $0 \leq q \leq k$.

Proof.

The inclusion $\iota : S^{k-1} \rightarrow \mathbb{R}^k \setminus \{\mathbf{0}\}$ defined by $\iota(\mathbf{u}) = \mathbf{u}$ and the retraction $\rho : \mathbb{R}^k \setminus \{\mathbf{0}\} \rightarrow S^{k-1}$ defined by $\rho(\mathbf{u}) = \|\mathbf{u}\|^{-1}\mathbf{u}$ satisfy $\rho \circ \iota = \text{Id} : S^{k-1} \rightarrow S^{k-1}$. However, $\iota \circ \rho : \mathbb{R}^k \setminus \{\mathbf{0}\} \rightarrow \mathbb{R}^k \setminus \{\mathbf{0}\}$ is not the identity but it is homotopic to the identity $\text{Id} : \mathbb{R}^k \setminus \{\mathbf{0}\} \rightarrow \mathbb{R}^k \setminus \{\mathbf{0}\}$. An homotopy $H : (\mathbb{R}^k \setminus \{\mathbf{0}\}) \times [0, 1] \rightarrow \mathbb{R}^k \setminus \{\mathbf{0}\}$ between $\iota \circ \rho$ and Id is given by $H(\mathbf{u}, t) = t\mathbf{u} + (1-t)(\iota \circ \rho)(\mathbf{u})$ for all $(\mathbf{u}, t) \in (\mathbb{R}^k \setminus \{\mathbf{0}\}) \times [0, 1]$. It follows from Proposition 7.2.25 that $(\iota \circ \rho)^\# = \text{Id}^\# : H^q(\mathbb{R}^k \setminus \{\mathbf{0}\}) \rightarrow H^q(\mathbb{R}^k \setminus \{\mathbf{0}\})$ for $0 \leq q \leq k$.

Hence, $\iota^\# \circ \rho^\# = (\rho \circ \iota)^\# = \text{Id}^\# : H^q(S^{k-1}) \rightarrow H^q(S^{k-1})$ and $\rho^\# \circ \iota^\# = (\iota \circ \rho)^\# = \text{Id}^\# : H^q(\mathbb{R}^k \setminus \{\mathbf{0}\}) \rightarrow H^q(\mathbb{R}^k \setminus \{\mathbf{0}\})$ for $0 \leq q \leq k$. Thus $\rho^\# : H^q(S^{k-1}) \rightarrow H^q(\mathbb{R}^k \setminus \{\mathbf{0}\})$ is the inverse of $\iota^\# : H^q(\mathbb{R}^k \setminus \{\mathbf{0}\}) \rightarrow H^q(S^{k-1})$. Therefore $\rho^\#$ provides an isomorphism between $H^q(S^{k-1})$ and $H^q(\mathbb{R}^k \setminus \{\mathbf{0}\})$. ■

The following result follows from Proposition 7.2.20 and is consistent with the previous proposition.

Corollary 7.2.27

$H^0(S^{k-1}) \cong H^0(\mathbb{R}^k \setminus \{\mathbf{0}\}) \cong \mathbb{R}$.

The next result follows directly from Proposition 7.2.26 because differential q -forms on a $(k-1)$ -dimensional manifold are null if $q > k-1$.

Corollary 7.2.28

$$H^k(\mathbb{R}^k \setminus \{0\}) \cong H^k(S^{k-1}) = 0.$$

Obviously, if S is a compact k -dimensional smooth manifold, then $H^k(S) = H_c^k(S)$. In particular, if S is a compact, connected and oriented k -dimensional smooth manifold without boundary, then we get from Theorem 7.2.9 that $H^k(S) = H_c^k(S) \cong \mathbb{R}$. The next result then follows from Proposition 7.2.26.

Corollary 7.2.29

$$H^{k-1}(\mathbb{R}^k \setminus \{0\}) \cong H^{k-1}(S^{k-1}) \cong \mathbb{R}.$$

To compute $H^q(\mathbb{R}^k \setminus \{0\}) \cong H^q(S^{k-1})$ for $0 < q < k-1$, we need to compute $H^k(S)$ for S a non-compact, connected and oriented k -dimensional smooth manifold without boundary. The proof is similar to the proof of Theorem 7.2.9. However, the fact that we are no longer limited to differential forms with compact supports complicates the proofs.

Many of the topological results that are presented below could be simplified and even skipped if we consider that all our manifolds are submanifolds of the metric space $\mathbb{R}^n$ and if we use all the properties of a metric space with a countable dense set. However, with just a little bit more work, we can present a general theory that may be useful to the reader if they wish to study more advanced topics of differential and topological geometry later.

We need the following preliminary information.

Definition 7.2.30

A topological space S is a **Hausdorff space** if, for every two distinct elements x_1 and x_2 in S , there exist disjoint open neighbourhoods U_1 and U_2 of x_1 and x_2 respectively.

All the manifolds that we consider are Hausdorff spaces because they are subsets of some Euclidean space $\mathbb{R}^n$ and their topology is induced from the topology on $\mathbb{R}^n$.

Definition 7.2.31

A Hausdorff space S is a **Lindelöf space** if every open cover of S admit a countable open subcover.

Again, every manifold S that we consider is a Lindelöf space. This result follows from the next proposition because S is a subset of some Euclidean space $\mathbb{R}^n$. Therefore, it admits a countable basis.

Proposition 7.2.32

If S is a Hausdorff space with a countable basis, then S is a Lindelöf space.

Proof.

Let $\{B_j\}_{j \in \mathbb{N}}$ be a countable basis of S ; namely, every open set in S can be expressed as the union of some B_j .

Suppose that $\{U_\alpha\}_{\alpha \in A}$ is an open cover of S . We have that each U_α is the union of some B_j for $j \in J_\alpha \subset \mathbb{N}$. For each set B_j with $j \in N = \bigcup_{\alpha \in A} J_\alpha \subset \mathbb{N}$, we associate one of the U_α containing B_j , say U_{α_j} . Since $S = \bigcup_{\alpha \in A} U_\alpha = \bigcup_{j \in N} B_j$, we have that $\{U_{\alpha_j}\}_{j \in N}$ is a (at most) countable open cover of S . ■

Definition 7.2.33

A Hausdorff space S is a **locally compact space** if, for every element x in S , there exists an open neighbourhood U of x in S such that $\overline{U}$ is compact.

Every manifold S is a locally compact space because, for every $\mathbf{u} \in S$, there exists a local chart (W, U, ϕ) such that $\mathbf{u} \in U$ and U is homeomorphic to $W \subset H_k$ which is locally compact.

Proposition 7.2.34

If S is a locally compact and Lindelöf space, then S is **σ -compact**; namely, S can be expressed as the union of at most countably many compact sets.

Proof.

For each $x \in S$, there exists an open set V_x such that $\overline{V_x}$ is compact. The collection $\{V_x\}_{x \in S}$ is an open cover of S . Since S is a Lindelöf space, there exists a countable open subcover $\{V_{x_i}\}_{i \in \mathbb{N}}$ of S . Hence $S = \bigcup_{i \in \mathbb{N}} \overline{V_{x_i}}$ where $\overline{V_{x_i}}$ is compact for all $i \in \mathbb{N}$. ■

An elementary proof of this result for connected and locally compact metric spaces is also given in [27].

Proposition 7.2.35

Suppose that S is a locally compact space. Given any compact set K and open set U of S such that $K \subset U$, there exist an open set V such that $K \subset V \subset \overline{V} \subset U$ and $\overline{V}$ is compact.

A proof of this proposition can be found in [5].

Proposition 7.2.36

Suppose that $\{B_\tau\}_{\tau \in T}$ is an open cover of a manifold S . There exists an atlas $\mathcal{J} = \{(W_j, U_j, \phi_j)\}_{j \in \mathbb{N}}$ on S such that each U_j is a subset of a B_{τ_j} for some $\tau_j \in T$ and $\{U_j\}_{j \in \mathbb{N}}$ is **locally finite**; namely, for every $\mathbf{u} \in S$, there exists an open neighbourhood $V_{\mathbf{u}}$ of $\mathbf{u}$ such that $U_j \cap V_{\mathbf{u}} = \emptyset$ for all but at most a finite number of $j \in \mathbb{N}$.

Proof.

Since S is a locally compact and Lindelöf space, we may use Proposition 7.2.34. So, we can express S as $S = \bigcup_{i \in \mathbb{N}} K_i$ where the K_i are compact sets.

From Proposition 7.2.35, there exists an open set V_0 such that $K_0 \subset V_0 \subset \bar{V}_0 \subset S$ and $\bar{V}_0$ is compact. It follows that $\bar{V}_0 \cup K_1$ is a compact set. Again, we have from Proposition 7.2.35 that there exists an open set V_1 such that $\bar{V}_0 \cup K_1 \subset V_1 \subset \bar{V}_1 \subset S$ and $\bar{V}_1$ is compact. In general, given $\bar{V}_{i-1}$ with $i > 0$ such that $\bar{V}_{i-1}$ is compact, we have that $\bar{V}_{i-1} \cup K_i$ is compact. We again use Proposition 7.2.35 to get an open set V_i such that $\bar{V}_{i-1} \cup K_i \subset V_i \subset \bar{V}_i \subset S$ and $\bar{V}_i$ is compact. We get a sequence of open sets V_i such that $V_1 \subset \bar{V}_1 \subset V_2 \subset \bar{V}_2 \subset V_3 \subset \dots$ and the $\bar{V}_i$ are compact sets and $K_i \subset V_i$ for all $i \in \mathbb{N}$.

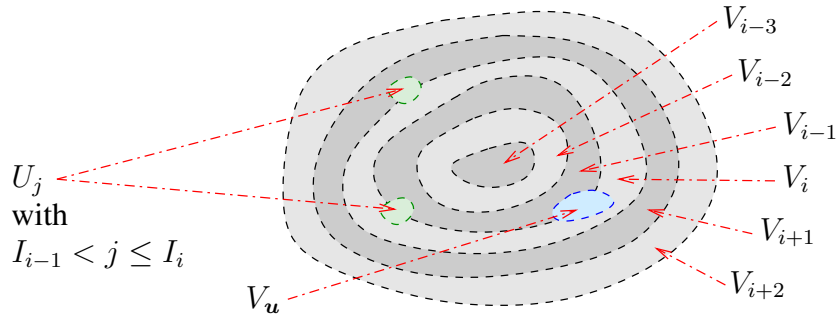
Since $S = \bigcup_{i \in \mathbb{N}} K_i \subset \bigcup_{i \in \mathbb{N}} V_i \subset S$, we get that $S = \bigcup_{i \in \mathbb{N}} V_i$. If we set $V_{-1} = \emptyset$, then we have that $S = \bigcup_{i \in \mathbb{N}} (\bar{V}_i \setminus V_{i-1})$ where each $\bar{V}_i \setminus V_{i-1}$ is a compact set because it is a closed set of the compact set $\bar{V}_i$.

Given $i \in \mathbb{N}$ and $\mathbf{u} \in \bar{V}_i \setminus V_{i-1}$, there exists a local chart (W, U, ϕ) and a set B_τ such that $\mathbf{u} \in U \cap B_\tau$. Let $V_{-2} = \emptyset$. If we set $U_{\mathbf{u}} = U \cap B_\tau \cap (V_{i+1} \setminus \bar{V}_{i-2})$, $W_{\mathbf{u}} = \phi^{-1}(U_{\mathbf{u}})$ and $\phi_{\mathbf{u}} = \phi$, then we get a local chart $(W_{\mathbf{u}}, U_{\mathbf{u}}, \phi_{\mathbf{u}})$ with $\mathbf{u} \in U_{\mathbf{u}} \subset (V_{i+1} \setminus \bar{V}_{i-2})$ and $U_{\mathbf{u}} \subset B_\tau$ for some $\tau \in T$. Repeating this for all $\mathbf{x} \in S$, we get an atlas $\{(W_{\mathbf{u}}, U_{\mathbf{u}}, \phi_{\mathbf{u}})\}_{\mathbf{u} \in S}$ on S such that $U_{\mathbf{u}} \subset (V_{i+1} \setminus \bar{V}_{i-2})$ for $\mathbf{u} \in \bar{V}_i \setminus V_{i-1}$ and $U_{\mathbf{u}} \subset B_\tau$ for some $\tau \in T$.

Since $\{U_{\mathbf{u}}\}_{\mathbf{u} \in \bar{V}_0}$ is an open cover of the compact set $\bar{V}_0$, there is a finite subcover $\{U_{\mathbf{u}_j}\}_{0 \leq j \leq J_0}$ for some $J_0 \in \mathbb{N}$. This yields a finite number of local charts $\{(W_j, U_j, \phi_j)\}_{0 \leq j \leq J_0}$ such that $\{U_j\}_{0 \leq j \leq J_0}$ is an open cover of $\bar{V}_0$ and $U_j \subset V_1$ for $1 \leq j \leq J_0$.

Let $I_{-1} = 0$. Suppose that we have a finite number of local charts $\{(W_j, U_j, \phi_j)\}_{J_{i-1}+1 \leq j \leq J_i}$ for some $J_{i-1} < J_i$ and $i \geq 0$ such that $\{U_j\}_{J_{i-1}+1 \leq j \leq J_i}$ is an open cover of $\bar{V}_i \setminus V_{i-1}$ and $U_j \subset (V_{i+1} \setminus \bar{V}_{i-2})$ for $J_{i-1} + 1 \leq j \leq J_i$.

Since $\{U_{\mathbf{u}}\}_{\mathbf{u} \in \bar{V}_{i+1} \setminus V_i}$ is an open cover of the compact set $\bar{V}_{i+1} \setminus V_i$, there is a finite subcover $\{U_{\mathbf{u}_j}\}_{J_i+1 \leq j \leq J_{i+1}}$ for some $J_{i+1} \in \mathbb{N}$. This yields a finite number of local charts $\{(W_j, U_j, \phi_j)\}_{J_i+1 \leq j \leq J_{i+1}}$ such that $\{U_j\}_{J_i+1 \leq j \leq J_{i+1}}$ is an open cover of $\bar{V}_{i+1} \setminus V_i$ and $U_j \subset (V_{i+2} \setminus \bar{V}_{i-1})$ for $J_i + 1 \leq j \leq J_{i+1}$.



By induction, we get the atlas $\mathcal{J} = \{(W_j, U_j, \phi_j)\}_{j \in \mathbb{N}}$ that we were looking for. By

construction, we have that U_j is a subset of a B_{τ_j} for some $\tau_j \in T$. To show that $\{U_j\}_{j \in \mathbb{N}}$ is locally finite, we consider $\mathbf{u} \in S$. Let $I_{-3} = 0$. We have that $\mathbf{u} \in \bar{V}_i \setminus V_{i-1}$ for some $i \in \mathbb{N}$. If we choose an open neighbourhood $V_{\mathbf{u}} \subset V_{i+1} \setminus \bar{V}_{i-2}$ of $\mathbf{u}$, then $U_j \cap V_{\mathbf{u}} = \emptyset$ for all $j \leq J_{i-3}$ and $j > J_{i+2}$. ■

Lemma 7.2.37

Suppose that $\{U_\alpha\}_{\alpha \in A}$ is a locally finite open cover of a topological space S . Given any compact set $K \subset S$, there exists an open set $V \supset K$ such that $V \cap U_\alpha = \emptyset$ for all but a finite number of $\alpha \in A$.

Proof.

Since the local cover is locally finite, there exists for each $\mathbf{x} \in K$ an open set $V_{\mathbf{x}}$ such that $U_\alpha \cap V_{\mathbf{x}} = \emptyset$ for all but a finite number of U_α . Since $\{V_{\mathbf{x}}\}_{\mathbf{x} \in K}$ is an open cover of the compact set K , there exists a finite subcover $\{V_{\mathbf{x}_i}\}_{1 \leq i \leq I}$ of K .

Let $V = \bigcup_{1 \leq i \leq I} V_{\mathbf{x}_i}$. We have that V is an open set such that $K \subset V$ and there exists only a finite number of U_α that intersect V because a set U_α that intersects V must also intersect one of the $V_{\mathbf{x}_i}$ with $1 \leq i \leq I$ and there are only a finite number of U_α that intersect each $V_{\mathbf{x}_i}$. ■

Theorem 7.2.38

If S is a non-compact, connected and oriented k -dimensional smooth manifold without boundary, then $H^k(S) = 0$.

Proof.

Since S is not compact, there exists an open cover $\{B_\tau\}_{\tau \in T}$ of S such that no finite subcover can cover S . Using Proposition 7.2.36, we get an atlas $\mathcal{A} = \{(W_i, U_i, \phi)\}_{i \in \mathbb{N}}$ on S such that each U_i is a subset of a B_{τ_i} for some $\tau_i \in T$ and $\{U_i\}_{i \in \mathbb{N}}$ is locally finite.

We have that $\{U_i\}_{i \in \mathbb{N}}$ is an open cover of S with the property that no finite subcover can cover S . If there were a finite subcover $\{U_{i_m}\}_{1 \leq m \leq M}$ of S , then $\{B_{\tau_{i_m}}\}_{1 \leq m \leq M}$ would be a finite subcover of S , contradicting our hypothesis.

Let $\{\psi_m\}_{m \in \mathbb{N}^+}$ be a partition of unity subordinate to $\mathcal{A}$. Given a differential k -form ω on S , we may express ω as $\omega = \sum_{m \in \mathbb{N}^+} \psi_m \omega$. Recall that the sum is well define because it is finite on any compact subset of S . If we prove for each $m \in \mathbb{N}^+$ that there exists a differential $(k-1)$ -form of S such that $\psi_m \omega = d\eta_m$, then we will have that $\omega = \sum_{m \in \mathbb{N}^+} d\eta_m = d\left(\sum_{m \in \mathbb{N}^+} \eta_m\right)$.

Hence $[\omega] = [0]$. This is what we plan to do.

Suppose that ω is a differential k -form ω on S with compact support in U_i for some $i \in \mathbb{N}$. Since no finite subcover of $\{U_i\}_{i \in \mathbb{N}}$ can cover S , we have that $U_i \subsetneq S$. Let $U_{i_0} = U_i$. Since S is connected, there is a set U_{i_1} such that $U_{i_0} \cap U_{i_1} \neq \emptyset$. Given $\{U_{i_m}\}_{0 \leq m \leq M}$ such that

$U_{i_m} \cap U_{i_{m-1}} \neq \emptyset$ for $1 \leq m \leq M$. we can again use the facts that no finite subcover of $\{U_i\}_{i \in \mathbb{N}}$ can cover S and that S is connected, to find a set $U_{i_{M+1}}$ such that $U_{i_{M+1}} \cap U_{i_M} \neq \emptyset$. Hence, by induction, we get an infinite sequence of sets $\{U_{i_m}\}_{m \in \mathbb{N}}$ such that $U_{i_m} \cap U_{i_{m-1}} \neq \emptyset$ for $m > 0$.

Choose a differential k -form ω_{i_1} with compact support in $U_{i_0} \cap U_{i_1}$. Since ω has a compact support in U_{i_0} , we may proceed as we did in the proof of Theorem 7.2.9 to get a differential $(k-1)$ -form η_{i_0} with support in U_{i_0} and $b_{i_1} \in \mathbb{R}$ such that $b_{i_1} \omega_{i_1} - \omega = d\eta_{i_0}$. For $m > 0$, choose a differential k -form $\omega_{i_{m+1}}$ with compact support in $U_{i_m} \cap U_{i_{m+1}}$. Since ω_{i_m} has a compact support in U_{i_m} , we may again proceed as we did in the proof of Theorem 7.2.9 to get a differential $(k-1)$ -form η_{i_m} with support in U_{i_m} and $b_{i_{m+1}} \in \mathbb{R}$ such that $b_{i_{m+1}} \omega_{i_{m+1}} - \omega_{i_m} = d\eta_{i_m}$.

We have

$$\begin{aligned}
 \omega &= b_{i_1} \omega_{i_1} - d\eta_{i_0} \\
 &= b_{i_1} b_{i_2} \omega_{i_2} - b_{i_1} d\eta_{i_1} - d\eta_{i_0} \\
 &= b_{i_1} b_{i_2} b_{i_3} \omega_{i_3} - b_{i_2} b_{i_1} d\eta_{i_2} - b_{i_1} d\eta_{i_1} - d\eta_{i_0} \\
 &= \dots \\
 &= \left(\prod_{s=1}^m b_{i_s} \right) \omega_{i_m} - \left(\prod_{s=1}^{m-1} b_{i_s} \right) d\eta_{i_{m-1}} - \left(\prod_{s=1}^{m-2} b_{i_s} \right) d\eta_{i_{m-2}} - b_{i_1} d\eta_{i_1} - d\eta_{i_0} \\
 &= \left(\prod_{s=1}^m b_{i_s} \right) \omega_{i_m} + d \left(- \left(\prod_{s=1}^{m-1} b_{i_s} \right) \eta_{i_{m-1}} - \left(\prod_{s=1}^{m-2} b_{i_s} \right) \eta_{i_{m-2}} - b_{i_1} \eta_{i_1} - \eta_{i_0} \right) \quad (7.2.4)
 \end{aligned}$$

for $m > 0$. Let

$$\eta = - \sum_{m=1}^{\infty} \left(\prod_{s=1}^m b_{i_s} \right) \eta_{i_m} - \eta_{i_0}.$$

We first note that this sum is finite on every compact subset K of S . We have that $\text{supp } \eta_{i_m} \subset U_{i_m}$ for all $m \in \mathbb{N}$. Since $\{U_i\}_{i \in \mathbb{N}}$ is a locally finite open cover of S , we get from Lemma 7.2.37 that there exists $M_K \geq 0$ such that $U_{i_m} \cap K = \emptyset$ for $m > M_K$. Thus $\eta_{i_m}|_K = 0$ for $m > M_K$. We get

$$\eta|_K = - \sum_{m=1}^{M_K} \left(\prod_{s=1}^m b_{i_s} \right) \eta_{i_m}|_K - \eta_{i_0}|_K.$$

Moreover, since $\text{supp } \omega_{i_m} \subset U_{i_m}$, we get from (7.2.4) that

$$\omega|_K = \left(\prod_{s=1}^{M_K+1} b_{i_s} \right) \underbrace{\omega_{i_{M_K+1}}|_K}_{=0} + d(\eta|_K) = d(\eta|_K)$$

Since S is σ -compact according to Proposition 7.2.34, we have that $S = \bigcup_{i \in \mathbb{N}} K_i$ where the

K_i are compact sets. The previous discussion applied to $K = \bigcup_{0 \leq i \leq I} K_i$ for all I shows that

$\omega = d\eta$. This completes the proof. ■

Lemma 7.2.39

Suppose that S is a k -dimensional smooth manifold. Then $H^q(S) \cong H^q(S \times \mathbb{R})$.

Proof.

This proof is similar to the proof of Proposition 7.2.26.

Consider the injection $\iota : S \rightarrow S \times \{0\}$ defined by $\iota(\mathbf{u}) = (\mathbf{u}, 0)$ and the restriction $\rho : S \times \mathbb{R} \rightarrow S$ defined by $\rho(\mathbf{u}, x) = \mathbf{u}$.

We have that $\rho \circ \iota = \text{Id} : S \rightarrow S$. We also have that $\iota \circ \rho : S \times \mathbb{R} \rightarrow S \times \mathbb{R}$ and $\text{Id} : S \times \mathbb{R} \rightarrow S \times \mathbb{R}$ are homotopic. An homotopy $H : (S \times \mathbb{R}) \times [0, 1] \rightarrow S \times \mathbb{R}$ between $\iota \circ \rho$ and Id is given by $H((\mathbf{u}, x), t) = t(\mathbf{u}, x) + (1 - t)(\iota \circ \rho)(\mathbf{u}, x)$ for all $(\mathbf{u}, x) \in S \times \mathbb{R}$. It follows from Proposition 7.2.25 that $(\iota \circ \rho)^\# = \text{Id}^\# : H^q(S \times \mathbb{R}) \rightarrow H^q(S \times \mathbb{R})$.

Hence, $\iota^\# \circ \rho^\# = (\rho \circ \iota)^\# = \text{Id}^\# : H^q(S) \rightarrow H^q(S)$ and $\rho^\# \circ \iota^\# = (\iota \circ \rho)^\# = \text{Id}^\# : H^q(S \times \mathbb{R}) \rightarrow H^q(S \times \mathbb{R})$ for $0 \leq q \leq k$. Thus $\rho^\# : H^q(S) \rightarrow H^q(S \times \mathbb{R})$ is the inverse of $\iota^\# : H^q(S \times \mathbb{R}) \rightarrow H^q(S)$. Therefore $\rho^\#$ provides an isomorphism between $H^q(S)$ and $H^q(S \times \mathbb{R})$. ■

We are now ready to complete the computations of $H^q(\mathbb{R}^k \setminus \{\mathbf{0}\})$ for the missing cases of $0 < q < k - 1$.

Proposition 7.2.40

$H^q(\mathbb{R}^k \setminus \{\mathbf{0}\}) \cong H^q(S^{k-1}) = 0$ for $0 < q < k - 1$.

Proof.

We prove by induction on $k > 2$ that $H^q(\mathbb{R}^k \setminus \{\mathbf{0}\}) = 0$ for $0 < q < k - 1$.

i) We first prove the hypothesis of induction for $k = 3$. Suppose that ω is a differential 1-form on $\mathbb{R}^3 \setminus \{\mathbf{0}\} = R_+ \cup R_-$ where $R_+ = \mathbb{R}^3 \setminus \{(0, 0, x_3) : x_3 \geq 0\}$ and $R_- = \mathbb{R}^3 \setminus \{(0, 0, x_3) : x_3 \leq 0\}$. Since R_+ and R_- are star-shaped (one may use $(0, 0, -1)$ as the centre of the star for R_+ and $(0, 0, 1)$ as the centre of the star for R_-), we get from Poincaré Lemma, Theorem 4.5.12, that there exist two differential 0-forms η_+ and η_- on R_+ and R_- respectively such that $\omega = d\eta_+$ on R_+ and $\omega = d\eta_-$ on R_- . Since η_+ and η_- are functions such that $d(\eta_+ - \eta_-) = d\eta_+ - d\eta_- = 0$ on $R_+ \cap R_-$, we have that $\eta_+ - \eta_- = c$, a constant, on $R_+ \cap R_-$. Let

$$\eta(\mathbf{u}) = \begin{cases} \eta_+(\mathbf{u}) - c & \text{if } \mathbf{u} \in R_+ \\ \eta_-(\mathbf{u}) & \text{if } \mathbf{u} \in R_- \setminus R_+ \end{cases}$$

This is a well defined continuous function on $\mathbb{R}^3 \setminus \{\mathbf{0}\} = R_+ \cup R_-$ because $\eta_+(\mathbf{u}) - c = \eta_-(\mathbf{u})$ for $\mathbf{u} \in R_- \cap R_+$. We have that $d\eta = \omega$ on $\mathbb{R}^3 \setminus \{\mathbf{0}\}$. Thus $[\omega] = [0]$ in $H^1(\mathbb{R}^3 \setminus \{\mathbf{0}\})$. This proves that $H^q(\mathbb{R}^3 \setminus \{\mathbf{0}\}) = 0$ for $0 < q < 2$.

ii) We assume that the hypothesis of induction is true for $k = j - 1$ with $j > 3$; namely, $H^q(\mathbb{R}^{j-1} \setminus \{\mathbf{0}\}) = 0$ for $0 < q < j - 2$.

iii) We now prove that the hypothesis of induction is true for $k = j$. Let $I_+ = \{(0, \dots, 0, x_j) \in \mathbb{R}^j : x_j \geq 0\}$ and $I_- = \{(0, \dots, 0, x_j) \in \mathbb{R}^j : x_j \leq 0\}$. Suppose that ω

is a differential 1-form on $\mathbb{R}^j \setminus \{\mathbf{0}\} = R_+ \cup R_-$ where $R_+ = \mathbb{R}^j \setminus I_+$ and $R_- = \mathbb{R}^j \setminus I_-$. As in (i), since R_+ and R_- are star-shaped, we get from Poincaré Lemma, Theorem 4.5.12, that there exist two differential 0-forms η_+ and η_- on R_+ and R_- respectively such that $\omega = d\eta_+$ on R_+ and $\omega = d\eta_-$ on R_- . Since η_+ and η_- are functions such that $d(\eta_+ - \eta_-) = d\eta_+ - d\eta_- = 0$ on $R_+ \cap R_-$, we have that $\eta_+ - \eta_- = c$, a constant, on $R_+ \cap R_-$. Let

$$\eta(\mathbf{u}) = \begin{cases} \eta_+(\mathbf{u}) - c & \text{if } \mathbf{u} \in R_+ \\ \eta_-(\mathbf{u}) & \text{if } \mathbf{u} \in R_- \setminus R_+ \end{cases}$$

This is a well defined continuous function on $\mathbb{R}^j \setminus \{\mathbf{0}\} = R_+ \cup R_-$ because $\eta_+(\mathbf{u}) - c = \eta_-(\mathbf{u})$ for $\mathbf{u} \in R_- \cap R_+$. We have that $d\eta = \omega$ on $\mathbb{R}^j \setminus \{\mathbf{0}\}$. Thus $[\omega] = [0]$ in $H^1(\mathbb{R}^j \setminus \{\mathbf{0}\})$. This proves that $H^1(\mathbb{R}^j \setminus \{\mathbf{0}\}) = 0$.

Suppose that ω is a differential q -form on $\mathbb{R}^j \setminus \{\mathbf{0}\} = R_+ \cup R_-$ with $1 < q < j - 1$. Again, we get from Poincaré Lemma that there exist two differential $(q - 1)$ -form η_+ and η_- on R_+ and R_- respectively such that $\omega = d\eta_+$ on R_+ and $\omega = d\eta_-$ on R_- . We have that $d(\eta_+ - \eta_-) = d\eta_+ - d\eta_- = 0$ on $R_+ \cap R_- = \{\mathbf{x} \in \mathbb{R}^j : x_j = 0\} = (\mathbb{R}^{j-1} \setminus \{\mathbf{0}\}) \times \mathbb{R}$. However, we get from Lemma 7.2.39 and our hypothesis of induction that $H^{q-1}((\mathbb{R}^{j-1} \setminus \{\mathbf{0}\}) \times \mathbb{R}) \cong H^{q-1}(\mathbb{R}^{j-1} \setminus \{\mathbf{0}\}) = 0$ for $1 < q < j - 1$. Therefore, there exists a differential $(q - 2)$ -form μ on $R_+ \cap R_-$ such that $\eta_+ - \eta_- = d\mu$ on $R_+ \cap R_-$.

Let $\{\psi_j\}_{j \in \mathbb{N}^+}$ be a partition of unity subordinate to the open cover $\{R_+, R_-\}$ of $\mathbb{R}^j \setminus \{\mathbf{0}\}$, and let $\psi_+ = \sum_{\text{supp } \psi_j \subset R_+} \psi_j$ and $\psi_- = \sum_{\text{supp } \psi_j \subset R_-} \psi_j$. Note that the two sums are finite on compact sets according to Proposition 3.4.4. We have that ψ_+ and ψ_- are two functions of class C^∞ such that $\text{supp } \psi_+ \subset R_+$ and $\text{supp } \psi_- \subset R_-$. Let

$$\tilde{\mu}_-(\mathbf{u}) = \begin{cases} \psi_+(\mathbf{u})\mu(\mathbf{u}) & \text{if } \mathbf{u} \in R_+ \cap R_- \\ 0 & \text{if } \mathbf{u} \in I_+ \end{cases}$$

and

$$\tilde{\mu}_+(\mathbf{u}) = \begin{cases} \psi_-(\mathbf{u})\mu(\mathbf{u}) & \text{if } \mathbf{u} \in R_+ \cap R_- \\ 0 & \text{if } \mathbf{u} \in I_- \end{cases}$$

By construction, $\tilde{\mu}_-$ is a differentiable $(q - 2)$ -form on R_- because $\text{supp } \tilde{\mu}_- \cap I_+ = \emptyset$. Similarly, $\tilde{\mu}_+$ is a differentiable $(q - 2)$ -form on R_+ because $\text{supp } \tilde{\mu}_+ \cap I_- = \emptyset$.

Since we have on $R_+ \cap R_-$ that

$$\begin{aligned} \eta_+ - d\tilde{\mu}_+ &= \eta_+ - \psi_- d\mu - d\psi_- \wedge \mu = \eta_+ - (1 - \psi_+) d\mu - d(1 - \psi_+) \wedge \mu \\ &= \eta_+ - d\mu - \psi_+ d\mu - d\psi_+ \wedge \mu = \eta_+ - d\mu - d\tilde{\mu}_- = \eta_- - d\tilde{\mu}_-, \end{aligned}$$

we may define the following differential $(q - 1)$ -form on $\mathbb{R}^j \setminus \{\mathbf{0}\}$.

$$\eta(\mathbf{u}) = \begin{cases} \eta_+(\mathbf{u}) - d\tilde{\mu}_+(\mathbf{u}) & \text{if } \mathbf{u} \in R_+ \\ \eta_-(\mathbf{u}) - d\tilde{\mu}_-(\mathbf{u}) & \text{if } \mathbf{u} \in R_- \setminus R_+ \end{cases}$$

We then have that $d\mu = d(\eta_+ - d\tilde{\mu}_+) = d\eta_+ = \omega$ on R_+ and $d\mu = d(\eta_- - d\tilde{\mu}_-) = d\eta_- = \omega$ on R_- . Thus $[\omega] = [0]$ in $H^q(\mathbb{R}^j \setminus \{\mathbf{0}\})$. This proves that $H^q(\mathbb{R}^j \setminus \{\mathbf{0}\}) = 0$. $\blacksquare$

We are now in a position to prove Theorem 7.2.11

Proof (of Theorem 7.2.11).

i) We first compute $H_c^0(\mathbb{R}^k)$. Let $f : \mathbb{R}^k \rightarrow \mathbb{R}$ be smooth function. We can proceed as in the proof of Proposition 7.2.20 to show that f is constant on $\mathbb{R}^k$. Namely, we have that $df = \sum_{i=1}^k \frac{\partial f}{\partial x_i} dx_i = 0$ on $\mathbb{R}^k$ implies $\frac{\partial f}{\partial x_i} = 0$ for $1 \leq i \leq k$ on $\mathbb{R}^k$. Thus f is constant. But since f has a compact support, we must have that $f = 0$ on $\mathbb{R}^k$. Thus $H_c^0(\mathbb{R}^k) = 0$.

ii) We compute $H_c^1(\mathbb{R}^k)$. Suppose that ω is a differential 1-form with compact support on $\mathbb{R}^k$. We have from Proposition 7.1.7 or Poincaré Lemma (Theorem 4.5.12 or Theorem 7.1.6) that there exists a differential 0-form f on $\mathbb{R}^k$ such that $\omega = df$.

Since ω has a compact support, there exists $r > 0$ such that $\text{supp } \omega \subset B_r(\mathbf{0})$. It follows from $df = \omega = 0$ on $\mathbb{R}^k \setminus B_r(\mathbf{0})$ that f is constant on $\mathbb{R}^k \setminus B_r(\mathbf{0})$. Suppose that $f(\mathbf{x}) = c$ for $\mathbf{x} \in \mathbb{R}^k \setminus B_r(\mathbf{0})$. Then $f - c$ is a differential 0-form on $\mathbb{R}^k$ with compact support in $B_r(\mathbf{0})$ such that $d(f - c) = df = \omega$. Thus, $\omega \sim 0$ and we obtain that $[\omega] = [0]$ in $H_c^1(\mathbb{R}^k)$. Since ω is an arbitrary 1-form with compact support, we get that $H_c^1(\mathbb{R}^k) = 0$.

iii) We now assume that $1 < q < k$. Suppose that ω is a differential q -form with compact support on $\mathbb{R}^k$. We have from Proposition 7.1.7 or Poincaré Lemma (Theorem 4.5.12 or Theorem 7.1.6) that there exists a differential $(q-1)$ -form η on $\mathbb{R}^k$ such that $\omega = d\eta$.

Since ω has a compact support, there exists $r > 0$ such that $\text{supp } \omega \subset B_r(\mathbf{0})$. Since $\mathbb{R}^k \setminus \overline{B_r(\mathbf{0})}$ is diffeomorphic to $\mathbb{R}^k \setminus \{\mathbf{0}\}$ ⁷, we get from the previous proposition that $H^{q-1}(\mathbb{R}^k \setminus \overline{B_r(\mathbf{0})}) \cong H^{q-1}(\mathbb{R}^k \setminus \{\mathbf{0}\}) \cong 0$ (see comments after Proposition 7.2.23). Thus, there exists a differential $(q-2)$ -form μ on $\mathbb{R}^k \setminus \overline{B_r(\mathbf{0})}$ such that $\eta = d\mu$.

Choose a function $\psi : \mathbb{R}^k \rightarrow \mathbb{R}$ of class C^∞ such that $\psi(\mathbf{x}) = 0$ for $\mathbf{x} \in B_{5r/4}(\mathbf{0})$ and $\psi(\mathbf{x}) = 1$ for $\mathbf{x} \in \mathbb{R}^k \setminus B_{2r}(\mathbf{0})$. Let $\tilde{\mu}$ be the differential $(q-2)$ -form on $\mathbb{R}^k$ defined by

$$\tilde{\mu}(\mathbf{x}) = \begin{cases} \psi(\mathbf{x})\mu(\mathbf{x}) & \text{if } \mathbf{x} \in \mathbb{R}^k \setminus \overline{B_r(\mathbf{0})} \\ 0 & \text{if } \mathbf{x} \in \overline{B_r(\mathbf{0})} \end{cases}$$

We have that $\text{supp}(\eta - d\tilde{\mu}) \subset B_{2r}(\mathbf{0})$ because $\eta - d\tilde{\mu} = \eta - d\mu = 0$ on $\mathbb{R}^k \setminus B_{2r}(\mathbf{0})$. Thus $\eta - d\tilde{\mu}$ is a differential $(q-1)$ -form with compact support on $\mathbb{R}^k$. Moreover, $d(\eta - d\tilde{\mu}) = d\eta = \omega$. Hence, $\omega \sim 0$ and we get that $[\omega] = [0]$ in $H_c^q(\mathbb{R}^k)$. Since ω is an arbitrary q -form with compact support, we conclude that $H_c^q(\mathbb{R}^k) = 0$. ■

7.2.3 Mayer-Vietoris Argument

The Mayer-Vietoris argument refers to a standard technique used in algebraic topology to compute the homology modules of some manifolds. We will have more to say on this technique in Section 8.4 on singular homology. As the reader will realize in the discussion

⁷A diffeomorphism is given by $g : \mathbb{R}^k \setminus \overline{B_r(\mathbf{0})} \rightarrow \mathbb{R}^k \setminus \{\mathbf{0}\}$ defined by $g(\mathbf{x}) = \mathbf{x} - r\|\mathbf{x}\|^{-1}\mathbf{x}$.

below, we have kind of used this technique in our presentation of cohomology. We will use Mayer-Vietoris argument to compute the q^{th} cohomology module of the sphere S^k from scratch.

Let $S_n^{k-1} = S^{k-1} \setminus \{-e_k\}$ and $S_s^{k-1} = S^{k-1} \setminus \{e_k\}$ be, respectively, the sphere in $\mathbb{R}^k$ minus the “south pole” and the sphere in $\mathbb{R}^k$ minus the “north pole”.

The following information about S_n^{k-1} and S_s^{k-1} will be useful later.

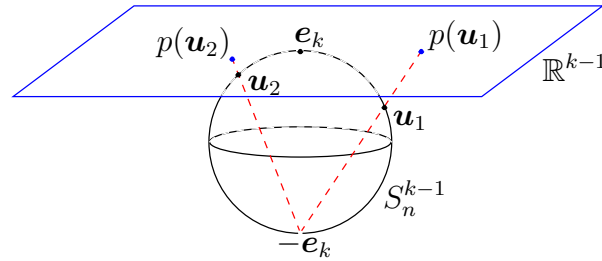
Lemma 7.2.41

S_n^{k-1} and S_s^{k-1} are smoothly contractible to a point.

Proof.

We prove that S_n^{k-1} is smoothly contractible to e_k and leave to the reader the task of proving that S_s^{k-1} is smoothly contractible to $-e_k$.

Let $p : S_n^{k-1} \rightarrow \mathbb{R}^{k-1}$ be the stereographic projection defined in the following figure.



We have represented $\mathbb{R}^{k-1}$ as the classical tangent plane $e_k + \{(x_1, x_2, \dots, x_{k-1}, 0) : x_i \in \mathbb{R} \text{ and } 1 \leq i \leq k-1\}$ to the sphere S^{k-1} at e_k . In particular, the origin in $\mathbb{R}^{k-1}$ is associated to e_k in the tangent plane.

Let $H : S_n^{k-1} \times [0, 1] \rightarrow S_n^{k-1}$ be the map defined by $H(u, t) = p^{-1}(t p(u))$ for $(u, t) \in S_n^{k-1} \times [0, 1]$. H is a smooth function such that $H(u, 0) = e_k$ and $H(u, 1) = u$ for all $u \in S_n^{k-1}$. ■

Proposition 7.2.42

$H^q(S^{k-1}) = H^q(S_n^{k-1} \cup S_s^{k-1}) = H^{q-1}(S_n^{k-1} \cap S_s^{k-1})$ for $q > 1$.

Proof.

The goal is to build a linear isomorphism between $H^q(S_n^{k-1} \cup S_s^{k-1})$ and $H^{q-1}(S_n^{k-1} \cap S_s^{k-1})$.

i) We first define a function that maps closed differential q -form on $S_n^{k-1} \cup S_s^{k-1}$ to closed differential $(q-1)$ -form on $S_n^{k-1} \cap S_s^{k-1}$.

Let ω be a closed differential q -form on $S_n^{k-1} \cup S_s^{k-1}$. Since S_n^{k-1} and S_s^{k-1} are smoothly contractible to a point according to Lemma 7.2.41, it follows from Poincaré lemma, Theorem 7.1.6, that there exist a differential $(q-1)$ -form η_n on S_n^{k-1} such that $\omega = d\eta_n$ and a differential $(q-1)$ -form η_s on S_s^{k-1} such that $\omega = d\eta_s$.

Let $\eta = \eta_n - \eta_s$. we have that η is a closed differential $(q-1)$ -form on $S_n^{k-1} \cap S_s^{k-1}$ because $d\eta = d\eta_n - d\eta_s = 0$ on $S_n^{k-1} \cap S_s^{k-1}$.

We now prove that the map $F : H^q(S_n^{k-1} \cup S_s^{k-1}) \rightarrow H^{q-1}(S_n^{k-1} \cap S_s^{k-1})$ defined by $F([\omega]) = [\eta]$ is well defined.

We first note that $[\eta]$ is independent of the choice of η_n and η_s . If $\tilde{\eta}_n$ and $\tilde{\eta}_s$ are two other differential $(q-1)$ -forms on S_n^{k-1} and S_s^{k-1} respectively such that $\omega = d\tilde{\eta}_n$ on S_n^{k-1} and $\omega = d\tilde{\eta}_s$ on S_s^{k-1} , then $d(\eta_n - \tilde{\eta}_n) = d\eta_n - d\tilde{\eta}_n = 0$ on S_n^{k-1} . It follows from Poincaré lemma that there exists a differential $(q-2)$ -form μ_n on S_n^{k-1} such that $\eta_n - \tilde{\eta}_n = d\mu_n$. Similarly, there exists a differential $(q-2)$ -form μ_s on S_s^{k-1} such that $\eta_s - \tilde{\eta}_s = d\mu_s$. Thus $(\eta_n - \eta_s) - (\tilde{\eta}_n - \tilde{\eta}_s) = d(\mu_n - \mu_s)$ on $S_n^{k-1} \cap S_s^{k-1}$. Hence $\eta_n - \eta_s \sim \tilde{\eta}_n - \tilde{\eta}_s$.

We also have that $[\eta]$ is independent of the chosen representative of the cohomology class $[\omega]$. Suppose that $\tilde{\omega} \sim \omega$. Then there exists a differential $(q-1)$ -form μ on S^{k-1} such that $\tilde{\omega} = \omega + d\mu$. Hence, the differential $(q-1)$ -form $\eta_n + \mu$ on S_n^{k-1} is such that $\tilde{\omega} = d(\eta_n + \mu)$ and the differential $(q-1)$ -form $\eta_s + \mu$ on S_s^{k-1} is such that $\tilde{\omega} = d(\eta_s + \mu)$. Thus, according to our procedure to define F , the closed differential q -form $\tilde{\omega}$ yields $(\eta_n + \mu) - (\eta_s + \mu) = \eta_n - \eta_s$.

ii) We define a function that maps closed differential $(q-1)$ -form on $S_n^{k-1} \cap S_s^{k-1}$ to closed differential q -form on $S_n^{k-1} \cup S_s^{k-1}$.

Let $\{\psi_n, \psi_s\}$ be a partition of unity subordinate to the open cover $\{S_n^{k-1}, S_s^{k-1}\}$ of S^{k-1} such that $\text{supp } \psi_n \subset S_n^{k-1}$ and $\text{supp } \psi_s \subset S_s^{k-1}$. Given a closed differential $(q-1)$ -form η on $S_n^{k-1} \cap S_s^{k-1}$, we define two closed differential $(q-1)$ -form η_n and η_s on S_n^{k-1} and S_s^{k-1} respectively as it follows.

$$\eta_n(\mathbf{u}) = \begin{cases} \psi_s(\mathbf{u})\eta(\mathbf{u}) & \text{if } \mathbf{u} \in S_n^{k-1} \setminus \{\mathbf{e}_k\} \\ 0 & \text{if } \mathbf{u} = \mathbf{e}_k \end{cases}$$

and

$$\eta_s(\mathbf{u}) = \begin{cases} -\psi_n(\mathbf{u})\eta(\mathbf{u}) & \text{if } \mathbf{u} \in S_s^{k-1} \setminus \{-\mathbf{e}_k\} \\ 0 & \text{if } \mathbf{u} = -\mathbf{e}_k \end{cases}$$

These two differential $(q-1)$ -forms are well defined and smooth because ψ_s is null in a neighbourhood of $\mathbf{e}_k$ since its compact support is in S_s^{k-1} and ψ_n is null in a neighbourhood of $-\mathbf{e}_k$ since its compact support is in S_n^{k-1} . We define a differential q -form ω on S^{k-1} as it follows.

$$\omega(\mathbf{u}) = \begin{cases} d\eta_n(\mathbf{u}) & \text{if } \mathbf{u} \in S_n^{k-1} \\ d\eta_s(\mathbf{u}) & \text{if } \mathbf{u} = -\mathbf{e}_k \end{cases}$$

This is a smooth differential q -form on S^{k-1} because $d\eta_n - d\eta_s = d(\eta_n + \eta_s) = d\eta = 0$ on $S_n^{k-1} \cap S_s^{k-1}$. It is obviously closed because $d\omega = d^2\eta_n = 0$ on S_n^{k-1} and $d\omega = -d^2\eta_s = 0$ on S_s^{k-1} .

We prove that the map $G : H^{q-1}(S_n^{k-1} \cap S_s^{k-1}) \rightarrow H^q(S_n^{k-1} \cup S_s^{k-1})$ defined by $G([\eta]) = [\omega]$ is well defined.

Suppose that $\tilde{\eta}$ is another closed differential $(q-1)$ -form on $S_n^{k-1} \cap S_s^{k-1}$ such that $\eta \sim \tilde{\eta}$. We then have that $\tilde{\eta} = \eta + d\tilde{\mu}$ for some differential $(q-2)$ -form $\tilde{\mu}$ on $S_n^{k-1} \cap S_s^{k-1}$. We follow

the procedure describe above. We set

$$\tilde{\eta}_n(\mathbf{u}) = \begin{cases} \psi_s(\mathbf{u})\eta(\mathbf{u}) + \psi_s(\mathbf{u}) d\tilde{\mu} & \text{if } \mathbf{u} \in S_n^{k-1} \setminus \{\mathbf{e}_k\} \\ 0 & \text{if } \mathbf{u} = \mathbf{e}_k \end{cases}$$

and

$$\tilde{\eta}_s(\mathbf{u}) = \begin{cases} -\psi_n(\mathbf{u})\eta(\mathbf{u}) - \psi_n(\mathbf{u}) d\tilde{\mu} & \text{if } \mathbf{u} \in S_s^{k-1} \setminus \{-\mathbf{e}_k\} \\ 0 & \text{if } \mathbf{u} = -\mathbf{e}_k \end{cases}$$

to get the differential q -form on S^{k-1} defined by

$$\tilde{\omega}(\mathbf{u}) = \begin{cases} d\tilde{\eta}_n(\mathbf{u}) & \text{if } \mathbf{u} \in S_n^{k-1} \\ d\tilde{\eta}_s(\mathbf{u}) & \text{if } \mathbf{u} = -\mathbf{e}_k \end{cases}$$

Let $\tilde{\mu}_n$ be the differential $(q-2)$ -form on S_n^{k-1} defined by

$$\tilde{\mu}_n(\mathbf{u}) = \begin{cases} \psi_s(\mathbf{u})\tilde{\mu}(\mathbf{u}) & \text{if } \mathbf{u} \in S_n^{k-1} \setminus \{\mathbf{e}_k\} \\ 0 & \text{if } \mathbf{u} = \mathbf{e}_k \end{cases}$$

and $\tilde{\mu}_s$ be the differential $(q-2)$ -form on S_s^{k-1} defined by

$$\tilde{\mu}_s(\mathbf{u}) = \begin{cases} -\psi_n(\mathbf{u})\tilde{\mu}(\mathbf{u}) & \text{if } \mathbf{u} \in S_s^{k-1} \setminus \{-\mathbf{e}_k\} \\ 0 & \text{if } \mathbf{u} = -\mathbf{e}_k \end{cases}$$

Since $\psi_s\tilde{\mu} = -\psi_n\tilde{\mu}$ on $S_n^{k-1} \cap S_s^{k-1}$, we may define a differential $(q-2)$ -form μ on S^{k-1} by

$$\mu(\mathbf{u}) = \begin{cases} \tilde{\mu}_n(\mathbf{u}) & \text{if } \mathbf{u} \in S_n^{k-1} \setminus \{-\mathbf{e}_k\} \\ \tilde{\mu}_s(\mathbf{u}) & \text{if } \mathbf{u} = -\mathbf{e}_k \end{cases}$$

We have that $\tilde{\omega} = \omega + d\tilde{\mu}$ because $\tilde{\omega} - \omega = d\tilde{\eta}_n$ on S_n^{k-1} and $\tilde{\omega} - \omega = d\tilde{\eta}_s$ on S_s^{k-1} . Thus $\tilde{\omega} \sim \omega$ on S^{k-1} .

iii) We have that $F = G^{-1}$. It is follows directly from the definitions of F and G that $F \circ G = \text{Id} : H^{q-1}(S_n^{k-1} \cap S_s^{k-1}) \rightarrow H^{q-1}(S_n^{k-1} \cap S_s^{k-1})$. To prove that $G \circ F = \text{Id} : H^q(S_n^{k-1} \cup S_s^{k-1}) \rightarrow H^q(S_n^{k-1} \cup S_s^{k-1})$, we consider a differential q -form ω on $H^q(S_n^{k-1} \cup S_s^{k-1})$.

According to the definition of F , starting with ω , we get a differential $(q-1)$ -form η_n on S_n^{k-1} such that $\omega = d\eta_n$ and a differential $(q-1)$ -form η_s on S_s^{k-1} such that $\omega = d\eta_s$, and set $\eta = \eta_n - \eta_s$. Thus $F([\omega]) = [\eta]$.

According to the definition of G , starting with η , we first define two differential $(q-1)$ -form $\tilde{\eta}_n$ and $\tilde{\eta}_s$ on S_n^{k-1} and S_s^{k-1} by

$$\tilde{\eta}_n(\mathbf{u}) = \begin{cases} \psi_s(\mathbf{u})\eta(\mathbf{u}) & \text{if } \mathbf{u} \in S_n^{k-1} \setminus \{\mathbf{e}_k\} \\ 0 & \text{if } \mathbf{u} = \mathbf{e}_k \end{cases}$$

and

$$\tilde{\eta}_s(\mathbf{u}) = \begin{cases} -\psi_n(\mathbf{u})\eta(\mathbf{u}) & \text{if } \mathbf{u} \in S_s^{k-1} \setminus \{-\mathbf{e}_k\} \\ 0 & \text{if } \mathbf{u} = -\mathbf{e}_k \end{cases}$$

We then define a differential q -form ω on S^{k-1} by

$$\tilde{\omega}(\mathbf{u}) = \begin{cases} d\tilde{\eta}_n(\mathbf{u}) & \text{if } \mathbf{u} \in S_n^{k-1} \\ d\tilde{\eta}_s(\mathbf{u}) & \text{if } \mathbf{u} = -\mathbf{e}_k \end{cases}$$

Since $\eta_s + \psi_n\eta = \eta_n - \psi_s\eta$ on $S_n^{k-1} \cap S_s^{k-1}$, we may define the differential $(q-1)$ -form μ on S^{k-1} by

$$\mu(\mathbf{u}) = \begin{cases} \eta_s(\mathbf{u}) + \psi_n(\mathbf{u})\eta(\mathbf{u}) & \text{if } \mathbf{u} \in S_n^{k-1} \\ \eta_n(\mathbf{u}) - \psi_s(\mathbf{u})\eta(\mathbf{u}) & \text{if } \mathbf{u} = -\mathbf{e}_k \end{cases}$$

We have that $\tilde{\omega} = \omega - d\mu$ on S^{k-1} because $\tilde{\omega} = d(\psi_s\eta) = d(\eta - \psi_n\eta) = d(\eta_n - \eta_s - \psi_n\eta) = \omega - d(\eta_s + \psi_n\eta)$ on S_n^{k-1} and $\tilde{\omega} = -d(\psi_n\eta) = -d(\eta - \psi_s\eta) = -d(\eta_n - \eta_s - \psi_s\eta) = \omega - d(\eta_n - \psi_s\eta)$ on S_s^{k-1} . Hence $\tilde{\omega} \sim \omega$. Thus $G([\eta]) = [\omega]$ and $G(F([\omega]) = [\omega]$ as desired. ■

Using the stereographic projection p defined in the proof of Lemma 7.2.41, we have that $S_n^{k-1} \cap S_s^{k-1}$ is diffeomorphic to $\mathbb{R}^{k-1} \setminus \{\mathbf{0}\}$. Using spherical coordinates defined by $\mathbf{x} \mapsto (\|\mathbf{x}\|^{-1}\mathbf{x}, \|\mathbf{x}\|)$ for $\mathbf{x} \in \mathbb{R}^{k-1} \setminus \{\mathbf{0}\}$, we have that $\mathbb{R}^{k-1} \setminus \{\mathbf{0}\}$ is diffeomorphic to $S^{k-2} \times]0, \infty[$. Therefore, $S_n^{k-1} \cap S_s^{k-1}$ is diffeomorphic to $S^{k-2} \times]0, \infty[$. It then follows from Lemma 7.2.39 (slightly modified) that $H^{q-1}(S_n^{k-1} \cap S_s^{k-1}) \cong H^{q-1}(S^{k-2} \times]0, \infty[) \cong H^{q-1}(S^{k-2})$. Combining this result with Proposition 7.2.42, we get

$$H^q(S^{k-1}) \cong H^{q-1}(S^{k-2}) \quad (7.2.5)$$

for $k > 1$ and $q > 1$.

Proposition 7.2.43

$H^1(S^1) \cong \mathbb{R}$ and $H^1(S^k) = 0$ for $k > 1$.

Proof.

i) We get from Example 4.7.4 that the dimension of $H^1(S^1)$ is non null. Since S^1 is a compact manifold without boundary, we may use Proposition 7.2.7 to prove that $H^1(S^1) \cong \mathbb{R}$.

Suppose that ω is a differential 1-form on S^1 such that $\int_{S^1} \omega = 0$.

Consider the path $\sigma : \mathbb{R} \rightarrow S^1$ defined by $\sigma(s) = (\cos(2\pi s), \sin(2\pi s))$ for $s \in \mathbb{R}$. Let $g(t) = \int_0^t \sigma^*(\omega)$ for $t \in \mathbb{R}$. We have by hypothesis that $g(0) = g(1) = \int_0^1 \sigma^*(\omega) = \int_{\sigma([0,1])} \omega = \int_{S^1} \omega = 0$. Thus $g = f \circ \sigma = \sigma^*(f)$ for some functions $f : S^1 \rightarrow \mathbb{R}$. If we derive both sides of $\sigma^*(f)(t) = g(t) = \int_0^t \sigma^*(\omega)$ with respect to t , then we get $\sigma^*(df) = d(\sigma^*(f)) = \sigma^*(\omega)$.

Thus $df = \omega$ and therefore $[\omega] = [0]$. This proves that the map M in Proposition 7.2.7 is an isomorphism.

ii) The proof is very similar to the proof of Proposition 6.2.10. The context is a little more general.

Suppose that ω is a closed differential 1-form on S^k with $k > 1$.

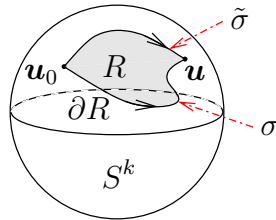
Choose $\mathbf{u}_0 \in S^k$. Since S^k is path connected, we can select for each $\mathbf{u} \in S^k$ a path $\sigma : [0, 1] \rightarrow S^k$ between $\mathbf{u}_0$ and $\mathbf{u}$. Let

$$f(\mathbf{u}) = \int_{\sigma} \omega = \int_{[0,1]} \sigma^*(\omega) .$$

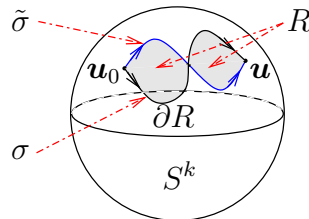
This function is well defined because the value of the integral is independent of the path chosen between $\mathbf{u}_0$ and $\mathbf{u}$. To prove this claim, suppose that $\tilde{\sigma} : [0, 1] \rightarrow S^k$ is another path between $\mathbf{u}_0$ and $\mathbf{u}$. Then $\sigma - \tilde{\sigma}$ is a closed path in S^k which is the boundary of a region $R \subset S^k$. Thus, according to Stokes' theorem,

$$\int_{\sigma} \omega - \int_{\tilde{\sigma}} \omega = \int_{\sigma - \tilde{\sigma}} \omega = \int_{\partial R} \omega = \int_R d\omega = 0 \quad (7.2.6)$$

because ω is closed. Thus $\int_{\sigma} \omega = \int_{\tilde{\sigma}} \omega$. As a little note of caution, we have assumed in (7.2.6) that $\sigma([0, 1]) \cap \tilde{\sigma}([0, 1]) = \{\mathbf{u}_0, \mathbf{u}\}$ and that the induced orientation on $\partial R = \sigma([0, 1]) \cup \tilde{\sigma}([0, 1])$ from R (and so S^k) matches the orientation provided by $\sigma - \tilde{\sigma}$.



We still get that $\int_{\sigma} \omega = \int_{\tilde{\sigma}} \omega$ even if $\sigma([0, 1])$ and $\tilde{\sigma}([0, 1])$ intersect at points other than $\mathbf{u}_0$ and $\mathbf{u}$ as in the following figure.



Stokes' theorem can be used by spitting the integral. Nevertheless, the conclusion is the same because $d\omega = 0$ on R .

We now shows that $df = \omega$. Since ω is arbitrary, this will prove that $[\omega] = [0]$ and so that $H^1(S^k) = 0$.

To compute the derivative of f at $\mathbf{u}_1 \in S^k$, we only need to know the value of f in an open neighbourhood U of $\mathbf{u}_1$. We may assume that (W, U, ϕ) is a local chart about $\mathbf{u}_1$. Let σ_1 be a path between $\mathbf{u}_0$ and $\mathbf{u}_1$. Given $\mathbf{u}_2 \in U$, let σ_2 be a path between $\mathbf{u}_1$ and $\mathbf{u}_2$ in U . Then $\sigma_1 + \sigma_2$ is a path between $\mathbf{u}_0$ and $\mathbf{u}_2$ that can be used to compute $f(\mathbf{u}_2)$. Hence

$$f(\mathbf{u}_2) - f(\mathbf{u}_1) = \int_{\sigma_1 + \sigma_2} \omega - \int_{\sigma_1} \omega = \int_{\sigma_2} \omega = \int_{[0,1]} \sigma_2^*(\omega) .$$

Since $\sigma_2([0, 1]) \subset U$, we get

$$f(\mathbf{u}_2) - f(\mathbf{u}_1) = \int_{[0,1]} \sigma_2^*(\omega) = \int_{[0,1]} (\phi^{-1} \circ \sigma_2)^* \phi^*(\omega) . \quad (7.2.7)$$

Suppose that the local representation of ω in W is given by $\phi^* \omega = \sum_{j=1}^k \tilde{\omega}_j dw_j$. If we show that $\phi^*(df) = d(\phi^*(f)) = \phi^*(\omega)$ on W , then we will have that $df = \omega$ because $\mathbf{u}_1$ and the local chart (W, U, ϕ) about $\mathbf{u}_1$ are arbitrary.

Suppose that $\mathbf{u}_1 = \phi(\mathbf{w}_1)$. To prove that $d(\phi^*(f))(\mathbf{w}_1) = \phi^*(\omega)(\mathbf{w}_1)$, we need to prove that $\frac{\partial}{\partial w_j}(\phi^*(f))(\mathbf{w}_1) = \tilde{\omega}_j(\mathbf{w}_1)$ for $1 \leq j \leq k$. We have that

$$\frac{1}{h}((\phi^*(f))(\mathbf{w}_1 + h\mathbf{e}_j) - \phi^*(f)(\mathbf{w}_1)) = \frac{1}{h}(f(\phi(\mathbf{w}_1 + h\mathbf{e}_j)) - f(\phi(\mathbf{w}_1))) = \frac{1}{h} \int_{[0,1]} \sigma_h^* \phi^*(\omega)$$

where $\sigma_h : [0, 1] \rightarrow W$ is the path from $\mathbf{w}_1$ to $\mathbf{w}_1 + h\mathbf{e}_j$ defined by $\sigma_h(s) = \mathbf{w}_1 + sh\mathbf{e}_j$ for $0 \leq s \leq 1$. We obviously assume that h is small enough. The path σ_2 in (7.2.7) is given by $\sigma_2 = \phi \circ \sigma_h$. We get

$$\begin{aligned} \frac{\partial}{\partial w_j}(\phi^*(f)) &= \lim_{h \rightarrow 0} \frac{1}{h} (\phi^*(f)(\mathbf{w}_1 + h\mathbf{e}_j) - \phi^*(f)(\mathbf{w}_1)) = \lim_{h \rightarrow 0} \frac{1}{h} \int_{[0,1]} \sigma_h^* \phi^*(\omega) \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \int_{[0,1]} \left(\sum_{j=1}^k \tilde{\omega}_j(\sigma_h(s)) \left(\frac{\partial}{\partial s}(\sigma_h(s))_j \right) ds \right) = \lim_{h \rightarrow 0} \frac{1}{h} \int_{[0,1]} \tilde{\omega}_j(\mathbf{w}_1 + sh\mathbf{e}_j) h ds \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \int_0^h \tilde{\omega}_j(\mathbf{w}_1 + t\mathbf{e}_j) dt = \tilde{\omega}_j(\mathbf{w}_1) . \quad \blacksquare \end{aligned}$$

It follows from (7.2.5) and Proposition 7.2.43 that $H^q(S^{k-1}) = 0$ for $1 < q+1 < k$ because $H^q(S^{k-1}) \cong H^{q-1}(S^{k-2}) \cong H^{q-2}(S^{k-3}) \cong \dots \cong H^1(S^{k-q}) = 0$ for $k-q > 1$ and $H^{k-1}(S^{k-1}) \cong \mathbb{R}$ for $k > 1$ because $H^{k-1}(S^{k-1}) \cong H^{k-2}(S^{k-2}) \cong \dots \cong H^1(S^1) \cong \mathbb{R}$ for $k > 1$. For the 0^{th} cohomology module on S^{k-1} , we have that $H^0(S^{k-1}) \cong \mathbb{R}$ for $k > 1$ according to Proposition 7.2.20.

7.3 Applications

There are many very important results that can be obtained using cohomology. We present a couple of them in this section.

Let S_1 and S_2 be two connected and oriented compact k -dimensional smooth manifolds without boundary. Suppose moreover that $f : S_1 \rightarrow S_2$ is a function of class C^∞ . Is there a relation between $\int_{S_2} \omega$ and $\int_{S_1} f^*(\omega)$ for all differential k -forms ω on S_2 ?

Choose a differential k -form ω_0 on S_2 such that $\int_{S_2} \omega_0 \neq 0$. Such a differential k -form exists according to Proposition 7.2.7 because $H^k(S_2) \cong \mathbb{R}$ according to Theorem 7.2.9. There is a number n such that

$$\int_{S_2} f^*(\omega_0) = n \int_{S_2} \omega_0. \quad (7.3.1)$$

For any other differential k -form ω on S_2 , we have that $\omega = a_\omega \omega_0 + d\mu_\omega$ for some differential $(k-1)$ -form μ_ω on S_2 and $a_\omega \in \mathbb{R}$. Hence, using Stokes' theorem, we find that

$$\int_{S_2} \omega = a_\omega \int_{S_2} \omega_0 + \int_{S_2} d\mu = a_\omega \int_{S_2} \omega_0 + \int_{\partial S_2} \mu = a_\omega \int_{S_2} \omega_0$$

and

$$\begin{aligned} \int_{S_1} f^*(\omega) &= a_\omega \int_{S_1} f^*(\omega_0) + \int_{S_1} f^*(d\mu) = a_\omega \int_{S_1} f^*(\omega_0) + \int_{S_1} d(f^*(\mu)) \\ &= a_\omega \int_{S_1} f^*(\omega_0) + \int_{\partial S_1} f^*(\mu) = a_\omega \int_{S_1} f^*(\omega_0) \end{aligned}$$

because $\partial S_i = \emptyset$ for $1 \leq i \leq 2$. Thus

$$\int_{S_1} f^*(\omega) = a_\omega \int_{S_1} f^*(\omega_0) = n a_\omega \int_{S_2} \omega_0 = n \int_{S_2} \omega$$

for all differential k -form ω on S_2 . The constant n depends only on f . This in itself is pretty amazing. But there is something more amazing to come.

The previous discussion is also valid if we consider connected and oriented k -dimensional smooth manifolds S_1 and S_2 without boundary, and assume that the function $f : S_1 \rightarrow S_2$ is of class C^∞ and **proper**; namely, $f^{-1}(K)$ is compact if K is compact. Obviously, we have in that context to consider $H_c^k(S_i) \cong \mathbb{R}$ for $1 \leq i \leq 2$ instead of $H^k(S_i) \cong \mathbb{R}$ for $1 \leq i \leq 2$.

Definition 7.3.1

The number n defined in (7.3.1) is called the **degree** of f and is denoted $\deg(f)$.

The surprising fact is that $\deg(f)$ is an integer number.

Theorem 7.3.2

Let S_1 and S_2 be two connected and oriented k -dimensional smooth manifolds without boundary, and $f : S_1 \rightarrow S_2$ be a proper function of class C^∞ . Suppose that $\mathbf{v}$ is a regular value of f and set, for each $\mathbf{u} \in f^{-1}(\{\mathbf{v}\})$, $\text{sgn}_{\mathbf{u}}(f) = 1$ if $f_* : T_{\mathbf{u}}S_1 \rightarrow T_{\mathbf{v}}S_2$ is orientation preserving and $\text{sgn}_{\mathbf{u}}(f) = -1$ if $f_* : T_{\mathbf{u}}S_1 \rightarrow T_{\mathbf{v}}S_2$ is not orientation

preserving. Then

$$\deg(f) = \sum_{\mathbf{u} \in f^{-1}(\{\mathbf{v}\})} \operatorname{sgn}_{\mathbf{u}}(f) .$$

Proof.

It follows from Proposition 3.7.2 that $f^{-1}(\{\mathbf{v}\})$ is the union of isolated points⁸. Since f is proper, $f^{-1}(\{\mathbf{v}\})$ is also a compact set. Therefore, $f^{-1}(\{\mathbf{v}\})$ is the finite union of isolated points.

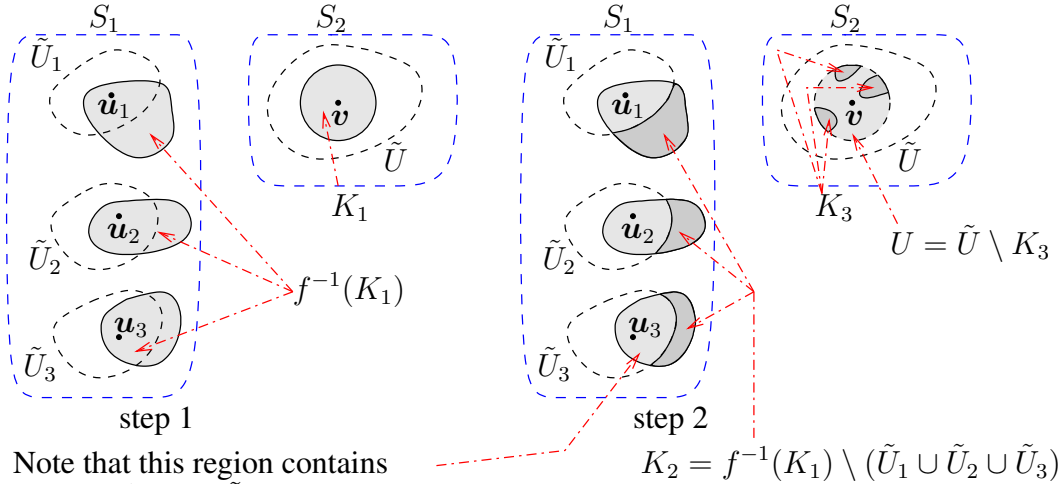
Suppose that $f^{-1}(\{\mathbf{v}\}) = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$. We now present one method (suggested in [27]) to select local charts $\{(W_i, U_i, \phi_i)\}_{1 \leq i \leq q}$ of S_1 and a local chart (W, U, ϕ) of S_2 such that $U_i \cap U_j = \emptyset$ for $i \neq j$, $\mathbf{u}_i \in U_i$ for $1 \leq i \leq q$, $\mathbf{v} \in U$, and $f^{-1}(U) = \bigcup_{1 \leq i \leq q} U_i$.

Let $\{(\tilde{W}_i, \tilde{U}_i, \tilde{\phi}_i)\}_{1 \leq i \leq q}$ be local charts of S_1 with $\mathbf{u}_i \in \tilde{U}_i$ for $1 \leq i \leq q$ and $(\tilde{W}, \tilde{U}, \tilde{\phi})$ be a local chart of S_2 such that $\mathbf{v} \in \tilde{U}$. Since the points $\mathbf{u}_i$ are isolated, we may assume by shrinking the sets $\tilde{U}_i$ if needed that $\tilde{U}_i \cap \tilde{U}_j = \emptyset$ if $i \neq j$.

According to Proposition 7.2.35, we may choose an open neighbourhood $\check{U}$ of $\mathbf{v}$ with compact closure $K_1 \subset \check{U}$. Since f is a proper function, $f^{-1}(K_1)$ is a compact subset of S_1 . Thus $K_2 = f^{-1}(K_1) \setminus (\bigcup_{1 \leq i \leq q} \tilde{U}_i) \subset f^{-1}(\check{U})$ is a compact subset of S_1 . Since f is continuous, $K_3 = f(K_2)$ is a compact subset of $\check{U}$. So K_3 is a closed subset of $\check{U}$ that does not contain $\mathbf{v}$.

The local charts that we were looking for are defined as it follows. Let $U = \check{U} \setminus K_3$ and $W = \phi^{-1}(U)$. The local chart of S_2 that contains $\mathbf{v}$ is given by (W, U, ϕ) . Let $U_i = \tilde{U}_i \cap f^{-1}(U)$ for $1 \leq i \leq q$. Note that $\mathbf{u}_i \in U_i$ for $1 \leq i \leq q$ because $\mathbf{v} \in U$. The local charts of S_1 are given by $(W_i, U_i, \phi_i) = (\phi_i^{-1}(U_i), U_i, \phi_i)$ for $1 \leq i \leq q$.

⁸We leave it to the reader to generalize Proposition 3.7.2 to a map between manifolds. Note that $\mathbf{u}_2 \in S_2$ is a regular value of a continuously differentiable function $f : S_1 \rightarrow S_2$ between two k -dimensional smooth manifolds S_1 and S_2 if $\phi_2^{-1}(\mathbf{u}_2)$ is a regular value of $\phi_2^{-1} \circ f \circ \phi_1$ for all local charts (W_1, U_1, ϕ_1) of S_1 and local charts (W_2, U_2, ϕ_2) of S_2 with $\phi_2^{-1}(\mathbf{u}_2) \in W_2$. The fact that countable atlases can be used according to Proposition 7.2.36 is also needed to generalize Proposition 3.7.2.



Note that this region contains $U_3 = f^{-1}(U) \cap \tilde{U}_3$ but is not equal to it because of boundary points are not included. We have a similar situation for U_1 and U_2 .

Let $g : W \rightarrow \mathbb{R}$ be a non null smooth function with compact support in W and such that $g(\mathbf{w}) \geq 0$ for all $\mathbf{w} \in W$. We consider the differential k -form ω on S_2 with the local representation $\phi^*(\omega) = g dw_1 \wedge dw_2 \wedge \dots \wedge dx_k$ on U and null on $S_2 \setminus U$. Since $\text{supp } f^*(\omega) \subset f^{-1}(\text{supp } \omega) \subset f^{-1}(U) = \bigcup_{1 \leq i \leq q} U_i$ and $U_i \cap U_j = \emptyset$ for $i \neq j$, we get

$$\int_{S_1} f^*(\omega) = \sum_{i=1}^q \int_{U_i} f^*(\omega) = \sum_{i=1}^q \left(\text{sgn}_{\mathbf{u}_i}(f) \int_U \omega \right) = \underbrace{\left(\sum_{i=1}^q \text{sgn}_{\mathbf{u}_i}(f) \right)}_{=n} \int_U \omega = n \int_{S^2} \omega$$

because $\int_{U_i} f^*(\omega) = \int_U \omega$ if f is orientation preserving and $\int_{U_i} f^*(\omega) = -\int_U \omega$ if f does not preserve the orientation (i.e. reverse the orientation) according to Proposition 5.6.8.

Note that ϕ and ϕ_i for $1 \leq i \leq q$ are orientation preserving, and $\det(\phi \circ f \circ \phi_i) \neq 0$ on W_i for $1 \leq i \leq q$. Therefore, either $f_* : T_{\mathbf{u}} S_1 \rightarrow T_{\mathbf{v}} S_2$ is orientation preserving for all $\mathbf{u} \in U_i$ or $f_* : T_{\mathbf{u}} S_1 \rightarrow T_{\mathbf{v}} S_2$ reverses the orientation for all $\mathbf{u} \in U_i$. To change the orientation on W_i , $\det(\phi \circ f \circ \phi_i)$ would have to change sign and so be null at some point in W_i . ■

Corollary 7.3.3

Let S_1 and S_2 be two connected and oriented compact k -dimensional smooth manifolds without boundary. If $f, g : S_1 \rightarrow S_2$ are two homotopic functions of class C^∞ , then $\deg(f) = \deg(g)$.

Proof.

Since f and g are homotopic, there exists a smooth map $H : S_1 \times [0, 1] \rightarrow S_2$ such that

$H(\mathbf{u}, 0) = f(\mathbf{u})$ and $H(\mathbf{u}, 1) = g(\mathbf{u})$ for all $\mathbf{u} \in S_1$. Let $H_t(\mathbf{u}) = H(\mathbf{u}, t)$ for all $(\mathbf{u}, t) \in S_1 \times [0, 1]$. If ω is a differential k -form on S_2 , then we have that

$$\int_{S_1} H_t^*(\omega) = n_t \int_{S_2} \omega$$

for each $t \in [0, 1]$ where $n_t \in \mathbb{Z}$. Since the value of the integral on the left side varies continuously with respect to t , we have that $t \mapsto n_t$ is an integer valued continuous function. Therefore n_t is constant. Hence $\deg(f) = n_0 = n_1 = \deg(g)$. ■

We have defined vector fields on a k -dimensional manifold at the beginning of Section 5.5 but we have not referred to them until now. The next result is very interesting. It is often humorously stated as saying that there is no way to comb our hair and not have one hair standing perpendicular. This is true if we assume that one's head is a ball fully covered with hair. But, other than for cartoon characters, we have not seen any such head yet.

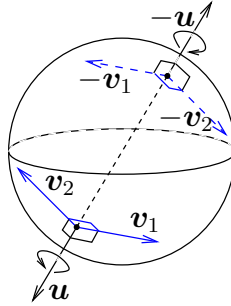
Proposition 7.3.4

There does not exist a nowhere null vector field on S^{k-1} if k is odd.

Proof.

If $g_0 : S^{k-1} \rightarrow S^{k-1}$ is the restriction to S^{k-1} of the identity map $\text{Id} : \mathbb{R}^k \rightarrow \mathbb{R}^k$, then $\deg(g_0) = 1$ because $\det \text{Id} = 1$.

If $g_1 : S^{k-1} \rightarrow S^{k-1}$ is the restriction to S^{k-1} of the map $-\text{Id} : \mathbb{R}^k \rightarrow \mathbb{R}^k$, so g_1 is the antipodal map defined by $g(\mathbf{u}) = -\mathbf{u}$ for all $\mathbf{u} \in S^{k-1}$, then $\deg(g_1) = -1$ because $\det(-\text{Id}) = -1$ for k odd.



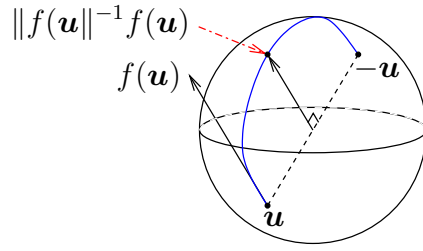
We have from Corollary 7.3.3 that g_1 cannot be smoothly homotopic to g_0 .

Suppose that there is a nowhere null vector field $F : S^{k-1} \rightarrow \bigcup_{\mathbf{u} \in S} T_{\mathbf{u}} S^{k-1}$. Let $F(\mathbf{u}) = (\mathbf{u}, f(\mathbf{u})) \in T_{\mathbf{u}} S^{k-1}$ for all $\mathbf{u} \in S^{k-1}$. We have that $f : S^{k-1} \rightarrow \mathbb{R}^k$ is a smooth function such that $f(\mathbf{u}) \neq \mathbf{0}$ for all $\mathbf{u} \in S^{k-1}$ because F is nowhere null.

Let $H : S^{k-1} \times [0, 1] \rightarrow S^{k-1}$ be the function defined by

$$H(\mathbf{u}, t) = \cos(\pi t)\mathbf{u} + \sin(\pi t) \|f(\mathbf{u})\|^{-1} f(\mathbf{u})$$

for all $(\mathbf{u}, t) \in S^{k-1}$. The function H is an homotopy between g_0 and g_1 . This is a contradiction.

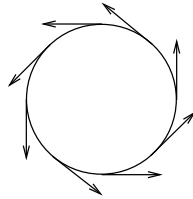


Note that $\|H(\mathbf{u}, t)\| = 1$ for all $(\mathbf{u}, t) \in S^{k-1}$ because $f(\mathbf{u})$ is perpendicular to $\mathbf{u}$ by definition of $T_{\mathbf{u}}S^{k-1}$. Hence, if we set $\mathbf{y} = \|f(\mathbf{u})\|^{-1}f(\mathbf{u})$ for $\mathbf{u} \in S^{k-1}$, then we get

$$\begin{aligned} \|H(\mathbf{u}, t)\|^2 &= \sum_{j=1}^k (\cos(\pi t)u_j + \sin(\pi t)y_j)^2 \\ &= \cos^2(\pi t) \underbrace{\|\mathbf{u}\|^2}_{=1} + 2\cos(\pi t)\sin(\pi t) \underbrace{\mathbf{u} \cdot \mathbf{y}}_{=0} + \sin^2(\pi t) \underbrace{\|\mathbf{y}\|^2}_{=1} = 1 \end{aligned}$$

for $t \in [0, 1]$. In fact, $t \mapsto H(\mathbf{u}, t)$ is a parametric representation of half the great circle from $\mathbf{u}$ to $-\mathbf{u}$ that goes through $\|f(\mathbf{u})\|^{-1}f(\mathbf{u})$. ■

If k is even, then it is always possible to define a nowhere null vector field of S^{k-1} . For instance, $F : S^{k-1} \rightarrow \bigcup_{\mathbf{u} \in S} T_{\mathbf{u}}S^{k-1}$ defined by $F(\mathbf{u}) = (\mathbf{u}, f(\mathbf{u}))$ for $\mathbf{u} \in S^{k-1}$ where $f(\mathbf{u}) = (-u_2, u_1, -u_4, u_3, \dots, -u_k, u_{k-1})$ is such a vector field. It is represented in the figure below for S^1 .



Chapter 8

Algebraic Topology

It would be a mistake to talk about cohomology theory without introducing homology theory. We have seen that it is quite challenging to compute the cohomology module of even some simple manifolds. Homology theory can greatly simplify the task of computing cohomology modules of some manifolds. In this chapter, we explain how the “homology module” of a manifold S can give us the cohomology module of S . We also provide a brief introduction to the major tools to find “homology modules.”

This chapter should not be treated as a full introduction to algebraic topology. We present only the important results about homology and cohomology theory to reach the goal mentioned in the previous paragraph. There are fundamentally two approaches to these two topics: “simplicial homology and cohomology” and “singular homology and cohomology”. The first approach is more geometric and is mostly associated to algebraic geometry while the second approach is purely algebraic and is associated to algebraic topology. We will briefly explore both approaches.

We present a very succinct introduction to simplicial homology and cohomology in Sections 8.2 and 8.3. We provide a little more complete introduction to singular homology and cohomology in Sections 8.4 and 8.5. The importance of these sections is due to De Rham’s theorem, Theorem 8.5.23, in Subsection 8.5.5 that provides a relation between singular cohomology and de Rham cohomology as presented in the previous chapter.

There is also a version of de Rham theorem, Theorem 8.3.15, relating de Rham cohomology and simplicial cohomology. We only sketch this proof. We have chosen to give a full proof of de Rham theorem in the context of singular cohomology only. In fact, as we show in Section 8.6, there is an equivalence between the simplicial and singular theory. Therefore, from a practical point of view, there is no need to study both approaches. From a pedagogical point of view, it is nice to study both. The readers whose main interest is the study of de Rham cohomology may choose to only browse through Sections 8.2 and 8.3 on simplicial homology and cohomology, and Section 8.6. These three sections are more interesting for those who plan to study algebraic topology later.

Readers interested in learning more about the fascinating subject of algebraic geometry should consult [25] for a good basic introduction. As for algebraic topology, [9] is excellent

references. The content of this chapter is mainly based on these two references and on [15, 21].

This chapter requires some basic knowledge of abstract algebra; in particular, of group theory. There are many good books in algebra. Two of them are [11, 12].

8.1 Fundamental Groups

All introductions to algebraic topology start with a study of fundamental groups associated to a topological spaces even if this is not the end goal. It is a very visual introduction to the subject of algebraic topology. We provide only a brief overview of this subject, leaving many of the stated results without proofs. As we said in the introduction to this chapter, the reader may find proofs of all the results that we state without proofs in [9, 21, 25] for instance.

The fundamental concept in this section is the notion of homotopy between continuous functions. We have already defined the concept of homotopic functions in Definition 7.2.24 in the context of maps between manifolds. We generalize this concept to any topological spaces.

Note 8.1.1

From now on, when we talk about a topological space, we assume that it is a Hausdorff space.

Definition 8.1.2

Let X and Y be two topological spaces and $f, g : X \rightarrow Y$ be two continuous functions. We say that f is **homotopic** to g if there exists a continuous map $H : X \times [0, 1] \rightarrow Y$ such that $H(x, 0) = f(x)$ and $H(x, 1) = g(x)$ for all $x \in X$. We write $f \sim g$. The function H is called an **homotopy** from f to g .

This definition of homotopic functions differs from the definition given in Definition 7.2.24 because we now require only that functions be continuous. Moreover, manifolds are replaced by any topological spaces. The concept of homotopy between continuous functions defines an equivalence relation. Namely, if $f, g, h : X \rightarrow Y$ are continuous functions between two topological spaces X and Y , then

1. $f \sim f$,
2. $f \sim g$ if and only if $g \sim f$, and
3. $f \sim g$ and $g \sim h$ imply that $f \sim h$.

Definition 8.1.3

A topological space X is **contractible** if there exists $x_0 \in X$ such that the identity function $\text{Id}_X : X \rightarrow X$ is homotopic to the constant function $f : X \rightarrow X$ defined by $f(x) = x_0$ for all $x \in X$.

As for the previous definition for homotopy between functions, we only require a continuous homotopy for the definition of contractible spaces instead of a smooth homotopy as in definition 7.1.1.

We also slightly modify our definition of path between two points given in Definition 6.2.1 for the context where we require only that functions be continuous.

Definition 8.1.4

Let X be a topological space. A **path from** $x_0 \in X$ **to** $x_1 \in X$ is a continuous function $\alpha : [0, 1] \rightarrow X$ such that $\alpha(0) = x_0$ and $\alpha(1) = x_1$.

Definition 8.1.5

Let X be a topological space. If α is a path from $x_0 \in X$ to $x_1 \in X$ and β is a path from $x_1 \in X$ to $x_2 \in X$, then the **product** of α and β , denoted $\alpha\beta$, is the path from x_0 to x_2 defined by

$$(\alpha\beta)(t) = \begin{cases} \alpha(2t) & \text{if } 0 \leq t \leq 1/2 \\ \beta(2t - 1) & \text{if } 1/2 < t \leq 1 \end{cases}$$

for $0 \leq t \leq 1$. Also, α^{-1} is the path from x_1 to x_0 defined by $\alpha^{-1}(t) = \alpha(1 - t)$ for $0 \leq t \leq 1$.

The product $\alpha\beta$ is what we informally called the union of the paths α and β in Chapter 6. In this section, we are really strict in preserving the interval $[0, 1]$ of definition of a path. We are not just interested in the image of α and β as in Chapter 6.

Definition 8.1.6

Let X be a topological space. Suppose that α and β are two paths from $x_0 \in X$ to $x_1 \in X$. We say that α is **homotopic** to β if there exists a continuous map $H : [0, 1] \times [0, 1] \rightarrow X$ such that $H(t, 0) = \alpha(t)$ and $H(t, 1) = \beta(t)$ for $0 \leq t \leq 1$, and $H(0, s) = x_0$ and $H(1, s) = x_1$ for $0 \leq s \leq 1$. We write $\alpha \underset{X}{\sim} \beta$. If there is no risk of confusion for the range X of H , then we may simply write $\alpha \sim \beta$.

The path α is deformed to the path β without changing the end points (Figure 8.1).

The homotopy of paths between two fixed points x_0 and x_1 of a topological space X defines an equivalence relation. Suppose that α , β and γ are paths from x_0 to x_1 . Then

1. $\alpha \sim \alpha$,

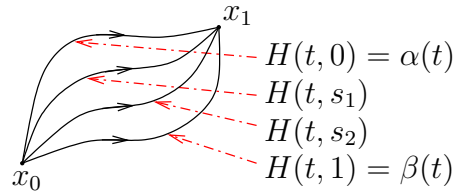


Figure 8.1: The path α from x_0 to x_1 is homotopic to the path β from x_0 to x_1 . We have $0 < s_1, s_2 < 1$ in the figure.

2. $\alpha \sim \beta$ if and only if $\beta \sim \alpha$, and
3. $\alpha \sim \beta$ and $\beta \sim \gamma$ imply that $\alpha \sim \gamma$.

We prove that last condition and left the proof of the two others to the reader. Since $\alpha \sim \beta$, there exists a continuous function $H_1 : [0, 1] \times [0, 1] \rightarrow X$ such that $H_1(t, 0) = \alpha(t)$ and $H_1(t, 1) = \beta(t)$ for $0 \leq t \leq 1$, and $H_1(0, s) = x_0$ and $H_1(1, s) = x_1$ for $0 \leq s \leq 1$. Since $\beta \sim \gamma$, there exists a continuous function $H_2 : [0, 1] \times [0, 1] \rightarrow X$ such that $H_2(t, 0) = \beta(t)$ and $H_2(t, 1) = \gamma(t)$ for $0 \leq t \leq 1$, and $H_2(0, s) = x_0$ and $H_2(1, s) = x_1$ for $0 \leq s \leq 1$. Let

$$H(t, s) = \begin{cases} H_1(t, 2s) & \text{if } 0 \leq s < 1/2 \text{ and } 0 \leq t \leq 1 \\ H_1(t, 2s - 1) & \text{if } 1/2 \leq s \leq 1 \text{ and } 0 \leq t \leq 1 \end{cases}$$

Then $H : [0, 1] \times [0, 1] \rightarrow X$ is a continuous function such that $H(t, 0) = \alpha(t)$ and $H(t, 1) = \gamma(t)$ for $0 \leq t \leq 1$, and $H(0, s) = x_0$ and $H(1, s) = x_1$ for $0 \leq s \leq 1$. Thus $\alpha \sim \gamma$.

It is not hard to prove that if α_1, α_2 are two paths from $x_0 \in X$ to $x_1 \in X$ such that $\alpha_1 \sim \alpha_2$, and β_1, β_2 are two paths from $x_1 \in X$ to $x_2 \in X$ such that $\beta_1 \sim \beta_2$, then $\alpha_1\beta_1$ and $\alpha_2\beta_2$ are two paths from x_0 to x_2 such that $\alpha_1\beta_1 \sim \alpha_2\beta_2$. Moreover, $\alpha_1^{-1}, \alpha_2^{-1}$ are two paths from $x_1 \in X$ to $x_0 \in X$ such that $\alpha_1^{-1} \sim \alpha_2^{-1}$.

Definition 8.1.7

Let X be a topological space and $x_0 \in X$. A closed path in X from x_0 to x_0 is called a **loop** at x_0 .

Let X be a topological space and $x_0 \in X$. The set of equivalence classes $[\alpha]$ of homotopic loops α at x_0 is denoted $\pi_1(X, x_0)$. It is a group where the product is defined by $[\alpha][\beta] = [\alpha\beta]$ for all loops α and β at x_0 , the inverse is defined by $[\alpha]^{-1} = [\alpha^{-1}]$ for all loops α at x_0 , and the identity element is defined by $[e_{x_0}]$ for the path e_{x_0} given by $e_{x_0}(t) = x_0$ for $0 \leq t \leq 1$. All these algebraic operations are well defined according to what we said in the paragraph preceding the definition above.

Definition 8.1.8

The group $\pi_1(X, x_0)$ is called the **fundamental group** or **1st homotopy group** of the pair (X, x_0) .

Proposition 8.1.9

Suppose that $f : X \rightarrow Y$ is a continuous function between two path-connected topological spaces and that $x_0 \in X$. Then the map $f_* : \pi_1(X, x_0) \rightarrow \pi_1(Y, f(x_0))$ defined by $f_*([\alpha]) = [f \circ \alpha]$ for all $[\alpha] \in \pi_1(X, x_0)$ is an homomorphism.

Proof.

i) The map f_* is well defined. If σ_0 and σ_1 are two loops in X at x_0 such that $\sigma_0 \sim \sigma_1$, then there exists $H : [0, 1] \times [0, 1] \rightarrow X$ such that $H(0, s) = H(1, s) = x_0$ for $0 \leq s \leq 1$, and $H(t, 0) = \sigma_0(t)$ and $H(t, 1) = \sigma_1(t)$ for $0 \leq t \leq 1$.

We have that $f \circ \sigma_0$ and $f \circ \sigma_1$ are two loops in Y at $f(x_0)$. Let $\tilde{H} = f \circ H$. Then $\tilde{H} : [0, 1] \times [0, 1] \rightarrow Y$ is such that $\tilde{H}(0, s) = f(H(0, s)) = f(x_0)$ and $\tilde{H}(1, s) = f(H(1, s)) = f(x_0)$ for $0 \leq s \leq 1$, and $\tilde{H}(t, 0) = f(H(t, 0)) = (f \circ \sigma_0)(t)$ and $\tilde{H}(t, 1) = f(H(t, 1)) = (f \circ \sigma_1)(t)$ for $0 \leq t \leq 1$. Thus $f \circ \sigma_0 \sim f \circ \sigma_1$.

Hence $f_*([\sigma])$ is independent of the representative path of $[\sigma]$ selected.

ii) The map f_* is an homomorphism because

$$\begin{aligned} f_*([\alpha_0][\alpha_1]) &= f_*([\alpha_0 \alpha_1]) = [f \circ (\alpha_0 \alpha_1)] = [(f \circ \alpha_0)(f \circ \alpha_1)] \\ &= [f \circ \alpha_0][f \circ \alpha_1] = f_*([\alpha_0])f_*([\alpha_1]) \end{aligned}$$

for all loops σ_0 and σ_1 in X at x_0 . ■

Suppose that X , Y and Z are three path-connected topological spaces, that $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ are continuous functions, and that $x_0 \in X$. Then it is easy to prove that $g_* \circ f_* = (g \circ f)_* : \pi_1(X, x_0) \rightarrow \pi_1(Z, g(f(x_0)))$.

Proposition 8.1.10

Suppose that $f : X \rightarrow Y$ is a homeomorphism between two path-connected topological spaces and that $x_0 \in X$. Then the map $f_* : \pi_1(X, x_0) \rightarrow \pi_1(Y, f(x_0))$ is an isomorphism.

Proof.

It follows from $f^{-1} \circ f = \text{Id}_X$ that $(f^{-1})_* \circ f_* = (f^{-1} \circ f)_* = (\text{Id}_X)_* = \text{Id}_{\pi_1(X, x_0)}$ and from $f \circ f^{-1} = \text{Id}_Y$ that $f_* \circ (f^{-1})_* = (f \circ f^{-1})_* = (\text{Id}_Y)_* = \text{Id}_{\pi_1(Y, f(x_0))}$. Thus f_* has an inverse given by $(f_*)^{-1} = (f^{-1})_*$. ■

Proposition 8.1.11

Suppose that X is a path-connected topological space and that $x_0, x_1 \in X$. Then $\pi_1(X, x_0) \cong \pi_1(X, x_1)$.

Proof.

Since X is path-connected, there exists a path β in X from x_0 to x_1 . Consider $\beta_\# : \pi_1(X, x_0) \rightarrow \pi_1(X, x_1)$ defined by $\beta_\#([\sigma]) = [\beta^{-1}\sigma\beta]$ for all $\sigma \in \pi_1(X, x_0)$.

i) We first note that $\beta_\#$ is well defined. If σ_0 and σ_1 are two loops in X at x_0 such that $\sigma_0 \sim \sigma_1$, then there exists $H : [0, 1] \times [0, 1] \rightarrow X$ such that $H(0, s) = H(1, s) = x_0$ for $0 \leq s \leq 1$, and $H(t, 0) = \sigma_0(t)$ and $H(t, 1) = \sigma_1(t)$ for $0 \leq t \leq 1$.

We have that $\beta^{-1}\sigma_0\beta$ and $\beta^{-1}\sigma_1\beta$ are two loops in X at x_1 . Let $\tilde{H} : [0, 1] \times [0, 1] \rightarrow X$ be the function defined by

$$\tilde{H}(t, s) = \begin{cases} \beta^{-1}(t) & \text{if } 0 \leq t \leq 1/4 \text{ and } 0 \leq s \leq 1 \\ H(4t - 1, s) & \text{if } 1/4 < t < 1/2 \text{ and } 0 \leq s \leq 1 \\ \beta(t) & \text{if } 1/2 \leq t \leq 1 \text{ and } 0 \leq s \leq 1 \end{cases}$$

Then $\tilde{H}(0, s) = \beta^{-1}(0) = x_1$ and $\tilde{H}(1, s) = \beta(1) = x_1$ for $0 \leq s \leq 1$, and

$$\begin{aligned} \tilde{H}(t, 0) &= \begin{cases} \beta^{-1}(t) & 0 \leq t \leq 1/4 \\ \sigma_0(4t - 1) & \text{if } 1/4 < t < 1/2 \\ \beta(t) & \text{if } 1/2 \leq t \leq 1 \end{cases} \\ &= (\beta^{-1}\sigma_0\beta)(t) \end{aligned}$$

and

$$\begin{aligned} \tilde{H}(t, 1) &= \begin{cases} \beta^{-1}(t) & 0 \leq t \leq 1/4 \\ \sigma_1(4t - 1) & \text{if } 1/4 < t < 1/2 \\ \beta(t) & \text{if } 1/2 \leq t \leq 1 \end{cases} \\ &= (\beta^{-1}\sigma_1\beta)(t) \end{aligned}$$

for $0 \leq t \leq 1$. Thus $\beta^{-1}\sigma_0\beta \sim \beta^{-1}\sigma_1\beta$.

ii) The map $\beta_\#$ is a homomorphism because

$$\begin{aligned} \beta_\#([\sigma_0][\sigma_1]) &= \beta_\#([\sigma_0\sigma_1]) = [\beta^{-1}\sigma_0\sigma_1\beta] = [\beta^{-1}\sigma_0 \underbrace{\beta\beta^{-1}}_{\sim e_{x_0}} \sigma_1\beta] \\ &= [\beta^{-1}\sigma_0\beta][\beta^{-1}\sigma_1\beta] = \beta_\#([\sigma_0])\beta_\#([\sigma_1]) \end{aligned}$$

for all loops σ_0 and σ_1 in X at x_0 .

iii) The map $\beta_\#$ is an isomorphism. In fact, it is easy to verify that $\beta_\#^{-1} = (\beta^{-1})_\#$. ■

If X is a path-connected topological space, then it follows from the previous proposition that the fundamental group of a pair (X, x_0) is, up to isomorphism, the same group independently of the base point $x_0 \in X$ chosen. Hence, the following definition is justified.

Definition 8.1.12

Let X be a path-connected topological space. The **fundamental group** of X , denoted $\pi_1(X)$, is the group $\pi_1(X, x_0)$ for any base point $x_0 \in X$.

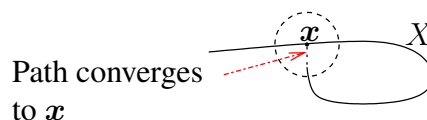
8.1.1 Covering Spaces

In this section, we prove only the results that will be used later. The missing proofs and more information about covering spaces can be found in [9, 25].

Definition 8.1.13

A topological space X is locally path-connected if for every $x \in X$ and open neighbourhood $V \subset X$ of x , there exists an open neighbourhood $U \subset V$ such that U is path-connected.

We insist on the fact that U is path-connected in the previous definition. Namely, given any $x_1, x_2 \in U$, there exists a path α in U from x_1 to x_2 . The fact that $\alpha(t) \in U$ for $0 \leq t \leq 1$ is crucial. A topological space could be path-connected without being locally path-connected as can be seen in the following figure.



The set X is a curve in $\mathbb{R}^2$ with the induced topology from $\mathbb{R}^2$. There is no open neighbourhood of x that is path-connected.

The notion of locally path-connected may be new to the reader but it is going to be essential for the next important concept.

Definition 8.1.14

Let X and Q be two path-connected and locally path-connected topological spaces. Suppose that $p : Q \rightarrow X$ is a continuous and surjective function such that, for all $x \in X$, there exists an open neighbourhood $U \subset X$ of x satisfying the following conditions:

1. $p^{-1}(U) = \bigcup_{\tau \in T} V_\tau$ with $V_\tau \subset Q$ open sets and $V_{\tau_1} \cap V_{\tau_2} = \emptyset$ for $\tau_1 \neq \tau_2$, and
2. $p|_{V_\tau} : V_\tau \rightarrow U$ is a homeomorphism for all τ .

Then (Q, p) is called a **covering** of X (Figure 8.2).

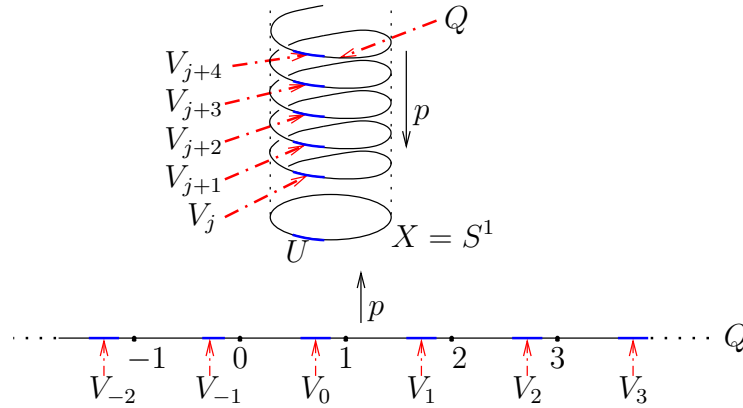


Figure 8.2: Two illustrations of the covering of S^1 . In the lower illustration, if we assume that S^1 is represented by a circle of radius 1 centred at the origin, then p is given by $p(t) = e^{2\pi ti}$ for $t \in \mathbb{R}$.

Proposition 8.1.15

If (Q, p) is a covering of a topological space X , then $p: Q \rightarrow X$ is an open mapping.

Proof.

Let W be an open subset of Q . Given $x = p(W)$, there exists an open neighbourhood $U \subset X$ of x satisfying Definition 8.1.14. Choose $y \in W$ such that $p(y) = x$. Then $y \in V_\tau$ for some τ . Since $W \cap V_\tau$ is an open subset of V_τ and $p|_{V_\tau}: V_\tau \rightarrow U$ is a homeomorphism, we have that $B = p(W \cap V_\tau)$ is an open neighbourhood of x in U and also in X because U is open. Thus B is an open neighbourhood of x with $B \subset p(W)$. Since $x \in p(W)$ is arbitrary, this proves that $p(W)$ is an open subset of X . ■

Proposition 8.1.16

Let (Q, p) be a covering of a topological space X . If $\alpha, \beta: Y \rightarrow Q$ are two continuous functions defined on a connected space Y such that $p \circ \alpha = p \circ \beta$ on Y and $\alpha(\tilde{y}) = \beta(\tilde{y})$ for a point $\tilde{y} \in Y$, then $\alpha = \beta$.

Proof.

Let $h: Y \rightarrow Q \times Q$ be the function defined by $h(y) = (\alpha(y), \beta(y))$ for $y \in Y$. We have that h is a continuous function.

Let $D = \{(q, q) : q \in Q\}$. Since D is a closed set¹, we have that $Z = \{y \in Y : \alpha(y) = \beta(y)\} = h^{-1}(D)$ is closed. We also have that $Z \neq \emptyset$ because $\tilde{y} \in Z$. Since Y is connected, it

¹In fact D is closed if and only if Q is hausdorff. One part of the proof consists in proving that Q Hausdorff implies that $(Q \times Q) \setminus D$ is open. The proof of the other direction is even more direct. See [5].

suffices to prove that Z is also open to conclude that $Z = Y$.

To prove that Z is open, choose $\check{y} \in Z$ and let $x = p(\alpha(\check{y})) = p(\beta(\check{y}))$. There exists an open neighbourhood $U \subset X$ of x satisfying Definition 8.1.14. Then $\alpha(\check{y}) = \beta(\check{y}) \in V_\tau$ for some τ . Since α is continuous, there exists an open neighbourhood $W_1 \subset \alpha^{-1}(V_\tau)$ of $\check{y}$. Similarly, since β is continuous, there exists an open neighbourhood $W_2 \subset \beta^{-1}(V_\tau)$ of $\check{y}$. Since $p|_{V_\tau} : V_\tau \rightarrow U$ is one-to-one and $p \circ \alpha = p \circ \beta$ on $W_1 \cap W_2$, we get that $\alpha = \beta$ on the open set $W_1 \cap W_2$. Thus $W_1 \cap W_2 \subset Z$. Since $\check{y} \in Z$ is arbitrary, this proves that Z is open. ■

The following theorem has deep consequences.

Theorem 8.1.17 (Covering Homotopy Theorem)

Suppose that (Q, p) is a covering of a topological space X and that Y is a compact and connected space. Moreover, suppose that $f, g : Y \rightarrow Q$ are two continuous functions and that $F : Y \times [0, 1] \rightarrow X$ is a homotopy between $p \circ f$ and $p \circ g$. Then there exists a homotopy $G : Y \times [0, 1] \rightarrow Q$ between f and g satisfying the following conditions.

1. $p \circ G = F$ on $Y \times [0, 1]$.
2. If there exists $y \in Y$ and an interval $E \subset [0, 1]$ such that $F(y, s)$ is constant for $s \in E$, then $G(y, s)$ is constant for $s \in E$.

Proof.

i) The set $Y \times [0, 1]$ is compact because Y and $[0, 1]$ are compact. For each $x \in X$, let $U_x \subset X$ be an open neighbourhood of x satisfying Definition 8.1.14. Since the collection $\{U_x\}_{x \in X}$ is an open cover of $F(Y \times [0, 1])$, we have that the collection $\{F^{-1}(U_x)\}_{x \in X}$ is an open cover of the compact set $Y \times [0, 1]$. Hence, there exists a finite subcover $\{F^{-1}(U_{x_j})\}_{1 \leq j \leq J}$ of $Y \times [0, 1]$.

We can cover $Y \times [0, 1]$ with open sets of the form $W_\alpha \times I_\alpha$ where $W_\alpha \subset Y$ is a connected open set, I_α is an open interval, and $W_\alpha \times I_\alpha \subset F^{-1}(U_{x_j})$ for some $j \in \{1, 2, \dots, J\}$. Again, because $Y \times [0, 1]$ is compact, there exists a finite subcover $\{W_{\alpha_k} \times I_{\alpha_k}\}_{1 \leq k \leq K}$ of $Y \times [0, 1]$. By splitting the intervals I_{α_k} , we get a finite subcover of $Y \times [0, 1]$ of the form $\{W_{\alpha_k} \times [s_{i-1}, s_i]\}_{1 \leq k \leq K, 1 \leq i \leq I}$ where $s_0 = 0 < s_1 < s_2 < \dots < s_I = 1$. We still have that $F(W_{\alpha_k} \times [s_{i-1}, s_i]) \subset U_{x_{j(k,i)}}$ for some $j(k, i) \in \{1, 2, \dots, J\}$.

ii) We define continuous function $G_i : Y \times [s_{i-1}, s_i] \rightarrow Q$ such that

$$p \circ G_i = F|_{Y \times [s_{i-1}, s_i]} \quad \text{and} \quad G_i|_{Y \times \{s_i\}} = G_{i-1}|_{Y \times \{s_i\}}$$

for $1 \leq i \leq I$. Because of the second condition, we will have that $G : Y \times [0, 1] \rightarrow Q$ defined by $G(y, s) = G_i(y, s)$ for $y \in Y$ and $t \in [s_{i-1}, s_i]$ is a well defined continuous function satisfying (1) in the statement of the theorem.

iii) Assume that $1 \leq i \leq I$. Let

$$G_i(y, s_{i-1}) = \begin{cases} f(y) & \text{if } i = 1 \\ G_{i-1}(y, s_{i-1}) & \text{if } 1 < i \leq I \end{cases}$$

for $y \in Y$. Since W_{α_k} is connected, we have that

$$G_i(W_{\alpha_k}, s_{i-1}) = \begin{cases} f(W_{\alpha_k}) & \text{if } i = 1 \\ G_{i-1}(W_{\alpha_k}, s_{i-1}) & \text{if } 1 < i \leq I \end{cases}$$

is connected. Since $F(W_{\alpha_k}, [s_{i-1}, s_i]) \subset U_{x_{j(k,i)}}$ for some $j(k,i) \in \{1, 2, \dots, J\}$ and

$$\begin{aligned} p \circ G_i|_{W_{\alpha_k} \times \{s_{i-1}\}} &= \begin{cases} p \circ f|_{W_{\alpha_k}} & \text{if } i = 0 \\ p \circ G_{i-1}|_{W_{\alpha_k} \times \{s_{i-1}\}} & \text{if } 1 < i \leq I \end{cases} \\ &= F|_{W_{\alpha_k} \times \{s_{i-1}\}}, \end{aligned}$$

we have that $G_i(W_{\alpha_k}, s_{i-1}) \subset V_{x_{j(k,i)}, \tau}$ for some $\tau \in T$, where $p^{-1}(U_{x_{j(k,i)}}) = \bigcup_{\tau \in T} V_{x_{j(k,i)}, \tau}$ with $V_{x_{j(k,i)}, \tau_1} \cap V_{x_{j(k,i)}, \tau_2} = \emptyset$ if $\tau_1 \neq \tau_2$ as given in Definition 8.1.14. More precisely, $G_i(W_{\alpha_k}, s_{i-1})$ is a connected subset of $p^{-1}(U_{x_{j(k,i)}}) = \bigcup_{\tau \in T} V_{x_{j(k,i)}, \tau}$ where the $V_{x_{j(k,i)}, \tau}$ are distinct open sets. Therefore $G_i(W_{\alpha_k}, s_{i-1})$ is a subset of only one of them.

Since $p|_{V_{x_{j(k,i)}, \tau}} : V_{x_{j(k,i)}, \tau} \rightarrow U_{x_{j(k,i)}}$ is a homeomorphism, we may set $G_{i,k}(y, s) = p^{-1}(F(y, s))$ for $(y, s) \in W_{\alpha_k} \times [s_{i-1}, s_i]$. We define $G_i : Y \times [s_{i-1}, s_i] \rightarrow Q$ by $G_i(y, s) = G_{i,k}(y, s)$ if $(y, s) \in W_{\alpha_k} \times [s_{i-1}, s_i]$. To prove that G_i is well defined and continuous on $Y \times [s_{i-1}, s_i]$, we prove in the next step that $G_{i,k_1}(y, s) = G_{i,k_2}(y, s)$ for $(y, s) \in (W_{\alpha_{k_1}} \cap W_{\alpha_{k_2}}) \times [s_{i-1}, s_i]$.

iv) Suppose that $W_{\alpha_{k_1}} \cap W_{\alpha_{k_2}} \neq \emptyset$. We have by definition of $G_{i,k}$ that

$$G_{i,k_m}(W_{\alpha_{k_m}}, [s_{i-1}, s_i]) \subset V_{x_{j(k_m,i)}, \tau_m}$$

for some $j(k_m, i) \in \{1, 2, \dots, J\}$ and $\tau_m \in T$ with $1 \leq m \leq 2$. We also have

$$G_{i,k_1}(y, s_{i-1}) = G_{i,k_2}(y, s_{i-1}) = \begin{cases} f(y) & \text{if } i = 1 \\ G_{i-1}(y, s_{i-1}) & \text{if } 1 < i \leq I \end{cases}$$

for all $(y, s) \in (W_{\alpha_{k_1}} \cap W_{\alpha_{k_2}}) \times [s_{i-1}, s_i]$. Thus $G_{i,k_1}(y, s_{i-1}) = G_{i,k_2}(y, s_{i-1}) \in V_{x_{j(k_1,i)}, \tau_1} \cap V_{x_{j(k_2,i)}, \tau_2}$ for all $(y, s) \in (W_{\alpha_{k_1}} \cap W_{\alpha_{k_2}}) \times [s_{i-1}, s_i]$.

Consider $(y, s) \in (W_{\alpha_{k_1}} \cap W_{\alpha_{k_2}}) \times [s_{i-1}, s_i]$ and the two functions

$$\begin{aligned} \sigma_1 : [0, 1] &\rightarrow Q \\ t &\mapsto G_{i,k_1}(y, ts + (1-t)s_{i-1}) \end{aligned}$$

and

$$\begin{aligned} \sigma_2 : [0, 1] &\rightarrow Q \\ t &\mapsto G_{i,k_2}(y, ts + (1-t)s_{i-1}) \end{aligned}$$

We have that $\sigma_1(0) = \sigma_2(0)$ and

$$\begin{aligned}(p \circ \sigma_1)(t) &= (p \circ G_{i,k_1}(y, ts + (1-t)s_{i-1})) = F(y, ts + (1-t)s_{i-1}) \\ &= (p \circ G_{i,k_2}(y, ts + (1-t)s_{i-1})) = (p \circ \sigma_2)(t)\end{aligned}$$

for $0 \leq t \leq 1$. It follows from Proposition 8.1.16 that $\sigma_1(t) = \sigma_2(t)$ for $0 \leq t \leq 1$. Thus $G_{i,k_1}(y, s) = \sigma_1(1) = \sigma_2(1) = G_{i,k_2}(y, s)$.

v) Repeating recursively (iii) and (iv) from $i = 1$ to $i = I$ yields the functions G_i in (ii).

vi) Recall that $G_{i,k}$ is defined by $G_{i,k}(y, s) = p^{-1}(F(y, s))$ for $(y, s) \in W_{\alpha_k} \times [s_{i-1}, s_i]$ where $p|_{V_{x_{j(k,i)},\tau}} : V_{x_{j(k,i)},\tau} \rightarrow U_{x_{j(k,i)}}$ is a homeomorphism. Therefore $G_{i,k}(y, s)$ is constant with respect to s if $F(y, s)$ is constant with respect to s because $p|_{V_{x_{j(k,i)},\tau}} : V_{x_{j(k,i)},\tau} \rightarrow U_{x_{j(k,i)}}$ is one-to-one. ■

It follows from Prop 8.1.16 that if (Q, p) is a covering of a topological space X and $q \in Q$, then $p_* : \pi_1(Q, q) \rightarrow \pi_1(X, p(q))$ is one-to-one. If α is a path in X such that $\alpha(0) = p(q)$, then there exists a unique path β in Q such that $p \circ \beta = \alpha$ and $\beta(0) = q$.

There are several interesting and important corollaries to this theorem. For instance, if $x \in X$ and $x = p(q)$, then there exists a one-to-one map from the quotient group $\pi_1(X, x)/p_*(\pi_1(Q, q))$ onto $p^{-1}(\{x\})$. To prove this result, we first construct a map M from $\pi_1(X, x)$ onto $p^{-1}(\{x\})$ as it follows. Given $[\alpha] \in \pi_1(X, x)$, we have from the result stated in the previous paragraph that there exists a unique path β in Q such that $p \circ \beta = \alpha$ with $\beta(0) = q$. We set $M([\alpha]) = \beta(1)$. The rest of the proof consists in proving that the definition of M is independent of the representative α chosen, that M is onto and that the kernel of M is $p_*(\pi_1(Q, q))$.

Definition 8.1.18

Let X be a path-connected and locally path-connected topological space. We say that X is **simply connected** if $\pi_1(X)$, the fundamental group of X , is trivial.

It follows that if X is simply connected and $x \in X$, then $\alpha \sim e_x$ for all loops α in X at x , where e_x is the path defined by $e_x(t) = x$ for $0 \leq t \leq 1$.

Definition 8.1.19

A topological space X is **locally simply connected** if for every $x \in X$ there exists an open neighbourhood U of x such $\alpha \underset{U}{\sim} e_x$ for all loops α in U at x .

We insist on the fact that the range of the homotopy in the previous definition is only U .

Definition 8.1.20

Let (Q, p) be a covering of a topological space X . We say that (Q, p) is a **universal covering** of X if Q is simply connected

Theorem 8.1.21

Let X be a path-connected, locally path-connected and locally simply connected topological space. Given $x \in X$, let H be a subgroup of $\pi_1(X, x)$. Then there exists a covering (Q, p) of X such that $p_*(\pi_1(Q, q)) = H$ for $q \in Q$ such that $p(q) = x$.

In particular, if $H = \{[e_x]\}$ for $x \in X$, then we get a universal covering (Q, p) of X because, as we stated before, $p_* : \pi_1(Q, q) \rightarrow \pi_1(X, p(q))$ is one-to-one.

It can be proved that if (Q, p) is a covering of a topological space X and $x \in X$, then $p_*(\pi_1(Q, q_1))$ and $p_*(\pi_1(Q, q_2))$ with $p(q_1) = p(q_2) = x$ are two conjugate subgroups in $\pi_1(X, x)$; namely, there exists $[\beta] \in \pi_1(X, x)$ such that $[\beta] p_*(\pi_1(Q, q_1)) [\beta]^{-1} = p_*(\pi_1(Q, q_2))$. In fact, if γ is a path in Q from q_1 to q_2 , then $\beta = p \circ \gamma$.

Definition 8.1.22

Two coverings (Q_1, p_1) and (Q_2, p_2) of a topological space X are **isomorphic** if there exists a homeomorphism $h : Q_1 \rightarrow Q_2$ such that $p_2 \circ h = p_1$.

Another result that can relatively easily be proved states that if (Q_1, p_1) and (Q_2, p_2) are two coverings of a locally simply connected space X such that $p_*(\pi_1(Q_1, q_1)) = p_*(\pi_1(Q_2, q_2))$ for any $q_1 \in Q_1$ and $q_2 \in Q_2$ with $p(q_1) = p(q_2)$, then (Q_1, p_1) and (Q_2, p_2) are isomorphic coverings.

Definition 8.1.23

Let (Q, p) be a covering of a topological space X . A **covering transformation** of (Q, p) is a homeomorphism $h : Q \rightarrow Q$ such that $p \circ h = p$.

The set of covering transformations of (Q, p) form a group under composition of functions. This group is denoted $\mathcal{G}(Q, p)$.

If U is an open set as in the definition of covering, Definition 8.1.14, then a covering transformation h permutes the open sets V_j .

Definition 8.1.24

A **regular covering** of a topological space X is a covering (Q, p) of X such that $p_*(\pi_1(Q, q))$ is a normal subgroup of $\pi_1(X, p(q))$ for $q \in Q$; namely, $[\beta] p_*(\pi_1(Q, q)) [\beta]^{-1} = p_*(\pi_1(Q, q))$ for all $[\beta] \in \pi_1(X, p(q))$.

It follows from the paragraph following Theorem 8.1.21 that the definition of regularity is independent of the point $q \in Q$ selected.

The next theorem is the fundamental theorem of this subsection.

Theorem 8.1.25

Suppose that (Q, p) is a regular covering of a locally simply connected space X and that $q \in Q$. Then $\mathcal{G}(Q, p)$ is group isomorphic to $\pi_1(X, p(q))/p_*(\pi_1(Q, q))$. In particular, if the covering is also universal, then $\mathcal{G}(Q, p)$ is group isomorphic to $\pi_1(X, p(q))$.

Proposition 8.1.26

$\pi_1(S^1, 1) \cong \mathbb{Z}$. In fact, since S^1 is path-connected, we have that $\pi_1(S^1) \cong \mathbb{Z}$.

Proof.

In the statement of the theorem, we assume the complex representation of S^1 ; namely, $S^1 = \{z \in \mathbb{C} : |z| = 1\}$.

Suppose that (Q, p) is the covering described in Figure 8.2. We have that $Q = \mathbb{R}$ and (Q, p) is an universal and regular covering of S^1 . If $h : \mathbb{R} \rightarrow \mathbb{R}$ is a covering transformation, then $p \circ h = p$ implies that

$$e^{2\pi h(q)i} = e^{2\pi qi} \iff e^{2\pi(h(q)-q)i} = 1$$

for all $q \in \mathbb{R}$. Thus $h(q) - q$ is a continuous function defined on $\mathbb{R}$, a connected set, whose range is $\mathbb{Z}$. This is possible only if there exists $n \in \mathbb{Z}$ such that $h(q) - q = n$ for all $q \in \mathbb{R}$. This proves that the covering transformations are all given by integer translations. Hence $\mathcal{G}(Q, p) \cong \mathbb{Z}$ and from the previous theorem $\pi_1(S^1) \cong \mathcal{G}(Q, p) \cong \mathbb{Z}$. ■

Proposition 8.1.27

Suppose that X and Y are path-connected topological spaces, and that $x_0 \in X$ and $y_0 \in Y$. Then $\pi_1(X \times Y, (x_0, y_0)) \cong \pi_1(X, x_0) \times \pi_1(Y, y_0)$

Proof.

It is easy to verify that the two projections $P : X \times Y \rightarrow X$ and $Q : X \times Y \rightarrow Y$ induces two homomorphism $P_* : \pi_1(X \times Y, (x_0, y_0)) \rightarrow \pi_1(X, x_0)$ defined by $P_*([\sigma]) = [P \circ \sigma]$ and $Q_* : \pi_1(X \times Y, (x_0, y_0)) \rightarrow \pi_1(Y, y_0)$ defined by $Q_*([\sigma]) = [Q \circ \sigma]$ for $[\sigma] \in \pi_1(X \times Y, (x_0, y_0))$. Thus $(P_*, Q_*) : \pi_1(X \times Y, (x_0, y_0)) \rightarrow \pi_1(X, x_0) \times \pi_1(Y, y_0)$ defined by $(P_*, Q_*)([\sigma]) = (P_*([\sigma]), Q_*([\sigma]))$ for $[\sigma] \in \pi_1(X \times Y, (x_0, y_0))$ is an homomorphism.

The map (P_*, Q_*) is also an isomorphism because it has an inverse. If σ_X is a loop in X at x_0 and σ_Y is a loop in Y at y_0 , then (σ_X, σ_Y) is a loop in $X \times Y$ at (x_0, y_0) . Let $R : \pi_1(X, x_0) \times \pi_1(Y, y_0) \rightarrow \pi_1(X \times Y, (x_0, y_0))$ be the map defined by $R([\sigma_X], [\sigma_Y]) = [(\sigma_X, \sigma_Y)]$ for $[\sigma_X] \in \pi_1(X, x_0)$ and $[\sigma_Y] \in \pi_1(Y, y_0)$. The map R is well defined because $\sigma_X \sim \tilde{\sigma}_X$ and $\sigma_Y \sim \tilde{\sigma}_Y$ imply $(\sigma_X, \sigma_Y) \sim (\tilde{\sigma}_X, \tilde{\sigma}_Y)$ ². Since all loops σ in $X \times Y$ at (x_0, y_0) are of the form $\sigma = (\sigma_X, \sigma_Y)$ where $\sigma_X = P \circ \sigma$ is a loop in X at x_0 and $\sigma_Y = Q \circ \sigma$ is a loop in Y at y_0 ,

²We leave it to the reader to construction the map $H : [0, 1] \times [0, 1] \rightarrow X \times Y$ needed to prove that $(\sigma_X, \sigma_Y) \sim (\tilde{\sigma}_X, \tilde{\sigma}_Y)$.

we have that R is onto. The map R is also one-to-one because $\sigma \sim \tilde{\sigma}$ in $X \times Y$ implies that $P \circ \sigma \sim P \circ \tilde{\sigma}$ in X and $Q \circ \sigma \sim Q \circ \tilde{\sigma}$ in Y ³.

It is now easy to verify that $R \circ (P_*, Q_*) = \text{Id}_{\pi_1(X \times Y, (x_0, y_0))}$ and $(P_*, Q_*) \circ R = \text{Id}_{\pi_1(X, x_0) \times \pi_1(Y, y_0)}$. ■

Example 8.1.28

The torus $T^2 \subset \mathbb{R}^3$ is isomorphic to $S^1 \times S^1$. This can be seen from the parametric representation of the torus. Hence, it follows from the previous proposition that $\pi_1(T^2) \cong \pi_1(S^1) \times \pi_1(S^1) \cong \mathbb{Z} \times \mathbb{Z}$ since T^2 is path-connected. ♣

To conclude this section, we plan to prove a limited version of the famous Brouwer fixed point theorem. Namely, a version limited to $\mathbb{R}^2$. We first need a new definition and a lemma.

Definition 8.1.29

Let Y be a subset of a topological space X . A **retraction** of X onto Y is a continuous function $r : X \rightarrow Y$ such that $r(x) = x$ for all $x \in Y$ or, equivalently, $r \circ \iota = \text{Id}_Y$ for the inclusion map $\iota : Y \rightarrow X$. We then say that Y is a **retract** of X .

Lemma 8.1.30

Let $D = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\| \leq 1\}$. There does not exist any retraction of D onto $S^1 = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\| = 1\}$.

Proof.

Suppose that there is a retraction $g : D \rightarrow S^1$. Let $\iota : S^1 \rightarrow D$ be the inclusion map. Then $g \circ \iota = \text{Id}_{S^1}$ and

$$g_* \circ \iota_* = (g \circ \iota)_* = \text{Id}_{\pi_1(S^1, \mathbf{e}_1)} . \quad (8.1.1)$$

However, $\pi_1(D, \mathbf{e}_1) = \{[e_{\mathbf{e}_1}]\}$ where $e_{\mathbf{e}_1}(t) = \mathbf{e}_1$ for $0 \leq t \leq 1$ because D is contractible. Thus $\iota_*(\pi_1(S^1, \mathbf{e}_1)) \subset \pi_1(D, \mathbf{e}_1) = \{[e_{\mathbf{e}_1}]\}$ implies that $(g_* \circ \iota_*)(\pi_1(S^1, \mathbf{e}_1)) = \{[e_{\mathbf{e}_1}]\}$. This is a contradiction of Proposition 8.1.26 because, according to (8.1.1), the image of $g_* \circ \iota_*$ should be $\pi_1(S^1, \mathbf{e}_1) \cong \mathbb{Z} \not\cong \{[e_{\mathbf{e}_1}]\}$. ■

Theorem 8.1.31 (Brouwer Fixed Point Theorem)

Let $D = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\| \leq 1\}$. Suppose that $f : D \rightarrow D$ is a continuous function. Then there exists $\mathbf{y} \in D$ such that $f(\mathbf{y}) = \mathbf{y}$; namely, f has a **fixed point** in D .

Proof.

Suppose that there does not exist $\mathbf{y} \in D$ such that $f(\mathbf{y}) = \mathbf{y}$. Let $g : D \rightarrow S^1$ be the function defined by $g(\mathbf{x})$ is the intersection with the circle S^1 of the line through $\mathbf{x}$ originating from $f(\mathbf{x})$. Such a line is well defined because $f(\mathbf{x}) \neq \mathbf{x}$ for all $\mathbf{x} \in D$. It is also a continuous

³Again, we leave it to the reader to construct the map $H_P : [0, 1] \times [0, 1] \rightarrow X$ and $H_Q : [0, 1] \times [0, 1] \rightarrow Y$ needed to prove that $P \circ \sigma \sim P \circ \tilde{\sigma}$ in X and $Q \circ \sigma \sim Q \circ \tilde{\sigma}$ in Y .

function because f is continuous. Moreover $g(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in S^1$. Thus, g is a retraction of D onto S^1 . This is a contradiction of Lemma 8.1.30. ■

The proof of the Brouwer fixed point theorem is standard and can be generalized to $\mathbb{R}^{n+1}$ with $n > 1$ using a group structure on S^n other than the fundamental group of S^n because $\pi_1(S^n, \mathbf{e}_1) = \{[e_{\mathbf{e}_1}]\}$ for $n > 1$.

Remark 8.1.32

It is possible to define higher order homotopy groups as it is explained in [9]. We however follow the tradition of instead introducing singular homology as we will do in Section 8.4. As we will see in Subsection 8.4.5 later, it is a natural way to generalize the concept of fundamental group. ♠

8.2 Simplicial Homology

As mentioned in the introduction of this chapter, readers may choose to only browse through this and the next sections if they are not very interested in learning about simplicial homology and cohomology. There are however some concepts and definitions in this two sections that the reader should learn because they are used in the next chapter. In particular, we are talking about the definitions of simplex and simplicial complex in Section 8.2.4, oriented simplex in Section 8.2.1, and smoothly triangulated manifold in Section 8.3.4. Sections 8.4 and 8.5 on singular homology and cohomology are the important sections in this chapter. In particular, Subsection 8.5.5 contains the statement and the proof of de Rham theorem, Theorem 8.5.23. which is the motivation for this chapter.

We present and sketch the proof of a version of de Rham theorem for simplicial cohomology in Section 8.3.4 for those who may only be interested in a brief study of simplicial homology and cohomology.

Our presentation of the simplicial homology and cohomology is more for a pedagogical purpose, to prepare the reader to the study of singular homology and cohomology. There is a lot of similarity between the results and techniques of proof used in simplicial theory and those used in singularity theory. This is not surprising as it is explained in Section 8.6. Nevertheless, there are really important and interesting results that can only be deduced from simplicial homology. Theorem 8.2.36 is one of them. Both simplicial and singular theories are important. It is not a question of which one is better than the other.

8.2.1 Simplicial Complexes

Definition 8.2.1

A subset $\{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ of $\mathbb{R}^n$ is **convex independent** if $\{\mathbf{x}_1 - \mathbf{x}_0, \mathbf{x}_2 - \mathbf{x}_0, \dots, \mathbf{x}_k - \mathbf{x}_0\}$ are linearly independent.

Definition 8.2.2

Let $s = \{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ be a convex independent subset of $\mathbb{R}^n$. The set $\left\{ \sum_{j=0}^k a_j \mathbf{x}_j : a_j \geq 0 \text{ and } \sum_{j=0}^k a_k = 1 \right\}$, denoted $[s]$ or $[\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k]$, is called a **closed k-simplex**. The **vertices** of $[s]$ are the points $\mathbf{x}_j$ for $0 \leq j \leq k$. The **closed faces** of $[s]$ are the closed m -simplices $[\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \mathbf{x}_{j_2}, \dots, \mathbf{x}_{j_m}]$ for $\{j_0, j_1, \dots, j_m\} \subset \{0, 1, 2, \dots, k\}$ and $0 \leq m \leq k$. The **dimension** of $[s]$, denoted $\dim [s]$, is k .

The reader should prove that $[s]$ is the smallest convex set containing s .

Definition 8.2.3

The coefficients a_j in the expression $\mathbf{x} = \sum_{j=0}^k a_j \mathbf{x}_j$ with $a_j \geq 0$ and $\sum_{j=0}^k a_k = 1$ are called **barycentric coordinates** of x .

Definition 8.2.4

Let $s = \{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ be a convex independent subset of $\mathbb{R}^n$. The set $\left\{ \sum_{j=0}^k a_j \mathbf{x}_j : a_j > 0 \text{ and } \sum_{j=0}^k a_k = 1 \right\}$, denoted (s) or $(\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k)$, is called an **open k-simplex**. The **dimension** of (s) , denoted $\dim (s)$, is k .

Definition 8.2.5

The **open faces** of the closed k -simplex $[\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_k]$ are the open m -simplices $(\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \mathbf{x}_{j_2}, \dots, \mathbf{x}_{j_m})$ for $\{j_0, j_1, \dots, j_m\} \subset \{0, 1, 2, \dots, k\}$ and $0 \leq m \leq k$.

A closed k -simplex $[\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_k]$ is the union of all its open faces (Figure 8.3). The reader must realize that the closed 0-complex $[\mathbf{x}_j]$ is equal to the open 0-complex $(\mathbf{x}_j)$. If (s_1) and (s_2) are two open faces of a closed k -simplex, then either $(s_1) = (s_2)$ or $(s_1) \cap (s_2) = \emptyset$.

Definition 8.2.6

A **simplicial complex** K in $\mathbb{R}^n$ is a finite set of open simplices such that:

1. If (s) is an open simplices in K , then all the open faces of the closed simplex $[s]$ are also in K .
2. If (s_1) and (s_2) are two open simplices in K , then either $(s_1) = (s_2)$ or $(s_1) \cap (s_2) = \emptyset$.

The **dimension** of the simplicial complex K , denoted $\dim K$, is defined as the dimension of the largest open simplex in K . The union of the open simplices in K is the

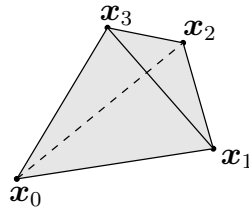


Figure 8.3: Illustration of a closed 3-simplex $[x_0, x_1, x_2, x_3] \subset \mathbb{R}^3$. This closed simplex is the union of the open faces (x_0, x_1, x_2, x_3) , (x_0, x_1, x_2) , (x_0, x_1, x_3) , (x_0, x_2, x_3) , (x_1, x_2, x_3) , (x_0, x_1) , (x_0, x_2) , (x_0, x_3) , (x_1, x_2) , (x_1, x_3) , (x_2, x_3) , (x_0) , (x_1) , (x_2) and (x_3) .

subset of $\mathbb{R}^n$ denoted $[K]$. The **vertices** of K are the elements $x \in [K]$ such that $(x) \in K$; namely, (x) is a 0-simplex in K .

We have that $[K] = \bigcup_{(s) \in K} (s) = \bigcup_{(s) \in K} [s]$ and $[K]$ is a compact subsets of $\mathbb{R}^n$ because it is a finite union of the compact sets $[s]$ for $(s) \in K$.

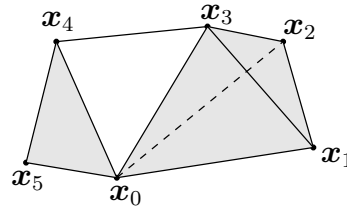


Figure 8.4: The simplicial complex K represented by the set $[K] \subset \mathbb{R}^3$ in the figure above is the union of the set of all the open faces of the closed 3-simplex $[x_0, x_1, x_2, x_3]$ listed in Figure 8.3 with the set $\{(x_0, x_4, x_5), (x_0, x_4), (x_0, x_5), (x_4, x_5), (x_4), (x_5)\}$. The dimension of K is 3.

All the following notions will be very useful to study simplicial complexes.

Definition 8.2.7

A subset L of a simplicial complex K is called a **subcomplex** of K if L is itself a simplicial complex.

Definition 8.2.8

Let K be a simplicial complex of dimension k . Given $0 \leq r \leq k$, the **r -skeleton** of K , denoted K^r , is the set $K^r = \{(s) : (s) \in K \text{ and } \dim(s) \leq r\}$.

Since a closed k -simplex $[s]$ is the union of all its open faces, we may treat s as the set of all the open faces of $[s]$; namely, we may treat s as a simplicial complex. Note that all the open faces of $[s]$ are open m -simplices with $m \leq \dim[s]$. By extension, the **r -skeleton** of s for $0 \leq r \leq k$, denoted s^r , is the set of all open faces of $[s]$ corresponding to open m -simplex with $m \leq r$. Moreover, $[s^r]$ will denote the union of all the open simplices in s^r .

Definition 8.2.9

Given $\mathbf{x} \in \mathbb{R}^n$ and $A \subset \mathbb{R}^n$, the pair $(\mathbf{x}, A)$ is in **general position** if $\mathbf{x} \notin A$ and, for any two distinct points $\mathbf{a}_1$ and $\mathbf{a}_2$ in A , we have that $[\mathbf{x}, \mathbf{a}_1] \cap [\mathbf{x}, \mathbf{a}_2] = \{\mathbf{x}\}$. If $(\mathbf{x}, A)$ is in general position, then the **cone** with base A and vertex $\mathbf{x}$, denoted $\mathbf{x} * A$, is the set $\mathbf{x} * A = \bigcup_{\mathbf{a} \in A} [\mathbf{x}, \mathbf{a}]$.

Given a simplicial complex K , the **cone** $\mathbf{x} * K$ is the simplicial complex consisting of all the open simplices of the form $(\mathbf{x}, \mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_k)$ with all their faces for all $(\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_m) \in K$ with $m \leq \dim K$.

The closed k -simplex $[s] = [\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k]$ can be expressed as $[s] = \mathbf{x} * [s^{k-1}]$ for any $\mathbf{x} \in (s)$.

8.2.2 Barycentric Subdivisions**Definition 8.2.10**

The **barycentre** of an open k -simplex $(s) = (\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_k)$ is the point $\mathbf{b}_{(s)} \in (s)$ defined by $\mathbf{b}_{(s)} = (k+1)^{-1} \sum_{j=0}^k \mathbf{x}_j$.

Definition 8.2.11

A **subdivision** of a simplicial complex K is another simplicial complex L such that $[L] = [K]$ and, if $(s) \in L$, then (s) is a subset of an open simplex in K .

Definition 8.2.12

A **partial order** on a simplicial complex K is defined as follows. Given (s_1) and (s_2) in K , we write that $(s_1) \leq (s_2)$ if (s_1) is an open face of $[s_2]$. If in addition $(s_1) \neq (s_2)$, then we write that $(s_1) < (s_2)$.

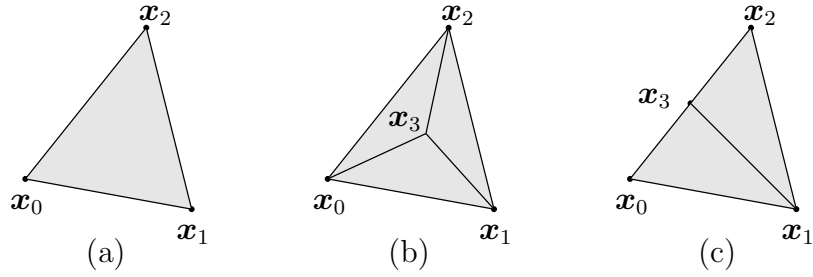


Figure 8.5: (b) and (c) illustrate two possible subdivisions of the simplicial complex illustrated in (a). (b) is not a subdivision of the simplicial complex illustrated in (c) and vice-versa.

The following result will play a fundamental role in the “simplicial approximation” of “simplicial function” in the next subsection. Its proof is not too complicated but it is long (see [25]). It uses the concept of general position introduced above.

Theorem 8.2.13

Suppose that K is a simplicial complex in $\mathbb{R}^n$. Let $K^{[1]}$ be the simplicial complex defined by

$$K^{[1]} = \{(\mathbf{b}_{(s_0)}, \mathbf{b}_{(s_1)}, \dots, \mathbf{b}_{(s_m)}) : (s_1) < (s_2) < \dots < (s_m) \text{ and } (s_j) \in K \text{ for } 0 \leq j \leq m \leq \dim K\}.$$

Then $K^{[1]}$ is a subdivision of K (Figure 8.6). Moreover, we have that $(\mathbf{b}_{(s_0)}, \mathbf{b}_{(s_1)}, \dots, \mathbf{b}_{(s_m)})$ is a subset of (s_m) for all sequences $(s_1) < (s_2) < \dots < (s_m)$ of open simplices in K .

It follows from the definition of $K^{[1]}$ that $\dim K^{[1]} = \dim K$.

Definition 8.2.14

Let K be a simplicial complex in $\mathbb{R}^n$. The simplicial complex $K^{[1]}$ is called the **first barycentric subdivision** of K . Inductively, we may define the **m^{th} barycentric subdivision** of K as the simplicial complex $K^{[m]} = (K^{[m-1]})^{[1]}$ for $m > 1$.

Recall that the diameter of a set $S \subset \mathbb{R}^n$ is defined by $\text{diam } S = \inf_{\mathbf{x}, \mathbf{y} \in S} \|\mathbf{x} - \mathbf{y}\|$ where $\|\cdot\|$ is a norm on $\mathbb{R}^n$, usually the Euclidean norm.

Lemma 8.2.15

If $[s] = [\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k]$ is a closed k -simplex, then $\text{diam } [s] = \|\mathbf{x}_{j_1} - \mathbf{x}_{j_2}\|$ for some $j_1, j_2 \in \{0, 1, 2, \dots, k\}$.

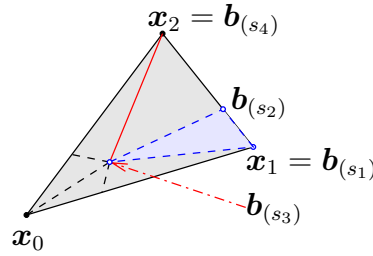
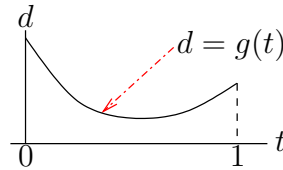


Figure 8.6: We have illustrated some of the open simplices in the simplicial complex $s^{[1]}$ where s is the simplicial complex associated to the closed 2-simplex $[s] = [x_0, x_1, x_2]$. We have drawn in blue the open 2-simplex $(b_{(s_1)}, b_{(s_2)}, b_{(s_3)})$ for $(s_1) = (x_1)$, $(s_2) = (x_1, x_2)$ and $(s_3) = (s)$. We effectively have that $(s_1) < (s_2) < (s_3)$. We have drawn in red the open 1-simplex $(b_{(s_4)}, b_{(s_3)})$ for $(s_4) = (x_2)$ and (s_3) . Again, we have that $(s_4) < (s_3)$. Note that the open 0-simplex $(b_{(t)})$ is associated to a 0-simplex (t) is (t) itself.

Proof.

Since $[s]$ is a compact set, there exist $\mathbf{y}_1$ and $\mathbf{y}_2$ in $[s]$ such that $\text{diam}[s] = \|\mathbf{y}_1 - \mathbf{y}_2\|$. Without loss of generality, suppose that $\mathbf{y}_2$ is not a vertex. Then we may express $\mathbf{y}_2$ as $\mathbf{y}_2 = p\mathbf{w}_1 + (1-p)\mathbf{w}_2$ for some $0 < p < 1$ and $\mathbf{w}_1, \mathbf{w}_2 \in [s]$. Let $g(t) = \|\mathbf{y}_1 - t\mathbf{w}_1 - (1-t)\mathbf{w}_2\|$ for $0 \leq t \leq 1$. We have that g is a convex function; namely, $g(tq_1 + (1-t)q_2) \leq tg(q_1) + (1-t)g(q_2)$ for $0 \leq q_1, q_2 \leq 1$ and $0 \leq t \leq 1$. Here is a possible graph of g .



Thus g reaches its maximum at one of the end points. This implies that $\text{diam}[s] = g(p) < \max\{g(0), g(1)\} = \max\{\|\mathbf{y}_1 - \mathbf{w}_1\|, \|\mathbf{y}_1 - \mathbf{w}_2\|\}$. This is a contradiction. ■

Definition 8.2.16

Let K be a simplicial complex. The **mesh** of K is defined as $\text{mesh } K = \max_{(s) \in K} \text{diam}[s]$.

Proposition 8.2.17

Suppose that K is a simplicial complex of dimension k . Then $\text{mesh } K^{[1]} \leq (k/(k+1)) \text{mesh } K$.

Proof.

By definitions of $K^{[1]}$ and the mesh of a simplicial complex, there exists $(s_1) < (s_2) < \dots <$

(s_m) in K with $m \leq k$ such that $\text{mesh } K^{[1]} = \text{diam} [\mathbf{b}_{(s_0)}, \mathbf{b}_{(s_1)}, \dots, \mathbf{b}_{(s_m)}]$. It follows from Lemma 8.2.15 that $\text{mesh}^{[1]} K = \|\mathbf{b}_{(s_p)} - \mathbf{b}_{(s_q)}\|$ for some $p, q \in \{0, 1, \dots, m\}$. Without loss of generality, we may re-index the vertices of K and assume that $p < q$, $(s_p) = (\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_p)$ and $(s_q) = (\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_q)$. Hence

$$\begin{aligned}
 \text{mesh } K^{[1]} &= \|\mathbf{b}_{(s_p)} - \mathbf{b}_{(s_q)}\| = \left\| \frac{1}{p+1} \sum_{j=0}^p \mathbf{x}_j - \frac{1}{q+1} \sum_{i=0}^q \mathbf{x}_i \right\| \\
 &= \left\| \left(\frac{1}{p+1} - \frac{1}{q+1} \right) \sum_{j=0}^p \mathbf{x}_j - \frac{1}{q+1} \sum_{i=p+1}^q \mathbf{x}_i \right\| = \left\| \frac{q-p}{(p+1)(q+1)} \sum_{j=0}^p \mathbf{x}_j - \frac{1}{q+1} \sum_{i=p+1}^q \mathbf{x}_i \right\| \\
 &= \frac{1}{q+1} \left\| \frac{q-p}{p+1} \sum_{j=0}^p \mathbf{x}_j - \sum_{i=p+1}^q \mathbf{x}_i \right\| = \frac{1}{q+1} \left\| \frac{q+1}{p+1} \sum_{j=0}^p \mathbf{x}_j - \sum_{i=0}^q \mathbf{x}_i \right\| \\
 &= \frac{1}{q+1} \left\| \sum_{i=0}^q \left(\frac{1}{p+1} \sum_{j=0}^p \mathbf{x}_j - \mathbf{x}_i \right) \right\| = \frac{1}{(q+1)(p+1)} \left\| \sum_{i=0}^q \left(\sum_{j=0}^p (\mathbf{x}_j - \mathbf{x}_i) \right) \right\| \\
 &\leq \frac{1}{(q+1)(p+1)} \sum_{i=0}^q \sum_{j=0}^p \|\mathbf{x}_j - \mathbf{x}_i\|,
 \end{aligned}$$

where we have used $(q-p) = (q+1) - (p+1)$ to get the fifth equality. There are $(p+1)(q+1) - (p+1) = q(p+1)$ terms in the last sum because $(p+1)$ of the terms are null, those for $i = j$. Moreover $\|\mathbf{x}_j - \mathbf{x}_i\| \leq \text{mesh } K$ for all i and j . thus

$$\text{mesh } K^{[1]} \leq \frac{q(p+1)}{(q+1)(p+1)} \text{mesh } K = \frac{q}{q+1} \text{mesh } K \leq \frac{k}{k+1} \text{mesh } K$$

because $q \leq k$. ■

It follows from the previous proposition that $\text{mesh } K^{[m]} \leq ((k/(k+1))^m \text{mesh } K \rightarrow 0$ as $m \rightarrow \infty$.

8.2.3 Simplicial Approximation

Definition 8.2.18

Let K and L be two simplicial complexes. A **simplicial map** between K and L is a function $\phi : [K] \rightarrow [L]$ that satisfies the following conditions.

1. $\phi(\mathbf{x})$ is a vertex of L if $\mathbf{x}$ is a vertex of K .
2. If $(s) = (\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_k) \in K$, then there exists $(t) \in L$ such that $\phi(\mathbf{x}_i) \in [t]$ for $0 \leq i \leq k$.

3. If $\mathbf{x} \in (s) = (\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_k)$ and $\mathbf{x} = \sum_{j=0}^k a_j \mathbf{x}_j$ with $a_j > 0$ and $\sum_{j=0}^k a_j = 1$, then

$$\phi(\mathbf{x}) = \sum_{j=0}^k a_j \phi(\mathbf{x}_j).$$

The set $\left\{ \phi(\mathbf{x}) : \mathbf{x} = \sum_{j=0}^k a_j \mathbf{x}_j \text{ with } a_j > 0 \text{ and } \sum_{j=0}^k a_j = 1 \right\}$ in (3) represents an open face of $[t]$ mentioned in (2). It may not be (t) itself because $\phi(\mathbf{x}_i)$ for $0 \leq i \leq k$ may not be all the vertices of $[t]$.

Definition 8.2.19

Let K be a simplicial complex and $\mathbf{x}$ be a vertex of K . The **star** of $\mathbf{x}$, denoted $\text{St}(\mathbf{x})$, is the set defined by $\text{St}(\mathbf{x}) = \bigcup_{\substack{(s) \in K \\ \mathbf{x} \in [s]}} (s)$ (Figure 8.7).

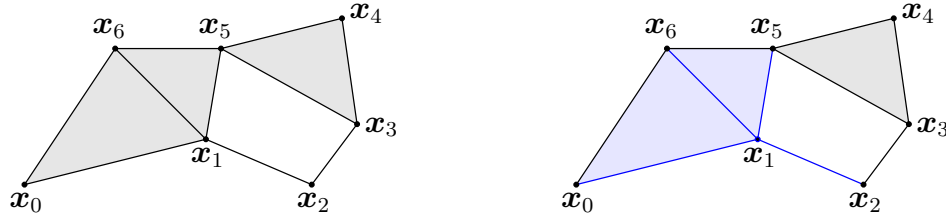


Figure 8.7: We have illustrated on the left a simplicial complex K where $[K]$ is the union of the closed simplices $[x_0, x_1, x_6]$, $[x_1, x_5, x_6]$, $[x_3, x_4, x_5]$, $[x_1, x_2]$ and $[x_2, x_3]$. We have illustrated on the right in blue the star of x_1 .

Given a vertex $\mathbf{x}$ of a simplicial complex K , we have that $\text{St}(\mathbf{x})$ is open in the induced topology on $[K]$ from $\mathbb{R}^n$. Moreover, $\mathbf{x}$ is the only vertex of K in $\text{St}(\mathbf{x})$ and $\{\text{St}(\mathbf{x}) : \mathbf{x} \text{ is a vertex of } K\}$ is an open cover of $[K]$ with respect to the induced topology on $[K]$ from $\mathbb{R}^n$.

Definition 8.2.20

Let K and L be two simplicial complexes. A simplicial map $\phi : [K] \rightarrow [L]$ is a **simplicial approximation** of a continuous function $f : [K] \rightarrow [L]$ if $f(\text{St}(\mathbf{x})) \subset \text{St}(\phi(\mathbf{x}))$ for all vertices $\mathbf{x}$ of K .

Proposition 8.2.21

Suppose that K and L are two simplicial complexes. If $\phi : [K] \rightarrow [L]$ is a simplicial approximation of a continuous function $f : [K] \rightarrow [L]$ and $\mathbf{x} \in [K]$, then $f(\mathbf{x})$ and $\phi(\mathbf{x})$ belongs to the same closed simple $[t]$ with $(t) \in L$.

Proof.

Given $\mathbf{x} \in [K]$, there exists an open simplex $(s) = (\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_m) \in K$ such that $\mathbf{x} \in (s)$. Since $\mathbf{x} \in (s) \subset \text{St}(\mathbf{x}_j)$ for all $0 \leq j \leq m$, we get from the definition of simplicial approximation that $f(\mathbf{x}) \in f(\text{St}(\mathbf{x}_j)) \subset \text{St}(\phi(\mathbf{x}_j))$ for $0 \leq j \leq m$.

There exists an open simplex $(t) \in L$ such that $f(\mathbf{x}) \in (t)$. Since $f(\mathbf{x}) \in (t) \cap \text{St}(\phi(\mathbf{x}_j))$ for $0 \leq j \leq m$, we get that $(t) \cap \text{St}(\phi(\mathbf{x}_j)) \neq \emptyset$ for $0 \leq j \leq m$. By definition of simplicial complex, we have that $(t) \subset \text{St}(\phi(\mathbf{x}_j))$ for $0 \leq j \leq m$. By definition of the star of a vertex, we have that $\phi(\mathbf{x}_j)$ for $0 \leq j \leq m$ are vertices of $[t]$. Note that they may not be all the vertices of $[t]$.

Since $\mathbf{x} \in (s)$, we have that $\mathbf{x} = \sum_{j=0}^m a_j \mathbf{x}_j$ for some $a_j > 0$ such that $\sum_{j=0}^m a_j = 1$. Hence,

we get from the definition of simplicial map that $\phi(\mathbf{x}) = \sum_{j=0}^m a_j \phi(\mathbf{x}_j)$ with $a_j > 0$ such that

$\sum_{j=0}^m a_j = 1$. Thus $\phi(\mathbf{x}) \in (\phi(\mathbf{x}_0), \phi(\mathbf{x}_1), \dots, \phi(\mathbf{x}_k)) \subset [t]$. Note that we may not have that $(\phi(\mathbf{x}_0), \phi(\mathbf{x}_1), \dots, \phi(\mathbf{x}_k)) = (t)$ because, as we said above, $\phi(\mathbf{x}_j)$ for $0 \leq j \leq m$ may not be all the vertices of $[t]$. Therefore $(\phi(\mathbf{x}_0), \phi(\mathbf{x}_1), \dots, \phi(\mathbf{x}_k))$ may be only an open face of $[t]$ not equal to (t) itself. ■

This proposition implies that $\|f - \phi\|_{[K]} = \max_{\mathbf{x} \in [K]} \|f(\mathbf{x}) - \phi(\mathbf{x})\| \leq \text{mesh } L$.

Proposition 8.2.22

Suppose that K and L are two simplicial complexes. If $\phi : [K] \rightarrow [L]$ is a simplicial approximation of a simplicial map $f : [K] \rightarrow [L]$, then $\phi = f$.

Proof.

For each vertex $\mathbf{x} \in K$, we have that $f(\mathbf{x}) \in f(\text{St}(\mathbf{x})) \subset \text{St}(\phi(\mathbf{x}))$. Since f is a simplicial map, $f(\mathbf{x})$ is a vertex of L . Since $\text{St}(\phi(\mathbf{x}))$ contains only one vertex of L , namely $\phi(\mathbf{x})$, we get that $f(\mathbf{x}) = \phi(\mathbf{x})$.

Since f and ϕ are simplicial maps that take the same value at all the vertices of K , then it follows from the third condition of the definition of simplicial maps, Definition 8.2.18, that $f = \phi$. ■

The following theorem provides a necessary and sufficient condition to create a simplicial approximation to a continuous function between two simplicial complexes.

Proposition 8.2.23

Suppose that K and L are two simplicial complexes and that $f : [K] \rightarrow [L]$ is a continuous function. If ϕ is a map defined on the set of vertices of K with values in the set of vertices of L , then ϕ can be extended to a simplicial approximation of f if

and only if $f(\text{St}(\mathbf{x})) \subset \text{St}(\phi(\mathbf{x}))$ for all vertices $\mathbf{x}$ of K .

Proof.

Obviously, if ϕ can be extended to a simplicial approximation of f , then we have from the definition of simplicial approximation that $f(\text{St}(\mathbf{x})) \subset \text{St}(\phi(\mathbf{x}))$ for all vertices $\mathbf{x}$ of K .

Suppose that $f(\text{St}(\mathbf{x})) \subset \text{St}(\phi(\mathbf{x}))$ for all vertices $\mathbf{x}$ of K . Consider a simplex $(s) = (\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_m) \in K$. Since $(s) \subset \text{St}(\mathbf{x}_j)$ for $0 \leq j \leq m$, we have that $f((s)) \subset f(\text{St}(\mathbf{x}_j)) \subset \text{St}(\phi(\mathbf{x}_j))$ for $0 \leq j \leq m$. Thus $\bigcap_{j=0}^m \text{St}(\phi(\mathbf{x}_j)) \neq \emptyset$. Therefore, there exists an open simplex $(t) \subset \text{St}(\phi(\mathbf{x}_j))$ for $0 \leq j \leq m$. We get that $\phi(\mathbf{x}_j)$ for $0 \leq j \leq m$ are vertices of $[t]$. Thus, we may define $\phi|_{[s]}$ by $\phi|_{[s]}(\mathbf{x}) = \sum_{j=0}^m a_j \phi(\mathbf{x}_j)$ for $\mathbf{x} = \sum_{j=0}^m a_j \mathbf{x}_j$ with $a_j > 0$ and $\sum_{j=0}^m a_j = 1$.

In particular, $\phi((s)) \subset [t]$.

Since $(s_1) = (s_2)$ or $(s_1) \cap (s_2) = \emptyset$ for all simplices (s_1) and (s_2) of K , and $\phi|_{[s_1]}(\mathbf{x}) = \phi|_{[s_2]}(\mathbf{x})$ for $\mathbf{x} \in [s_1] \cap [s_2]$, we have that ϕ is a well defined simplicial map between K and L . ■

Theorem 8.2.24

Suppose that K and L are two simplicial complexes and that $f : [K] \rightarrow [L]$ is a continuous function. Moreover, suppose that $\{K_m\}_{m \in \mathbb{N}}$ is a collection of subdivision of K such that $\text{mesh } K_m \rightarrow 0$ as $m \rightarrow \infty$. Then there exists $M > 0$ such that there exists a simplicial approximation $\phi_m : [K_m] \rightarrow [L]$ of f for each $m \geq M$.

Proof.

Since $\{\text{St}(\mathbf{y}) : \mathbf{y} \text{ is a vertex of } L\}$ is an open cover of $[L]$ with respect to the induced topology on $[L]$ and $f : [K] \rightarrow [L]$ is continuous, we have that $\{f^{-1}(\text{St}(\mathbf{y})) : \mathbf{y} \text{ is a vertex of } L\}$ is an open cover of $[K]$ with respect to the induced topology on $[K]$. Since $[K]$ is compact, there exists a number $\epsilon > 0$, called the Lebesgue number, such that any open ball in $[K]$ of radius less than or equal to ϵ is a subset $f^{-1}(\text{St}(\mathbf{y}))$ for some vertex $\mathbf{y}$ of L ⁴.

Choose M such that $\text{mesh } K_m < \epsilon/2$ for $m \geq M$. Assume that $m \geq M$. If $s \in K_m$, then $\text{diam}[s] \leq \text{mesh } K_m < \epsilon/2$. Hence, for each vertex $\mathbf{x}$ of K_m , we have that $\text{St}(\mathbf{x}) \subset B_\epsilon(\mathbf{x}) \subset f^{-1}(\text{St}(\mathbf{y}))$ for some vertex $\mathbf{y}$ of L ; namely,

$$f(\text{St}(\mathbf{x})) \subset \text{St}(\mathbf{y}) \quad (8.2.1)$$

for some vertex $\mathbf{y}$ of L . If there are more than one vertex $\mathbf{y}$ of L satisfying the previous statement, select one of them. Let $\phi_m(\mathbf{x}) = \mathbf{y}$. This defines a map from the set of vertices of

⁴Since the number of vertices of L is finite, we could have proved the existence of ϵ without making use of the theory of compact sets.

K_m to the set of vertices of L . We also have from (8.2.1) that $f(\text{St}(\mathbf{x})) \subset \text{St}(\phi(\mathbf{x}))$. We can then use Proposition 8.2.23 to get a simplicial approximation $\phi_m : [K_m] \rightarrow [L]$ of f . ■

The previous theorem ensures that we can always find a simplicial approximation as closed as we want to a continuous function between two simplicial complexes.

Corollary 8.2.25

Suppose that K and L are two simplicial complexes and that $f : [K] \rightarrow [L]$ is a continuous function. Given $\epsilon > 0$, there exist subdivisions $\tilde{K}$ and $\tilde{L}$ of K and L respectively, and a simplicial approximation $\phi : [\tilde{K}] \rightarrow [\tilde{L}]$ of f , such that $\|f - \phi\|_{[K]} < \epsilon$.

Proof.

As remarked after Proposition 8.2.17, if $L_j = L^{[j]}$ for $j \in \mathbb{N}^+$, then we get a sequence of barycentric subdivisions of L such that $\text{mesh } L_j \rightarrow 0$ as $j \rightarrow \infty$. Thus there exists $j > 0$ such that $\text{mesh } L_j < \epsilon$.

Similarly, if $K_m = K^{[m]}$ for $m \in \mathbb{N}^+$, then we get a sequence of barycentric subdivisions of M such that $\text{mesh } K_m \rightarrow 0$ as $m \rightarrow \infty$. If we apply Theorem 8.2.24 to $f : [K] \rightarrow [L_j] = [L]$ using the collection of subdivisions $\{K_m\}_{m \in \mathbb{N}^+}$ of K , then we find that there exists $m > 0$ and a simplicial approximation $\phi : [K_m] \rightarrow [L_j]$ of f . As remarked after Proposition 8.2.21, we then have that $\|f - \phi\|_{[K]} \leq \text{mesh } L_j < \epsilon$. We take $\tilde{K} = K_m$ and $\tilde{L} = L_j$. ■

8.2.4 Oriented Simplicial Homology

From now on, when we refer to a simplicial simplex K , we assume that a fixed indexing of the vertices of K has been selected. More specifically, we assume that $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots$ are the vertices of K .

Definition 8.2.26

Let $[s] = [\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_k}]$ be a closed k -simplex. An **oriented k -simplex** associated to $[s]$ is the pair $([s], \sigma)$ where $\sigma : \{j_0, j_1, j_2, \dots, j_k\} \rightarrow \{j_0, j_1, j_2, \dots, j_k\}$ is a permutation.

Definition 8.2.27

Let $[s] = [\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_k}]$ be a closed k -simplex. We say that two oriented k -simplices $([s], \sigma_1)$ and $([s], \sigma_2)$ are **equivalent** if there exists an even permutation $\beta : \{j_0, j_1, j_2, \dots, j_k\} \rightarrow \{j_0, j_1, j_2, \dots, j_k\}$ such that $\beta \circ \sigma_1 = \sigma_2$; namely, $\beta(\sigma_1(j_i)) = \sigma_2(j_i)$ for $0 \leq i \leq k$.

The notion of equivalence defined in the previous definition is an equivalence relation that partitions the oriented k -simplices associated to a closed k -simplex into two equivalence classes.

Definition 8.2.28

Let $[s] = [\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_k}]$ be a closed k -simplex. We use the notations $[[\mathbf{x}_{\sigma(j_0)}, \mathbf{x}_{\sigma(j_1)}, \dots, \mathbf{x}_{\sigma(j_k)}]]$ or simply $[[s_\sigma]]$ to denote the equivalence class associated to $([s], \sigma)$.

If $[s] = [\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_k}]$, then the two equivalence classes of oriented k -simplices associate to $[s]$ are $[[\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \mathbf{x}_{j_2}, \dots, \mathbf{x}_{j_k}]]$ and $[[\mathbf{x}_{j_1}, \mathbf{x}_{j_0}, \mathbf{x}_{j_2}, \dots, \mathbf{x}_{j_k}]]$. Note that $[[\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \mathbf{x}_{j_2}, \dots, \mathbf{x}_{j_k}]]$ is the equivalence class associated to $([s], \sigma_1)$ with $\sigma_1 = \text{Id}$ and $[[\mathbf{x}_{j_1}, \mathbf{x}_{j_0}, \mathbf{x}_{j_2}, \dots, \mathbf{x}_{j_k}]]$ is the equivalence class associated to $([s], \sigma_2)$ with σ_2 is defined by

$$\sigma_2(j_m) = \begin{cases} j_1 & \text{if } m = 0 \\ j_0 & \text{if } m = 1 \\ j_m & \text{otherwise} \end{cases}$$

Definition 8.2.29

Let K be a simplicial complex and R be an integral domain. The **module of k -chains** of K , denoted $C_k(K; R)$, is the free abelian group generated by the equivalence classes of oriented k -simplices $[[s_\sigma]]$ for all $(s) \in K$ where, given $(s) = (\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_k})$, we set $[[s_{\sigma_1}]] = -[[s_{\sigma_2}]]$ if $\beta \circ \sigma_1 = \sigma_2$ for an odd permutation $\beta : \{j_0, j_1, j_2, \dots, j_k\} \rightarrow \{j_0, j_1, j_2, \dots, j_k\}$. The elements of $C_k(K; R)$ are called **k -chains** of K .

For us, the integral domain R will either be $\mathbb{Z}$ or $\mathbb{R}$.

The elements $c \in C_k(K; R)$ are finite sums of the form $c = \sum_{\substack{(s) \in K \\ \dim(s)=k}} a_{(s)} [[s]]$ with $a_{(s)} \in R$,

or more explicitly

$$c = \sum_{(\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_k}) \in K} a_{(\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_k})} [[\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_k}]]$$

for $a_{(\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_k})} \in R$.

Definition 8.2.30

Let K be a simplicial complex and R be an integral domain. If $(s) = (\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_k}) \in K$ with $k > 0$, then the **boundary** of $[[s]] \in C_k(K; R)$, denoted $\partial_k [[s]]$, is defined by

$$\partial_k [[s]] = \sum_{i=0}^k (-1)^i [[\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \dots, \widehat{\mathbf{x}_{j_i}}, \dots, \mathbf{x}_{j_k}]] \in C_{k-1}(K; R).$$

(Figure 8.8). If $(s) = (\mathbf{x}_j) \in K$, then $\partial_0 [[\mathbf{x}_j]] = 0$.

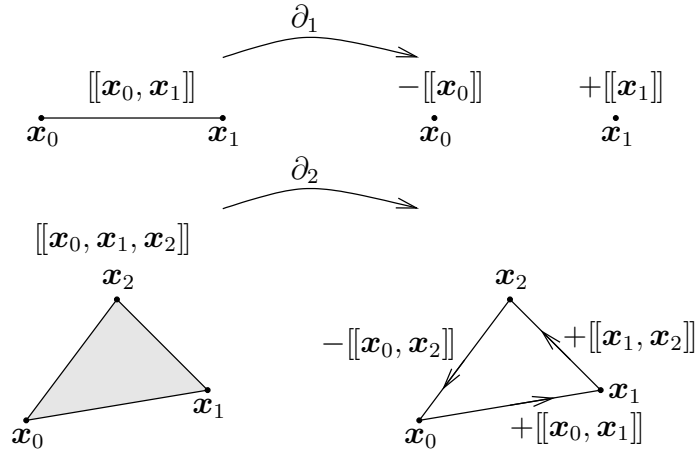


Figure 8.8: Examples of the boundary of k -chains for the simplest k -chains. For the 1-chain $[[x_0, x_1]]$, we have that $\partial_1[[x_0, x_1]] = [[x_0]] - [[x_1]]$. For the 2-chain $[[x_0, x_1, x_2]]$, we have that $\partial_2[[x_0, x_1, x_2]] = [[x_1, x_2]] - [[x_0, x_2]] + [[x_0, x_1]] = [[x_0, x_1]] + [[x_1, x_2]] + [[x_2, x_0]]$ as we have illustrated in the figure above.

Definition 8.2.31

Let K be a simplicial complex and R an integral domain. The **boundary map** $\partial_k : C_k(K; R) \rightarrow C_{k-1}(K; R)$ is defined by $\partial_k \left(\sum_{\substack{(s) \in K \\ \dim(s)=k}} a_{(s)} [[s]] \right) = \sum_{\substack{(s) \in K \\ \dim(s)=k}} a_{(s)} \partial_k [[s]]$ or more explicitly

$$\begin{aligned} \partial_k \left(\sum_{(x_{j_0}, x_{j_1}, \dots, x_{j_k}) \in K} a_{(x_{j_0}, x_{j_1}, \dots, x_{j_k})} [[x_{j_0}, x_{j_1}, \dots, x_{j_k}]] \right) \\ = \sum_{(x_{j_0}, x_{j_1}, \dots, x_{j_k}) \in K} a_{(x_{j_0}, x_{j_1}, \dots, x_{j_k})} \partial_k [[x_{j_0}, x_{j_1}, \dots, x_{j_k}]] . \end{aligned}$$

A simple computation yields $\partial_{k-1}(\partial_k(c)) = 0$ for all k -chains c . Note that $\partial_{k-1}(\partial_k(c))$ is the sum of expressions of the form

$$\begin{aligned} (-1)^{j_s + j_r} a_{[[x_0, x_1, \dots, x_k]]} [[x_{j_0}, x_{j_1}, \dots, \widehat{x_{j_r}}, \dots, \widehat{x_{j_s}}, \dots, x_{j_k}]] \\ + (-1)^{j_s + j_r - 1} a_{[[x_0, x_1, \dots, x_k]]} [[x_{j_0}, x_{j_1}, \dots, \widehat{x_{j_r}}, \dots, \widehat{x_{j_s}}, \dots, x_{j_k}]] \end{aligned}$$

that cancel each other.

Definition 8.2.32

Let K be a simplicial complex and R an integral domain,
 $Z_k(K; R) = \{c \in C_k(K; R) : \partial_k(c) = 0\}$ and $B_k(K; R) = \{\partial_{k+1}(c) : c \in C_{k+1}(K; R)\}.$

The **(simplicial) k^{th} homology module** of K is the free abelian quotient group $H_k(K; R) = Z_k(K; R)/B_k(K; R)$. The elements of $Z_k(K; R)$ are called **cycles** and those of $B_k(K; R)$ are called **boundaries**. The equivalence class of $H_k(K; R)$ associated to $c \in Z_k(K; R)$ is denoted $[c]_K$.

Suppose that K and L are two simplicial complexes and that $f : [K] \rightarrow [L]$ is a homeomorphism that maps open simplices in K to open simplices in L . The map $C(f) : C_k(K; R) \rightarrow C_k(L; R)$ defined by $C(f)\left(\sum_{\substack{(s) \in K \\ \dim(s)=k}} a_{(s)}[[s]]\right) = \sum_{\substack{(s) \in K \\ \dim(s)=k}} a_{(s)}[[f((s))]]$ is an isomorphism. Since $C(f)$ commutes with ∂_k for all k , we may use $C(f)$ to induce a map $H(f) : H_k(K; R) \rightarrow H_k(L; R)$ because $C(f)(Z_k(K; R)) \subset Z_k(L; R)$ and $C(f)(B_k(K; R)) \subset B_k(L; R)$.

Proposition 8.2.33

Let K be a simplicial complex and R be an integral domain. The dimension of $H_0(K; R)$ is equal to the number of (connected) components of $[K]$. Namely, the number of independent generators of $H_0(K; R)$ is equal to the number of (connected) components of $[K]$.

Proof (Sketch).

Since $Z_0(K; R) = C_0(K; R)$ and all elements of $C_0(K; R)$ are of the form $c = \sum_{(x_j) \in K} a_{(x_j)}[[x_j]]$

with $a_{(x_j)} \in R$, it therefore suffices to consider the linear independence of the equivalence classes in $H_0(K; R)$ associated to $[[x_j]] \in C_0(K; R)$ for $(x_j) \in K$ to determine the dimension of $H_0(K; R)$.

If $(x_{j_0}, x_{j_1}) \in K$, then $\partial_1[[x_{j_0}, x_{j_1}]] = [[x_{j_1}]] - [[x_{j_0}]]$. Hence $[[x_{j_0}]]$ and $[[x_{j_1}]]$ are in the same equivalence classes of $H_0(K; R)$. Inductively, we find that two equivalence classes of $H_0(K; R)$, one associated to $[[x_{j_0}]]$ and one associated $[[x_{j_1}]]$, are distinct if there is no sequence $(x_{j_0}, x_{k_1}), (x_{k_1}, x_{k_2}), (x_{k_2}, x_{k_3}), \dots, (x_{k_{m-1}}, x_{k_m}), (x_{k_m}, x_{j_1})$ of open 1-simplices in K . This is only possible if x_{j_0} and x_{j_1} belongs to two distinct connected components of $[K]$. ■

Example 8.2.34

We consider the simplicial complex

$$K = \{(x_0), (x_1), (x_2), (x_0, x_1), (x_1, x_2), (x_0, x_2), (x_0, x_1, x_2)\}.$$

a) It follows from Proposition 8.2.33 that $H_0(K^1; \mathbb{Z}) \cong \mathbb{Z}$.

b) We have

$$C_1(K^1; \mathbb{Z}) = \{a_{(x_0, x_1)}[[x_0, x_1]] + a_{(x_1, x_2)}[[x_1, x_2]] + a_{(x_0, x_2)}[[x_0, x_2]] : a_{(x_0, x_1)}, a_{(x_1, x_2)}, a_{(x_0, x_2)} \in \mathbb{Z}\} \cong \mathbb{Z}^3$$

and

$$\begin{aligned} \partial_1 & (a_{(x_0, x_1)} [[x_0, x_1]] + a_{(x_1, x_2)} [[x_1, x_2]] + a_{(x_0, x_2)} [[x_0, x_2]]) \\ &= a_{(x_0, x_1)} ([[x_1]] - [[x_0]]) + a_{(x_1, x_2)} ([[x_2]] - [[x_1]]) + a_{(x_0, x_2)} ([[x_2]] - [[x_0]]) \\ &= (a_{(x_0, x_1)} - a_{(x_1, x_2)}) [[x_1]] - (a_{(x_0, x_2)} + a_{(x_0, x_1)}) [[x_0]] + (a_{(x_1, x_2)} + a_{(x_0, x_2)}) [[x_2]] = 0 \end{aligned}$$

if and only if $a_{(x_0, x_1)} = a_{(x_1, x_2)} = -a_{(x_0, x_2)}$. Hence

$$Z_1(K^1; \mathbb{Z}) = \{a ([[x_0, x_1]] + [[x_1, x_2]] - [[x_0, x_2]]) : a \in \mathbb{Z}\} \cong \mathbb{Z}.$$

We also have that $B_1(K^1; \mathbb{Z}) = \{0\}$ because $C_2(K^1; \mathbb{Z}) = \{0\}$. Therefore $H_1(K^1; \mathbb{Z}) = Z_1(K^1; \mathbb{Z})/B_1(K^1; \mathbb{Z}) \cong \mathbb{Z}$.

c) It follows from Proposition 8.2.33 that $H_0(K; \mathbb{Z}) \cong \mathbb{Z}$.

d) As before, we have

$$Z_1(K; \mathbb{Z}) = \{a ([[x_0, x_1]] + [[x_1, x_2]] - [[x_0, x_2]]) : a \in \mathbb{Z}\} \cong \mathbb{Z}.$$

We have

$$C_2(K; \mathbb{Z}) = \{a [[x_0, x_2, x_3]] : a \in \mathbb{Z}\} \cong \mathbb{Z}.$$

Since

$$\partial_2 (a [[x_0, x_1, x_2]]) = a ([[x_1, x_2]] - [[x_0, x_2]] + [[x_0, x_1]]) \quad (8.2.2)$$

for $a \in \mathbb{Z}$, we get

$$B_1(K; \mathbb{Z}) = \{a ([[x_1, x_2]] - [[x_0, x_2]] + [[x_0, x_1]]) : a \in \mathbb{Z}\} = Z_1(K^2; \mathbb{Z}) \cong \mathbb{Z}.$$

Thus $H_1(K; \mathbb{Z}) = Z_1(K; \mathbb{Z})/B_1(K; \mathbb{Z}) = 0$.

e) Finally, it follows from (8.2.2) that $\partial_3 (a [[x_1, x_2, x_3]]) = 0$ if and only if $a = 0$. So $Z_2(K; \mathbb{Z}) = \{0\}$ and therefore $H_2(K; \mathbb{Z}) = 0$. ♣

If $R = \mathbb{R}$, then $C_k(K, \mathbb{R})$, $B_k(K, \mathbb{R})$ can be treated as vector spaces over $\mathbb{R}$.

We now define the Euler characteristic of a simplicial complex and deduce some very interesting properties associated to the Euler characteristic of a simplicial complex.

Definition 8.2.35

Let K be a simplicial complex. The **j^{th} Betti number** of K , denoted β_j , is the non-negative integer defined by $\beta_j = \dim(H_j(K; \mathbb{R}))$. The **Euler characteristic** of

K , denote $\mathcal{X}(K)$, is the integer defined by $\mathcal{X}(K) = \sum_{j=0}^{\dim(K)} (-1)^j \beta_j$.

Some of the readers may have recognized the famous Euler characteristic formula introduced in undergraduate geometry courses [4]. Suppose that K is the triangulation of the unit sphere $S^2 \subset \mathbb{R}^3$. The Euler characteristic of this triangulation K of S^2 is given by $V - E + F$ where V is the number of vertices, E is the number of edges, and F is the number of faces (i.e. triangles) of K . It is a value independent of the triangulation of S^2 .

Theorem 8.2.36

Let K be a simplicial complex and α_j be the number of open j -simplices in K . Then

$$\mathcal{X}(K) = \sum_{j=0}^{\dim(K)} (-1)^j \alpha_j.$$

Proof.

Let $C_{-1}(K; \mathbb{R}) = \{0\}$. Since $\partial_0([[\mathbf{x}_j]]) = 0$ for all vertices $\mathbf{x}_j$ of K , we may also set $B_{-1}(K; \mathbb{R}) = \{0\}$. Since $\partial_i : C_i(K; \mathbb{R}) \rightarrow C_{i-1}(K; \mathbb{R})$ is a linear operator for $0 \leq i \leq \dim(K)$, we have

$$\alpha_i = \dim(C_i(K; \mathbb{R})) = \dim(\mathcal{N}(\partial_i)) + \dim(\mathcal{R}(\partial_i)) = \dim(Z_i(K; \mathbb{R})) + \dim(B_{i-1}(K; \mathbb{R}))$$

for $0 \leq i \leq \dim(K)$. We also have

$$\beta_i = \dim(H_i(K; \mathbb{R})) = \dim(Z_i(K; \mathbb{R})/B_i(K; \mathbb{R})) = \dim(Z_i(K; \mathbb{R})) - \dim(B_i(K; \mathbb{R}))$$

for $0 \leq i \leq \dim(K)$. Hence

$$\begin{aligned} \mathcal{X}(K) &= \sum_{i=0}^{\dim(K)} (-1)^i \beta_i = \sum_{i=0}^{\dim(K)} (-1)^i (\dim(Z_i(K; \mathbb{R})) - \dim(B_i(K; \mathbb{R}))) \\ &= \sum_{i=0}^{\dim(K)} (-1)^i \dim(Z_i(K; \mathbb{R})) + \sum_{i=0}^{\dim(K)} (-1)^{i+1} \dim(B_i(K; \mathbb{R})) \\ &= \sum_{i=0}^{\dim(K)} (-1)^i \dim(Z_i(K; \mathbb{R})) + \sum_{i=1}^{1+\dim(K)} (-1)^i \dim(B_{i-1}(K; \mathbb{R})) \\ &= (\dim(Z_0(K; \mathbb{R})) + \underbrace{\dim(B_{-1}(K; \mathbb{R}))}_{=0}) \\ &\quad + \sum_{i=1}^{\dim(K)} (-1)^i (\dim(Z_i(K; \mathbb{R})) + \dim(B_{i-1}(K; \mathbb{R}))) \\ &\quad + (-1)^{1+\dim(K)} \underbrace{\dim(B_{\dim(K)}(K; \mathbb{R}))}_{=0} \\ &= \sum_{i=0}^{\dim(K)} (-1)^i (\dim(Z_i(K; \mathbb{R})) + \dim(B_{i-1}(K; \mathbb{R}))) = \sum_{i=0}^{\dim(K)} (-1)^i \alpha_i. \end{aligned}$$

Note that $\dim(B_{\dim(K)}(K; \mathbb{R})) = 0$ because $C_{1+\dim(K)}(K; \mathbb{R}) = \{0\}$. ■

8.2.5 Reduced and Relative Homology

Let K be a simplicial complex and R be an integral domain. We have by definition that $\partial_0(c) = 0$ for all $c \in C_0(K; R)$. Namely, we are assuming that the image of ∂_0 is in the space $C_{-1}(K; R) = \{0\}$.

We replace $C_{-1}(K; R) = \{0\}$ and the boundary operator $\partial_0 : C_0(K; R) \rightarrow C_{-1}(K; R)$ to obtain a slightly new collection of homology modules. Let

$$C_k^\sharp(K; R) = \begin{cases} R & \text{if } k = -1 \\ C_k(K; R) & \text{if } k \neq -1 \end{cases}$$

Also, let $\partial_k^\sharp = \partial_k$ for $k \neq 0$ and

$$\partial_0^\sharp \left(\sum_{(\mathbf{x}_j) \in K} a_{(\mathbf{x}_j)} [[\mathbf{x}_j]] \right) = \sum_{(\mathbf{x}_j) \in K} a_{(\mathbf{x}_j)}$$

for $a_{(\mathbf{x}_j)} \in R$. It is not difficult to prove that $\partial_0^\sharp \circ \partial_1^\sharp = 0$. We can then proceed as we did before to define $Z_k^\sharp(X; R)$, $B_k^\sharp(X; R)$ and $H_k^\sharp(X; R)$.

Definition 8.2.37

Let K be a simplicial complex and R be an integral domain. The quotient $H_k^\sharp(K; R) = Z_k^\sharp(X; R)/B_k^\sharp(X; R)$ is called the **reduced (simplicial) k^{th} homology module** of X .

Again, let K be a simplicial complex and R be an integral domain. Suppose that L is a simplicial subcomplex of K . Recall that this means that $L \subset K$ and L is itself a simplicial complex. Assuming that we keep the same indexing of the vertices of L as the one used for K , then $C_k(L; R)$ is a submodule of $C_k(K; R)$.

Definition 8.2.38

Let L be a simplicial subcomplex of a simplicial complex K and R be an integral domain. The **module of relative chains** of $K \bmod L$, denoted $C_k(K, L; R)$, is the free abelian group defined as $C_k(K, L; R) = C_k(K; R)/C_k(L; R)$.

Since $\partial_k : C_k(L; R) \rightarrow C_{k-1}(L; R)$, we may define the operator $\bar{\partial}_k : C_k(K, L; R) \rightarrow C_{k-1}(K, L; R)$ as $\bar{\partial}_k([[c]]_{K,L}) = [[\partial_k c]]_{K,L}$ for all $[[c]]_{K,L} \in C_k(K, L; R)$ where $[[c]]_{K,L}$ denotes the equivalence class in $C_k(K, L; R)$ associated to $c \in C_k(K; R)$. We get the following commutative diagram.

$$\begin{array}{ccc} C_k(K; R) & \xrightarrow{\pi_k} & C_k(K, L; R) \\ \downarrow \partial_k & & \downarrow \bar{\partial}_k \\ C_{k-1}(K; R) & \xrightarrow{\pi_{k-1}} & C_{k-1}(K, L; R) \end{array}$$

where $\pi_k : C_k(K; R) \rightarrow C_k(K, L; R)$ is the projection defined by $\pi_k(c) = [[c]]_{K,L}$ for $c \in C_k(K; R)$.

Definition 8.2.39

Let L be a simplicial subcomplex of a simplicial complex K and R be an integral domain. The **relative (simplicial) k^{th} homology module** of $K \bmod L$ is the free

abelian group $H_k(K, L; R) = \mathcal{N}(\bar{\partial}_k)/\mathcal{R}(\bar{\partial}_{k-1})$. The equivalence class in $H_k(K, L; R)$ associated to $[[c]]_{K,L} \in Z_k(K, L; R) = \mathcal{N}(\bar{\partial}_k)$ is denoted $[c]_{K,L}$.

It is possible to develop the theory of reduced and relative homology as it is done in [21] but we will only do so in the context of singular homology later on. Since singular homology seems to be more accessible and user friendly than simplicial homology (though the latest is more visual), we will only study reduced and relative homology in the context of singular homology in Sections 8.4.

8.3 Simplicial Cohomology

We only provide a very brief introduction to simplicial cohomology. We give just enough information to reach our goal of providing a relation between the de Rham cohomology on a manifold S and the simplicial cohomology on a simplicial complex K associated (in some way to be defined) to S .

8.3.1 Cohomology Modules

Definition 8.3.1

Let K be a simplicial complex. For $0 \leq k \leq \dim(K)$, let $C^k(K; \mathbb{R})$ denote the dual space of $C_k(K; \mathbb{R})$. Namely, $C^k(K; \mathbb{R})$ is the vector space of all linear functionals on $C_k(K; \mathbb{R})$. The elements of $C^k(K; \mathbb{R})$ are called **cochains**.

Definition 8.3.2

Let K be a simplicial complex. The adjoint of the boundary map $\partial_k : C_k(K; \mathbb{R}) \rightarrow C_{k-1}(K; \mathbb{R})$ is the map $d_{k-1} : C^{k-1}(K; \mathbb{R}) \rightarrow C^k(K; \mathbb{R})$ defined by

$$(d_{k-1}(\phi))(c) = \phi(\partial_k(c)) \quad (8.3.1)$$

for all $c \in C_k(K; \mathbb{R})$ and $\phi \in C^{k-1}(K; \mathbb{R})$. The map d_k is called the **coboundary operator**.

It is clear that d_k is well and uniquely defined by (8.3.1). It is easy to prove using $\partial_{k+1} \circ \partial_{k+2} = 0$ that $d_{k+1} \circ d_k = 0$.

Definition 8.3.3

Let K be a simplicial complex, $Z^k(K; \mathbb{R}) = \{\phi \in C^k(K; \mathbb{R}) : d_k \phi = 0\}$ and $B^k(K; \mathbb{R}) = \{d_{k-1} \phi : \phi \in C^{k-1}(K; \mathbb{R})\}$. The **k^{th} cohomology module** of K is the quotient $H^k(K; \mathbb{R}) = Z^k(K; \mathbb{R})/B^k(K; \mathbb{R})$. The elements of $Z^k(K; \mathbb{R})$ are called **cocycles**

and those of $B^k(K; \mathbb{R})$ are called **coboundaries**. The equivalence class in $H^k(K; \mathbb{R})$ associated to $\phi \in Z^k(K; \mathbb{R})$ is denoted $[\phi]_K$.

Theorem 8.3.4

Let K be a simplicial complex. Then $H^k(K; \mathbb{R})$ is isomorphic to $(H_k(K; \mathbb{R}))^*$.

Proof.

The map $h : H^k(K; \mathbb{R}) \rightarrow (H_k(K; \mathbb{R}))^*$ defined by $h([\phi]_K) = \phi$ for $\phi \in Z^k(K; \mathbb{R})$ is well defined because

$$\begin{aligned} \phi \in Z^k(K; \mathbb{R}) &\Rightarrow d_k(\phi) = 0 \Rightarrow \phi(\partial_{k+1}(c)) = (d_k\phi)(c) = 0 \text{ for all } c \in C_{k+1}(K; \mathbb{R}) \\ &\Rightarrow \phi = 0 \text{ on } B_k(K; \mathbb{R}) \Rightarrow \phi \in (H_k(K; \mathbb{R}))^* . \end{aligned}$$

We have that h is one-to-one. Suppose that $[\phi_1]_K = [\phi_2]_K$. Then $\phi_1 - \phi_2 = \psi \in B^k(K; \mathbb{R})$. Hence $h([\phi_1]_K) = h([\phi_2]_K) + h([\psi]_K) = h([\phi_2]_K)$ because

$$\begin{aligned} \psi \in B^k(K; \mathbb{R}) &\Rightarrow \psi = d_{k-1}\eta \text{ for some } \eta \in C^{k-1}(K; \mathbb{R}) \\ &\Rightarrow \psi(c) = (d_{k-1}\eta)(c) = \eta(\partial_k(c)) = 0 \text{ for all } c \in Z_k(K; \mathbb{R}) \\ &\Rightarrow \psi = 0 \text{ on } Z_k(K; \mathbb{R}) \Rightarrow h([\psi]_K) = 0 . \end{aligned}$$

We have that h is onto. Given $\phi \in (H_k(K; \mathbb{R}))^*$, then $[\phi]_K \in H^k(K; \mathbb{R})$ because

$$\begin{aligned} \phi(c) = 0 \text{ for all } c \in B_k(K; \mathbb{R}) &\Rightarrow (d_k\phi)(c) = \phi(\partial_{k+1}(c)) = 0 \text{ for all } c \in C_{k+1}(K; \mathbb{R}) \text{ since} \\ &\quad \partial_{k+1}(c) \in B_k(K; \mathbb{R}) \text{ for all } c \in C_{k+1}(K; \mathbb{R}) \\ &\Rightarrow d_k\phi = 0 \Rightarrow \phi \in Z^k(K; \mathbb{R}) . \end{aligned}$$

Thus $h([\phi]_K) = \phi$.

We leave it to the reader to prove that h is a linear mapping.

Hence, the map h is an isomorphism between $H^k(K; \mathbb{R})$ and $(H_k(K; \mathbb{R}))^*$. ■

It is possible to describe explicitly the action of the coboundary operator d_k . Suppose that K is a simplicial complex. Given an open k -simplex (s) in K , we defined $\phi_{[s]} \in C^k(K; \mathbb{R})$ by

$$\phi_{[s]}([t]) = \begin{cases} 1 & \text{if } [t] = [s] \\ -1 & \text{if } [t] = -[s] \\ 0 & \text{otherwise} \end{cases} \quad (8.3.2)$$

for all open k -simplices (t) in K . If $\{[s_1], [s_2], \dots, [s_m]\}$ is a basis of $C_k(K; \mathbb{R})$, then $\{\phi_{[s_1]}, \phi_{[s_2]}, \dots, \phi_{[s_m]}\}$ is a basis of $C^k(K; \mathbb{R})$.

Proposition 8.3.5

We have for all open k -simplices $(s) = (x_0, x_1, \dots, x_k) \in K$ that

$$d_k(\phi_{[x_0, x_1, \dots, x_k]}) = \sum_{\substack{(y) \in K \\ (y, x_0, x_1, \dots, x_k) \in K}} \phi_{[y, x_0, x_1, \dots, x_k]} . \quad (8.3.3)$$

Proof.

Since $C_{k+1}(K; \mathbb{R})$ is generated by oriented $(k+1)$ -simplices, we only have to prove (8.3.3) for an arbitrary ordered $(k+1)$ -simplex $[[z_0, z_1, \dots, z_{k+1}]]$.

By definition,

$$\begin{aligned} (d_k(\phi_{[x_0, x_1, \dots, x_k]}))([z_0, z_1, \dots, z_{k+1}]) &= \phi_{[x_0, x_1, \dots, x_k]}(\partial_{k+1}([z_0, z_1, \dots, z_{k+1}])) \\ &= \phi_{[x_0, x_1, \dots, x_k]} \left(\sum_{j=0}^{k+1} (-1)^j [[z_0, z_1, \dots, \widehat{z}_j, \dots, z_{k+1}]] \right) \\ &= \sum_{j=0}^{k+1} (-1)^j \phi_{[x_0, x_1, \dots, x_k]}([z_0, z_1, \dots, \widehat{z}_j, \dots, z_{k+1}]) . \end{aligned}$$

However

$$\begin{aligned} &\phi_{[x_0, x_1, \dots, x_k]}([z_0, z_1, \dots, \widehat{z}_j, \dots, z_{k+1}]) \\ &= \begin{cases} 1 & \text{if } [[z_0, z_1, \dots, \widehat{z}_j, \dots, z_{k+1}]] = [[x_0, x_1, \dots, x_k]] \\ -1 & \text{if } [[z_0, z_1, \dots, \widehat{z}_j, \dots, z_{k+1}]] = -[[x_0, x_1, \dots, x_k]] \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} 1 & \text{if } [[z_0, z_1, \dots, z_{k+1}]] = (-1)^j [[z_j, x_0, x_1, \dots, x_k]] \\ -1 & \text{if } [[z_0, z_1, \dots, z_{k+1}]] = (-1)^{j+1} [[z_j, x_0, x_1, \dots, x_k]] \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} (-1)^j \phi_{[z_j, x_0, x_1, \dots, x_k]}([z_0, z_1, \dots, z_{k+1}]) & \text{if } [[z_j, x_0, x_1, \dots, x_k]] \\ &= \pm [[z_0, z_1, \dots, z_{k+1}]] \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

Thus (8.3.3) is true for $[[z_0, z_1, \dots, z_{k+1}]]$. ■

8.3.2 Relative Cohomology

It is also natural to define relative k -cohomology modules. This is another topic that we only briefly introduce but that could be greatly developed. We will not do so in the context of simplicial theory but will treat it a little bit more seriously in Subsection 8.5.3 in the context of singular theory.

Definition 8.3.6

Let L be a simplicial subcomplex of a simplicial complex K , and R be an integral domain. The adjoint of the boundary operator $\bar{\partial}_k : C_k(K, L; R) \rightarrow C_{k-1}(K, L; R)$ is the operator $\bar{d}_{k-1} : (C_{k-1}(K, L; R))^* \rightarrow (C_k(K, L; R))^*$ defined

$$(\bar{d}_{k-1}(\phi))([c]) = \phi(\bar{\partial}_k([c])) \quad (8.3.4)$$

for all $[c]_{K,L} \in C_k(K, L; R)$ and $\phi \in (C_{k-1}(K, L; R))^*$.

It is clear that $\bar{d}_k$ is well and uniquely defined by (8.3.4).

Definition 8.3.7

Let L be a simplicial subcomplex of a simplicial complex K , and R be an integral domain. The **relative k^{th} cohomology module** of $K \bmod L$ is the quotient space $H^k(K, L; R) = \mathcal{N}(\bar{d}_k)/\mathcal{R}(\bar{d}_{k-1})$.

8.3.3 Cup Product

There is a feature of cohomology theory that is equivalent to the wedge product of differential forms on manifolds. It is called the cup product.

Definition 8.3.8

Let K be a simplicial complex and R be an integral domain. The **cup product** of $\phi_1 \in C^{k_1}(K; R)$ with $\phi_2 \in C^{k_2}(K; R)$ is the linear functional $\phi_1 \cup \phi_2 \in C^{k_1+k_2}(K; R)$ defined first by $(\phi_1 \cup \phi_2)([\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{k_1+k_2}]) = \phi_1([\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{k_1}]) \phi_2([\mathbf{x}_{k_1}, \mathbf{x}_{K_1+1}, \dots, \mathbf{x}_{k_1+k_2}])$ for all $(\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{k_1+k_2}) \in K$ and after extended linearly to $C^{k_1+k_2}(K; R)$.

It is not hard to prove that the cup product is a bilinear map, that it is associative and that the identity element is the linear functional $\delta \in C^0(K; R)$ defined by $\delta([\mathbf{x}_i]) = 1$ for all vertices $\mathbf{x}_i$ of K . We define the cup product of $[\phi_1]_K \in H^{k_1}(K; R)$ and $[\phi_2]_K \in H^{k_2}(K; R)$ as $[\phi_1]_K \cup [\phi_2]_K = [\phi_1 \cup \phi_2]_K$.

We will not justify these definitions and deduce their properties. We will prove later that simplicial homology and simplicial cohomology on the simplex K are respectively isomorphic to singular homology and singular cohomology on $[K]$. It is in the context of singular theory that we will prove that the cup product is well defined, and present other results about the cup product. There is effectively a lot more that can be said about the cup product.

8.3.4 Relation Between Simplicial Cohomology and de Rham Cohomology

Definition 8.3.9

A triple (S, K, h) is a **smoothly triangulated manifold** if

1. S is a manifold of class C^∞ ,
2. K is a simplicial complex,
3. $h : [K] \rightarrow S$ is a homeomorphism,
4. there exist, for every $(s) = (\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_p}) \in K$, an open subset $U_{[s]}$ of the k -dimensional affine space $A_{[s]} = \{\mathbf{x}_{j_0} + \sum_{i=1}^p y_i(\mathbf{x}_{j_i} - \mathbf{x}_{j_0}) : y_i \in \mathbb{R} \text{ for } 1 \leq i \leq p\}$ and a map $h_{[s]} : U_{[s]} \rightarrow S$ of class C^∞ such that $[s] \subset U_{[s]}$ and $h_{[s]}(\mathbf{x}) = h(\mathbf{x})$ for all $\mathbf{x} \in [s]$ (Figure 8.9), and
5. $h_{[s]}(U_{[s]})$ is a submanifold of S of class C^∞ .

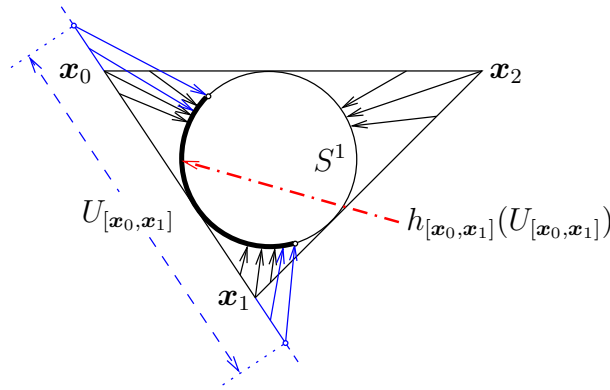


Figure 8.9: We consider the simplicial complex $K = \{(\mathbf{x}_0), (\mathbf{x}_1), (\mathbf{x}_2), (\mathbf{x}_0, \mathbf{x}_1), (\mathbf{x}_0, \mathbf{x}_2), (\mathbf{x}_1, \mathbf{x}_2)\}$ and the circle S^1 inscribed in $[K]$. The maps h is represented by the black arrows. A possible open set $U_{[x_0, x_1]}$ and map $h_{[x_0, x_1]}$ is provided with $h_{[x_0, x_1]}$ on $U_{[x_0, x_1]} \setminus [\mathbf{x}_0, \mathbf{x}_1]$ represented by the blue arrows.

Definition 8.3.10

Let K be a simplicial complex and $\mathbf{x}_j$ for $0 \leq j \leq q$ be the vertices of K . The $\mathbf{j}^{\text{th}}$

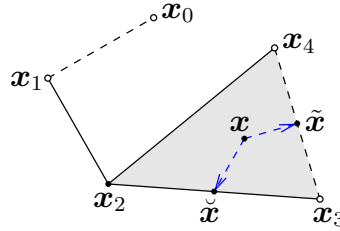
barycentric coordinates of $\mathbf{x} \in [K]$, denoted $b_j(\mathbf{x})$, is defined by

$$b_j(\mathbf{x}) = \begin{cases} 0 & \text{if } \mathbf{x} \notin \text{St}(\mathbf{x}_j) \\ \text{The } j^{\text{th}} \text{ barycentric} & \text{if } \mathbf{x} \in (s) \subset \text{St}(\mathbf{x}_j) \\ \text{coordinate of } \mathbf{x} \in (s) & \end{cases}$$

Remark 8.3.11

The j^{th} barycentric coordinates given in Definition 8.3.10 have several interesting properties.

1. We have that $b_j : [K] \rightarrow \mathbb{R}$ is continuous. For instance, suppose that K is the simplicial complex consisting of all the faces of $[\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_5]$ and the open simplices $(\mathbf{x}_0)$, $(\mathbf{x}_1)$, $(\mathbf{x}_0, \mathbf{x}_1)$ and $(\mathbf{x}_1, \mathbf{x}_2)$ as represented in the following figure.



$\text{St}(\mathbf{x}_2)$ is represented by the continuous lines and the shaded area. Given

$\mathbf{x} \in (\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4)$, we have that $\mathbf{x} = \sum_{j=2}^4 a_j \mathbf{x}_j$ with $\sum_{j=2}^4 a_j = 1$. The function b_2 is continuous with respect to the a_j in the open set $(\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4)$. Therefore, if $\mathbf{x}$ converges to $\tilde{\mathbf{x}} \in [\mathbf{x}_3, \mathbf{x}_4]$, then a_3 and a_4 converge to $\tilde{a}_3$ and $\tilde{a}_4$ respectively where $\tilde{\mathbf{x}} = \sum_{j=3}^4 \tilde{a}_j \mathbf{x}_j$

with $\sum_{j=3}^4 \tilde{a}_j = 1$. Thus $a_3 + a_4 \rightarrow 1$ and so $b_2(\mathbf{x}) = a_2$ converges to 0. Hence b_2 is continuous along $[\mathbf{x}_3, \mathbf{x}_4]$. Likewise, if $\mathbf{x}$ converges to $\check{\mathbf{x}} \in [\mathbf{x}_2, \mathbf{x}_3]$, then a_2 and a_3 converge to $\check{a}_2$ and $\check{a}_3$ respectively where $\check{\mathbf{x}} = \sum_{j=2}^3 \check{a}_j \mathbf{x}_j$ with $\sum_{j=2}^3 \check{a}_j = 1$. Thus $b_2(\mathbf{x}) = a_2$ converges to $\check{a}_2 = b_2(\check{\mathbf{x}})$. Hence b_2 is continuous along $[\mathbf{x}_2, \mathbf{x}_3]$. Proceeding in a similar way, we can prove that b_2 is continuous on all of $[K]$. Likewise, b_1 and b_3 are continuous on all of $[K]$.

2. For all $\mathbf{x} \in [K]$, we have that $\mathbf{x} = \sum_{j=0}^q b_j(\mathbf{x}) \mathbf{x}_j$ with $b_j(\mathbf{x}) \geq 0$ and $\sum_{j=0}^q b_j(\mathbf{x}) = 1$. More precisely, since $\mathbf{x} \in (\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_p}) \in K$ for some $j_i \in \{0, 1, 2, \dots, q\}$ and $0 \leq i \leq p$, we get that $\mathbf{x} = \sum_{i=0}^p b_{j_i}(\mathbf{x}) \mathbf{x}_{j_i}$ with $b_{j_i}(\mathbf{x}) > 0$ and $\sum_{i=0}^p b_{j_i}(\mathbf{x}) = 1$. Moreover, $b_j(\mathbf{x}) = 0$ if $j \notin \{j_0, j_1, \dots, j_p\}$.


Definition 8.3.12

Let K be a simplicial complex and (s) be an open simplex in K . The **star** of (s) , denoted $\text{St}((s))$, is defined by $\text{St}((s)) = \bigcup_{\substack{(t) \in K \text{ such that} \\ (s) \text{ is a face of } [t]}} (t)$.

Recall that, by definition of a simplicial complex, if $(s) \in K$, then all the faces of $[s]$ are also in K ; in particular, $(s) \subset \text{St}((s))$ because (s) is a face of $[s]$.

Remark 8.3.13

Below are some of the properties of the star of an open simplex as given in Definition 8.3.12. Some these properties are easy to prove and proving them is left to the reader.

1. For a 0-simplex $(\mathbf{x}) \in K$, we have that $\text{St}((\mathbf{x})) = \text{St}(\mathbf{x})$, the star of the vertex $\mathbf{x}$ of K as given in Definition 8.2.19.
2. If (s) is an open simplex of K , then $\text{St}((s))$ is an open set in $[K]$. This follows from the fact that $\text{St}((s))$ is the union of open simplices of $[K]$.
3. If $(s) = (\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_p}) \in K$ where the $\mathbf{x}_{j_i}$ for $0 \leq i \leq p$ are vertices of K , then $\mathbf{x} \in \text{St}((s))$ if and only if $b_{j_i}(\mathbf{x}) \neq 0$ for $0 \leq i \leq p$. To verify this claim, it suffices to note that if $\mathbf{x} \in \text{St}((s))$ then $\mathbf{x} \in (t)$ such that (s) is a face of $[t]$. Thus $(t) = (\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_q})$ with $q \geq p$ and $j_i \notin \{j_0, j_1, \dots, j_p\}$ for $p < i \leq q$. Hence $\mathbf{x} = \sum_{i=0}^q b_{j_i}(\mathbf{x}) \mathbf{x}_{j_i}$ with $b_{j_i}(\mathbf{x}) > 0$ for $0 \leq i \leq q$ and $\sum_{i=0}^q b_{j_i}(\mathbf{x}) = 1$.
4. It follows from (3) that $[K] \setminus \text{St}((s)) = \{\mathbf{x} \in K : b_{j_i}(\mathbf{x}) = 0 \text{ for some } i \in \{0, 1, \dots, p\}\}$.
5. If (s_1) and (s_2) are two open q -simplices of a simplicial complex K , and $(s_1) \neq (s_2)$, then $[s_2] \subset [K] \setminus \text{St}((s_1))$. This follows from the fact that all faces of a closed q -simplex, other than the open q -simplex itself, are open j -simplices with $j < q$. So (s_1) cannot be a face of $[s_2]$.



Suppose that (S, K, h) is a smoothly triangulated manifold. We would like to find a relation between the k^{th} cohomology module of S defined in Definition 7.2.14 and the k^{th} cohomology group $H^k(K; \mathbb{R})$ of K . For that, we will define a map $F_k : C^\infty(S, \Omega^k(S)) \rightarrow C^k(K, \mathbb{R})$ such that the following diagram commutes.

$$\begin{array}{ccc}
 C^\infty(S, \Omega^k(S)) & \xrightarrow{d} & C^\infty(S, \Omega^{k+1}(S)) \\
 F_k \downarrow & & \downarrow F_{k+1} \\
 C^k(K; \mathbb{R}) & \xrightarrow{d_k} & C^{k+1}(K; \mathbb{R})
 \end{array}$$

Suppose that ω is a differential k -form on S . To define $F_k(\omega) \in C^k(K; \mathbb{R})$, it is enough to define $F_k(\omega)$ on the equivalence classes of oriented k -simplices of K and use the linearity of $F_k(\omega)$ to extend its definition to all of $C_k(K; \mathbb{R})$. Given $[[s]] \in C_k(K; \mathbb{R})$ with $(s) = (\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_k}) \in K$, there exists an open subset $U_{[s]}$ and a smooth mapping $h_{[s]} : U_{[s]} \rightarrow S$ that satisfy Definition 8.3.9. We set

$$F_k(\omega)([[s]]) = \int_{[s]} h_{[s]}^*(\omega) .$$

We may consider that $[s]$ is a subset of the k -dimensional manifold $U_{[s]}$. This manifold can be described by only one local chart; namely, $(W_{[s]}, U_{[s]}, \phi_{[s]})$ where

$$\begin{aligned} \phi_{[s]} : \mathbb{R}^k &\rightarrow \{\mathbf{x}_{j_0} + \sum_{i=1}^k y_i(\mathbf{x}_{j_i} - \mathbf{x}_{j_0}) : y_i \in \mathbb{R} \text{ for } 1 \leq i \leq k\} \\ \mathbf{y} &\mapsto \mathbf{x}_{j_0} + \sum_{i=1}^k y_i(\mathbf{x}_{j_i} - \mathbf{x}_{j_0}) \end{aligned}$$

and $W_{[s]} = \phi_{[s]}^{-1}(U_{[s]})$. Hence, the local representation of $h_{[s]}^*(\omega)$ is given by $\phi_{[s]}^*(h_{[s]}^*(\omega)) = g_{[s]} dy_1 \wedge dy_2 \wedge \dots \wedge dy_k$ where $g_{[s]} : W_{[s]} \rightarrow \mathbb{R}$ is of class C^∞ and

$$F_k(\omega)([[s]]) = \int_{W_{[s]}} g_{[s]} dy_1 \wedge dy_2 \wedge \dots \wedge dy_k .$$

Proposition 8.3.14

We have that $d_k \circ F_k = F_{k+1} \circ d$.

Proof.

It suffices to use Stokes' theorem. Given a differential k -form on S and $[[s]] \in C_k(K; \mathbb{R})$, we have

$$\begin{aligned} F_{k+1}(d\omega)([[s]]) &= \int_{[s]} h_{[s]}^*(d\omega) = \int_{[s]} d(h_{[s]}^*(\omega)) = \int_{\partial[s]} h_{[s]}^*(\omega) \\ &= F_k(\omega)(\partial_k[[s]]) = d_k(F_k(\omega))([s]) . \end{aligned}$$

Stokes' theorem, Theorem 4.7.5, was used for the third equality. Moreover, $\partial[s]$ must be interpreted according to Stokes' theorem. Hence the orientation on $\partial[s]$ is the induced orientation from the orientation on $[s]$. It is in fact the orientation given by $\partial_k[[s]]$. ■

The next theorem is the main result of this section.

Theorem 8.3.15 (De Rham's Theorem)

Let (S, K, h) be a smoothly triangulated manifold and $\tilde{F}_k : H^k(S) \rightarrow H^k(K; \mathbb{R})$ be the map defined by $\tilde{F}_k([\omega])([c]) = [F_k(\omega)(c)]_K$ for the closed differential k -form ω on S and k -chains c in $Z_k(K; \mathbb{R})$. Then $\tilde{F}_k$ is an isomorphism for $0 \leq k \leq \dim(S)$.

The proof of this theorem is very long but it is not hard. We only sketch the proof of this famous theorem. The reader can find the detailed proof in [25].

Proof (Sketch).

Claim 1) There exist linear maps $G_k : C^k(K; \mathbb{R}) \rightarrow C^\infty(S, \Omega^k(S))$ for $0 \leq k \leq \dim(S)$ such that:

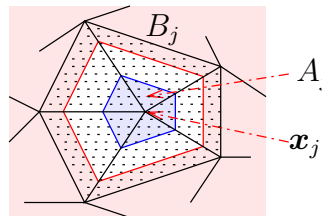
1. $d \circ G_k = G_{k+1} \circ d_k$ for $0 \leq k < \dim(S)$.
2. $F_k \circ G_k = \text{Id}$.
3. If $\phi \in C^0(K, \mathbb{R})$ satisfies $\phi([x]) = 1$ for all 0-simplices $(x) \in K$, then $G_0(\phi)$ is the differential 0-form (i.e. a function on S) defined by $G_0(\phi)(x) = 1$ for all $x \in S$.
4. If $[s]$ is an oriented k -simplex of K , then $G_k(\phi_{[s]})$ is null in an open set containing $[K] \setminus \text{St}([s])$, where $\phi_{[s]}$ is defined in (8.3.2).

i) Suppose that K is a simplicial complex of dimension q and that $x_0, x_1, \dots, x_p$ are the vertices of K . We define a partition of unity subordinated to the open cover $\{\text{St}(x_j)\}_{0 \leq j \leq p}$ of $[K]$.

Let $A_j = \{x \in [K] : b_j(x) \geq 1/(q+1)\}$ and $B_j = \{x \in [K] : b_j(x) \leq 1/(q+2)\}$ for $0 \leq j \leq p$, where b_j is the barycentric coordinate function defined in Definition 8.3.10.

Since A_j is a closed subset of the compact set $[K]$, we have that A_j is compact for $0 \leq j \leq p$.

We have that $[K] \setminus \text{St}(x_j) \subset B_j$ and $A_j \subset C_j = [K] \setminus B_j \subset \text{St}(x_j)$ for $0 \leq j \leq p$ as can be seen in the figure below.



In the figure above, the dotted area represents $\text{St}(x_j)$. The collection $\{A_j\}_{0 \leq j \leq p}$ is a closed cover of $[K]$. Given $x \in [K]$, we have that x is an element of an open m -simplex $(s) = (x_{j_1}, x_{j_2}, \dots, x_{j_m})$ of K with $m \leq q$. By definition of the barycentric coordinates, we have

that $b_{j_i}(\mathbf{x}) > 0$ for $0 \leq i \leq m$ and $\sum_{i=0}^m b_{j_i}(\mathbf{x}) = 1$. Therefore, we must have that $b_{j_i}(\mathbf{x}) \geq 1/(m+1) \geq 1/(q+1)$ for some $0 \leq i \leq m$. Thus $\mathbf{x} \in A_{j_i}$.

Given $0 \leq j \leq p$, there exists according to Proposition 7.2.35 an open set $V_j \subset \mathbb{R}^n$ such that $A_j \subset V_j \subset \overline{V_j} \subset \mathbb{R}^n \setminus B_j$. Using partition of unity for instance, we can select a function $f_j \in C^\infty(\mathbb{R}^n)$ such that $0 \leq f_j(\mathbf{x}) \leq 1$ for all $\mathbf{x} \in \mathbb{R}^n$, $f_j(\mathbf{x}) = 1$ for all $\mathbf{x} \in A_j$ and $f_j(\mathbf{x}) = 0$ for all $\mathbf{x} \notin V_j$. It follows from the previous paragraph that, given any $\mathbf{x} \in [K]$, there exists $0 \leq j \leq p$ such that $f_j(\mathbf{x}) = 1$. Let $g = \sum_{j=0}^p f_j$ on $\mathbb{R}^n$. We have that $g(\mathbf{x}) \geq 1$ for all $\mathbf{x} \in [K]$. Thus $g(\mathbf{x}) \geq 1/2 > 0$ for all $\mathbf{x}$ in an open set $W \supset [K]$. Let $g_j = f_j/g$ on W for $0 \leq j \leq p$. Hence $g_j \in C^\infty(W)$ and $\sum_{j=0}^p g_j = 1$ on $[K]$. Moreover, $\text{supp } g_j|_{[K]} \subset \text{supp } f_j|_{[K]} \subset C_j \subset \text{St}(\mathbf{x}_j)$ for $0 \leq j \leq p$.

The collection $\{g_j\}_{0 \leq j \leq p}$ is the partition of unity subordinated to the open cover $\{\text{St}(\mathbf{x}_j)\}_{0 \leq j \leq p}$ of $[K]$ that we are looking for.

ii) Since G_k is linear, we only have to define G_k for $\phi_{[s]}$ where $[s]$ is an oriented k -simplices of K . Suppose that $[s] = [\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_k}]$. We set

$$G_k(\phi_{[s]}) = k! \sum_{i=0}^k (-i)^i g_{j_i} dg_{j_0} \wedge dg_{j_1} \wedge \dots \wedge \widehat{dg_{j_i}} \wedge \dots \wedge dg_{j_k}.$$

Note that $G_k(\phi_{[s]})$ defines a differential k -form on $W \supset [K]$. We can use the maps $h_{[s]}$ provided by the smoothly triangulated manifold (S, K, h) to transport $G_k(\phi_{[s]})$ to S ⁵.

The proof given in [25] that G_k satisfies all four conditions listed above requires, among several results, Stokes' theorem on chains.

Claim 2) If ω is a closed differential k -form on S such that $F_k(\omega) = d_k(\phi)$ for some $\phi \in C^{k-1}(K; R)$, then there exists a differential $(k-1)$ -form ρ on S such that $F_{k-1}(\rho) = \phi$ and $d\rho = \omega$.

The proof of this claim is in two parts.

First Part) Given an open q -simplex $(s) \in K$, let $B_{(s)} = \bigcup_{\substack{(t) \text{ an open face of } [s] \\ \text{with } \dim(t) < q}} (t).$

We prove the following statements.

1. Assuming $q > 0$, if ω is a closed differential k -form defined in an open neighbourhood of $B_{(s)}$ in $\mathbb{R}^n$ where $(s) \in K$ is an open q -simplex. then there exists a closed differ-

⁵In [25], they simplify the writing of the proof by assuming that $[K] = S$. In other words, they project K on S using the maps h_s .

ential k -form ρ defined in an open neighbourhood of $[s]$ such that $\rho = \omega$ in the open neighbourhood of $B_{(s)}$. If $q = k + 1$, then we must also require that $\int_{\partial[s]} \omega = 0$.

2. Assuming $q, k > 0$, if ω is a closed differential k -form defined in an open neighbourhood of $[s]$ in $\mathbb{R}^n$ where $(s) \in K$ is an open q -simplex, and ρ is a differential $(k - 1)$ -form defined in an open neighbourhood of $B_{(s)}$ and such that $d\rho = \omega$ in that neighbourhood of $B_{(s)}$, then there exists a differential $(k - 1)$ -form ν defined in an open neighbourhood of $[s]$ such that $\nu = \rho$ in an open neighbourhood of $B_{(s)}$ and $d\nu = \omega$ in a neighbourhood of $[s]$. If $q = k$, then we must also require that $\int_{\partial[s]} \rho = \int_{[s]} \omega$.

Note that the two extra conditions in the two previous statement are required because we get from Stokes' theorem that

$$\int_{\partial[s]} \omega = \int_{\partial[s]} \rho = \int_{[s]} d\rho = \int_{[s]} 0 = 0$$

if ρ exists in the first statement and $q = k + 1$, and

$$\int_{[s]} \omega = \int_{[s]} d\nu = \int_{\partial[s]} \nu = \int_{\partial[s]} \rho$$

if ν exists in the second statement and $q = k$.

The proof of the two statements above is an iterative proof. We first prove that (1) is true for $k = 0$. Next, we prove that if (1) is true for $k = i$, then (2) is true for $k = i + 1$. Finally, we prove that if (2) is true for $k = i + 1$, then (1) is true for $k = i + 1$. Hence, iterating from $i = 0$ up to $i = \dim(S)$ shows that (1) and (2) are true for all values of k .

To prove (1) for $k = 0$, we note that ω is a function on an open neighbourhood of $B_{(s)}$ such that $\frac{\partial \omega}{\partial x_i} = 0$ in this open neighbourhood for all i . Thus ω is constant on the components of $B_{(s)}$. If $q > 1$, then $B_{(s)}$ is connected and so $\omega(\mathbf{x}) = c$, a constant, for all $\mathbf{x} \in B_{(s)}$. Therefore, we may set $\rho = c$ in an open neighbourhood of $[s]$. If $q = 1$, then $[s] = [\mathbf{x}_{j_1}, \mathbf{x}_{j_2}]$ for some vertices $\mathbf{x}_{j_1}$ and $\mathbf{x}_{j_2}$ of K , and $B_{(s)} = \{\mathbf{x}_{j_1}, \mathbf{x}_{j_2}\}$. But since $q = k + 1$, we have that $0 = \int_{\partial[s]} \omega = \omega(\mathbf{x}_{j_1}) - \omega(\mathbf{x}_{j_2})$. Thus $\omega(\mathbf{x}_{j_1}) = \omega(\mathbf{x}_{j_2}) = c$, a constant. Therefore, we may again set $\rho = c$ in an open neighbourhood of $[s]$.

To prove that (1) for $k = i$ implies (2) for $k = i + 1$, we may assume that ω is a closed differential k -form in a connected open set $U \supset [s]$ because $[s]$ is connected. It then follows from Corollary 7.2.22 that $\omega = d\mu_1$ for some differential $(k - 1)$ -form μ_1 on U ⁶. Since μ_1 may not be equal to ρ in a neighbourhood of $B_{(s)}$, we consider $\mu_2 = \mu_1 - \rho$. We have that $d\mu_2 = d\mu_1 - d\rho = \omega - \omega = 0$ in a neighbourhood of $B_{(s)}$ and, if $q = k$, we have from Stokes' theorem that

$$\int_{\partial[s]} \mu_2 = \int_{\partial[s]} \mu_1 - \int_{\partial[s]} \rho = \int_{[s]} d\mu_1 - \int_{\partial[s]} \rho = \int_{[s]} \omega - \int_{\partial[s]} \rho = 0$$

⁶We are using the fact that $H^k(U) = H^k(\mathbb{R}^n) = 0$ for $k > 0$ because U is diffeomorphic to $\mathbb{R}^n$.

by assumption. Therefore, we may use (1) with μ_2 to get a closed differential k -form μ_3 defined in an open neighbourhood of $[s]$ such that $\mu_3 = \mu_2$ in a neighbourhood of $B_{(s)}$. Then $\nu = \mu_1 - \mu_3$ because $\nu = \rho$ and $d\nu = d\mu_1 - d\mu_3 = \omega$ in a neighbourhood of $B_{(s)}$.

To prove that (2) for $k = i + 1$ implies (1) for $k = i + 1$ is a little bit trickier. Suppose that $(s) = (\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_q})$. Let $(t) = (\mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_q})$ and $F_{(s)} = B_{(s)} \setminus (t)$. Since $F_{(s)}$ is star-shaped, we may assume that the open neighbourhood of $F_{(s)}$ where ω is defined is also star-shaped. Hence, we may use Poincaré Lemma to get a differential $(k - 1)$ -form μ defined in this neighbourhood such that $d\mu = \omega$.

If $q > 1$, then we can show that (2) can be used with ρ replaced by μ and (s) by (t) to get a differential $(k - 1)$ -form ν defined in an open neighbourhood of $[t]$ such that $\nu = \mu$ in an open neighbourhood of $B_{(t)}$ and $d\nu = \omega$ in a neighbourhood of $[t]$. Consider the differential k -form defined by

$$\tau = \begin{cases} \mu & \text{on the neighbourhood of } F_{(s)} \\ \nu & \text{on the neighbourhood of } [t] \end{cases}$$

where the neighbourhoods may have to be shrunk a little if necessary. We have that $d\tau = \omega$ in a neighbourhood of $B_{(s)}$. If $q = 1$, then we also have that there is a differential $k - 1$ -form τ defined in $B_{(s)}$ such that $d\tau = \omega$. In this case, we use the fact that the neighbourhood of $B_{(s)}$ may be chosen to be the union of two disconnect open sets; one about each of the two vertices in $B_{(s)}$.

To complete the proof, we select a real valued function $f \in C^\infty(\mathbb{R}^n)$ such that $f = 1$ in an open neighbourhood of $B_{(s)}$ and $f = 0$ outside the open neighbourhood where τ is defined. Then $\rho = d(f\tau)$ satisfies (1) because it is obviously exact and

$$d\rho = df \wedge \tau + f \wedge d\tau = d\tau = \omega$$

in the open neighbourhood where $f = 1$.

Second Part) The second part of the proof of Claim 2 used the second statement of the previous part to construct a sequence of differential $(k - 1)$ -forms $\rho_0, \rho_1, \rho_2, \dots, \rho_{\dim(S)}$ such that ρ_j is defined in an open neighbourhood of the j -skeleton $[K^j]$ of K , $d\rho_j = \omega$ in this open neighbourhood of $[K^j]$, $\rho_j = \rho_{j-1}$ in an open neighbourhood of $[K^{j-1}]$ for $0 < j \leq \dim(S)$, and $F_j(\rho_j) = \phi$ if $j = k - 1$. We have that ρ in Claim 2 is given by $\rho = \rho_{\dim(S)}$.

The construction of the ρ_j is done by induction on the order j of the skeletons of K .

For $j = 0$, the skeleton K^0 is a set of discrete points. We may assume that the open neighbourhood of $[K^0]$ is the union of disjoint open balls B_j with $\mathbf{x}_j$ being the only vertex of K in B_j . Using Poincaré Lemma on each open ball, we get a differential $(k - 1)$ -form $\tilde{\rho}_0$ on the open neighbourhood of $[K^0]$ such that $d\tilde{\rho}_0 = \omega$. If $k - 1 > 0$, then we may take $\rho_0 = \tilde{\rho}_0$. If $k - 1 = 0$, then we have to make sure that $F_0(\rho_0) = \phi$ on $[K^0]$. For each vertex $\mathbf{x}_j$ of K , set $a_j = \phi((\mathbf{x}_j)) - \int_{(\mathbf{x}_j)} \tilde{\rho}_0 = \phi((\mathbf{x}_j)) - \tilde{\rho}_0(\mathbf{x}_j)$. Then ρ_0 defined in the open neighbourhood of $[K^0]$ by $\rho_0|_{B_j} = \tilde{\rho}_0 + a_j$ for all j satisfies all the required conditions. In particular, $F_0(\rho_0)((\mathbf{x}_j)) = \int_{(\mathbf{x}_j)} (\tilde{\rho}_0 + a_j) = \phi((\mathbf{x}_j))$ for all $\mathbf{x}_j$.

We assume that we have constructed ρ_{j-1} according to the requirements.

We have that ω is a closed differential k -form defined in an open neighbourhood of a closed j -simplex $[s]$ and ρ_{j-1} is a differential $(j-1)$ -form defined in an open neighbourhood of $B_{(s)}$ such that $d\rho_{j-1} = \omega$ in this open neighbourhood of $B_{(s)}$. If $k = j$, then

$$\int_{[s]} \omega = F_k(\omega)([s]) = d_k(\phi)([s]) = \phi(\partial_k[s]) = F_{j-1}(\rho_{j-1})(\partial_k[s]) = \int_{\partial_k[s]} \rho_{j-1}$$

where the second equality comes from the assumption in Claim 2 and the fourth equality from our hypothesis of induction. Hence the second statement of the first part can be used to get a differential $(k-1)$ -form $\tilde{\rho}_{(s)}$ in an open neighbourhood $U_{(s)}$ of $[s]$ such that $\tilde{\rho}_{[s]} = \rho_{j-1}$ in an open neighbourhood of $B_{(s)}$ and $d\tilde{\rho}_{(s)} = \omega$ in $U_{(s)}$. We define $\tilde{\rho}_j$ in a neighbourhood of $[K^j]$ by $\tilde{\rho}_j|_{U_{(s)}} = \tilde{\rho}_{(s)}$ for all open j -simplices (s) in K . The differential $(k-1)$ -form $\tilde{\rho}_j$ is well defined by because $\tilde{\rho}_{(s_1)} = \tilde{\rho}_{(s_2)} = \rho_{j-1}$ in an open neighbourhood of the intersection of $[s_1]$ and $[s_2]$ for all open j -simplices (s_1) and (s_2) in K .

If $k-1 \neq j$, then we may define ρ_j in a neighbourhood of $[K^j]$ as $\rho_j = \tilde{\rho}_j$.

If $k-1 = j$, then we use a trick similar to the one that we have used in the case $j = 0$ to ensure that $F_j(\rho_j) = \phi$ on $C^j(K; R)$. Let $\alpha_j = \phi - F_{k-1}(\tilde{\rho}_{k-1}) \in C^j(K; R)$. We define ρ_j in a neighbourhood U of $[K^j]$ as $\rho_j = \tilde{\rho}_j|_U + G_{k-1}(\alpha_j)|_U$.

We have from Claim 1 that $G_q(\phi_{[t]}) = 0$ in an open neighbourhood of $K \setminus \text{St}((t))$ for all open q -simplices (t) in K . Thus $G_q(\phi_{[t]}) = 0$ in a neighbourhood of $[K^{q-1}]$ for all open q -simplices (t) in K . Since all elements of $C^q(K; R)$ are linear combinations of elements of the form $\phi_{[t]}$ for (t) an open q -simplex in K , we get that $G_q(\phi) = 0$ in a neighbourhood of $[K^{q-1}]$ for all $\phi \in C^q(K; R)$. Thus, for $q = k-1 = j$, we get that

$$d\rho_j = d\tilde{\rho}_j + d(G_{k-1}(\alpha_j)) = d\tilde{\rho}_j + G_k(d_{k-1}(\alpha_j)) = d\tilde{\rho}_j = \omega$$

in a neighbourhood of $[K^j]$, and

$$\rho_j = \tilde{\rho}_j + G_{k-1}(\alpha_j) = \tilde{\rho}_j = \rho_{j-1}$$

in a neighbourhood of $[K^{j-1}]$. Moreover,

$$F_j(\rho_j)([s]) = F_j(\tilde{\rho}_j)([s]) + F_j(G_{k-1}(\alpha_j)([s])) = \phi([s]) - \alpha_j([s]) + \alpha_j([s]) = \phi([s])$$

for all j -simplices $([s])$, where we have use (2) of Claim 1 to get the second equality.

Conclusion It follows from (1) of Claim 1 that $G_k(Z^k(K; \mathbb{R})) \subset Z^k(S)$ where $Z^k(S)$ is the set of closed differential k -forms, and $G_k(B^k(K; \mathbb{R})) \subset B^k(S)$ where $B^k(S)$ is the set of exact differential k -forms. Thus G_k defines a map, that we denote $\tilde{G}_k$, from $H^k(K; \mathbb{R})$ to $H^k(S)$. Hence, it follows from (2) that $\tilde{G}_k$ is a right inverse for $\tilde{F}_k$ and so $\tilde{F}_k$ is onto.

It follows from Claim 2 that, if $\omega \in Z^k(S)$ and $F_k(\omega) \in B^k(K; \mathbb{R})$, then $\omega \in B^k(S)$. Since F_k is a linear map, this proves that $\tilde{F}_k([\omega]) = [0]_K$ implies that $[\omega] = [0]$. Thus F_k is one-to-one. This completes the proof that $\tilde{F}_k$ is an isomorphism. ■

8.4 Singular Homology

We present the singular approach to homology theory. We generally follow the approach of [9] with a few exceptions.

8.4.1 Homotopy Modules

Definition 8.4.1

For $k \geq 0$, the **standard k -simplex**, denoted Δ_k , is defined by

$$\Delta_k = \left\{ \sum_{j=0}^k a_j \mathbf{e}_j : a_j \geq 0 \text{ and } \sum_{j=0}^k a_j = 1 \right\} \text{ where } \mathbf{e}_0 = \mathbf{0} \in \mathbb{R}^\infty \text{ and } \mathbf{e}_j \in \mathbb{R}^\infty \text{ for } 1 \leq j \leq k \text{ is the unit vector with 1 for the } j^{\text{th}} \text{ component and 0 everywhere else.}$$

Using the notation of the section on simplicial homology, we have that Δ_k is the closed k -simplex $[\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k]$.

Definition 8.4.2

For $k \geq 0$, a **singular k -simplex** in a topological space X is a continuous map $\sigma : \Delta_k \rightarrow X$. Given an integral domain R , the free module over R generated by all the singular k -simplices in X is denoted $S_k(X; R)$. The elements of $S_k(X; R)$ are called **singular k -chain**. The R -module $S_k(X; R)$ is simply denoted $S_k(X)$ when the choice of the integral domain is clear. It is also convenient to set $S_k(X; R) = \{0\}$ for $k < 0$.

The two integral domains that we will use in all our applications are $R = \mathbb{R}$ or $R = \mathbb{Z}$.

By definition of a module over an integral domain generated by a set of elements, a singular k -chain $c \in S_k(X; R)$ can be written as $c = \sum_{j \in J} a_j \sigma_j$ where $J \subset \mathbb{N}$ is a finite set, $a_j \in R$

and σ_j is a singular k -simplex for $j \in J$. The sum of two singular k -chains $c_1 = \sum_{j \in J} a_j \sigma_j$

and $c_2 = \sum_{j \in J} b_j \sigma_j$ is therefore the singular k -chain given by $c_1 + c_2 = \sum_{j \in J} (a_j + b_j) \sigma_j$. We

may always assume that the two singular k -chains c_1 and c_2 are sums over the same set J by adding terms of the form $0\sigma_j$ to the sum where they are not present. The product of a

singular k -chain $c = \sum_{j \in J} a_j \sigma_j$ by $r \in R$ is the singular k -chain given by $rc = \sum_{j \in J} (ra_j) \sigma_j$.

The next known definition is important to review since it is going to play a very important role in this section.

Definition 8.4.3

An **affine space** is a subset X of $\mathbb{R}^n$ such that $t\mathbf{x}_1 + (1-t)\mathbf{x}_2 \in X$ for all $\mathbf{x}_1, \mathbf{x}_2 \in X$ and $0 \leq t \leq 1$. A function $f : X \rightarrow Y$ between two affine spaces X and Y is **affine** if $f(t\mathbf{x}_1 + (1-t)\mathbf{x}_2) = tf(\mathbf{x}_1) + (1-t)f(\mathbf{x}_2)$ for all $\mathbf{x}_1, \mathbf{x}_2 \in X$ and $0 \leq t \leq 1$.

There is a more general definition of affine space than the one that we have given. We will only use affine spaces that are subsets $\mathbb{R}^n$ for some $n > 0$. More precisely, we will basically only consider the affine space Δ_k for $k \in \mathbb{N}$.

Suppose that X and Y are two affine spaces. It follows from the definition of an affine function that if $f : X \rightarrow Y$ is an affine function, then $f\left(\sum_{j \in J} a_j \mathbf{x}_j\right) = \sum_{j \in J} a_j f(\mathbf{x}_j)$ for all $\mathbf{x}_j \in X$ and $a_j \in \mathbb{R}$ such that $\sum_{j \in J} a_j = 1$ with $a_j \geq 0$.

Proposition 8.4.4

If $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is affine, then there exist a linear mapping $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $\tilde{\mathbf{x}} \in \mathbb{R}^m$ such that $f(\mathbf{x}) = L(\mathbf{x}) + \tilde{\mathbf{x}}$ for all $\mathbf{x} \in \mathbb{R}^n$.

Proof.

Let $\tilde{\mathbf{x}} = f(\mathbf{0})$ and $L(\mathbf{x}) = f(\mathbf{x}) - \tilde{\mathbf{x}}$ for $\mathbf{x} \in \mathbb{R}^n$. It suffices to prove that $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a linear map. We have

$$\begin{aligned} L(a\mathbf{x}) &= f(a\mathbf{x}) - \tilde{\mathbf{x}} = f(a\mathbf{x} + (1-a)\mathbf{0}) - \tilde{\mathbf{x}} = af(\mathbf{x}) + (1-a)f(\mathbf{0}) - \tilde{\mathbf{x}} \\ &= af(\mathbf{x}) + (1-a)\tilde{\mathbf{x}} - \tilde{\mathbf{x}} = a(f(\mathbf{x}) - \tilde{\mathbf{x}}) = aL(\mathbf{x}) \end{aligned} \quad (8.4.1)$$

for all $a \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^n$. Moreover

$$\begin{aligned} L(\mathbf{x}_1 + \mathbf{x}_2) &= L\left(2\left(\frac{1}{2}\mathbf{x}_1 + \frac{1}{2}\mathbf{x}_2\right)\right) = 2L\left(\frac{1}{2}\mathbf{x}_1 + \frac{1}{2}\mathbf{x}_2\right) = 2\left(f\left(\frac{1}{2}\mathbf{x}_1 + \frac{1}{2}\mathbf{x}_2\right) - \tilde{\mathbf{x}}\right) \\ &= 2\left(\frac{1}{2}f(\mathbf{x}_1) + \frac{1}{2}f(\mathbf{x}_2) - \tilde{\mathbf{x}}\right) = (f(\mathbf{x}_1) - \tilde{\mathbf{x}}) + (f(\mathbf{x}_2) - \tilde{\mathbf{x}}) = L(\mathbf{x}_1) + L(\mathbf{x}_2) \end{aligned}$$

for all $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$, where we have used (8.4.1) to get the second equality. ■

It follows from the previous proposition that an affine function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is completely determined by its values at $\mathbf{v}_j$ for $0 \leq j \leq n$ if $\mathbb{R}^n = \text{span}\{\mathbf{v}_j : 0 \leq j \leq n\}$. Moreover, we

have that $f\left(\sum_{j=0}^n a_j \mathbf{v}_j\right) = \sum_{j=0}^n a_j f(\mathbf{v}_j)$ if $\sum_{j=0}^n a_j = 1$ for $a_j \in \mathbb{R}$.

For $k > 0$ and $0 \leq j \leq k$, let $F_k^j : \Delta_{k-1} \rightarrow \Delta_k$ be the affine function defined by

$$F_k^j(\mathbf{e}_i) = \begin{cases} \mathbf{e}_i & \text{if } i < j \\ \mathbf{e}_{i+1} & \text{if } i \geq j \end{cases}$$

for $0 \leq i \leq k-1$.

Definition 8.4.5

Let X be a topological space and R be an integral domain. Given a singular k -simplex $\sigma : \Delta_k \rightarrow X$ in X with $k > 0$. The **j-face** of σ is the singular $(k-1)$ -simplex defined by $\sigma^{(j)} = \sigma \circ F_k^j : \Delta_{k-1} \rightarrow X$. The **boundary** of σ is defined by $\partial_k(\sigma) = \sum_{j=0}^k (-1)^j \sigma^{(j)}$. The **boundary operator** $\partial_k : S_k(X; R) \rightarrow S_{k-1}(X; R)$ for $k > 0$ is defined by linearity; namely, given a singular k -chain $c = \sum_{j \in J} a_j \sigma_j$, we set $\partial_k c = \sum_{j \in J} a_j \partial_k(\sigma_j)$. It is convenient to define $\partial_k = 0$ for $k \leq 0$.

Proposition 8.4.6

Let X be a topological space. Then $\partial_{k-1} \circ \partial_k = 0$ for all $k \in \mathbb{Z}$.

Proof.

The result is obviously true for $k \leq 1$. Thus, we may assume that $k > 1$.

i) We first note that $F_k^i \circ F_{k-1}^j = F_k^j \circ F_{k-1}^{i-1}$ for $0 \leq j < i \leq k$. To prove this statement, it suffices to prove it for all e_q for $1 \leq q \leq k-1$ because we are working with affine functions. We have from

$$F_{k-1}^j(e_q) = \begin{cases} e_q & \text{if } q < j \\ e_{q+1} & \text{if } j \leq q \end{cases}$$

that

$$F_k^i(F_{k-1}^j(e_q)) = \begin{cases} e_q & \text{if } q < j \\ e_{q+1} & \text{if } j \leq q < i-1 \\ e_{q+2} & \text{if } i-1 \leq q \end{cases}$$

Moreover, we have from

$$F_{k-1}^{i-1}(e_q) = \begin{cases} e_q & \text{if } q < i-1 \\ e_{q+1} & \text{if } i-1 \leq q \end{cases}$$

that

$$F_k^j(F_{k-1}^{i-1}(e_q)) = \begin{cases} e_q & \text{if } q < j \\ e_{q+1} & \text{if } j \leq q < i-1 \\ e_{q+2} & \text{if } i-1 \leq q \end{cases}$$

Thus $F_k^i(F_{k-1}^j(e_q)) = F_k^j(F_{k-1}^{i-1}(e_q))$.

ii) By linearity, it is enough to prove $\partial_{k-1} \circ \partial_k = 0$ for singular k -simplices. Given a singular k -simplex σ , we have that

$$\partial_{k-1}(\partial_k(\sigma)) = \partial_{k-1} \left(\sum_{i=0}^k (-1)^i \sigma \circ F_k^i \right) = \sum_{j=0}^{k-1} \sum_{i=0}^k (-1)^j (-1)^i \sigma \circ F_k^i \circ F_{k-1}^j$$

$$\begin{aligned}
&= \sum_{i=1}^k \sum_{j=0}^{i-1} (-1)^{j+i} \sigma \circ F_k^i \circ F_{k-1}^j + \sum_{i=0}^{k-1} \sum_{j=i}^{k-1} (-1)^{j+i} \sigma \circ F_k^i \circ F_{k-1}^j \\
&= \sum_{i=1}^k \sum_{j=0}^{i-1} (-1)^{j+i} \sigma \circ F_k^j \circ F_{k-1}^{i-1} + \sum_{j=0}^{k-1} \sum_{i=j+1}^k (-1)^{j+i-1} \sigma \circ F_k^j \circ F_{k-1}^{i-1} = 0
\end{aligned}$$

where, for the fourth equality, we have use (i) to get the first sum and the substitution $(i, j) \mapsto (j, i-1)$ to get the second sum. Note that each term in the first sum appears with the opposite sign in the second sum, and we have $0 \leq j < i \leq k$ in both sums. ■

Definition 8.4.7

Let X be a topological space. Two k -cycles $c_1, c_2 \in S_k(X; R)$ are **homologous** if there exists $\tilde{c} \in S_{k+1}(X; R)$ such that $c_1 - c_2 = \partial_{k+1} \tilde{c}$. We write $c_1 \stackrel{h}{\sim} c_2$.

The reader should verify that the homologous relation defines an equivalence relation on $S_k(X; R)$.

Definition 8.4.8

Let X be a topological space. A **k -cycle** in $S_k(X; R)$ is a singular k -chain c such that $\partial_k c = 0$. The set of all k -cycles in $S_k(X; R)$ is a R -module denoted $Z_k(X; R)$.

A **k -boundary** in $S_k(X; R)$ is a singular k -chain c with the property that there exists a singular $(k+1)$ -chain $\tilde{c}$ such that $\partial_{k+1} \tilde{c} = c$. The set of all k -boundaries in $S_k(X; R)$ is a R -module denoted $B_k(X; R)$.

The **singular k^{th} homology module** of X is the R -module defined as $H_k(X; R) = Z_k(X; R)/B_k(X; R)$. The equivalence class of $H_k(X; R)$ associated to $c \in Z_k(X; R)$ is denoted $[c]_X$.

Example 8.4.9

The first example is a simple but important example. Let $X = \{x\}$, a set with a single element. A singular k -simplex $\sigma : \Delta_k \rightarrow X$ is therefore a constant function defined by $\sigma(\mathbf{x}) = x$ for all $\mathbf{x} \in \Delta_k$. Hence

$$\partial_k \sigma = \sum_{j=0}^k (-1)^j \sigma^{(j)} = \begin{cases} 0 & \text{if } k \text{ is odd} \\ \sigma^{(0)} & \text{if } k \text{ is even} \end{cases}$$

because $\sigma^{(j)} : \Delta_{k-1} \rightarrow X$ is the constant function defined by $\sigma^{(j)}(\mathbf{x}) = x$ for all $\mathbf{x} \in \Delta_{k-1}$ for $0 \leq j \leq k$.

It follows that

$$Z_k(X; R) = \begin{cases} S_k(X; R) & \text{if } k \text{ is odd or } k = 0 \\ \{0\} & \text{if } k \text{ is even} \end{cases}$$

and

$$B_k(X; R) = \begin{cases} S_k(X; R) & \text{if } k \text{ is odd} \\ \{0\} & \text{if } k \text{ is even or } k = 0 \end{cases}$$

Therefore

$$H_k(X; R) = \begin{cases} S_0(X; R) & \text{if } k = 0 \\ \{[0]_X\} & \text{if } k > 0 \end{cases}$$

From now on, we adopt the traditional convention of writing $H_k(X; R) = 0$ instead of $H_k(X; R) = \{[0]_X\}$. ♣

Proposition 8.4.10

If the topological space X is the union of distinct (connected) components X_j for $j \in J$, then $H_k(X; R) = \bigoplus_{j \in J} H_k(X_j; R)$.

Proof (Sketch).

If $\sigma \in S_k(X; R)$, then $\sigma(\Delta_k) \subset X_j$ for some $j \in J$ because $\sigma(\Delta_k)$ is connected. Therefore $S_k(X; R) = \bigoplus_{j \in J} S_k(X_j; R)$. Moreover $\partial_q : S_q(X_j; R) \rightarrow S_{q-1}(X_j; R)$ for all $q \in \mathbb{Z}$ and all $j \in J$. Thus $Z_k(X; R) = \bigoplus_{j \in J} Z_k(X_j; R)$ and $B_k(X; R) = \bigoplus_{j \in J} B_k(X_j; R)$. The conclusion of the proof follows from a standard result about the quotient space of direct sums of modules in algebra. ■

Proposition 8.4.11

Let X be a path-connected topological space. Then $H_0(X; R) \cong R$.

Proof.

Let $e_x : \Delta_0 \rightarrow X$ be the singular 0-simplex defined by $e_x(e_0) = x$ for $x \in X$.

i) We define a map $h : S_0(X; R) \rightarrow R$ as it follows. Given $c \in S_0(X; R)$, we have that $c = \sum_{j \in J} a_j e_{x_j}$ where $J \subset \mathbb{N}$ is a finite set, $a_j \in R$ and $x_j \in X$ for all $j \in J$. We

set $\epsilon(c) = \sum_{j \in J} a_j$. We prove that if c_1 and c_2 are two homologous singular 0-chains, then

$\epsilon(c_1) = \epsilon(c_2)$. If c_1 and c_2 are two homologous singular 0-chains, then $c_1 - c_2 = \partial_1 c$ for some $c \in S_1(X; R)$. Hence $\epsilon(c_1) - \epsilon(c_2) = \epsilon(c_1 - c_2) = \epsilon(\partial_1 c)$. To prove that $\epsilon(\partial_1 c) = 0$, it suffices to prove that $\epsilon(\partial_1 \sigma) = 0$ for any singular 1-simplex in X . We effectively have that $\epsilon(\partial_1 \sigma) = \epsilon(\sigma \circ F_1^0 - \sigma \circ F_1^1) = \epsilon(e_{\sigma(e_1)} - e_{\sigma(e_0)}) = 1 - 1 = 0$.

We may therefore define a map $\tilde{\epsilon} : H_0(X; R) \rightarrow R$ as it follows. Given a singular 0-chain $c = \sum_{j \in J} a_j e_{x_j}$ where $J \subset \mathbb{N}$ is a finite set, $a_j \in R$ and $x_j \in X$ for all $j \in J$, we set

$$\tilde{\epsilon}([c]_X) = \sum_{j \in J} a_j.$$

ii) Let c be a singular 0-chain. We have that $c = \sum_{j \in J} a_j e_{x_j}$ where $J \subset \mathbb{N}$ is finite, $a_j \in R$ and $x_j \in X$ for all $j \in J$. Choose $\tilde{x} \in X$. Given any $x \in X$, there exists a path $\tau_x : [0, 1] \rightarrow X$

such that $\tau_x(0) = \tilde{x}$ and $\tau_x(1) = x$ because X is path-connected. Let $\hat{\tau}_x : \Delta_1 \rightarrow X$ be the singular 1-simplex defined by $\hat{\tau}_x((1-t)e_0 + te_1) = \tau_x(t)$ for $t \in [0, 1]$. Since $\partial_1 \hat{\tau}_x = e_x - e_{\tilde{x}}$, we have that $e_x \stackrel{h}{\sim} e_{\tilde{x}}$. Therefore $c \stackrel{h}{\sim} \left(\sum_{j \in J} a_j \right) e_{\tilde{x}}$.

iii) The map $\tilde{\epsilon}$ is a module homomorphism. If $r \in R$ and $c = \sum_{j \in J} a_j e_{x_j}$ for some $a_j \in R$ and $x_j \in X$, then $rc = \sum_{j \in J} ra_j e_{x_j}$. Hence $\tilde{\epsilon}([rc]_X) = \sum_{j \in J} ra_j = r \sum_{j \in J} a_j = r \tilde{\epsilon}([c]_X)$. If $c_1 = \sum_{j \in J} a_j e_{x_j}$ and $c_2 = \sum_{j \in J} b_j e_{x_j}$ for some $a_j, b_j \in R$ and $x_j \in X$, then $c_1 + c_2 = \sum_{j \in J} (a_j + b_j) e_{x_j}$. Therefore $\tilde{\epsilon}([c_1 + c_2]_X) = \sum_{j \in J} (a_j + b_j) = \sum_{j \in J} a_j + \sum_{j \in J} b_j = \tilde{\epsilon}([c_1]_X) + \tilde{\epsilon}([c_2]_X)$.

The map $\tilde{\epsilon}$ is obviously onto. It is also injective. Suppose that $c_1 = \sum_{j \in J} a_j e_{x_j}$ and $c_2 = \sum_{j \in J} b_j e_{x_j}$ for some $a_j, b_j \in R$ and $x_j \in X$. If $\tilde{\epsilon}([c_1]_X) = \tilde{\epsilon}([c_2]_X)$, then $\sum_{j \in J} a_j = \sum_{j \in J} b_j$. Hence, we get from (ii) that $c_1 \stackrel{h}{\sim} \sum_{j \in J} a_j e_{\tilde{x}} = \sum_{j \in J} b_j e_{\tilde{x}} \stackrel{h}{\sim} c_2$. Thus $[c_1]_X = [c_2]_X$. ■

The proof of the next proposition is left to the reader.

Proposition 8.4.12

Let X and Y be two topological spaces, and $f : X \rightarrow Y$ be a continuous function. The map $S_k(f) : S_k(X; R) \rightarrow S_k(Y; R)$ defined by $S_k(f) \left(\sum_{j \in J} a_j \sigma_j \right) = \sum_{j \in J} a_j f \circ \sigma_j$ for is a module homomorphism.

Proposition 8.4.13

Let X, Y and Z be three topological spaces. If $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ are two continuous functions, and $k \in \mathbb{Z}$, then

1. $S_k(g \circ f) = S_k(g) \circ S_k(f)$ and
2. $\partial_k(S_k(f)c) = S_{k-1}(f)(\partial_k c)$ for all $c \in S_k(X; R)$.

Proof.

The proof of (1) is left to the reader. As usual, we only need to prove (2) for singular k -simplices. Let σ be a singular k -simplex, then

$$\begin{aligned} S_{k-1}(f)(\partial_k(\sigma)) &= S_{k-1}(f) \left(\sum_{i=0}^k (-1)^i \sigma \circ F_k^i \right) = \sum_{i=0}^k (-1)^i f \circ \sigma \circ F_k^i \\ &= \partial_k(f \circ \sigma) = \partial_k(S_k(f)(\sigma)) . \end{aligned}$$

■

Corollary 8.4.14

Let X and Y be two topological spaces, and $f : X \rightarrow Y$ be a continuous function. Then $S_k(f)(Z_q(X; R)) \subset Z_k(Y; R)$ and $S_k(f)(B_q(X; R)) \subset B_k(Y; R)$. Moreover, $H_k(f) : H_k(X; R) \rightarrow H_k(Y; R)$ defined by $H_k(f)([c]_X) = [S_k(f)(c)]_Y$ for all $c \in H_k(X; R)$ is a module homomorphism.

Proof.

Since $\partial_k(S_k(f)c) = S_{k-1}(f)(\partial_k c)$ for all $c \in S_k(X; R)$, we have that $S_k(f)(Z_q(X; R)) \subset Z_k(Y; R)$ and $S_k(f)(B_q(X; R)) \subset B_k(Y; R)$.

To prove that $H_k(f)([c]_X)$ is well defined, we need to show that $[S_k(f)(c)]_Y$ is independent of the representative c of the equivalence class $[c]_X$ that is used. Suppose that $\tilde{c} \stackrel{h}{\sim} c$. Then $\tilde{c} = c + \partial_{k+1}b$ for some $b \in S_{k+1}(X; R)$. Hence $S_k(f)(\tilde{c}) = S_k(f)(c) + S_k(f)(\partial_{k+1}b) = S_k(f)(c) + \partial_{k+1}(S_k(f)(b))$ with $S_k(f)(b) \in S_{k+1}(Y; R)$. Thus $[S_k(f)(\tilde{c})]_Y = [S_k(f)(c)]_Y$.

It is easy to verify that $H_k(f)([c_1]_X + [c_2]_X) = H_k(f)([c_1]_X) + H_k(f)([c_2]_X)$ and $H_k(f)(\alpha[c]_X) = \alpha H_k(f)([c]_X)$ for $[c]_X, [c_1]_X, [c_2]_X \in H_k(X; R)$ and $\alpha \in R$. This proves that $H_k(f)$ is a module homomorphism. ■

The following results follow from Proposition 8.4.13.

Corollary 8.4.15

If X, Y and Z are topological spaces and $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ are two continuous functions, then $H_k(g \circ f) = H_k(g) \circ H_k(f)$.

Corollary 8.4.16

Suppose that X and Y are two topological spaces. If X and Y are homeomorphic, then $H_k(X; R) \cong H_k(Y; R)$.

8.4.2 Exact Sequences**Definition 8.4.17**

Let X_j be modules over an integral domain R and $f_j : X_j \rightarrow X_{j+1}$ be continuous functions. We say that the sequence

$$\xrightarrow{f_{j-1}} X_j \xrightarrow{f_j} X_{j+1} \xrightarrow{f_{j+1}} X_{j+2} \xrightarrow{f_{j+2}}$$

is a **(long) exact sequence** if the kernel of a map in the sequence equal the image of the previous map in the sequence.

We often have to deal with the following type of sequences.

Definition 8.4.18

Let X, Y and Z be modules over an integral domain R . Moreover let $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ be continuous functions. We say that the sequence

$$0 \longrightarrow X \xrightarrow{f} Y \xrightarrow{g} Z \longrightarrow 0$$

is a **short exact sequence** if the kernel of a map in the sequence equal the image of the previous map (if any) in the sequence.

By tradition, the expression $0 \rightarrow X$ denotes the map from the set $\{0\}$ to X that maps 0 to the null element in X . Similarly, the expression $Z \rightarrow 0$ denotes the map from Z to $\{0\}$ that map all elements in Z to 0.

Definition 8.4.19

Let X, Y and Z be modules over an integral domain R . We say that a short exact sequence

$$0 \longrightarrow X \xrightarrow{f} Y \xrightarrow{g} Z \longrightarrow 0$$

is **split** if there exists a function $h : Y \rightarrow X$ such that $h \circ f = \text{Id}_X$.

The next result is kind of surprising.

Proposition 8.4.20

Let X, Y and Z be modules over an integral domain R . A short exact sequence

$$0 \longrightarrow X \xrightarrow{f} Y \xrightarrow{g} Z \longrightarrow 0$$

is split if and only if there exists a function $k : Z \rightarrow Y$ such that $g \circ k = \text{Id}_Z$.

Proof.

$\Rightarrow$) Since the short exact sequence is split, there exists $h : Y \rightarrow X$ such that $h \circ f = \text{Id}_X$. To define $k : Z \rightarrow Y$, consider $z \in Z$. Since g is onto, there exists $y \in Y$ such that $g(y) = z$. Let $k(z) = y - f(h(y))$.

We first show that k is well defined. Suppose that $\tilde{y} \in Y$ is such that $g(\tilde{y}) = z$. Then $g(y - \tilde{y}) = g(y) - g(\tilde{y}) = 0$. So $y - \tilde{y} \in \mathcal{N}(g) = \mathcal{R}(f)$. Thus there exists $x \in X$ such that $f(x) = y - \tilde{y}$. We get

$$\begin{aligned} (y - f(h(y))) - (\tilde{y} - f(h(\tilde{y}))) &= y - \tilde{y} - f(h(y - \tilde{y})) \\ &= f(x) - f(h(f(x))) = f(x) - f(x) = 0 \end{aligned}$$

because $h \circ f = \text{Id}_X$.

We also have that $g(k(z)) = g(y - f(h(y))) = g(y) - g(f(h(y))) = g(y) = z$ for all $z \in Z$ because $\mathcal{N}(g) = \mathcal{R}(f)$, where $y \in Y$ is the y in the definition of k at $z \in Z$. Thus $g \circ k = \text{Id}_Z$.

We must not forget to verify that k is a homomorphism of modules. Given $z, \tilde{z} \in Z$. Suppose that $y, \tilde{y}$ are such that $g(y) = z$ and $g(\tilde{y}) = \tilde{z}$. Then $g(y + \tilde{y}) = z + \tilde{z}$ and $k(z + \tilde{z}) = y + \tilde{y} + f(h(y + \tilde{y})) = (y + f(h(y))) + (\tilde{y} + f(h(\tilde{y}))) = k(z) + k(\tilde{z})$. Proceeding similarly, we also get that $k(rz) = rk(z)$ for all $z \in Z$ and $r \in R$.

$\Leftarrow$) We assume that there exists $k : Z \rightarrow Y$ such that $k \circ g = \text{Id}_Z$. To define $h : Y \rightarrow X$, consider $y \in Y$. Since $g(y - k(g(y))) = g(y) - g(k(g(y))) = g(y) - g(y) = 0$ because $k \circ g = \text{Id}_Z$, we have that $y - k(g(y)) \in \mathcal{N}(g)$. Since $\mathcal{R}(f) = \mathcal{N}(g)$, there exists $x \in X$ such that $f(x) = y - k(g(y))$. We set $h(y) = x$. The function h is well defined because f is one-to-one. We also have that $h(f(x)) = x$ for all $x \in X$ because $f(x) - k(g(f(x))) = f(x)$ since $\mathcal{N}(g) = \mathcal{R}(f)$ and f is one-to-one. Thus $h \circ f = \text{Id}_X$. Finally, it is easy to show that h is an homomorphism of modules because f is one. ■

Remark 8.4.21

It should be pointed out that h and k in the previous proposition yields the following short exact sequence

$$0 \longleftarrow X \xleftarrow{h} Y \xleftarrow{k} Z \longleftarrow 0$$

We have that $h \circ f = \text{Id}_X$ ensures that h is onto and $g \circ k = \text{Id}_Z$ ensures that k is one-to-one. We also have that $\mathcal{R}(k) = \mathcal{N}(h)$. To prove that $\mathcal{R}(k) \subset \mathcal{N}(h)$, consider $\tilde{y} = k(z)$ for $z \in Z$. Then $\tilde{y} = y - f(h(y))$ for some $y \in Y$. Hence $h(\tilde{y}) = h(y - f(h(y))) = h(y) - h(f(h(y))) = h(y) - h(y) = 0$ because $h \circ f = \text{Id}_X$. Thus $\tilde{y} \in \mathcal{N}(h)$. To prove that $\mathcal{R}(k) \supset \mathcal{N}(h)$, suppose that $y \in \mathcal{N}(h)$. Then $k(g(y)) = y - f(h(y)) = y$ because $h(y) = 0$. Thus $y \in \mathcal{R}(k)$. ♠

Proposition 8.4.22

A short exact sequence

$$0 \longrightarrow X \xrightarrow{f} Y \xrightarrow{g} Z \longrightarrow 0 \quad (8.4.2)$$

is split if and only if $Y \cong X \oplus Z$.

Proof.

i) Suppose that (8.4.2) is split. We then have that

$$0 \longrightarrow X \begin{array}{c} \xrightarrow{f} \\ \xleftarrow{h} \end{array} Y \begin{array}{c} \xrightarrow{g} \\ \xleftarrow{k} \end{array} Z \longrightarrow 0$$

is exact in both direction.

Given $y \in f(X) \cap k(Z)$, then $y = f(x) = k(z)$ for some $x \in X$ and $z \in Z$. Thus $0 = g(f(x)) = g(y) = g(k(z)) = z$ because $\mathcal{N}(g) = \mathcal{R}(f)$. Hence $y = k(z) = k(0) = 0$. We get that $f(X) \cap k(Z) = \{0\}$.

Given $y \in Y$, it follows from $g(y - k(g(y))) = g(y) - g(k(g(y))) = g(y) - g(y) = 0$ that $y - k(g(y)) = f(x)$ for some $x \in X$ because $\mathcal{N}(g) = \mathcal{R}(f)$. Thus $y = f(x) + k(g(y)) \in f(X) + k(Z)$. We get that $Y \subset f(X) + k(Z)$.

Therefore $Y = f(X) \oplus k(Z)$. Since $X \cong f(X)$ and $Z \cong k(Z)$ because f and k are isomorphism, we have that $Y \cong X \oplus Z$.

ii) Suppose that $Y \cong X \oplus Z$. So there exists an isomorphism $\Phi : Y \rightarrow X \oplus Z$. Let $\tilde{h} : X \oplus Z \rightarrow X$ be the function defined by $\tilde{h}(x, z) = x$ for all $(x, z) \in X \oplus Z$, $\tilde{g} : X \oplus Z \rightarrow Z$ be the function defined by $\tilde{g}(x, z) = z$ for all $(x, z) \in X \oplus Z$, $\tilde{f} : X \rightarrow X \oplus Z$ be the function defined by $\tilde{f}(x) = (x, 0)$ for all $x \in X$ and $\tilde{k} : Z \rightarrow X \oplus Z$ be the function defined by $\tilde{k}(z) = (0, z)$ for all $z \in Z$. Then

$$0 \longrightarrow X \begin{array}{c} \xrightarrow{\tilde{f}} \\ \xleftarrow{\tilde{h}} \end{array} X \oplus Z \begin{array}{c} \xrightarrow{\tilde{g}} \\ \xleftarrow{\tilde{k}} \end{array} Z \longrightarrow 0$$

is exact in both direction. Let $f = \Phi^{-1} \circ \tilde{f}$, $g = \tilde{g} \circ \Phi$ and $h = \tilde{h} \circ \Phi$. We get that

$$0 \longrightarrow X \xrightarrow{f} Y \xrightarrow{g} Z \longrightarrow 0$$

is a short exact sequence and it is split because $h \circ f = (\tilde{h} \circ \Phi) \circ (\Phi^{-1} \circ \tilde{f}) = \tilde{h} \circ \tilde{f} = \text{Id}_X$. ■

8.4.3 Chain Complexes

Definition 8.4.23

Let R be an integral domain. A **chain complex** is a sequence $\mathcal{C} = \{(C_k, \partial_k)\}_{k \in \mathbb{Z}}$ where the sets C_k are R -modules and the maps $\partial_k : C_k \rightarrow C_{k-1}$ are homomorphism such that $\partial_k \circ \partial_{k+1} = 0$ for $k \in \mathbb{Z}$.

A diagram representing a chain complex is given below.

$$\xrightarrow{\partial_{k+2}} C_{k+1} \xrightarrow{\partial_{k+1}} C_k \xrightarrow{\partial_k} C_{k-1} \xrightarrow{\partial_{k-1}}$$

Remark 8.4.24

Unless otherwise stated, we always assume that $C_k = \{0\}$ for $k < 0$. This is sufficient for our applications. ♠

We have already worked with chain complexes before. One such example is given by the sequence of pairs $(C_k, \partial_k) = (S_k(X; R), \partial_k)$ for $k \in \mathbb{Z}$ defined at the beginning of the previous section.

Definition 8.4.25

A **chain map** between two chain complexes given by $\mathcal{C} = \{(C_k, \partial_k)\}_{k \in \mathbb{Z}}$ and $\tilde{\mathcal{C}} = \{(\tilde{C}_k, \tilde{\partial}_k)\}_{k \in \mathbb{Z}}$ is a sequence $\mathcal{F} = \{f_k\}_{k \in \mathbb{Z}}$ where the maps $f_k : C_k \rightarrow \tilde{C}_k$ are homomorphism such that $\tilde{\partial}_k \circ f_k = f_{k-1} \circ \partial_k$ for $k \in \mathbb{Z}$.

The previous definition yields the following commutative diagram.

$$\begin{array}{ccccccc}
 & \xrightarrow{\partial_{k+2}} & C_{k+1} & \xrightarrow{\partial_{k+1}} & C_k & \xrightarrow{\partial_k} & C_{k-1} \xrightarrow{\partial_{k-1}} \\
 & & \downarrow f_{k+1} & & \downarrow f_k & & \downarrow f_{k-1} \\
 & \xrightarrow{\tilde{\partial}_{k+2}} & \tilde{C}_{k+1} & \xrightarrow{\tilde{\partial}_{k+1}} & \tilde{C}_k & \xrightarrow{\tilde{\partial}_k} & \tilde{C}_{k-1} \xrightarrow{\tilde{\partial}_{k-1}}
 \end{array}$$

As we did before, we may define k -cycles, k -boundaries and k -singular homology modules.

Definition 8.4.26

Consider a chain complex given by $\mathcal{C} = \{(C_k, \partial_k)\}_{k \in \mathbb{Z}}$. A **k-cycle** in $\mathcal{C}$ is an element $c \in C_k$ such that $\partial_k(c) = 0$. The set of all k -cycles in C_k is a R -module denoted $Z_k(\mathcal{C})$. A **k-boundary** in $\mathcal{C}$ is an element $c \in C_k$ with the property that there exists an element $\tilde{c} \in C_{k+1}$ such that $\partial_{k+1}(\tilde{c}) = c$. The set of all k -boundaries in C_k is a R -module denoted $B_k(\mathcal{C})$.

The **singular kth homology module** of $\mathcal{C}$ is the quotient space $H_k(\mathcal{C}) = Z_k(\mathcal{C})/B_k(\mathcal{C})$. The equivalence class of $H_k(\mathcal{C})$ associated to $c \in Z_k(\mathcal{C})$ is denoted $[c]$.

It follows from the definition that a chain map $\mathcal{F} = \{f_k\}_{k \in \mathbb{Z}}$ between two chain complexes $\mathcal{C} = \{(C_k, \partial_k)\}_{k \in \mathbb{Z}}$ and $\tilde{\mathcal{C}} = \{(\tilde{C}_k, \tilde{\partial}_k)\}_{k \in \mathbb{Z}}$ sends k -cycles in $\mathcal{C}$ to k -cycles in $\tilde{\mathcal{C}}$, and k -boundaries in $\mathcal{C}$ to k -boundaries in $\tilde{\mathcal{C}}$. Hence, as we did in Corollary 8.4.14, we may define $H_k(\mathcal{F}) : H_k(\mathcal{C}) \rightarrow H_k(\tilde{\mathcal{C}})$ by $H_k(\mathcal{F})([c]) = [f_k(c)]$ for all $c \in Z_k(\mathcal{C})$.

The next result is traditionally called a lemma but it really should be a theorem because it is used in very many proofs in algebraic topology.

Lemma 8.4.27 (Zig-Zag Lemma)

Suppose that $\mathcal{F} = \{f_k\}_{k \in \mathbb{Z}}$ is a chain map between the chain complexes $\mathcal{A} = \{(A_k, \partial_k^{\mathcal{A}})\}_{k \in \mathbb{Z}}$ and $\mathcal{B} = \{(B_k, \partial_k^{\mathcal{B}})\}_{k \in \mathbb{Z}}$, and $\mathcal{G} = \{g_k\}_{k \in \mathbb{Z}}$ is a chain map between the chain complexes $\mathcal{B} = \{(B_k, \partial_k^{\mathcal{B}})\}_{k \in \mathbb{Z}}$ and $\mathcal{C} = \{(C_k, \partial_k^{\mathcal{C}})\}_{k \in \mathbb{Z}}$. If

$$0 \longrightarrow A_k \xrightarrow{f_k} B_k \xrightarrow{g_k} C_k \longrightarrow 0$$

is an exact sequence for all k , then there exists a long exact sequence

$$\xrightarrow{\tilde{\partial}_{k+1}} H_k(\mathcal{A}) \xrightarrow{H_k(\mathcal{F})} H_k(\mathcal{B}) \xrightarrow{H_k(\mathcal{G})} H_k(\mathcal{C}) \xrightarrow{\tilde{\partial}_k} H_{k-1}(\mathcal{A}) \xrightarrow{H_{k-1}(\mathcal{F})}$$

where $\tilde{\partial}_k$ is induced by $\partial_k^{\mathcal{B}}$.

Proof.

The proof is based on the following commutative diagram.

$$\begin{array}{ccccccc}
 0 & \longrightarrow & A_{k+1} & \xrightarrow{f_{k+1}} & B_{k+1} & \xrightarrow{g_{k+1}} & C_{k+1} \longrightarrow 0 \\
 & & \downarrow \partial_{k+1}^A & & \downarrow \partial_{k+1}^B & & \downarrow \partial_{k+1}^C \\
 0 & \longrightarrow & A_k & \xrightarrow{f_k} & B_k & \xrightarrow{g_k} & C_k \longrightarrow 0 \\
 & & \downarrow \partial_k^A & & \downarrow \partial_k^B & & \downarrow \partial_k^C \\
 0 & \longrightarrow & A_{k-1} & \xrightarrow{f_{k-1}} & B_{k-1} & \xrightarrow{g_{k-1}} & C_{k-1} \longrightarrow 0
 \end{array}$$

i) Our first task is to define $\tilde{\partial}_k$ on $H_k(\mathcal{C})$. Given $[c] \in H_k(\mathcal{C})$, there exists $b \in B_k$ such that $g_k(b) = c$ because g_k is onto. We have that $\partial_k^B(b) \in B_{k-1} \cap \mathcal{R}(f_{k-1})$ because $\partial_k^B(b) \in \mathcal{N}(g_{k-1}) = \mathcal{R}(f_{k-1})$ since $g_{k-1}(\partial_k^B(b)) = \partial_k^C(g_k(b)) = \partial_k^C(c) = 0$ for $[c] \in H_k(\mathcal{C})$. There exists $a \in A_{k-1}$ such that $f_{k-1}(a) = \partial_k^B(b)$. In fact, a is unique because f_{k-1} is one-to-one.

We set $\tilde{\partial}_k([c]) = [a] \in H_{k-1}(\mathcal{A})$.

We have to show that $\tilde{\partial}_k([c])$ is well defined; namely, that it is independent of the representative of the equivalence class $[c] \in H_k(\mathcal{C})$ that we choose to find $a \in A_{k-1}$. Suppose that $\tilde{c} \in Z_k(\mathcal{C})$ is such that $[\tilde{c}] = [c]$. As before, there exists $\tilde{b} \in B_k$ such that $g_k(\tilde{b}) = \tilde{c}$, and $\tilde{a} \in A_{k-1}$ such that $f_{k-1}(\tilde{a}) = \partial_k^B(\tilde{b})$.

Since $[c] = [\tilde{c}]$, we have that $c - \tilde{c} = \partial_{k+1}^C(\check{c})$ for some $\check{c} \in C_{k+1}$. Choose $\check{b} \in B_{k+1}$ such that $g_{k+1}(\check{b}) = \check{c}$. This is possible because g_{k+1} maps B_{k+1} onto C_{k+1} . Then

$$g_k(b - \tilde{b} - \partial_{k+1}^B(\check{b})) = g_k(b) - g_k(\tilde{b}) - \partial_{k+1}^C(g_{k+1}(\check{b})) = c - \tilde{c} - \partial_{k+1}^C(\check{c}) = 0.$$

Since $b - \tilde{b} - \partial_{k+1}^B(\check{b}) \in \mathcal{N}(g_k) = \mathcal{R}(f_k)$, there exists $\check{a} \in A_k$ such that $f_k(\check{a}) = b - \tilde{b} - \partial_{k+1}^B(\check{b})$. Moreover, since

$$\begin{aligned}
 f_{k-1}(\partial_k^A(\check{a})) &= \partial_k^B(f_k(\check{a})) = \partial_k^B(b - \tilde{b} - \partial_{k+1}^B(\check{b})) = \partial_k^B(b) - \partial_k^B(\tilde{b}) \\
 &= f_{k-1}(a) - f_{k-1}(\tilde{a}) = f_{k-1}(a - \tilde{a})
 \end{aligned}$$

and f_{k-1} is one-to-one, we get that $a - \tilde{a} = \partial_k^A(\check{a})$ with $\check{a} \in A_k$. Therefore $[a] = [\tilde{a}]$.

To prove that $\tilde{\partial}_k$ is an homomorphism, suppose that $[c], [\tilde{c}] \in H_k(\mathcal{C})$. We have by definition of $\tilde{\partial}_k$ that $\tilde{\partial}_k([c]) = [a]$ and $\tilde{\partial}_k([\tilde{c}]) = [\tilde{a}]$ where $f_{k-1}(a) = \partial_k^B(b)$ with $g_k(b) = c$ for some $b \in B_k$ and $f_{k-1}(\tilde{a}) = \partial_k^B(\tilde{b})$ with $g_k(\tilde{b}) = \tilde{c}$ for some $\tilde{b} \in B_k$. Since $f_{k-1}(a + \tilde{a}) = \partial_k^B(b + \tilde{b})$ with $g_k(b + \tilde{b}) = c + \tilde{c}$, we get

$$\tilde{\partial}_k([c] + [\tilde{c}]) = \tilde{\partial}_k([c + \tilde{c}]) = [a + \tilde{a}] = [a] + [\tilde{a}] = \tilde{\partial}_k([c]) + \tilde{\partial}_k([\tilde{c}]).$$

Similarly, $\tilde{\partial}_k(r[c]) = r\tilde{\partial}_k([c])$ for all $r \in R$.

ii) We prove that $\mathcal{R}(H_k(\mathcal{F})) = \mathcal{N}(H_k(\mathcal{G}))$. Since $g_k \circ f_k = 0$ on A_k , we get that $(H_k(\mathcal{G}) \circ H_k(\mathcal{F}))([c]) = [(g_k \circ f_k)(c)] = [0]$ for all $[c] \in H_k(\mathcal{A})$. Thus $\mathcal{R}(H_k(\mathcal{F})) \subset \mathcal{N}(H_k(\mathcal{G}))$.

Suppose that $[b] \in \mathcal{N}(H_k(\mathcal{G}))$. We therefore have that $H_k(\mathcal{G})([b]) = [g_k(b)] = [0]$. Thus $g_k(b) = \partial_{k+1}^C(c)$ for some $c \in C_{k+1}$. Since g_{k+1} is onto, there exists $\tilde{b} \in B_{k+1}$ such that $g_{k+1}(\tilde{b}) = c$. We get

$$g_k(b - \partial_{k+1}^B(\tilde{b})) = g_k(b) - \partial_{k+1}^C(g_{k+1}(\tilde{b})) = g_k(b) - \partial_{k+1}^C(c) = 0.$$

Since $\mathcal{R}(f_k) = \mathcal{N}(g_k)$, there exists $a \in A_k$ such that $f_k(a) = b - \partial_{k+1}^{\mathcal{B}}(\tilde{b})$.

We have that $a \in Z_k(\mathcal{A})$ because $b \in Z_k(\mathcal{B})$ implies that

$$f_{k-1}(\partial_k^{\mathcal{A}}(a)) = \partial_k^{\mathcal{B}}(f_k(a)) = \partial_k^{\mathcal{B}}(b - \partial_{k+1}^{\mathcal{B}}(\tilde{b})) = \partial_k^{\mathcal{B}}(b) = 0$$

and f_{k-1} being one-to-one implies that $\partial_k^{\mathcal{A}}(a) = 0$. Thus $[a] \in H_k(\mathcal{A})$. Moreover, $H_k(\mathcal{F})([a]) = [f_k(a)] = [b - \partial_{k+1}^{\mathcal{B}}(\tilde{b})] = [b]$. Thus $[b] \in \mathcal{R}(H_k(\mathcal{F}))$. This proves that $\mathcal{N}(H_k(\mathcal{G})) \subset \mathcal{R}(H_k(\mathcal{F}))$.

iii) We prove that $\mathcal{R}(H_k(\mathcal{G})) = \mathcal{N}(\tilde{\partial}_k)$. Given $[c] \in H_k(\mathcal{C})$, we have that $\tilde{\partial}_k([c]) = [a]$ where $f_{k-1}(a) = \partial_k^{\mathcal{B}}(b)$ with $g_k(b) = c$ for some $b \in B_k$.

If $[c] \in \mathcal{R}(H_k(\mathcal{G}))$, then $[c] = H_k(\mathcal{G})([b]) = [g_k(b)]$ for some $[b] \in H_k(\mathcal{B})$. In particular, $b \in Z_k(\mathcal{B})$. Thus $f_{k-1}(a) = \partial_k^{\mathcal{B}}(b) = 0$ implies that $a = 0$ because f_{k-1} is one-to-one. Thus $\tilde{\partial}_k([c]) = [0]$; namely, $[c] \in \mathcal{N}(\tilde{\partial}_k)$. We have proved that $\mathcal{R}(H_k(\mathcal{G})) \subset \mathcal{N}(\tilde{\partial}_k)$.

Suppose that $[c] \in \mathcal{N}(\tilde{\partial}_k)$. Then $a = \partial_k^{\mathcal{A}}(\tilde{a})$ for some $\tilde{a} \in A_k$. We have that $b - f_k(\tilde{a}) \in Z_k(\mathcal{B})$ because

$$\partial_k^{\mathcal{B}}(b - f_k(\tilde{a})) = \partial_k^{\mathcal{B}}(b) - f_{k-1}(\partial_k^{\mathcal{A}}(\tilde{a})) = \partial_k^{\mathcal{B}}(b) - f_{k-1}(a) = 0$$

and $[c] = H_k(\mathcal{G})([b - f_k(\tilde{a})])$ because

$$g_k(b - f_k(\tilde{a})) = g_k(b) - g_k(f_k(\tilde{a})) = g_k(b) - 0 = c.$$

Thus $[c] \in \mathcal{R}(H_k(\mathcal{G}))$. This proves that $\mathcal{N}(\tilde{\partial}_k) \subset \mathcal{R}(H_k(\mathcal{G}))$.

iv) We prove that $\mathcal{R}(\tilde{\partial}_k) = \mathcal{N}(H_{k-1}(\mathcal{F}))$. If $[a] \in \mathcal{R}(\tilde{\partial}_k)$, then $f_{k-1}(a) = \partial_k^{\mathcal{B}}(b)$ for some $b \in B_k$. Hence $H_{k-1}(\mathcal{F})([a]) = [f_{k-1}(a)] = [\partial_k^{\mathcal{B}}(b)] = [0]$. Thus $[a] \in \mathcal{N}(H_{k-1}(\mathcal{F}))$. Hence $\mathcal{R}(\tilde{\partial}_k) \subset \mathcal{N}(H_{k-1}(\mathcal{F}))$.

Suppose that $[a] \in \mathcal{N}(H_{k-1}(\mathcal{F}))$. Since $H_{k-1}(\mathcal{F})([a]) = [f_{k-1}(a)] = [0]$, we get that $f_{k-1}(a) = \partial_k^{\mathcal{B}}(b)$ for some $b \in B_k$. Let $c = g_k(b)$. We have that $c \in Z_k(\mathcal{C})$ because

$$\partial_k^{\mathcal{C}}(c) = \partial_k^{\mathcal{C}}(g_k(b)) = g_{k-1}(\partial_k^{\mathcal{B}}(b)) = g_{k-1}(f_{k-1}(a)) = 0.$$

We also have by definition of $\tilde{\partial}_k$ that $\tilde{\partial}_k([c]) = [a]$. Thus $[a] \in \mathcal{R}(\tilde{\partial}_k)$. This proves that $\mathcal{N}(H_{k-1}(\mathcal{F})) \subset \mathcal{R}(\tilde{\partial}_k)$. ■

Definition 8.4.28

Let $\mathcal{F} = \{f_k\}_{k \in \mathbb{Z}}$ and $\mathcal{G} = \{g_k\}_{k \in \mathbb{Z}}$ be two chain maps between the chain complexes $\mathcal{C} = \{(C_k, \partial_k)\}_{k \in \mathbb{Z}}$ and $\tilde{\mathcal{C}} = \{(\tilde{C}_k, \tilde{\partial}_k)\}_{k \in \mathbb{Z}}$. A **chain homotopy** between $\mathcal{F}$ and $\mathcal{G}$ is a sequence $\mathcal{D} = \{D_k\}_{k \in \mathbb{Z}}$ where the maps $D_k : C_k \rightarrow \tilde{C}_{k+1}$ are homomorphism such that $\tilde{\partial}_{k+1} \circ D_k + D_{k-1} \circ \partial_k = f_k - g_k$ for $k \in \mathbb{Z}$.

A commutative diagram representing a chain homotopy is given below.

$$\begin{array}{ccccccc}
 & \xrightarrow{\partial_{k+2}} & C_{k+1} & \xrightarrow{\partial_{k+1}} & C_k & \xrightarrow{\partial_k} & C_{k-1} & \xrightarrow{\partial_{k-1}} & \\
 & \searrow & \downarrow g_{k+1} & \swarrow & \downarrow g_k & \swarrow & \downarrow g_{k-1} & \swarrow & \\
 & D_{k+1} & & D_k & & D_{k-1} & & D_{k-2} & \\
 & \swarrow & \downarrow f_{k+1} & \swarrow & \downarrow f_k & \swarrow & \downarrow f_{k-1} & \swarrow & \\
 \tilde{C}_{k+2} & \xrightarrow{\tilde{\partial}_{k+2}} & \tilde{C}_{k+1} & \xrightarrow{\tilde{\partial}_{k+1}} & \tilde{C}_k & \xrightarrow{\tilde{\partial}_k} & \tilde{C}_{k-1} & \xrightarrow{\tilde{\partial}_{k-1}} &
 \end{array}$$

Proposition 8.4.29

If we have a chain homotopy as in Definition 8.4.28, then $H_k(\mathcal{F}) = H_k(\mathcal{G})$ for all k .

Proof.

Given $[c] \in H_k(C)$. Then $\partial_k(c) = 0$. It follows from $\tilde{\partial}_{k+1}(D_k(c)) + D_{k-1}(\partial_k(c)) = f_k(c) - g_k(c)$ that $f_k(c) - g_k(c) = \tilde{\partial}_{k+1}(D_k(c))$ with $D_k(c) \in \tilde{C}_{k+1}$. Thus $H_k(\mathcal{F})([c]) = [f_k(c)] = [g_k(c)] = H_k(\mathcal{G})([c])$. ■

Definition 8.4.30

Let $\mathcal{C} = \{(C_k, \partial_k)\}_{k \in \mathbb{Z}}$ be a chain complex. A surjective linear operator $\epsilon : C_0 \rightarrow R$ such that $\epsilon \circ \partial_1 = 0$ is called an **augmentation** of the chain complex $\mathcal{C}$.

Suppose that ϵ is an augmentation of the $\mathcal{C} = \{(C_k, \partial_k)\}_{k \in \mathbb{Z}}$. If we set

$$C_k^\# = \begin{cases} C_k & \text{if } k \neq -1 \\ R & \text{if } k = -1 \end{cases} \quad \text{and} \quad \partial_k^\# = \begin{cases} \partial_k & \text{if } k \neq 0 \\ \epsilon & \text{if } k = 0 \end{cases}$$

then we get a new chain complex $\mathcal{C}^\# = \{(C_k^\#, \partial_k^\#)\}_{k \in \mathbb{Z}}$.

Definition 8.4.31

The chain complex $\mathcal{C}^\# = \{(C_k^\#, \partial_k^\#)\}_{k \in \mathbb{Z}}$ defined above is called the **reduced chain complex** of the chain complex $\mathcal{C}$ with augmentation ϵ ⁷.

Suppose that $\tilde{\mathcal{C}} = \{(\tilde{C}_k, \tilde{\partial}_k)\}_{k \in \mathbb{Z}}$ is another chain complex and $\tilde{\mathcal{C}}^\# = \{(\tilde{C}_k^\#, \tilde{\partial}_k^\#)\}_{k \in \mathbb{Z}}$ is a reduced chain complex of $\tilde{\mathcal{C}}$ with augmentation $\tilde{\epsilon}$. Moreover, suppose that $\mathcal{F} = \{f_k\}_{k \in \mathbb{Z}}$ is a chain map between $\mathcal{C}$ and $\tilde{\mathcal{C}}$.

⁷We have not found any reason why an “augmentation” of a chain complex leads to a “reduced” chain complex. It is perhaps due to Proposition 8.4.45. The choice of terminology in mathematics is not always obvious or even logic.

Definition 8.4.32

We say that the chain map $\mathcal{F}$ defined above is **augmentation-preserving** if $\tilde{\epsilon} \circ f_0 = \epsilon$.

If we set $f_{-1} = \text{Id}_R$, then $\mathcal{F} = \{f_k\}_{k \in \mathbb{Z}}$ becomes a chain map between $\mathcal{C}^\sharp$ and $\tilde{\mathcal{C}}^\sharp$.

Proposition 8.4.33

Let $\mathcal{C}^\sharp = \{(C_k^\sharp, \partial_k^\sharp)\}_{k \in \mathbb{Z}}$ be a reduced chain complex of the chain complex $\mathcal{C} = \{(C_k, \partial_k)\}_{k \in \mathbb{Z}}$ with augmentation ϵ , and $\tilde{\mathcal{C}}^\sharp = \{(\tilde{C}_k^\sharp, \tilde{\partial}_k^\sharp)\}_{k \in \mathbb{Z}}$ be a reduced chain complex of the chain complex $\tilde{\mathcal{C}} = \{(\tilde{C}_k, \tilde{\partial}_k)\}_{k \in \mathbb{Z}}$ with augmentation $\tilde{\epsilon}$. Suppose that $\mathcal{F} = \{f_k\}_{k \in \mathbb{Z}}$ is an augmentation-preserving chain map between $\mathcal{C}$ and $\tilde{\mathcal{C}}$. Then $H_k(\mathcal{F}) : H_k(\mathcal{C}^\sharp) \rightarrow H_k(\tilde{\mathcal{C}}^\sharp)$ is an isomorphism if and only if $H_k(\mathcal{F}) : H_k(\mathcal{C}) \rightarrow H_k(\tilde{\mathcal{C}})$ is an isomorphism.

Proof.

The result is obvious for $k \neq 0$ because $H_k(\mathcal{C}^\sharp) = H_k(\mathcal{C})$ and $H_k(\tilde{\mathcal{C}}^\sharp) = H_k(\tilde{\mathcal{C}})$ for $k \neq 0$.

For $k = 0$, we have the following commutative diagram

$$\begin{array}{ccccccc} 0 & \longrightarrow & \mathcal{N}(\epsilon) & \xrightarrow{\iota} & C_0 & \begin{array}{c} \xrightarrow{\epsilon} \\ \xleftarrow{g} \end{array} & R \longrightarrow 0 \\ & & \downarrow f_0 & & \downarrow f_0 & & \downarrow \text{Id}_R \\ 0 & \longrightarrow & \mathcal{N}(\tilde{\epsilon}) & \xrightarrow{\iota} & \tilde{C}_0 & \begin{array}{c} \xrightarrow{\tilde{\epsilon}} \\ \xleftarrow{\tilde{g}} \end{array} & R \longrightarrow 0 \end{array} \quad (8.4.3)$$

where the two horizontal sequence are exact. The map ι is the inclusion map and $g : R \rightarrow C_0$ is a map such that $\epsilon \circ g = \text{Id}_R$. A map g exists because ϵ is onto. Let $\tilde{g} = f_0 \circ g \circ \text{Id}_R$.

It follows from Proposition 8.4.20 that both horizontal sequence are split. Hence, we get from Proposition 8.4.22 that $C_0 = \mathcal{N}(\epsilon) \oplus \mathcal{R}(g)$ and $\tilde{C}_0 = \mathcal{N}(\tilde{\epsilon}) \oplus \mathcal{R}(\tilde{g})$.

Since $C_0 = \mathcal{N}(\partial_0)$ because we assume that $C_{-1} = \{0\}$, we have

$$H_0(\mathcal{C}) = C_0/\mathcal{R}(\partial_1) = (\mathcal{N}(\epsilon) \oplus \mathcal{R}(g))/\mathcal{R}(\partial_1) \cong (\mathcal{N}(\epsilon)/\mathcal{R}(\partial_1)) \oplus \mathcal{R}(g)$$

because $\mathcal{R}(\partial_1) \subset \mathcal{N}(\epsilon)$ and $\mathcal{N}(\epsilon) \cap \mathcal{R}(g) = \{0\}$. More precisely, we have that $\mathcal{R}(g)/\mathcal{R}(\partial_1) \cong \mathcal{R}(g)$.

Similarly, since $\tilde{C}_0 = \mathcal{N}(\tilde{\partial}_0)$ because we assume that $\tilde{C}_{-1} = \{0\}$, we have

$$H_0(\tilde{\mathcal{C}}) = \tilde{C}_0/\mathcal{R}(\tilde{\partial}_1) = (\mathcal{N}(\tilde{\epsilon}) \oplus \mathcal{R}(\tilde{g}))/\mathcal{R}(\tilde{\partial}_1) \cong (\mathcal{N}(\tilde{\epsilon})/\mathcal{R}(\tilde{\partial}_1)) \oplus \mathcal{R}(\tilde{g})$$

because $\mathcal{R}(\tilde{\partial}_1) \subset \mathcal{N}(\tilde{\epsilon})$ and $\mathcal{N}(\tilde{\epsilon}) \cap \mathcal{R}(\tilde{g}) = \{0\}$. Again, we have that $\mathcal{R}(\tilde{g})/\mathcal{R}(\tilde{\partial}_1) \cong \mathcal{R}(\tilde{g})$.

Using the notation of the reduced chain complexes, we get

$$H_0(\mathcal{C}) \cong (\mathcal{N}(\partial_0^\sharp)/\mathcal{R}(\partial_1^\sharp)) \oplus \mathcal{R}(g) = H_0(\mathcal{C}^\sharp) \oplus \mathcal{R}(g) \quad (8.4.4)$$

and

$$H_0(\tilde{\mathcal{C}}) \cong (\mathcal{N}(\tilde{\partial}_0^\sharp)/\mathcal{R}(\tilde{\partial}_1^\sharp)) \oplus \mathcal{R}(\tilde{g}) \cong H_0(\tilde{\mathcal{C}}^\sharp) \oplus \mathcal{R}(\tilde{g}). \quad (8.4.5)$$

Since $\mathcal{R}(g) \cong R$ and $\mathcal{R}(\tilde{g}) \cong R$ because the two horizontal sequences in (8.4.3) are split, we get that $H_0(\mathcal{C}^\sharp) \cong H_0(\tilde{\mathcal{C}}^\sharp)$ if and only if $H_0(\mathcal{C}) \cong H_0(\tilde{\mathcal{C}})$.

We have that $f_0 : \mathcal{R}(g) \rightarrow \mathcal{R}(\tilde{g})$ is an isomorphism. Suppose that $f_0(a) = f_0(b)$ with $a, b \in \mathcal{R}(g)$. Then $\epsilon(a) = \tilde{\epsilon}(f_0(a)) = \tilde{\epsilon}(f_0(b)) = \epsilon(b)$. Thus $a - b \in \mathcal{N}(\epsilon) \cap \mathcal{R}(g) = \{0\}$. Hence $a = b$. This proves that $f_0 : \mathcal{R}(g) \rightarrow \mathcal{R}(\tilde{g})$ is one-to-one. To prove that it is onto, consider $b \in \mathcal{R}(\tilde{g})$. Therefore there exists $r \in R$ such that $b = \tilde{g}(r) = f_0(g(r))$ and $g(r) \in \mathcal{R}(g)$.

It then follows from (8.4.4) and (8.4.5) that $H_0(\mathcal{F}) : H_0(\mathcal{C}) \rightarrow H_0(\tilde{\mathcal{C}})$ is an isomorphism if and only if $H_0(\mathcal{F}) : H_0(\mathcal{C}^\sharp) \rightarrow H_0(\tilde{\mathcal{C}}^\sharp)$ is an isomorphism because $H_0(\mathcal{F}) : \mathcal{R}(g)/\mathcal{R}(\partial_1) \cong \mathcal{R}(g) \rightarrow \mathcal{R}(\tilde{g})/\mathcal{R}(\tilde{\partial}_1) \cong \mathcal{R}(\tilde{g})$ is an isomorphism. ■

Definition 8.4.34

A chain complex $\{(C_k, \partial_k)\}_{k \in \mathbb{Z}}$ is **acyclic** if $H_k(\mathcal{C}) = 0$ for all k .

Definition 8.4.35

A topological space X is **aspherical** if every continuous map $f : S^k \rightarrow X$ can be extended to a continuous map $F : B_{k+1} \rightarrow X$ where $B_{k+1} = \overline{B_1(\mathbf{0})} \subset \mathbb{R}^{k+1}$ is the closed unit ball centred at the origin.

Remark 8.4.36

1. Convex subsets $X \subset \mathbb{R}^n$ are aspherical. Suppose that $f : S^k \rightarrow X$ is a continuous map. Choose $y \in X$. Since $\overline{B_1(\mathbf{0})} \cong (S^k \times [0, 1]) / (S^k \times \{0\})$, we define $F : \overline{B_1(\mathbf{0})} \rightarrow X$ as it follows. $F([x, t]) = (1 - t)y + tf(x)$ for all $[x, t] \in \overline{B_1(\mathbf{0})}$. Note that $F([x, t]) \in X$ for all $x \in X$ and $t \in [0, 1]$ because X is convex. We also have that $F([x, 1]) = f(x)$ for all $x \in S^k$. For the readers who may be wondering why we had to quotient $S^k \times [0, 1]$ by $S^k \times \{0\}$, they should remember that we always have this singularity at the origin when using polar coordinates. The quotient by $S^k \times \{0\}$ eliminates this problem.
2. Since $\Delta_k \times [0, 1]$ is convex, it is aspherical.



Theorem 8.4.37

Suppose that X is an aspherical space and that R is an integral domain. Then a reduced chain complex $\mathcal{C}^\sharp = \{(S_k^\sharp(X; R), \partial_k^\sharp)\}_{k \in \mathbb{Z}}$ of the chain complex $\mathcal{C} = \{(S_k(X; R), \partial_k)\}_{k \in \mathbb{Z}}$ with augmentation ϵ is acyclic.

Proof.

The proof is in two steps.

I) Define the linear map $\eta : R \rightarrow S_0(X; R)$ by $\eta(s) = s e_x$ for some $x \in X$ such that $\epsilon(e_x) = 1$. Recall that $e_x : \Delta_0 \rightarrow X$ is the singular 0-simplex defined by $e_x(e_0) = x$. Hence, η is a right inverse of ϵ ; namely, $\epsilon \circ \eta = \text{Id}_R$.

Let $\mathcal{F} = \{f_k\}_{k \in \mathbb{Z}}$ be the chain map between $\mathcal{C}$ and itself defined by $f_0 = \eta \circ \epsilon$ and $f_k = 0$ otherwise. Moreover, Let $\mathcal{G} = \{g_k\}_{k \in \mathbb{Z}}$ be another chain map between $\mathcal{C}$ and itself defined by $g_k = \text{Id}_{S_k(X; R)}$ for $k \geq 0$ and $g_k = 0$ otherwise.

We prove that if there is a chain homotopy $\mathcal{D} = \{D_k\}_{k \in \mathbb{Z}}$ between $\mathcal{F}$ and $\mathcal{G}$, then $\mathcal{C}^\sharp$ is acyclic. We have that $D_k : S_k(X; R) \rightarrow S_{k+1}(X; R)$ are homomorphism such that

$$\partial_{k+1} \circ D_k + D_{k-1} \circ \partial_k = g_k - f_k \quad (8.4.6)$$

for $k \in \mathbb{Z}$.

Suppose that $\sigma \in Z_k(S; R)$. It follows from (8.4.6) that

$$\partial_{k+1}(D_k(\sigma)) = \begin{cases} \sigma & \text{if } k \neq 0 \\ \sigma - \eta(\epsilon(\sigma)) & \text{if } k = 0 \end{cases}$$

We have that $\sigma \in B_k(S; R)$ for $k > 0$. Thus $[\sigma] = 0$. Since $\sigma \in Z_k(S; R)$ is arbitrary, we get that $H_k^\sharp(S; R) = H_k(S; R) = 0$ for $k > 0$. If $k = 0$, then we have

$$\sigma - \eta(\epsilon(\sigma)) = \partial_1(D_0(\sigma)). \quad (8.4.7)$$

By definition of the augmentation ϵ , we have that $\mathcal{N}(\epsilon) = Z_0^\sharp(S; R)$. We need to prove that $\mathcal{N}(\epsilon) \subset \mathcal{R}(\partial_1) = B_0^\sharp(S; R)$ to obtain that $H_0^\sharp(S; R) = 0$. Consider $\sigma \in \mathcal{N}(\epsilon)$. It follows from (8.4.7) that $\sigma = \partial_1(D_0(\sigma))$. Thus $\sigma \in \mathcal{R}(\partial_1) = B_0^\sharp(S; R)$. Hence $[\sigma] = 0 \in H_0^\sharp(S; R)$. Since $\sigma \in \mathcal{N}(\epsilon)$ is arbitrary, we get that $H_0^\sharp(S; R) = 0$.

II) We prove that there is a chain homotopy $\mathcal{D} = \{D_k\}_{k \in \mathbb{Z}}$ between the chain maps $\mathcal{F}$ and $\mathcal{G}$ defined in (i).

Let $\dot{\Delta}_{k+1} = \bigcup_{i=0}^{k+1} F_{k+1}^i(\Delta_k)$. It is the topological boundary of Δ_{k+1} .

Given a singular k -simplex, we proceed by induction to define $D_k : S_k(X; R) \rightarrow S_{k+1}(X; R)$. Obviously, we may set $D_k = 0$ for $k < 0$.

II.i) Given a singular 0-form σ in X , we define $\beta_\sigma : \dot{\Delta}_1 \rightarrow X$ such that $\beta_\sigma \circ F_1^0 = \sigma$ and $\beta_\sigma \circ F_1^1 = (\eta \circ \epsilon)(\sigma)$. Since $\dot{\Delta}_1 = \{\mathbf{e}_0, \mathbf{e}_1\}$, two isolated points, we have that β_σ is obviously continuous on $\dot{\Delta}_0$. Thus, the map β_σ can be extended to Δ_1 because X is aspherical to give $D_0(\sigma) : \Delta_1 \rightarrow X$. Note that B_k is homeomorphic to Δ_k for all $k \geq 0$. Now that D_0 is defined for arbitrary singular 0-forms in X , we may extend D_0 by linearity to $S_0(X; R)$.

II.ii) Given a singular 1-simplex σ in X , we define $\beta_\sigma : \dot{\Delta}_2 \rightarrow X$ such that $\beta_\sigma \circ F_2^0 = \sigma$ and $\beta_\sigma \circ F_2^i = D_0(\sigma \circ F_1^{i-1})$ for $i = 1, 2$. To prove that β_σ is continuous on $\dot{\Delta}_2$, we have to prove that β_σ is continuous on the intersection of the distinct faces of Δ_2 . Let $K_{i,j} = F_2^i(\Delta_1) \cap F_2^j(\Delta_1)$ for $0 \leq j < i \leq 2$.

We have proved in the context of the proof of Proposition 8.4.6 that $F_2^i \circ F_1^j = F_2^j \circ F_1^{i-1}$ for $0 \leq j < i \leq 2$. Thus $K_{i,j} = (F_2^i \circ F_1^j)(\Delta_0) = (F_2^j \circ F_1^{i-1})(\Delta_0)$ for $0 \leq j < i \leq 2$. Hence $K_{0,1} = \{\mathbf{e}_2\} = \{(F_2^1 \circ F_1^0)(\mathbf{e}_0)\}$, $K_{0,2} = \{\mathbf{e}_1\} = \{(F_2^2 \circ F_1^0)(\mathbf{e}_0)\}$ and $K_{1,2} = \{\mathbf{e}_0\} = \{(F_2^2 \circ F_1^1)(\mathbf{e}_0)\}$. Since the intersection of the faces is only one point, we obviously have

that β_σ is continuous on $\dot{\Delta}_1$. Therefore, the map β_σ can be extended to Δ_2 because X is aspherical to give $D_1(\sigma) : \Delta_2 \rightarrow X$. Now that D_1 is defined for arbitrary singular 1-forms in X , we may extend D_1 by linearity to $S_1(X; R)$.

II.iii) We assume that $D_{k-1} : S_{k-1}(X; R) \rightarrow S_k(X; R)$ has been defined. Given a singular k -simplex σ in X , we define $\beta_\sigma : \dot{\Delta}_{k+1} \rightarrow X$ such that $\beta_\sigma \circ F_{k+1}^0 = \sigma$ and $\beta_\sigma \circ F_{k+1}^i = D_{k-1}(\sigma \circ F_k^{i-1})$ for $1 \leq i \leq k+1$. To prove that β_σ is continuous on $\dot{\Delta}_{k+1}$, we have to prove that β_σ is continuous on the intersection of the distinct faces of Δ_{k+1} .

Let $K_{i,j} = F_{k+1}^i(\Delta_k) \cap F_{k+1}^j(\Delta_k)$ for $0 \leq j < i \leq k+1$. As we mentioned before, we have proved in the context of the proof of Proposition 8.4.6 that $F_{k+1}^i \circ F_k^j = F_{k+1}^j \circ F_k^{i-1}$ for $0 \leq j < i \leq k+1$. Thus $K_{i,j} = (F_{k+1}^i \circ F_k^j)(\Delta_{k-1}) = (F_{k+1}^j \circ F_k^{i-1})(\Delta_{k-1})$ for $0 \leq j < i \leq k+1$. Since $i > j \geq 0$, we get from the definition of D_{k-1} that

$$\begin{aligned} \beta_\sigma \circ F_{k+1}^i \circ F_k^j &= (\beta_\sigma \circ F_{k+1}^i) \circ F_k^j = D_{k-1}(\sigma \circ F_k^{i-1}) \circ F_k^j \\ &= \begin{cases} \sigma \circ F_k^{i-1} & \text{if } j = 0 \\ D_{k-2}(\sigma \circ F_k^{i-1} \circ F_{k-1}^{j-1}) & \text{if } 0 < j < i \end{cases} \end{aligned} \quad (8.4.8)$$

To get the last equality, we use the fact that $D_{k-1}(\sigma \circ F_k^{i-1})|_{\dot{\Delta}_k} = \beta_{\sigma \circ F_k^{i-1}} : \dot{\Delta}_k \rightarrow X$ for the definition of D_{k-1} .

Moreover, we get from the definition of β_σ and D_{k-1} that

$$\begin{aligned} \beta_\sigma \circ F_{k+1}^j \circ F_k^{i-1} &= (\beta_\sigma \circ F_{k+1}^j) \circ F_k^{i-1} = \begin{cases} \sigma \circ F_k^{i-1} & \text{if } j = 0 \\ D_{k-1}(\sigma \circ F_k^{j-1}) \circ F_k^{i-1} & \text{if } 0 < j < i \end{cases} \\ &= \begin{cases} \sigma \circ F_k^{i-1} & \text{if } j = 0 \\ D_{k-2}(\sigma \circ F_k^{j-1} \circ F_{k-1}^{i-2}) & \text{if } 0 < j < i \end{cases} \end{aligned} \quad (8.4.9)$$

To get the last equality, we again use the fact that $D_{k-1}(\sigma \circ F_k^{i-1})|_{\dot{\Delta}_k} = \beta_{\sigma \circ F_k^{i-1}} : \dot{\Delta}_k \rightarrow X$ for the definition of D_{k-1} .

Since $F_k^{i-1} \circ F_{k-1}^{j-1} = F_k^{j-1} \circ F_{k-1}^{i-2}$ for $1 \leq j < i \leq k+1$, we get from (8.4.8) and (8.4.9) that $\beta_\sigma \circ F_{k+1}^j \circ F_k^{i-1} = \beta_\sigma \circ F_{k+1}^j \circ F_k^{i-1}$; namely, β_σ on the j^{th} face of the i^{th} face of $\dot{\Delta}_{k+1}$ is equal to β_σ on the $(i-1)^{th}$ face of the j^{th} face of $\dot{\Delta}_{k+1}$. Thus $\beta_\sigma : \dot{\Delta}_k \rightarrow X$ is continuous. Therefore, the map β_σ can be extended to Δ_{k+1} because X is aspherical to give $D_k(\sigma) : \Delta_{k+1} \rightarrow X$. Now that D_k is defined for arbitrary singular k -forms in X , we may extend D_k by linearity to $S_k(X; R)$.

II.iv) It is left to prove that $\mathcal{D} = \{D_k\}_{k \in \mathbb{Z}}$ is a chain homotopy between the chain maps $\mathcal{F}$ and $\mathcal{G}$ defined in (i); namely, that

$$\partial_{k+1} \circ D_k + D_{k-1} \circ \partial_k = g_k - f_k \quad (8.4.10)$$

for $k \in \mathbb{Z}$. The equation is trivially true for $k < 0$. For $k = 0$, we have from the definition of D_0 that

$$\partial_1(D_0(\sigma)) = D_0(\sigma) \circ F_1^0 - D_0(\sigma) \circ F_1^1 = \sigma - (\eta \circ \epsilon)(\sigma) = g_0(\sigma) - f_0(\sigma)$$

for all singular 0-simplices in X . By linearity, (8.4.10) is true on $S_0(X; R)$.

For $k > 0$, we have

$$\begin{aligned}
 \partial_{k+1}(D_k(\sigma)) + D_{k-1}(\partial_k(\sigma)) &= \sum_{i=0}^{k+1} (-1)^i D_k(\sigma) \circ F_{k+1}^i + D_{k-1} \left(\sum_{i=0}^k (-1)^i \sigma \circ F_k^i \right) \\
 &= (D_k(\sigma)) \circ F_{k+1}^0 + \sum_{i=1}^{k+1} (-1)^i D_k(\sigma) \circ F_{k+1}^i + \sum_{i=0}^k (-1)^i D_{k-1}(\sigma \circ F_k^i) \\
 &= \underbrace{D_k(\sigma) \circ F_{k+1}^0}_{=\sigma} + \sum_{i=0}^k (-1)^{i+1} \underbrace{D_k(\sigma) \circ F_{k+1}^{i+1}}_{=D_{k-1}(\sigma \circ F_k^i)} + \sum_{i=0}^k (-1)^i D_{k-1}(\sigma \circ F_k^i) \\
 &= \sigma = g_k(\sigma) - f_k(\sigma)
 \end{aligned}$$

for all singular k -simplices in X . By linearity, (8.4.10) is true on $S_k(X; R)$ for $k > 0$. ■

Proposition 8.4.38

Let R be an integral domain. Then $H_k(\Delta_k \times [0, 1]; R) = 0$ for $k > 0$.

Proof.

According to the Remark 8.4.36, $\Delta_k \times [0, 1]$ is aspherical. It follows from the previous theorem that the reduced chain complex $\mathcal{C}^\# = \{(S_k^\#(\Delta_k \times [0, 1]; R), \partial_k^\#)\}_{k \in \mathbb{Z}}$ of the chain complex $\mathcal{C} = \{(S_k(\Delta_k \times [0, 1]; R), \partial_k)\}_{k \in \mathbb{Z}}$ with augmentation ϵ is acyclic. Therefore $H_k(\Delta_k \times [0, 1]; R) = H_k^\#(\Delta_k \times [0, 1]; R) = 0$ for $k > 0$. ■

The proof of the next theorem is based on the previous proposition.

Theorem 8.4.39

Let X and Y be two topological spaces, and $f, g : X \rightarrow Y$ be two homotopic functions. Consider chain maps $\mathcal{F} = \{f_k\}_{k \in \mathbb{Z}}$ and $\mathcal{G} = \{g_k\}_{k \in \mathbb{Z}}$ between the chain complexes $\mathcal{C} = \{(S_k(X; R), \partial_k)\}_{k \in \mathbb{Z}}$ and $\mathcal{C} = \{(S_k(Y; R), \partial_k)\}_{k \in \mathbb{Z}}$ defined by $f_k = S_k(f)$ and $g_k = S_k(g)$ for all k . There exists a chain homotopy $\mathcal{D} = \{D_k\}_{k \in \mathbb{Z}}$ between $\mathcal{F}$ and $\mathcal{G}$.

Proof.

i) Let $\iota_y : X \rightarrow X \times [0, 1]$ be the functions defined by $\iota_y(x) = (x, y)$ for $x \in X$ and $y \in [0, 1]$. Moreover, let $\mathcal{S}_0 = \{S_k(\iota_0)\}_{k \in \mathbb{Z}}$ and $\mathcal{S}_1 = \{S_k(\iota_1)\}_{k \in \mathbb{Z}}$ be chain maps between the chain complexes $\mathcal{C} = \{(S_k(X; R), \partial_k)\}_{k \in \mathbb{Z}}$ and $\mathcal{C} = \{(S_k(X \times [0, 1]; R), \partial_k)\}_{k \in \mathbb{Z}}$. We prove that there exists a chain homotopy $\tilde{\mathcal{D}} = \{\tilde{D}_k\}_{k \in \mathbb{Z}}$ between $\mathcal{S}_0$ and $\mathcal{S}_1$. To be more precise, we prove by induction the following statement.

P_k: There exists a homomorphism $\tilde{D}_k : S_k(X; R) \rightarrow S_{k+1}(X \times [0, 1]; R)$ such that:

$$1. \partial_{k+1} \circ \tilde{D}_k + \tilde{D}_{k-1} \circ \partial_k = S_k(\iota_1) - S_k(\iota_0).$$

2. $\tilde{D}_k \circ S_k(h) = S_{k+1}(h \times \text{Id}_{[0,1]}) \circ \tilde{D}_k$ for all continuous map $h : X \rightarrow Y$ ⁸.
3. $\partial_{k+1} \circ (S_{k+1}(\iota_1) - S_{k+1}(\iota_0) - \tilde{D}_k \circ \partial_{k+1}) = 0$.

Note that $h \times \text{Id}_{[0,1]} : X \times [0, 1] \rightarrow \tilde{X} \times [0, 1]$ is defined by $(h \times \text{Id}_{[0,1]})(x, t) = (h(x), t)$ for all $(x, t) \in X \times [0, 1]$.

We may set $\tilde{D}_k = 0$ for $k < 0$. Hence, we only have to consider $k \geq 0$.

If σ is a singular k -simplex in X , then we also note that the singular k -simplex $S_k(\iota_y)(\sigma)$ in $X \times [0, 1]$ is defined by $(S_k(\iota_y)(\sigma))(\mathbf{x}) = (\sigma, e_y)(\mathbf{x}) = (\sigma(\mathbf{x}), y)$ for all $\mathbf{x} \in \Delta_k$, where the singular k -simplex $e_y : \Delta_k \rightarrow [0, 1]$ is defined by $e_y(\mathbf{x}) = y$ for all $\mathbf{x} \in \Delta_k$.

We prove that P_0 is true. Given a singular 0-simplex σ in X , we define $\tilde{D}_0(\sigma) \in S_1(X \times [0, 1]; R)$ as $\tilde{D}_0(\sigma)(\mathbf{x}) = (\rho_\sigma(\mathbf{x}), \rho_I(\mathbf{x}))$ for all $\mathbf{x} \in \Delta_1$, where $\rho_\sigma : \Delta_1 \rightarrow X$ and $\rho_I : \Delta_1 \rightarrow [0, 1]$ are defined by $\rho_\sigma(\mathbf{x}) = \sigma(\mathbf{e}_0)$ and $\rho_I(\mathbf{x}) = \mathbf{x} \cdot \mathbf{e}_1$ for all $\mathbf{x} \in \Delta_1$. We extend $\tilde{D}_0$ to $S_1(X \times [0, 1]; R)$ by linearity. Moreover, because of the linearity, it is enough to verify the three properties of P_0 for arbitrary singular simplices in X .

Suppose that σ is an arbitrary singular 0-simplex in X . Since $\tilde{D}_{-1} = 0$, we have

$$(\partial_1 \circ \tilde{D}_0)(\sigma) + (\tilde{D}_{-1} \circ \partial_0)(\sigma) = (\partial_1 \circ \tilde{D}_0)(\sigma) = S_0(\iota_1)(\sigma) - S_0(\iota_0)(\sigma)$$

because

$$\begin{aligned} (\partial_1(\tilde{D}_0(\sigma)))(\mathbf{e}_0) &= (\tilde{D}_0(\sigma) \circ F_1^0)(\mathbf{e}_0) - (\tilde{D}_0(\sigma) \circ F_1^1)(\mathbf{e}_0) = \tilde{D}_0(\sigma)(\mathbf{e}_1) - \tilde{D}_0(\sigma)(\mathbf{e}_0) \\ &= (\sigma(\mathbf{e}_0), 1) - (\sigma(\mathbf{e}_0), 0) = S_0(\iota_1)(\sigma)(\mathbf{x}_0) - S_0(\iota_0)(\sigma)(\mathbf{x}_0). \end{aligned}$$

So (1) of P_0 is satisfied.

Since

$$\begin{aligned} \tilde{D}_0(S_0(h)(\sigma)) &= \tilde{D}_0(h \circ \sigma) = (\rho_{h \circ \sigma}, \rho_I) = (h \circ \rho_\sigma, \rho_I) = (h \times \text{Id}_{[0,1]})(\rho_\sigma, \rho_I) \\ &= S_1(h \times \text{Id}_{[0,1]})(\tilde{D}_0(\sigma)), \end{aligned}$$

we get that (2) of P_0 is satisfied.

Suppose that σ is an arbitrary singular 1-simplex in X . To prove (3) of P_0 , namely $\partial_1 \circ (S_1(\iota_1) - S_1(\iota_0) - \tilde{D}_0 \circ \partial_1) = 0$, we note that

$$\begin{aligned} (\partial_1(\tilde{D}_0(\partial_1(\sigma))))(\mathbf{e}_0) &= ((\tilde{D}_0(\partial_1(\sigma))) \circ F_1^0 - (\tilde{D}_0(\partial_1(\sigma))) \circ F_1^1)(\mathbf{e}_0) \\ &= (\tilde{D}_0(\partial_1(\sigma)))(\mathbf{e}_1) - (\tilde{D}_0(\partial_1(\sigma)))(\mathbf{e}_0) = (\partial_1(\sigma)(\mathbf{e}_0), 1) - (\partial_1(\sigma)(\mathbf{e}_0), 0) \\ &= ((\sigma \circ F_1^0)(\mathbf{e}_0), 1) - ((\sigma \circ F_1^1)(\mathbf{e}_0), 1) - ((\sigma \circ F_1^0)(\mathbf{e}_0), 0) + ((\sigma \circ F_1^1)(\mathbf{e}_0), 0) \\ &= (\sigma(\mathbf{e}_1), 1) - (\sigma(\mathbf{e}_0), 1) - (\sigma(\mathbf{e}_1), 0) + (\sigma(\mathbf{e}_0), 0) \\ &= S_1(\iota_1)(\sigma)(\mathbf{e}_1) - S_1(\iota_1)(\sigma)(\mathbf{e}_0) - S_1(\iota_0)(\sigma)(\mathbf{e}_1) + S_1(\iota_0)(\sigma)(\mathbf{e}_0) \end{aligned}$$

⁸This property is called **naturality** and is really important in category theory. In particular, we can show that the definition of $\tilde{D}_k$ is independent of the topological space homeomorphic to X . We leave this subject to those in pure algebra and logic.

$$\begin{aligned}
&= (S_1(\iota_1)(\sigma) \circ F_1^0 - S_1(\iota_1)(\sigma) \circ F_1^1)(e_0) - (S_1(\iota_0)(\sigma) \circ F_1^0 - S_1(\iota_0)(\sigma) \circ F_1^1)(e_0) \\
&= \partial_1(S_1(\iota_1)(\sigma))(e_0) - \partial_1(S_1(\iota_0)(\sigma))(e_0)
\end{aligned}$$

We assume that P_{k-1} is true and prove that P_k is true.

It follows from (3) of P_{k-1} with X replaced by Δ_k that

$$\partial_k \circ ((S_k(\iota_1) - S_k(\iota_0) - \tilde{D}_{k-1} \circ \partial_k)(\delta)) = 0$$

where $\delta : \Delta_k \rightarrow \Delta_k$ is defined by $\delta(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in \Delta_k$. Note that everything that we have said so far is true for any topological space X . This is the reason why we may use $X = \Delta_k$.

Thus $(S_k(\iota_1) - S_k(\iota_0) - \tilde{D}_{k-1} \circ \partial_k)(\delta) \in Z_k(\Delta_k \times [0, 1]; R)$. Since $H_k(\Delta_k \times [0, 1]; R) = 0$ for $k > 0$ according to Proposition 8.4.38, there exists $\tilde{\delta} \in S_{k+1}(\Delta_k \times [0, 1]; R)$ such that

$$\partial_{k+1}\tilde{\delta} = (S_k(\iota_1) - S_k(\iota_0) - \tilde{D}_{k-1} \circ \partial_k)(\delta) . \quad (8.4.11)$$

We set $\tilde{D}_k(\delta) = \tilde{\delta}$.

Given a singular k -simplex σ in X , we set $\tilde{D}_k(\sigma) = S_{k+1}(\sigma \times \text{Id}_{[0,1]})(\tilde{\delta}) \in S_{k+1}(X \times [0, 1]; R)$. Recall that $(\sigma \times \text{Id}_{[0,1]}) : \Delta_k \times [0, 1] \rightarrow \Delta_k \times [0, 1]$ is defined by $(\sigma \times \text{Id}_{[0,1]})(\mathbf{x}, t) = (\sigma(\mathbf{x}), t)$ for all $(\mathbf{x}, t) \in \Delta_k \times [0, 1]$. We extend $\tilde{D}_k$ to $S_k(X; R)$ by linearity.

As for P_0 , because of the linearity, it is enough to verify the three properties of P_k for arbitrary singular simplices in X .

Suppose that σ is an arbitrary singular k -simplex in X . (2) of P_k follows from

$$\begin{aligned}
\tilde{D}_k(S_k(h)(\sigma)) &= \tilde{D}_k(h \circ \sigma) = S_{k+1}((h \circ \sigma) \times \text{Id}_{[0,1]})(\tilde{\delta}) = S_{k+1}(h \times \text{Id})(S_{k+1}(\sigma \times \text{Id})(\tilde{\delta})) \\
&= S_{k+1}(h \times \text{Id}_{[0,1]})(\tilde{D}_k(\sigma)) .
\end{aligned}$$

To prove (1) of P_k , we get from (8.4.11) that

$$\partial_{k+1}(\tilde{D}_k(\delta)) + \tilde{D}_{k-1}(\partial_k(\delta)) = \partial_{k+1}(\tilde{\delta}) + \tilde{D}_{k-1}(\partial_k(\delta)) = S_k(\iota_1)(\delta) - S_k(\iota_0)(\delta) \quad (8.4.12)$$

and note that $\sigma = S_k(\sigma)(\delta)$. Hence

$$\begin{aligned}
\partial_{k+1}(\tilde{D}_k(\sigma)) + \tilde{D}_{k-1}(\partial_k(\sigma)) &= \partial_{k+1}(S_{k+1}(\sigma \times \text{Id}_{[0,1]})(\tilde{D}_k(\delta))) + \tilde{D}_{k-1}(\partial_k(S_k(\sigma)(\delta))) \\
&= S_k(\sigma \times \text{Id}_{[0,1]})(\partial_{k+1}(\tilde{D}_k(\delta))) + \tilde{D}_{k-1}(S_{k-1}(\sigma)(\partial_k(\delta))) \\
&= S_k(\sigma \times \text{Id}_{[0,1]})(\partial_{k+1}(\tilde{D}_k(\delta))) + S_k(\sigma \times \text{Id}_{[0,1]})(\tilde{D}_{k-1}(\partial_k(\delta))) \\
&= S_k(\sigma \times \text{Id}_{[0,1]})(\partial_{k+1}(\tilde{D}_k(\delta)) + \tilde{D}_{k-1}(\partial_k(\delta))) = S_k(\sigma \times \text{Id}_{[0,1]})(S_k(\iota_1)(\delta) - S_k(\iota_0)(\delta)) \\
&= S_k(\sigma \times \text{Id}_{[0,1]})((\delta, e_1) - (\delta, e_0)) = (\sigma, e_1) - (\sigma, e_0) = S_k(\iota_1)(\sigma) - S_k(\iota_0)(\sigma)
\end{aligned}$$

where the second equality comes from (2) of Proposition 8.4.13, the third equality comes from (2) for P_{k-1} and the fifth equality comes from (8.4.12).

Finally, (3) of (P_k) follows from (1) because

$$\partial_{k+1} \circ (S_{k+1}(\iota_1) - S_{k+1}(\iota_0)) = (S_k(\iota_1) - S_k(\iota_0)) \circ \partial_{k+1} = (\partial_{k+1} \circ \tilde{D}_k + \tilde{D}_{k+1} \circ \partial_k) \circ \partial_{k+1}$$

$$= \partial_{k+1} \circ \tilde{D}_k \circ \partial_{k+1}$$

where again we have use (2) of Proposition 8.4.13 to obtain the first equality. This completes the proof by induction.

ii) Since $f \sim g$, there exists a continuous function $H : X \times [0, 1] \rightarrow X$ such that $H(x, 0) = f(x)$ and $H(x, 1) = g(x)$ for all $x \in X$. Thus $f = H \circ \iota_0$ and $g = H \circ \iota_1$.

It follows from (i) that $\partial_{k+1} \circ \tilde{D}_k + \tilde{D}_{k-1} \circ \partial_k = S_k(i_1) - S_k(i_0)$ for all k . Hence,

$$\begin{aligned} \partial_{k+1} \circ (S_{k+1}(H) \circ \tilde{D}_k) + (S_k(H) \circ \tilde{D}_{k-1}) \circ \partial_k &= S_k(H) \circ (\partial_{k+1} \circ \tilde{D}_k) + S_k(H) \circ (\tilde{D}_{k-1} \circ \partial_k) \\ &= S_k(H) (\partial_{k+1} \circ \tilde{D}_k + \tilde{D}_{k-1} \circ \partial_k) = S_k(H) \circ (S_k(i_1) - S_k(i_0)) \\ &= S_k(H \circ i_1) - S_k(H \circ i_0) = S_k(g) - S_k(f) \end{aligned}$$

where the first and fourth equality come from Proposition 8.4.13, and the third equality comes from (i). This complete the proof of the theorem if we take $D_k = S_{k+1}(H) \circ \tilde{D}_k$ for all k . ■

The next corollary is a consequence of Theorem 8.4.39 and Proposition 8.4.29.

Corollary 8.4.40

Suppose that X and Y are two topological spaces. If $f, h : X \rightarrow Y$ are homotopic maps, then $H_k(f) = H_k(g)$ for all k .

Corollary 8.4.41

If X is a contractible topological space and R is an integral domain, then $H_k(X; R) = 0$ for $k > 0$ and $H_0(X; R) \cong R$.

Proof.

It follows from the definition of contractible spaces that there exists $x_0 \in X$ such that the identity function $\text{Id}_X : X \rightarrow X$ is homotopic to the constant function $f : X \rightarrow X$ defined by $f(x) = x_0$ for all $x \in X$. From the previous corollary, we have that $H_k(\text{Id}_X) = H_k(f)$. It follows from Example 8.4.9 that

$$H_k(X; R) = H_k(\text{Id}_X)(H_k(X; R)) = H_k(f)(H_k(X; R)) \cong \begin{cases} 0 & \text{if } k > 0 \\ R & \text{if } k = 0 \end{cases}$$

because $H_k(f)(H_k(X; R)) = H_k(\{x_0\}; R)$. ■

Definition 8.4.42

Let X and Y be two topological spaces. We say that a continuous map $f : X \rightarrow Y$ is a **homotopic equivalence** if there exists a continuous map $h : Y \rightarrow X$ such that $h \circ f \sim \text{Id}_X$ and $f \circ h \sim \text{Id}_Y$. In that situation, we say that X and Y are **homotopic equivalent**.

Proposition 8.4.43

Suppose that X and Y are two homotopic equivalent topological spaces. Then $H_k(X; R) \cong H_k(Y; R)$.

Proof.

The previous proposition follows from Corollaries 8.4.15 and 8.4.40.

Suppose that $f : X \rightarrow Y$ and $h : Y \rightarrow X$ are two continuous maps such that $h \circ f \sim \text{Id}_X$ and $f \circ h \sim \text{Id}_Y$. Then $H_k(h) \circ H_k(f) = H_k(h \circ f) = H_k(\text{Id}_X) = \text{Id}_{H_k(X; R)}$ and $H_k(f) \circ H_k(h) = H_k(f \circ h) = H_k(\text{Id}_Y) = \text{Id}_{H_k(Y; R)}$. Thus $H_k(h)$ is the inverse of $H_k(f)$. Hence $H_k(f)$ is an isomorphism between $H_k(X; R)$ and $H_k(Y; R)$. ■

8.4.4 Reduced Homology

Let X be a topological space and R be an integral domain. This section presents an example of an augmentation for the chain complex $\mathcal{C} = \{(S_k(X; R), \partial_k)\}_{k \in \mathbb{Z}}$. This example plays a crucial role later.

We consider the reduced chain complex $\mathcal{C}^\# = \{(S_k^\#(X; R), \partial_k^\#)\}_{k \in \mathbb{Z}}$ of the chain complex $\mathcal{C}$ with augmentation

$$\partial_0^\# \left(\sum_{j \in J} a_j \sigma_j \right) = \sum_{j \in J} a_j ,$$

where $J \subset \mathbb{N}$ is a finite set, $a_j \in R$ and σ_j is a singular 0-simplex for $j \in J$. According to the definition of a reduced chain complex, Definition 8.4.31, we have

$$S_k^\#(X; R) = \begin{cases} R & \text{if } k = -1 \\ S_k(X; R) & \text{if } k \neq -1 \end{cases}$$

and $\partial_k^\# = \partial_k$ for $k \neq 0$. We have already seen this augmentation in the proof of Proposition 8.4.11 where $\partial_0^\#$ was denoted ϵ . We also proved that $\partial_0^\# \circ \partial_1^\# = 0$. We can then proceed as we did before to define $Z_k^\#(X; R)$, $B_k^\#(X; R)$ and $H_k^\#(X; R)$.

Definition 8.4.44

Let X be a topological space X and R be an integral domain. The R -module $H_k^\#(X; R)$ is called the **reduced (singular) k^{th} homology module** of X .

In fact $H_k^\#(X; R) \cong H_k(X; R)$ for $k \neq 0$. Only $H_0^\#(X; R)$ and $H_0(X; R)$ may be different. For instance, we saw that in Example 8.4.9 that $H_0(X, R) = R$ if $X = \{x\}$. However, we have that $H_0^\#(X, R) = 0$ if $X = \{x\}$. In general, we have the following result.

Proposition 8.4.45

Let X be a topological space X and R be an integral domain. Then $H_0(X; R) \cong H_0^\sharp(X; R) \oplus R$.

Proof.

We first note that $Z_0(X; R) = S_0(X; R)$. Moreover, $B_0(X; R) = B_0^\sharp(X; R)$ because $\partial_1^\sharp = \partial_1$. Since $\partial_0^\sharp \circ \partial_1 = 0$, we also have that $B_0(X; R) \subset \mathcal{N}(\partial_0^\sharp) = Z_0^\sharp(X; R)$.

Given $b \in S_0(X; R)$ such that $\partial_0^\sharp(b) = 1$, we prove that

$$Z_0(X; R) = Z_0^\sharp(X; R) \oplus \{rb : r \in R\}.$$

Suppose that $c \in Z_0^\sharp(X; R) \cap \{rb : r \in R\}$. Then $c = qb$ for some $q \in R$. But $q = q\partial_0^\sharp(b) = \partial_0^\sharp(qb) = \partial_0^\sharp(c) = 0$. Thus $c = 0$ and so $Z_0^\sharp(X; R) \cap \{rb : r \in R\} = \{0\}$. Given $c \in S_0(X; R)$, suppose that $\partial_0^\sharp(c) = q$. Then $c = (c - qb) + qb$ with $c - qb \in \mathcal{N}(\partial_0^\sharp) = Z_0^\sharp(X; R)$ and $qb \in \{rb : r \in R\}$. Thus $S_0(X; R) = Z_0(X; R) = Z_0^\sharp(X; R) + \{rb : r \in R\}$.

Hence

$$\begin{aligned} H_0(X; R) &= Z_0(X; R)/B_0(X; R) = (Z_0^\sharp(X; R) \oplus \{rb : r \in R\})/B_0^\sharp(X; R) \\ &\cong (Z_0^\sharp(X; R)/B_0^\sharp(X; R)) \oplus \{rb : r \in R\} \cong H_0^\sharp(X; R) \oplus R. \end{aligned}$$

Note that $b \notin B_0^\sharp(X; R)$ because $\partial_0^\sharp(b) \neq 0$. ■

In general, the results about singular homology and reduced singular homology differ only at the level of $S_0(X; R)$.

8.4.5 Relation Between $\pi_1(X, x_0)$ and $H_1(X; R)$.

We take a short pause in our study of singular homology to present the relation between the concept of fundamental groups introduced in Section 8.1 and the singular 1st homology module.

Theorem 8.4.46

Let X be a topological space and $x_0 \in X$. There exists a natural group homomorphism $\mathcal{X}$ from $\pi_1(X, x_0)$ to $H_1(X; \mathbb{Z})$. Moreover, if X is path connected, then $\mathcal{X}$ is onto and the kernel of $\mathcal{X}$, denoted $\mathcal{N}(\mathcal{X})$, is the commutator subgroup of $\pi_1(X, x_0)$.

Proof.

The map $\mathcal{X} : \pi_1(X, x_0) \rightarrow H_1(X; \mathbb{Z})$ is defined by $\mathcal{X}([\alpha]) = [\check{\alpha}]$ where $\check{\alpha}$ is the singular 1-simplex defined by $\check{\alpha}((1-t)\mathbf{e}_0 + t\mathbf{e}_1) = \alpha(t)$ for $0 \leq t \leq 1$.

i) We first show that $\mathcal{X}$ is well defined. Let α_0 and α_1 be two loops at x_0 such that $\alpha_0 \sim \alpha_1$. Therefore, according to the definition of homotopic paths, Definition 8.1.6, there exists a

continuous function $H : [0, 1] \times [0, 1] \rightarrow X$ such that $H(t, 0) = \alpha_0(t)$ and $H(t, 1) = \alpha_1(t)$ for $0 \leq t \leq 1$ and $H(0, s) = H(1, s) = x_0$ for $0 \leq s \leq 1$.

In the context of this proof, we will associate to each path $\alpha : [0, 1] \rightarrow X$ the singular 1-simplex $\check{\alpha} : \Delta_1 \rightarrow X$ defined by $\check{\alpha}((1-t)\mathbf{e}_0 + t\mathbf{e}_1) = \alpha(t)$ for $0 \leq t \leq 1$, and vice-versa. In particular, singular simplices and singular chains are identified with the symbol $\check{}$.

We define a singular 2-simplex $\beta : \Delta_2 \rightarrow X$ in X as

$$\check{\beta}(t(s\mathbf{e}_2 + (1-s)\mathbf{x}_1) + (1-t)\mathbf{e}_0) = H(t, s)$$

for $t, s \in [0, 1]$. The only issue with this definition may be at $\mathbf{e}_0$ because $\mathbf{e}_0 = 0(s\mathbf{e}_2 + (1-s)\mathbf{x}_1) + \mathbf{e}_0$ for $0 \leq s \leq 1$. However, this is not an issue because $H(0, s) = x_0$ for $0 \leq s \leq 1$. We also have that $\check{\beta}$ is continuous.

We have

$$\partial_2 \check{\beta} = \beta \circ F_2^0 - \check{\beta} \circ F_2^1 + \check{\beta} \circ F_2^2 \quad (8.4.13)$$

where

$$\begin{aligned} \check{\beta}(F_2^2((1-t)\mathbf{e}_0 + t\mathbf{e}_1)) &= \check{\beta}((1-t)\mathbf{e}_0 + t\mathbf{e}_1) = \check{\beta}(t(0\mathbf{e}_2 + \mathbf{e}_1) + (1-t)\mathbf{e}_0) = H(t, 0) \\ &= \alpha_0(t) = \check{\alpha}_0((1-t)\mathbf{e}_0 + t\mathbf{e}_1) \end{aligned}$$

for $0 \leq t \leq 1$,

$$\begin{aligned} \check{\beta}(F_2^1((1-t)\mathbf{e}_0 + t\mathbf{e}_1)) &= \check{\beta}((1-t)\mathbf{e}_0 + t\mathbf{e}_2) = \check{\beta}(t(\mathbf{e}_2 + 0\mathbf{e}_1) + (1-t)\mathbf{e}_0) = H(t, 1) \\ &= \alpha_1(t) = \check{\alpha}_1((1-t)\mathbf{e}_0 + t\mathbf{e}_1) \end{aligned}$$

for $0 \leq t \leq 1$, and

$$\begin{aligned} \check{\beta} \circ F_2^0((1-s)\mathbf{e}_0 + s\mathbf{e}_1) &= \check{\beta}((1-s)\mathbf{e}_1 + s\mathbf{e}_2) = \check{\beta}(s\mathbf{e}_2 + (1-s)\mathbf{e}_1 + 0\mathbf{e}_0) = H(0, s) \\ &= x_0 = \check{\epsilon}_{x_0}((1-s)\mathbf{e}_0 + s\mathbf{e}_1) \end{aligned}$$

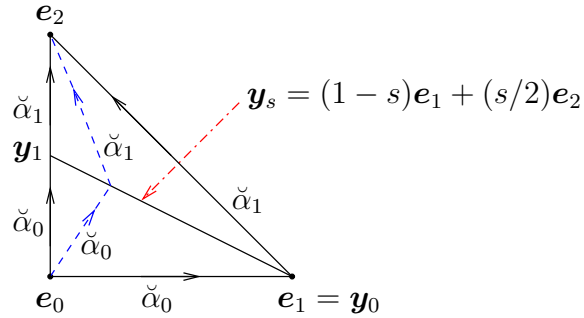
for $0 \leq s \leq 1$. In the previous expression, $\check{\epsilon}_{x_0} : \Delta_1 \rightarrow X$ is defined by $\check{\epsilon}_{x_0}(\mathbf{x}) = x_0$ for $\mathbf{x} \in \Delta_1$. Since $\check{\beta} \circ F_2^0 = \check{\epsilon}_{x_0} = \partial_2(\check{\gamma})$ where the singular 2-simplex $\check{\gamma}$ in X is defined by $\check{\gamma}(\mathbf{x}) = x_0$ for all $\mathbf{x} \in \Delta_2$, we get from (8.4.13) that $\check{\alpha}_0 - \check{\alpha}_1 = \check{\beta} \circ F_2^2 - \check{\beta} \circ F_2^1 \in B_1(X; \mathbb{Z})$. Thus $\mathcal{X}([\alpha_0]) = [\check{\alpha}_0] = [\check{\alpha}_1] = \mathcal{X}([\alpha_1])$ in $H_1(X; \mathbb{Z})$.

ii) We show that $\mathcal{X}$ is an homomorphism. Let α_0 and α_1 be two loops at x_0 . We need to show that $\mathcal{X}([\alpha_0][\alpha_1]) = \mathcal{X}([\alpha_0]) + \mathcal{X}([\alpha_1])$. Since $[\alpha_0][\alpha_1] = [\alpha_0\alpha_1]$, we need to show that $[\alpha_0\check{\alpha}_1] = [\check{\alpha}_0] + [\check{\alpha}_1]$ in $H_1(X; \mathbb{Z})$.

Let $\mathbf{y}_s = (1-s)\mathbf{e}_1 + (s/2)\mathbf{e}_2$ for $0 \leq s \leq 1$ and $\check{\beta}$ be the singular 2-simplex in X defined by

$$\check{\beta}(\mathbf{x}) = \begin{cases} \check{\alpha}_0((1-2t(2-s)^{-1})\mathbf{e}_0 + 2t(2-s)^{-1}\mathbf{e}_1) & \text{if } \mathbf{x} = (1-2t(2-s)^{-1})\mathbf{e}_0 \\ & + 2t(2-s)^{-1}\mathbf{y}_s \text{ with} \\ & 0 \leq s \leq 1 \text{ and } 0 \leq t \leq (2-s)/2 \\ \check{\alpha}_1(2(1-t)(2-s)^{-1}\mathbf{e}_0 \\ + (1-2(1-t)(2-s)^{-1})\mathbf{e}_1) & \text{if } \mathbf{x} = (1-2(1-t)(2-s)^{-1})\mathbf{e}_2 \\ & + 2(1-t)(2-s)^{-1}\mathbf{y}_s \text{ with} \\ & 0 \leq s \leq 1 \text{ and } s/2 < t \leq 1 \end{cases}$$

Note that $\check{\beta}$ is well defined because $\check{\beta}(\mathbf{x}) = \check{\alpha}_0(\mathbf{e}_1) = \check{\alpha}_1(\mathbf{e}_0) = x_0$ for $\mathbf{x} = \mathbf{y}_s$ with $0 \leq s \leq 1$. The following figure illustrates the singular 2-simplex $\check{\beta}$.



We have

$$\partial\check{\beta} = \check{\beta} \circ F_2^0 - \check{\beta} \circ F_2^1 + \check{\beta} \circ F_2^2 \quad (8.4.14)$$

where

$$(\check{\beta} \circ F_2^2)((1-t)\mathbf{e}_0 + t\mathbf{e}_1) = \check{\beta}((1-t)\mathbf{e}_0 + t\mathbf{e}_1) = \check{\beta}((1-t)\mathbf{e}_0 + t\mathbf{y}_0) = \check{\alpha}_0((1-t)\mathbf{e}_0 + t\mathbf{e}_1)$$

for $0 \leq t \leq 1$,

$$(\check{\beta} \circ F_2^0)((1-t)\mathbf{e}_0 + t\mathbf{e}_1) = \check{\beta}((1-t)\mathbf{e}_1 + t\mathbf{e}_2) = \check{\beta}((1-t)\mathbf{y}_0 + t\mathbf{e}_2) = \check{\alpha}_1((1-t)\mathbf{e}_0 + t\mathbf{e}_1)$$

for $0 \leq t \leq 1$, and

$$\begin{aligned} (\check{\beta} \circ F_2^1)((1-t)\mathbf{e}_0 + t\mathbf{e}_1) &= \check{\beta}((1-t)\mathbf{e}_0 + t\mathbf{e}_2) \\ &= \begin{cases} \check{\beta}((1-2t)\mathbf{e}_0 + 2t\mathbf{y}_1) & \text{if } 0 \leq t \leq 1/2 \\ \check{\beta}(1-2(1-t))\mathbf{e}_2 + 2(1-t)\mathbf{y}_1) & \text{if } 1/2 \leq t \leq 1 \end{cases} \\ &= \begin{cases} \check{\alpha}_0((1-2t)\mathbf{e}_0 + 2t\mathbf{e}_1) & \text{if } 0 \leq t \leq 1/2 \\ \check{\alpha}_1(2(1-t)\mathbf{e}_0 + (1-2(1-t))\mathbf{e}_1) & \text{if } 1/2 \leq t \leq 1 \end{cases} \\ &= \alpha_0\check{\alpha}_1((1-t)\mathbf{e}_0 + t\mathbf{e}_1) \end{aligned}$$

for $0 \leq t \leq 1$. It then follows from (8.4.14) that $\partial\check{\beta} = \check{\alpha}_1 - \alpha_0\check{\alpha}_1 + \check{\alpha}_0$. Thus $\check{\alpha}_1 - \alpha_0\check{\alpha}_1 + \check{\alpha}_0 \in B_1(X; \mathbb{Z})$. Hence $\mathcal{X}([\alpha_0\alpha_1]) = [\alpha_0\check{\alpha}_1] = [\check{\alpha}_0] + [\check{\alpha}_1] = \mathcal{X}([\alpha_0]) + \mathcal{X}([\alpha_1])$ in $H_1(X; \mathbb{Z})$.

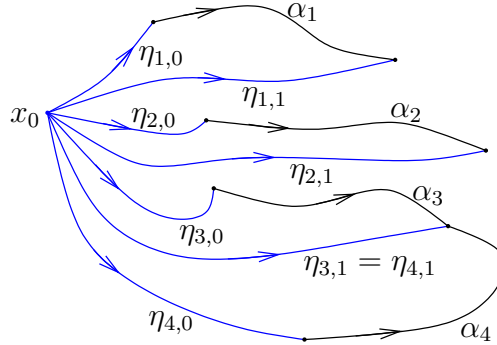
iii) We assume that X is path-connected and show that $\mathcal{X}$ is onto. Suppose that $\check{c} \in$

$Z_1(X; \mathbb{Z})$. We can write $\check{c}$ as $\check{c} = \sum_{j=1}^J a_j \check{\alpha}_j$ where $a_j \in \mathbb{Z}$ and $\check{\alpha}_j$ is singular 1-simplex for $1 \leq j \leq J$. Hence

$$\partial_1(\check{c}) = \sum_{j=1}^J a_j (\check{\alpha}_j \circ F_1^0 - \check{\alpha}_j \circ F_1^1) = 0. \quad (8.4.15)$$

Since X is path-connected, we can choose for each $\check{\alpha}_j$ in the summation above, a path $\eta_{j,1}$ in X from x_0 to $(\check{\alpha}_j \circ F_1^0)(\mathbf{e}_0) = \check{\alpha}_j(\mathbf{e}_1)$ and a path $\eta_{j,0}$ in X from x_0 to $(\check{\alpha}_j \circ F_1^1)(\mathbf{e}_0) = \check{\alpha}_j(\mathbf{e}_0)$ that satisfy the following conditions:

1. If $(\check{\alpha}_{j_1} \circ F_1^0)(e_0) = (\check{\alpha}_{j_2} \circ F_1^0)(e_0)$ for some $j_1 \neq j_2$, then $\eta_{j_1,1} = \eta_{j_2,1}$.
2. If $(\check{\alpha}_{j_1} \circ F_1^0)(e_0) = (\check{\alpha}_{j_2} \circ F_1^1)(e_0)$ for some $j_1 \neq j_2$, then $\eta_{j_1,1} = \eta_{j_2,0}$.
3. If $(\check{\alpha}_{j_1} \circ F_1^1)(e_0) = (\check{\alpha}_{j_2} \circ F_1^1)(e_0)$ for some $j_1 \neq j_2$, then $\eta_{j_1,0} = \eta_{j_2,0}$.
4. If $(\check{\alpha}_j \circ F_1^1)(e_0) = x_0$, then $\eta_{j,0} = e_{x_0}$ where $e_{x_0}(t) = x_0$ for $0 \leq t \leq 1$.
5. If $(\check{\alpha}_j \circ F_1^0)(e_0) = x_0$, then $\eta_{j,1} = e_{x_0}$.



It follows from (8.4.15) that

$$\sum_{j=1}^J a_j (\check{\eta}_{j,1} - \check{\eta}_{j,0}) = 0 \quad (8.4.16)$$

where $\check{\eta}_{j,i}((1-t)e_0 + te_1) = \eta_{j,i}(t)$ for $0 \leq t \leq 1$ because

$$\sum_{j=1}^J a_j (\check{\eta}_{j,1}(e_1) - \check{\eta}_{j,0}(e_1)) = \sum_{j=1}^J a_j ((\check{\alpha}_j \circ F_1^0)(e_0) - (\check{\alpha}_j \circ F_1^1)(e_0)) = 0$$

and $\check{\eta}_{j_1,1} \neq \check{\eta}_{j_2,1}$ if and only if $\check{\eta}_{j_1,1}(e_1) \neq \check{\eta}_{j_2,1}(e_1)$, $\check{\eta}_{j_1,1} \neq \check{\eta}_{j_2,0}$ if and only if $\check{\eta}_{j_1,1}(e_1) \neq \check{\eta}_{j_2,0}(e_1)$, and $\check{\eta}_{j_1,0} \neq \check{\eta}_{j_2,0}$ if and only if $\check{\eta}_{j_1,0}(e_1) \neq \check{\eta}_{j_2,0}(e_1)$ by construction.

Consider the loops at x_0 given by $\beta_j = \eta_{j,0} \alpha_j \eta_{j,1}^{-1}$ for all j . We have from (ii) that

$$\begin{aligned} \mathcal{X}([\beta_j]) &= \mathcal{X}([\eta_{j,0}] [\alpha_j] [\eta_{j,1}^{-1}]) = \mathcal{X}([\eta_{j,0}]) + \mathcal{X}([\alpha_j]) - \mathcal{X}([\eta_{j,1}]) \\ &= [\check{\eta}_{j,0}] + [\check{\alpha}_j] - [\check{\eta}_{j,1}] \end{aligned} \quad (8.4.17)$$

for all j . In the previous equality, we have used the fact that $\eta_{j,i} \eta_{j,i}^{-1} \sim e_{x_0}$ for $0 \leq i \leq 1$ and all j implies that

$$\mathcal{X}([\eta_{j,i}]) + \mathcal{X}([\eta_{j,i}^{-1}]) = \mathcal{X}([\eta_{j,i}] [\eta_{j,i}^{-1}]) = \mathcal{X}([\eta_{j,i} \eta_{j,i}^{-1}]) = \mathcal{X}([e_{x_0}]) = [\check{e}_{x_0}] = 0$$

in $H_1(X; \mathbb{Z})$ for $0 \leq i \leq 1$ and all j . Recall that $e_{x_0} : [0, 1] \rightarrow X$ is the loop at x_0 defined by $e_{x_0}(t) = x_0$ for $0 \leq t \leq 1$. Thus $\mathcal{X}([\eta_{j,i}^{-1}]) = -\mathcal{X}([\eta_{j,i}])$ for $0 \leq i \leq 1$ and all j .

Consider the loop at x_0 given by $\gamma = \prod_{j=1}^J \beta_j^{a_j}$. It follows from (8.4.16) and (8.4.17) that

$$\begin{aligned} \mathcal{X}([\gamma]) &= \mathcal{X}\left(\prod_{j=1}^J [\beta_j^{a_j}]\right) = \sum_{j=1}^J \mathcal{X}([\beta_j^{a_j}]) = \sum_{j=1}^J \mathcal{X}([\beta_j]^{a_j}) = \sum_{j=1}^J a_j \mathcal{X}([\beta_j]) \\ &= \sum_{j=1}^J a_j ([\check{\eta}_{j,0}] + [\check{\alpha}_j] - [\check{\eta}_{j,1}]) = \sum_{j=1}^J a_j [\check{\alpha}_j] + \sum_{j=1}^J a_j ([\check{\eta}_{j,0}] - [\check{\eta}_{j,1}]) = \sum_{j=1}^J a_j [\check{\alpha}_j] = [\check{c}]. \end{aligned}$$

iv) All elements in the commutator subgroup G of $\pi_1(X, x_0)$ are of the form $\alpha\beta\alpha^{-1}\beta^{-1}$ where α and β are two loops in X at x_0 . For these elements, we have

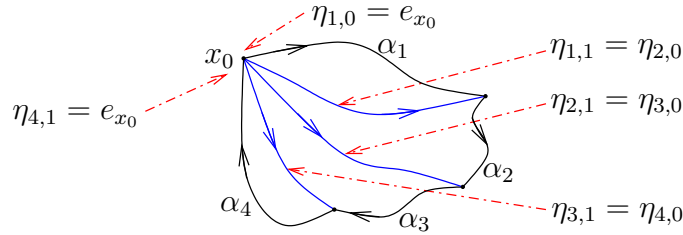
$$\mathcal{X}([\alpha\beta\alpha^{-1}\beta^{-1}]) = \mathcal{X}([\alpha][\beta][\alpha^{-1}][\beta^{-1}]) = \mathcal{X}([\alpha]) + \mathcal{X}([\beta]) - \mathcal{X}([\alpha]) - \mathcal{X}([\beta]) = 0.$$

Thus $G \subset \mathcal{N}(\mathcal{X})$.

To prove that $G = \mathcal{N}(\mathcal{X})$, we need to introduce a little tool. Given a loop γ in X at x_0 , suppose that $\gamma = \prod_{j=1}^J \alpha_j^{\epsilon_j}$ where the α_j are paths in X (they may not be loops) and the ϵ_j are either 1 or -1 . We set $\exp_\gamma(\alpha_i) = \sum_{\alpha_j = \alpha_i} \epsilon_j$ for each path α_i in γ .

Claim: If $\exp_\gamma(\alpha_i) = 0$ for all α_i in γ , then $[\gamma] \in G$.

To prove this claim, we choose the paths $\eta_{j,i}$ for $0 \leq i \leq 1$ and $1 \leq j \leq J$ as we have done in (iii).



Hence $\gamma \sim \prod_{j=1}^J (\eta_{j,0} \alpha_j \eta_{j,1}^{-1})^{\epsilon_j}$ because $\eta_{j,1}^{-1} \eta_{j+1,0} \sim e_{x_0}$ for $1 \leq j < J$, $\eta_{1,0} = e_{x_0}$ and $\eta_{J,1} = e_{x_0}$.

If we set $\beta_j = \eta_{j,0} \alpha_j \eta_{j,1}^{-1}$ for all j , then we have that $\gamma \sim \prod_{j=1}^J \beta_j^{\epsilon_j}$ where the β_j are loops at x_0 .

We consider the quotient group $\pi_1(X, x_0)/G$. Let $[\gamma]_G$ and $[\beta_j]_G$ denote the equivalence classes of $[\gamma]$ and $[\beta_j]$ in $\pi_1(X, x_0)/G$. We note that $\pi_1(X, x_0)/G$ is a commutative group because $(\sigma_1 \sigma_2)(\sigma_2 \sigma_1)^{-1} = (\sigma_1 \sigma_2)(\sigma_1^{-1} \sigma_2^{-1}) \in G$ for any loops σ_1, σ_2 in X at x_0 . Thus $[\sigma_1]_G [\sigma_2]_G = [\sigma_1 \sigma_2]_G = [\sigma_2 \sigma_1]_G = [\sigma_2]_G [\sigma_1]_G$.

Hence, we may reorder the factors in the product for $[\gamma]_G$ to get

$$[\gamma]_G = \prod_{i=1}^I [\beta_{j_i}]_G^{\exp_\gamma(\alpha_i)} = \prod_{i=1}^I [\beta_{j_i}]_G^0 = [e_{x_0}]_G$$

in $\pi_1(X, x_0)/G$ where each $[\beta_{j_i}]_G$ in the product above represent a distinct element of $\pi_1(X, x_0)/G$. Namely, each β_{j_i} in the sum is a representative of all the β_j which are homotopic to β_{j_i} , or equivalently, α_{j_i} is a representative of all the α_j which are homotopic to α_{j_i} . We therefore have that $[\gamma] \in G$.

Now that this little claim is proved, we go back to the proof that $\mathcal{N}(\mathcal{X}) = G$.

If $\mathcal{X}([\gamma]) = 0$, then $\check{\gamma} \in B_1(X; \mathbb{R})$. Thus

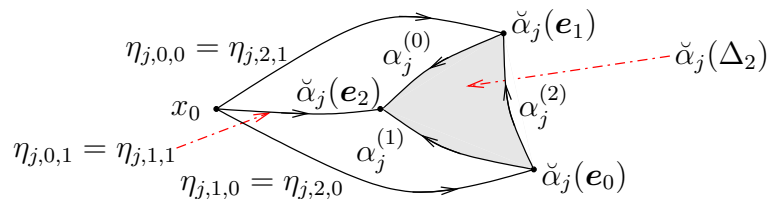
$$\begin{aligned} \check{\gamma} &= \partial_2 \left(\sum_{j=1}^J b_j \check{\sigma}_j \right) = \sum_{j=1}^J b_j \left(\underbrace{\check{\alpha}_j \circ F_2^0}_{=\check{\alpha}_j^{(0)}} - \underbrace{\check{\alpha}_j \circ F_2^1}_{=\check{\alpha}_j^{(1)}} + \underbrace{\check{\alpha}_j \circ F_2^2}_{=\check{\alpha}_j^{(2)}} \right) \\ &= \sum_{j=1}^J b_j \left(\check{\alpha}_j^{(0)} - \check{\alpha}_j^{(1)} + \check{\alpha}_j^{(2)} \right) \end{aligned} \quad (8.4.18)$$

for some singular 2-simplices $\check{\alpha}_j$ and $b_j \in \mathbb{Z}$. It follows that the sum of the coefficients of $\check{\gamma}$ in the last sum in (8.4.18) is 1 and the sum of the coefficients for the other singular 1-simplices in the last sum in (8.4.18) is 0.

We choose the paths $\eta_{j,k,i}$ for $0 \leq i \leq 1$, $0 \leq k \leq 2$ and $1 \leq j \leq J$ as we have done in (iii); namely, $\eta_{j,k,1}$ is a path in X from x_0 to $(\check{\alpha}_j^{(k)} \circ F_1^0)(e_0)$ and $\eta_{j,k,0}$ is a path in X from x_0 to $(\check{\alpha}_j^{(k)} \circ F_1^1)(e_0)$ for $0 \leq k \leq 2$ and $1 \leq j \leq J$. Moreover, paths η_{j_1,k_1,i_1} and η_{j_2,k_2,i_2} that end at the same point are the same path, and paths $\eta_{j,k,i}$ that end at x_0 are in fact equal to the constant path e_{x_0} . Note that

$$\begin{aligned} \eta_{j,0,1}(1) &= (\check{\alpha}_j^{(0)} \circ F_1^0)(e_0) = \check{\alpha}_j^{(0)}(e_1) = (\check{\alpha}_j \circ F_2^0)(e_1) = \check{\alpha}_j(e_2) , \\ \eta_{j,0,0}(1) &= (\check{\alpha}_j^{(0)} \circ F_1^1)(e_0) = \check{\alpha}_j^{(0)}(e_0) = (\check{\alpha}_j \circ F_2^0)(e_0) = \check{\alpha}_j(e_1) , \\ \eta_{j,1,1}(1) &= (\check{\alpha}_j^{(1)} \circ F_1^0)(e_0) = \check{\alpha}_j^{(1)}(e_1) = (\check{\alpha}_j \circ F_2^1)(e_1) = \check{\alpha}_j(e_2) , \\ \eta_{j,1,0}(1) &= (\check{\alpha}_j^{(1)} \circ F_1^1)(e_0) = \check{\alpha}_j^{(1)}(e_0) = (\check{\alpha}_j \circ F_2^1)(e_0) = \check{\alpha}_j(e_0) , \\ \eta_{j,2,1}(1) &= (\check{\alpha}_j^{(2)} \circ F_1^0)(e_0) = \check{\alpha}_j^{(2)}(e_1) = (\check{\alpha}_j \circ F_2^2)(e_1) = \check{\alpha}_j(e_1) \text{ and} \\ \eta_{j,2,0}(1) &= (\check{\alpha}_j^{(2)} \circ F_1^1)(e_0) = \check{\alpha}_j^{(2)}(e_0) = (\check{\alpha}_j \circ F_2^2)(e_0) = \check{\alpha}_j(e_0) . \end{aligned}$$

All this information is summarized in the following figure.



Let $\beta_{j,k} = \eta_{j,k,0} \alpha_j^{(k)} \eta_{j,k,1}^{-1}$ for $0 \leq k \leq 2$ and $1 \leq j \leq J$, and $\beta_j = \beta_{j,0} \beta_{j,1}^{-1} \beta_{j,2}$ for $1 \leq j \leq J$. We have

$$\begin{aligned} \beta_j &= \beta_{j,0} \beta_{j,1}^{-1} \beta_{j,2} = (\eta_{j,0,0} \alpha_j^{(0)} \eta_{j,0,1}^{-1}) (\eta_{j,1,0} \alpha_j^{(1)} \eta_{j,1,1}^{-1})^{-1} (\eta_{j,2,0} \alpha_j^{(2)} \eta_{j,2,1}^{-1}) \\ &= \eta_{j,0,0} \alpha_j^{(0)} (\eta_{j,0,1}^{-1} \eta_{j,1,1}) (\alpha_j^{(1)})^{-1} (\eta_{j,1,0}^{-1} \eta_{j,2,0}) \alpha_j^{(2)} \eta_{j,2,1}^{-1} \\ &\sim \eta_{j,0,0} \alpha_j^{(0)} (\alpha_j^{(1)})^{-1} \alpha_j^{(2)} \eta_{j,2,1}^{-1} \sim e_{x_0} \end{aligned} \quad (8.4.19)$$

because $\eta_{j,0,1}^{-1} \eta_{j,1,1} \sim e_{x_0}$ and $\eta_{j,1,0}^{-1} \eta_{j,2,0} \sim e_{x_0}$. We also use the fact that $\hat{\alpha}_j(\Delta_2)$ is contractible to $\check{\alpha}_j(\mathbf{e}_1)$, because Δ_2 is contractible to $\mathbf{e}_1$, to obtain the second homotopy in (8.4.19).

Hence $\xi = \prod_{j=1}^J \beta_j^{b_j} \sim e_{x_0}$. It follows that $\gamma \xi^{-1} \sim \gamma$ and therefore $[\gamma] [\xi^{-1}] = [\gamma]$ in $\pi_1(X, x_0)$.

However, if α is one of the factors of γ , then we have that $\exp_\xi(\alpha) = 1$ if $\alpha = \gamma$ and $\exp_\xi(\alpha) = 0$ otherwise. This follows from the fact that the sum of the coefficients of $\check{\alpha}$ in the last sum in (8.4.18) is equal to $\exp_\xi(\alpha)$ because $\mathcal{X} : \pi_1(X, x_0) \rightarrow H_1(X; \mathbb{Z})$ is a group homomorphism, and the paragraph following (8.4.18). Moreover $\exp_\xi(\alpha) = 0$ if α is one of the $\eta_{j,k,i}$ because $\eta_{j,k,i}^{-1}$ is present in ξ if $\eta_{j,k,i}$ is present in ξ . Therefore $\exp_{\gamma\xi^{-1}}(\sigma) = 0$ for all factors α in $\gamma\xi^{-1}$. It follows from the claim that $[\gamma] = [\gamma\xi^{-1}]$ is in the commutator subgroup G of $\pi_1(X, x_0)$. ■

Corollary 8.4.47

Suppose that X is a path-connected topological space and $x_0 \in X$, then the map $\mathcal{X}$ defined in the previous theorem is a group isomorphism between $\pi_1(X)$ and $H_1(X; \mathbb{Z})$ if and only if $\pi_1(X)$ is an abelian group.

This corollary follows from the fact that a group is commutative if its commutator subgroup is trivial; namely, it is composed of the unit or identity element of the group only.

It follows from the previous corollary and Proposition 8.1.26 that $H_1(S^1; \mathbb{Z}) \cong \pi_1(S^1) \cong \mathbb{Z}$. Moreover, we get from Example 8.1.28 that $H_1(\mathbb{T}^2; \mathbb{Z}) \cong \pi_1(\mathbb{T}^2) \cong \mathbb{Z}^2$ where, as usual, $\mathbb{T}^2$ denotes the torus in $\mathbb{R}^3$.

8.4.6 Relative Homology

Though our goal is not to provide a full introduction to algebraic topology, we feel that we cannot at least mention some of the important concepts and results of this subject. The readers should consider this subsection as an invitation to pursue their study of algebraic topology.

Let X be a topological space and R be an integral domain. Moreover, suppose that A is a subset of X . We may therefore consider the R -module $S_k(A; R)$. It is the submodule of $S_k(X; R)$ consisting of all the singular k -chains that are linear combinations of singular k -simplices with images in A .

Since the boundary operator $\partial_k : S_k(X; R) \rightarrow S_{k-1}(X; R)$ maps $S_k(A; R)$ into $S_{k-1}(A; R)$, we can define the operator $\bar{\partial}_k : S_k(X; R)/S_k(A; R) \rightarrow S_{k-1}(X; R)/S_{k-1}(A; R)$ as $\bar{\partial}_k([c]_{X,A}) = [\partial_k c]_{X,A}$ for all $[c]_{X,A} \in S_k(X; R)/S_k(A; R)$, where $[c]_{X,A}$ denotes the equivalence class in $S_k(X; R)/S_k(A; R)$ associated to $c \in S_k(X; R)$. We get the following commutative diagram.

$$\begin{array}{ccc} S_k(X; R) & \xrightarrow{\pi_k} & S_k(X; R)/S_k(A; R) \\ \downarrow \partial_k & & \downarrow \bar{\partial}_k \\ S_{k-1}(X; R) & \xrightarrow{\pi_{k-1}} & S_{k-1}(X; R)/S_{k-1}(A; R) \end{array}$$

where $\pi_q : S_q(X; R) \rightarrow S_q(X; R)/S_q(A; R)$ is the projection defined by $\pi_q(c) = [c]_{X,A}$ for $c \in S_q(X; R)$.

It follows from $\partial_{k-1}\partial_k = 0$ that $\bar{\partial}_{k-1}\bar{\partial}_k = 0$.

Definition 8.4.48

Let A be a subset of a topological space X and R be an integral domain. The **relative (singular) k^{th} homology module** of $X \bmod A$ is the R -module defined as $H_k(X, A; R) = \mathcal{N}(\bar{\partial}_k)/\mathcal{R}(\bar{\partial}_{k+1})$. The equivalence class in $H_k(X, A; R)$ associated to $c \in S_k(X; R)$ is denoted $[c]_{X,A}$.

It is possible to describe more explicitly the elements of $H_k(X, A; R)$. For that, we need a couple of definitions.

Definition 8.4.49

Let A be a subset of a topological space X and R be an integral domain. Given $c_1, c_2 \in S_k(X; R)$, we say that c_1 and c_2 are **homologous in $X \bmod A$** if there exists $\tilde{c} \in S_{k+1}(X; R)$ such that $c_1 - c_2 - \partial_{k+1}(\tilde{c}) \in S_k(A; R)$. We then write that $c_1 \stackrel{h}{\sim} c_2$ in $X \bmod A$.

Definition 8.4.50

Elements of $Z_k(X, A; R) = \{c \in S_k(X; R) : \partial_k(c) \in S_{k-1}(A; R)\}$ are called **relative k -cycle** of $X \bmod A$ and elements of $B_k(X, A; R) = \{c \in S_k(X; R) : c \stackrel{h}{\sim} 0 \text{ in } X \bmod A\}$ are called **relative k -boundary** of $X \bmod A$.

We have that $Z_k(X, A; R) \supset Z_k(X; R)$ and $B_k(X, A; R) \supset B_k(X; R) \cup S_k(A; R)$. The reader may be wondering why the boundary operator is not used in the definition of $B_k(X, A; R)$. In fact, it is used when referring to homologous equivalence in $X \bmod A$ as it is illustrated in the proof of the following proposition.

Proposition 8.4.51

Let A be a subset of a topological space X and R be an integral domain. Then $H_k(X, A; R) \cong Z_k(X, A; R)/B_k(X, A; R)$.

Proof.

We have

$$\begin{aligned}\mathcal{N}(\bar{\partial}_k) &= \{[[c]]_{X,A} : \bar{\partial}_k([[c]]_{X,A}) = [[\partial_k c]]_{X,A} = [[0]]_{X,A}\} \\ &= \{[[c]]_{X,A} : c \in S_k(X, R) \text{ and } \partial_k(c) \in S_k(A; R)\} = Z_k(X, A; R)/S_k(A; R)\end{aligned}$$

and

$$\begin{aligned}\mathcal{R}(\bar{\partial}_{k+1}) &= \{[[c]]_{X,A} : [[c]]_{X,A} = \bar{\partial}_{k+1}([[b]]_{X,A}) \text{ for some } [[b]]_{X,A} \in S_{k+1}(X; R)/S_{k+1}(A; R)\} \\ &= \{[[c]]_{X,A} : [[c - \partial_{k+1}(b)]]_{X,A} = [[0]]_{X,A} \text{ for some } [[b]]_{X,A} \in S_{k+1}(X; R)/S_{k+1}(A; R)\} \\ &= \{[[c]]_{X,A} : c \in S_k(X; R) \text{ and } c - \partial_{k+1}(b) \in S_k(A; R) \text{ for some } b \in S_{k+1}(X; R)\} \\ &= \{[[c]]_{X,A} : c \in S_k(X; R) \text{ and } c \stackrel{h}{\sim} 0 \text{ in } X \pmod{A}\} = B_k(X, A; R)/S_k(A; R) .\end{aligned}$$

Hence

$$\begin{aligned}\mathcal{N}(\bar{\partial}_k)/\mathcal{R}(\bar{\partial}_k) &= (Z_k(X, A; R)/S_k(A; R))/(B_k(X, A; R)/S_k(A; R)) \\ &\cong Z_k(X, A; R)/B_k(X, A; R) .\end{aligned}$$

Definition 8.4.52

Let A be a subset of a topological space X and B be a subset of a topological space Y . If $f : X \rightarrow Y$ is a function such that $f(A) \subset B$, then we write $f : (X, A) \rightarrow (Y, B)$.

Suppose that $f : (X, A) \rightarrow (Y, B)$ is a continuous function and R is an integral domain. Then $S_k(f) : S_k(X; R) \rightarrow S_k(X; R)$ and $S_k(f) : S_k(A; R) \rightarrow S_k(B; R)$. This implies that $\tilde{S}_k(f) : S_k(X; R)/S_k(A; R) \rightarrow S_k(X; R)/S_k(B; R)$ defined by $\tilde{S}_k(f)([[c]]_{X,A}) = [[c]]_{Y,B}$ for $c \in S_k(X; R)$ is well defined. The following proposition can then be easily proved by the reader.

Proposition 8.4.53

Suppose that A is a subset of a topological space X , B is a subset of a topological space Y and R is an integral domain. If $f : (X, A) \rightarrow (Y, B)$ is continuous function, then $H_k(f) : H_k(X, A; R) \rightarrow H_k(Y, B; R)$ defined by $H_k(f)([c]_{X,A}) = [f \circ c]_{Y,B}$ for all $[c]_{X,A} \in H_k(X, A; R)$ is a homomorphism that satisfies the following statements.

1. $H_{k-1}(f) \circ \bar{\partial}_k = \bar{\partial}_k \circ H_k(f)$.
2. If C is a subset of a topological space Z and $g : (Y, B) \rightarrow (Z, C)$ is a continuous function, then $H_k(g) \circ H_k(f) = H_k(g \circ f)$.

Suppose that A and B are two subsets of a topological space X such that $A \subset B$, and R is an integral domain. The function $\text{Id}_X : X \rightarrow X$ is a continuous function such that $\text{Id}_X : (X, \emptyset) \rightarrow (X, A)$. Hence $H_k(\text{Id}_X) : H_k(X, A; R) \rightarrow H_k(X, B; R)$ is well defined by $H_k(\text{Id}_X)([c]_{X,A}) = [\text{Id}_X \circ c]_{X,B} = [c]_{X,B}$ for all $[c]_{X,A} \in H_k(X, A; R)$. If $A = \emptyset$, then we get that $H_k(\text{Id}_X) : H_k(X; R) = H_k(X, \emptyset; R) \rightarrow H_k(X, B; R)$. Another function that will play

a significant role in this section is the inclusion map $\iota : A \rightarrow X$ for $A \subset X$. We have the following result.

Proposition 8.4.54

Suppose that A is a subset of a topological space X and R is an integral domain. Then the sequence.

$$H_k(A; R) \xrightarrow{H_k(\iota)} H_k(X; R) \xrightarrow{H_k(\text{Id}_X)} H_k(X, A; R)$$

is exact.

Proof.

Since $\text{Id}_X \circ \iota : A \rightarrow X$ is a continuous function such that $(\text{Id}_X \circ \iota)(\emptyset) = \emptyset$, we have that $H_k(\text{Id}_X \circ \iota) : H_k(A; R) = H_k(A, \emptyset; R) \rightarrow H_k(X, A; R)$ is well defined. Moreover, $H_k(\text{Id}_X \circ \iota)([c]_A) = [0]_{X,A} \in H_k(X, A; R)$ for all $[c]_A \in H_k(A; R)$ because

$$c \in S_k(A; R) \Rightarrow c \stackrel{h}{\sim} 0 \text{ in } X \text{ mod } A \Rightarrow c \in B_k(X, A; R) \Rightarrow [[c]]_{X,A} \in \mathcal{R}(\bar{\partial}_{k+1}) .$$

Thus $H_k(\text{Id}_X) \circ H_k(\iota) = H_k(\text{Id}_X \circ \iota) = 0$. Hence $\mathcal{R}(H_k(\iota)) \subset \mathcal{N}(H_k(\text{Id}_X))$.

To prove the reverse inclusion, consider $[c]_X \in \mathcal{N}(H_k(\text{Id}_X))$. We get from $H_k(\text{Id}_X)([c]_X) = [c]_{X,A} = [0]_{X,A} \in H_k(X, A; R)$ that $[[c]]_{X,A} \in \mathcal{R}(\bar{\partial}_{k+1})$. Thus $[[c]]_{X,A} = \overline{\partial_{k+1}}([b])_{X,A} = [[\partial_{k+1}b]]_{X,A}$ for some $b \in S_{k+1}(X; R)$. Hence $c - \partial_{k+1}(b) \in S_k(A; R)$ for some $b \in S_{k+1}(X; R)$. Since $\partial_k(c) = 0$ because $[c]_X \in H_k(X; R)$, we have that $c - \partial_{k+1}(b)$ is a k -cycle in $S_k(A; R)$. Thus $[c - \partial_{k+1}(b)]_A \in H_k(A; R)$ and $H_k(\iota)([c - \partial_{k+1}(b)]_A) = [c - \partial_{k+1}(b)]_X = [c]_X \in H_k(X; R)$ proving that $[c]_X \in \mathcal{R}(H_k(\iota))$. ■

Many of the results about singular homology modules have matching results for regular homology modules. For instance, Proposition 8.4.10 becomes the following proposition.

Proposition 8.4.55

Suppose that A is a subset of a topological space X . If the topological space X is the union of distinct path-connected components X_j for $j \in J$ and $A_j = A \cap X_j$ for all $j \in J$, then $H_k(X, A; R) = \bigoplus_{j \in J} H_k(X_j, A_j; R)$.

However, there are some differences. For instance, Proposition 8.4.11 is replaced by the following proposition.

Proposition 8.4.56

Suppose that A is a non-empty subset of a path-connected topological space X . Then $H_0(X, A; R) = 0$.

Proof.

Choose $x_0 \in A$ and, for each $x \in X$, choose a path $\tilde{\sigma}_x$ from x_0 to x . For $x \in X$, let σ_x be

the singular 1-simplex defined by $\sigma_x((1-t)e_0 + te_1) = \tilde{\sigma}_x(t)$ for $0 \leq t \leq 1$. As usual, we define the singular 0-simplex e_x by $e_x(e_0) = x$.

Suppose that c is a relative 0-cycle in $X \bmod A$. We have that $c = \sum_{j \in J} a_j e_{x_j}$ for some $a_j \in R$ and $x_j \in X$, where J is a finite subset of $\mathbb{N}^+$. Since

$$\partial_1 \left(\sum_{j \in J} a_j \sigma_{x_j} \right) = \sum_{j \in J} a_j (\sigma_{x_j} \circ F_1^0 - \sigma_{x_j} \circ F_1^1) = \sum_{j \in J} a_j (e_{x_j} - e_{x_0}) = c - \sum_{j \in J} a_j e_{x_0},$$

we have that $c \stackrel{h}{\sim} \sum_{j \in J} a_j e_{x_0} \in S_0(A; R)$ in $X \bmod A$. Thus c is a relative k -boundary in $X \bmod A$. Therefore $[c]_{X,A} = [0]_{X,A} \in H_k(X, A; R)$. ■

We extend the definition of homotopic maps between two topological spaces to the context of relative homology modules.

Definition 8.4.57

Let A be a subset of a topological space X , B be a subset of a topological space Y and R be an integral domain. Suppose that $f : (X, A) \rightarrow (Y, B)$ and $g : (X, A) \rightarrow (Y, B)$ are two continuous functions. We say that f is **homotopic** to g if there exists a continuous mapping $H : X \times [0, 1] \rightarrow Y$ such that $H(x, 0) = f(x)$ and $H(x, 1) = g(x)$ for all $x \in X$, and $H(A, t) \subset B$ for $0 \leq t \leq 1$. We write $f \stackrel{rel}{\sim} g$.

The next proposition could be proved using the same technique used to prove Corollary 8.4.40. This obviously requires the use of chain complexes for relative homology modules.

Proposition 8.4.58

Let A be a subset of a topological space X , B be a subset of a topological space Y and R be an integral domain. Suppose that $f : (X, A) \rightarrow (Y, B)$ is homotopic to $g : (X, A) \rightarrow (Y, B)$. Then $H_k(f) = H_k(g)$.

There are many interesting results about relative homology modules. We are going to present some of these results. The reader should consult [9, 21] for more information.

The map $C_k : H_k(X, A; R) \rightarrow H_{k-1}(A; R)$ defined by $C_k([c]_{X,A}) = [\partial_k(c)]_A$ is well defined because

$$[c]_{X,A} \in H_k(X, A; R) \Rightarrow [[c]]_{X,A} \in \mathcal{N}(\bar{\partial}_k) \Rightarrow \partial_k(c) \in S_{k-1}(A; R).$$

Thus $\partial_k(c) \in Z_{k-1}(A; R)$ because $\partial_k(c) \in S_{k-1}(A; R)$ and $\partial_{k-1}(\partial_k(c)) = 0$. Hence, $[\partial_k(c)]_A \in H_{k-1}(A; R)$.

Definition 8.4.59

The map C_k defined above is called a **connecting homomorphism**

This definition is justified by the following result.

Proposition 8.4.60

Let A be a subset of a topological space X and R be an integral domain. Then the sequence

$$\begin{aligned} & \xrightarrow{C_{k+1}} H_k(A; R) \xrightarrow{H_k(\iota)} H_k(X; R) \xrightarrow{H_k(\text{Id}_X)} H_k(X, A; R) \\ & \xrightarrow{C_k} H_{k-1}(A; R) \xrightarrow{H_{k-1}(\iota)} H_{k-1}(X; R) \xrightarrow{H_{k-1}(\text{Id}_X)} \end{aligned} \quad (8.4.20)$$

is a long exact sequence.

Proof.

We have already from Proposition 8.4.54 that $\mathcal{R}(H_k(\iota)) = \mathcal{N}(H_k(\text{Id}_X))$. The only things that we have to prove is that $\mathcal{R}(C_k) = \mathcal{N}(H_{k-1}(\iota))$ and $\mathcal{R}(H_k(\text{Id}_X)) = \mathcal{N}(C_k)$.

i) $\mathcal{R}(H_k(\text{Id}_X)) \subset \mathcal{N}(C_k)$. Consider $[c]_X \in H_k(X; R)$. We have that $\partial_k(c) = 0$. Thus $H_k(\text{Id}_X)([c]_X) = [c]_{X,A}$ with $\partial_k(c) = 0$. Hence $C_k([c]_{X,A}) = [\partial_k(c)]_A = [0]_A \in H_{k-1}(A; R)$.

ii) $\mathcal{R}(H_k(\text{Id}_X)) \supset \mathcal{N}(C_k)$. Suppose that $[c]_{X,A} \in H_k(X, A; R)$ is such that $C_k([c]_{X,A}) = [\partial_k(c)]_A = [0]_A \in H_{k-1}(A; R)$. Therefore $\partial_k(c) \in B_{k-1}(A; R)$; namely, $\partial_k(c)$ is a $(k-1)$ -boundary in $S_{k-1}(A; R)$. Hence, there exists $b \in S_k(A; R)$ such that $\partial_k(b) = \partial_k(c)$. Since $c - b \in S_k(X; R)$ with $\partial_k(c - b) = \partial_k(c) - \partial_k(b) = 0$, we have that $c - b$ is a k -cycle in $S_k(X; R)$ and so $[c - b]_X \in H_k(X; R)$. We finally have that $H_k(\text{Id}_X)([c - b]_X) = [c - b]_{X,A} = [c]_{X,A}$ because $b \in S_k(A; R)$.

iii) $\mathcal{R}(C_k) \subset \mathcal{N}(H_{k-1}(\iota))$. Consider $[c]_{X,A} \in H_k(X, A; R)$. Then $C_k([c]_{X,A}) = [\partial_k(c)]_A \in H_{k-1}(A; R)$. Hence

$$H_{k-1}(\iota)([\partial_k(c)]_A) = [S_{k-1}(\iota)(\partial_k(c))]_X = [\partial_k(S_k(\iota)(c))]_X = [0]_X \in H_{k-1}(X; R)$$

because $\partial_k(S_k(\iota)(c)) \in B_{k-1}(X; R)$ since $S_k(\iota)(c) \in S_k(X; R)$.

iv) $\mathcal{R}(C_k) \supset \mathcal{N}(H_{k-1}(\iota))$. Suppose that $[c]_A \in H_{k-1}(A; R)$ is such that $H_{k-1}(\iota)([c]_A) = [S_{k-1}(\iota)(c)]_X = [0]_X \in H_{k-1}(X; R)$. Since $S_{k-1}(\iota)(c) = c \in B_{k-1}(X; R)$, there exists $b \in S_k(X; R)$ such that $\partial_k(b) = c$. Since $c \in S_{k-1}(A; R)$, we have that $\partial_k(b) \in S_{k-1}(A; R)$. Thus $b \in Z_k(X, A; R)$. Finally, we get that $C_k([b]_{X,A}) = [\partial_k(b)]_A = [c]_A$. ■

Note that there is a much shorter proof of the previous proposition using the Zig-Zag Lemma. The reader can find it in (ii) the proof of Lemma 8.4.85 given later. However, we choose to give this stand alone proof since it illustrates nicely the basic techniques used in algebraic topology.

Definition 8.4.61

The sequence in (8.4.20) is called a **homology sequence** for the pair (X, A) .

If B is a subset of a topological space Y and $f : (X, A) \rightarrow (Y, B)$ is a continuous function, then the following diagram commutes.

$$\begin{array}{ccccccc}
 \xrightarrow{C_{k+1}} & H_k(A; R) & \xrightarrow{H_k(\iota)} & H_k(X; R) & \xrightarrow{H_k(\text{Id}_X)} & H_k(X, A; R) & \xrightarrow{C_k} \\
 & \downarrow H_k(f) & & \downarrow H_k(f) & & \downarrow H_k(f) & \\
 \xrightarrow{C_{k+1}} & H_k(B; R) & \xrightarrow{H_k(\iota)} & H_k(Y; R) & \xrightarrow{H_k(\text{Id}_X)} & H_k(Y, B; R) & \xrightarrow{C_k}
 \end{array}$$

Proposition 8.4.62

Suppose that A is a retract of a topological space X and R is an integral domain. Then the sequence.

$$0 \longrightarrow H_k(A; R) \xrightarrow{H_k(\iota)} H_k(X; R) \xrightarrow{H_k(\text{Id}_X)} H_k(X, A; R) \longrightarrow 0$$

is a short exact sequence.

Proof.

We have already from Proposition 8.4.54 that $\mathcal{R}(H_k(\iota)) = \mathcal{N}(H_k(\text{Id}_X))$.

Since A is a retract of X , there exists a map $r : X \rightarrow A$ such that $r \circ \iota = \text{Id}_A$ where $\iota : A \rightarrow X$ is the inclusion map. It follows that $H_k(r) \circ H_k(\iota) = H_k(\text{Id}_A) = \text{Id}_{H_k(A; R)}$. Thus $\mathcal{N}(H_k(\iota)) = \{[0]_A\}$. It is therefore justify to introduce $0 \rightarrow H_k(A; R)$.

If we consider the long exact sequence of Proposition 8.4.60 and apply the previous reasoning to $k - 1$, then we get that $\mathcal{R}(C_k) = \mathcal{N}(H_{k-1}(\iota)) = \{[0]\}$. Thus, we may also introduce $H_k(X, A; R) \rightarrow 0$. ■

The fact that $\mathcal{R}(H_k(\iota)) = \mathcal{N}(H_k(\text{Id}_X))$ in the previous proposition is not unique to A a retract of X . We have already proved it in Proposition 8.4.54 in the general case of A a subset of X .

We have shown in the proof of Proposition 8.4.62 that

$$0 \longrightarrow H_k(A; R) \xrightarrow{H_k(\iota)} H_k(X; R) \xrightarrow{H_k(\text{Id}_X)} H_k(X, A; R) \longrightarrow 0$$

is in fact a split short exact sequence if A is a retract of X because $H_k(r) \circ H_k(\iota) = \text{Id}_{H_k(A; R)}$, where $r : X \rightarrow A$ is the map such that $r \circ \iota = \text{Id}_A$ by definition of a retract. Therefore $H_k(X; R) \cong H_k(A; R) \oplus H_k(X, A; R)$.

Proposition 8.4.63

Suppose that A is a contractible subset a path-connected topological space X and R is an integral domain. Then the sequence.

$$0 \longrightarrow H_k(X; R) \xrightarrow{H_k(\text{Id}_X)} H_k(X, A; R) \longrightarrow 0$$

is an exact sequence for $k > 0$. In particular, $H_k(X; R) \cong H_k(X, A; R)$ for $k > 0$.

Proof.

$k > 1$) Since A is contractible, we get from Corollary 8.4.41 that $H_k(A; R) = 0$. It follows from Proposition 8.4.60 that

$$0 \xrightarrow{H_k(\iota)} H_k(X; R) \xrightarrow{H_k(\text{Id}_X)} H_k(X, A; R) \xrightarrow{C_k} 0$$

is exact. Hence $\mathcal{N}(H_k(\text{Id}_X)) = \mathcal{R}(H_k(\iota)) = \{[0]_X\}$ and $\mathcal{R}(H_k(\text{Id}_X)) = \mathcal{N}(C_k) = H_k(X, A; R)$. Thus $H_k(X; R) \cong H_k(X, A; R)$.

$k = 1$) We still have that $H_1(A; R) = 0$ because A is contractible. We also have from Proposition 8.4.56 that $H_0(X, A; R) = 0$ because X is path-connected. Hence, it follows from Proposition 8.4.60 that

$$0 \xrightarrow{H_1(\iota)} H_1(X; R) \xrightarrow{H_1(\text{Id}_X)} H_1(X, A; R) \xrightarrow{C_1} H_0(A; R) \xrightarrow{H_0(\iota)} H_0(X; R) \xrightarrow{H_0(\text{Id}_X)} 0$$

Since $\mathcal{N}(H_1(\text{Id}_X)) = \mathcal{R}(H_1(\iota)) = \{[0]_X\}$, we have that $H_1(\text{Id}_X)$ is one-to-one. We have that $H_0(A; R) \cong R$ according to Corollary 8.4.41 and $H_0(X; R) \cong R$ according to Propositions 8.4.11. Therefore, since $\mathcal{R}(H_0(\iota)) = \mathcal{N}(H_0(\text{Id}_X)) = H_0(X; R) \cong R$, we get that $\mathcal{N}(H_0(\iota)) = \{[0]_A\}$. It follows that $\mathcal{R}(C_1) = \mathcal{N}(H_0(\iota)) = \{[0]_A\}$. Hence, we get that

$$0 \xrightarrow{H_1(\iota)} H_1(X; R) \xrightarrow{H_1(\text{Id}_X)} H_1(X, A; R) \xrightarrow{C_1} 0$$

is exact. As in (i), this implies that $H_1(X; R) \cong H_1(X, A; R)$. ■

The previous result is also true for $k = 0$ if we consider reduced homology.

Proposition 8.4.64

Suppose that A is a contractible subset of a topological space X and R is an integral domain. Then the sequence.

$$0 \longrightarrow H_k^\#(X; R) \xrightarrow{H_k(\text{Id}_X)} H_k^\#(X, A; R) \longrightarrow 0$$

is an exact sequence for $k \geq 0$. In particular, $H_k^\#(X; R) \cong H_k^\#(X, A; R)$ for $k \geq 0$.

Proof.

Since A is contractible, we have that from Corollary 8.4.41 that $H_k^\#(A; R) = H_k(A; R) = 0$ for $k > 0$. In the reduced homology, we also have that $H_0^\#(A; R) = 0$.

It follows from Proposition 8.4.60 (which is also true for reduced homology) that

$$0 \xrightarrow{H_k(\iota)} H_k^\#(X; R) \xrightarrow{H_k(\text{Id}_X)} H_k^\#(X, A; R) \xrightarrow{C_k} 0$$

is exact for all k . Hence $\mathcal{N}(H_k(\text{Id}_X)) = \mathcal{R}(H_k(\iota)) = \{[0]_X\}$ and $\mathcal{R}(H_k(\text{Id}_X)) = \mathcal{N}(C_k) = H_k^\#(X, A; R)$ for all k . Thus $H_k^\#(X; R) \cong H_k^\#(X, A; R)$. ■

8.4.7 Excision

As for the previous subsection, the readers should consider this subsection as an invitation to pursue their study of algebraic topology.

The following definition has very profound applications.

Definition 8.4.65

Let A be a subset of a topological space X , U be a subset of A and R be an integral domain. We say that the inclusion map $\iota : (X \setminus U, A \setminus U) \rightarrow (X, A)$ is an **excision** if $H_k(\iota) : H_k(X \setminus U, A \setminus U; R) \rightarrow H_k(X, A; R)$ is an isomorphism for all k . We then say that U can be **excised** from (X, A) .

The following theorem is one of the fundamental theorems concerning excisions.

Theorem 8.4.66 (Excision)

Let A be a subset of a topological space X , U be a subset of A and R be an integral domain. If $\overline{U} \subset A^\circ$, then U can be excised from (X, A) .

Recall that $\overline{U}$ denotes the closure of the set U and A° denotes the interior of the set A .

It is interesting to note that the proof of Theorem 8.4.66 requires a singular version of the theory of simplicial approximation introduced in Section 8.2.3. We briefly review this theory and then prove Theorem 8.4.66. The first concept to generalize is the concept of barycentric subdivision.

Definition 8.4.67

Let X be an affine space and $\delta : \Delta_k \rightarrow X$ be an affine function. Given $b \in X$, the **join** $J_{k,b}(\delta)$ is the affine function from Δ_{k+1} to X defined by

$$J_{k,b}(\delta) \left(\sum_{i=0}^{k+1} a_i e_i \right) = a_0 b + \sum_{i=1}^{k+1} a_i \delta(e_{i-1})$$

for $\sum_{i=0}^{k+1} a_i = 1$ with $a_i \geq 0$ for all i (Figure 8.10). Given a linearly combination $c = \sum_{j \in J} a_j \delta_j$ where $J \subset \mathbb{N}$ is a finite set, $\delta_j : \Delta_k \rightarrow X$ is an affine function and $a_j \in \mathbb{R}$

for all $j \in J$, we define $J_{k,b}(c)$ as $J_{k,b}(c) = \sum_{j \in J} a_j J_{k,b}(\delta_j)$.

If $\delta : \Delta_0 \rightarrow X$, then

$$\begin{aligned} (\partial_1(J_{0,b}(\delta)))(e_0) &= (J_{0,b}(\delta) \circ F_1^0 - J_{0,b}(\delta) \circ F_1^1)(e_0) = J_{0,b}(\delta)(e_1) - J_{0,b}(\delta)(e_0) \\ &= \delta(e_0) - b. \end{aligned}$$

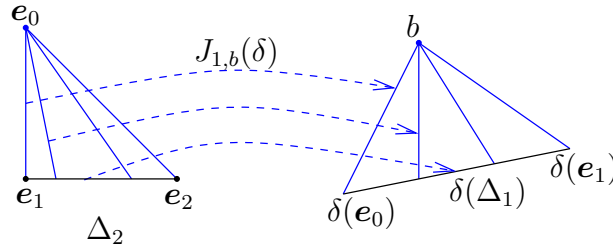


Figure 8.10: Graphical representation of $J_{1,b}(\delta)$ for an affine function singular $\delta : \Delta_1 \rightarrow X$.

Thus

$$\partial_1(J_{0,b}(\delta)) = \delta - e_b, \quad (8.4.21)$$

where $e_b : \Delta_0 \rightarrow X$ is defined by $e_b(e_0) = b$. Hence, if $c = \sum_{j \in J} a_j \delta_j$ where $\delta_j : \Delta_0 \rightarrow X$ is an affine function and $a_j \in \mathbb{R}$ for all $j \in J$, then $\partial_1(J_{0,b}(c)) = \sum_{j \in J} a_j \partial_1(J_{0,b}(\delta_j)) = \sum_{j \in J} a_j (\delta_j - e_b) = c - \left(\sum_{j \in J} a_j \right) e_b$.

If $\delta : \Delta_k \rightarrow X$ for $k > 0$ is an affine function, then

$$\begin{aligned} (\partial_{k+1}(J_{k,b}(\delta))) \left(\sum_{i=0}^k a_i e_i \right) &= \sum_{j=0}^{k+1} (-1)^j (J_{k,b}(\delta) \circ F_{k+1}^j) \left(\sum_{i=0}^k a_i e_i \right) \\ &= \sum_{j=0}^{k+1} (-1)^j J_{k,b}(\delta) \left(\sum_{i=0}^{j-1} a_i e_i + \sum_{i=j}^k a_i e_{i+1} \right) \\ &= \sum_{i=0}^k a_i \delta(e_i) + \sum_{j=1}^{k+1} (-1)^j \left(a_0 b + \sum_{i=1}^{j-1} a_i \delta(e_{i-1}) + \sum_{i=j}^k a_i \delta(e_i) \right) \\ &= \delta \left(\sum_{i=0}^k a_i e_i \right) + \sum_{j=1}^{k+1} (-1)^j \left(a_0 b + \sum_{i=1}^{j-1} a_i \delta(e_{i-1}) + \sum_{i=j}^k a_i \delta(e_i) \right) \end{aligned}$$

and

$$\begin{aligned} J_{k-1,b}(\partial_k(\delta)) \left(\sum_{i=0}^k a_i e_i \right) &= J_{k-1,b} \left(\sum_{j=0}^k (-1)^j (\delta \circ F_k^j) \right) \left(\sum_{i=0}^k a_i e_i \right) \\ &= \sum_{j=0}^k (-1)^j J_{k-1,b}(\delta \circ F_k^j) \left(\sum_{i=0}^k a_i e_i \right) = \sum_{j=0}^k (-1)^j \left(a_0 b + \sum_{i=1}^k a_i (\delta \circ F_k^j)(e_{i-1}) \right) \\ &= \sum_{j=0}^k (-1)^j \left(a_0 b + \sum_{i=1}^j a_i \delta(e_{i-1}) + \sum_{i=j+1}^k a_i \delta(e_i) \right) \end{aligned}$$

$$= - \sum_{j=1}^{k+1} (-1)^j \left(a_0 b + \sum_{i=1}^{j-1} a_i \delta(\mathbf{e}_{i-1}) + \sum_{i=j}^k a_i \delta(\mathbf{e}_i) \right).$$

Thus

$$\partial_{k+1}(J_{k,b}(\delta)) = \delta - J_{k-1,b}(\partial_k(\delta)) \quad (8.4.22)$$

for $k > 0$. Hence, if $c = \sum_{j \in J} a_j \delta_j$ where $J \subset \mathbb{N}$ is a finite set, $\delta_j : \Delta_k \rightarrow X$ is an affine function and $a_j \in \mathbb{R}$ for all $j \in J$, then $\partial_{k+1}(J_{k,b}(c)) = \sum_{j \in J} a_j \partial_{k+1}(J_{k,b}(\delta_j)) = \sum_{j \in J} a_j (\delta_j - J_{k-1,b}(\partial_k(\delta_j))) = c - J_{k-1,b}(\partial_k(c))$.

If the reader has read Subsection 8.2.4 about oriented simplicial homology, then there is a nice way to simplify all the computations that we have done above. We can use the following associative diagram.

$$\begin{array}{ccc} \text{affine functions} & & \text{oriented simplices} \\ \gamma : \Delta_k \rightarrow X & \longleftrightarrow & [[\gamma(\mathbf{e}_0), \gamma(\mathbf{e}_1), \gamma(\mathbf{e}_k)]] \\ \downarrow & & \downarrow \\ J_{k,b}(\gamma) & \longleftrightarrow & [[b, \gamma(\mathbf{e}_0), \gamma(\mathbf{e}_1), \dots, \gamma(\mathbf{e}_k)]] \\ \downarrow & & \downarrow \\ \partial_{k+1}(J_{k,b}(\gamma)) & \longleftrightarrow & \partial_{k+1}[[b, \gamma(\mathbf{e}_0), \gamma(\mathbf{e}_1), \dots, \gamma(\mathbf{e}_k)]] \end{array}$$

The relations $\partial_1(J_{0,b}(\delta)) = \delta - e_b$ and $\partial_{k+1}(J_{k,b}(\delta)) = \delta - J_{k-1,b}(\partial_k(\delta))$ for $k > 0$ then follow easily from

$$\begin{aligned} \partial_{k+1}(J_{k,b}(\delta)) &= \partial_{k+1}[[b, \gamma(\mathbf{e}_0), \gamma(\mathbf{e}_1), \dots, \gamma(\mathbf{e}_k)]] \\ &= [[\gamma(\mathbf{e}_0), \gamma(\mathbf{e}_1), \dots, \gamma(\mathbf{e}_k)]] + \sum_{j=1}^{k+1} (-1)^j [[b, \gamma(\mathbf{e}_0), \gamma(\mathbf{e}_1), \dots, \widehat{\gamma(\mathbf{e}_{j-1})}, \dots, \gamma(\mathbf{e}_k)]] \\ &= [[\gamma(\mathbf{e}_0), \gamma(\mathbf{e}_1), \dots, \gamma(\mathbf{e}_k)]] - \sum_{j=0}^k (-1)^j [[b, \gamma(\mathbf{e}_0), \gamma(\mathbf{e}_1), \dots, \widehat{\gamma(\mathbf{e}_j)}, \dots, \gamma(\mathbf{e}_k)]] \\ &= [[\gamma(\mathbf{e}_0), \gamma(\mathbf{e}_1), \dots, \gamma(\mathbf{e}_k)]] - J_{k-1,b}(\partial_k[[\gamma(\mathbf{e}_0), \gamma(\mathbf{e}_1), \dots, \gamma(\mathbf{e}_k)]])) = \delta - J_{k-1,b}(\partial_k(\delta)). \end{aligned}$$

The concept of barycentric subdivision introduced in Definition 8.2.14 is replaced by the following operator.

Definition 8.4.68

The **barycentric subdivision operator** on the space of affine functions from Δ_k to itself, denoted Sd_k , is the operator defined by

$$\text{Sd}_k(\delta) = \begin{cases} \delta & \text{if } k = 0 \\ J_{k-1,b}(\text{Sd}_{k-1}(\partial_k(\delta))) & \text{if } k > 0 \end{cases}$$

for all affine functions $\delta : \Delta_k \rightarrow \Delta_k$ where $\mathbf{b}_{(k)} = (1+k)^{-1} \sum_{j=0}^k \mathbf{e}_j$ is the barycentre of Δ_k . If $c = \sum_{j \in J} a_j \delta_j$ where $J \subset \mathbb{N}$ is a finite set, $\delta_j : \Delta_k \rightarrow \Delta_k$ is an affine function and $a_j \in \mathbb{R}$ for all $j \in J$, then $\text{Sd}_k(c)$ is defined as $\text{Sd}_k(c) = \sum_{j \in J} a_j \text{Sd}_k(\delta_j)$.

Example 8.4.69

We have sketched in Figure 8.11 the action of the Barycentric subdivision operators $\text{Sd}_1(\delta)_1$ and $\text{Sd}_2(\delta_2)$ where δ_i is the identity map on Δ_i for $i = 1, 2$. In the first case

$$\text{Sd}_1(\delta_1) = J_{0,\delta_1(\mathbf{b}_{(1)})}(\text{Sd}_0(\delta_1 \circ F_1^0) - \text{Sd}_0(\delta_1 \circ F_1^1)) = J_{0,\mathbf{b}_{(1)}}(F_1^0) - J_{0,\mathbf{b}_{(1)}}(F_1^1)$$

In the second case

$$\begin{aligned} \text{Sd}_2(\delta_2) &= J_{1,\delta_2(\mathbf{b}_{(2)})}(\text{Sd}_1(\delta_2 \circ F_2^0)) - J_{1,\delta_2(\mathbf{b}_{(2)})}(\text{Sd}_1(\delta_2 \circ F_2^1)) + J_{1,\delta_2(\mathbf{b}_{(2)})}(\text{Sd}_1(\delta_2 \circ F_2^2)) \\ &= J_{1,\mathbf{b}_{(2)}}(\text{Sd}_1(F_2^0)) - J_{1,\mathbf{b}_{(2)}}(\text{Sd}_1(F_2^1)) + J_{1,\mathbf{b}_{(2)}}(\text{Sd}_1(F_2^2)) \\ &= J_{1,\mathbf{b}_{(2)}}(J_{0,F_2^0(\mathbf{b}_{(1)})}(\text{Sd}_0(\partial_1 F_2^0))) - J_{1,\mathbf{b}_{(2)}}(J_{0,F_2^1(\mathbf{b}_{(1)})}(\text{Sd}_0(\partial_1 F_2^1))) \\ &\quad + J_{1,\mathbf{b}_{(2)}}(J_{0,F_2^2(\mathbf{b}_{(1)})}(\text{Sd}_0(\partial_1 F_2^2))) \\ &= J_{1,\mathbf{b}_{(2)}}(J_{0,F_2^0(\mathbf{b}_{(1)})}(F_2^0 \circ F_1^0)) - J_{1,\mathbf{b}_{(2)}}(J_{0,F_2^0(\mathbf{b}_{(1)})}(F_2^0 \circ F_1^1)) \\ &\quad - J_{1,\mathbf{b}_{(2)}}(J_{0,F_2^1(\mathbf{b}_{(1)})}(F_2^1 \circ F_1^0)) + J_{1,\mathbf{b}_{(2)}}(J_{0,F_2^1(\mathbf{b}_{(1)})}(F_2^1 \circ F_1^1)) \\ &\quad + J_{1,\mathbf{b}_{(2)}}(J_{0,F_2^2(\mathbf{b}_{(1)})}(F_2^2 \circ F_1^0)) - J_{1,\mathbf{b}_{(2)}}(J_{0,F_2^2(\mathbf{b}_{(1)})}(F_2^2 \circ F_1^1)) . \end{aligned}$$

♣

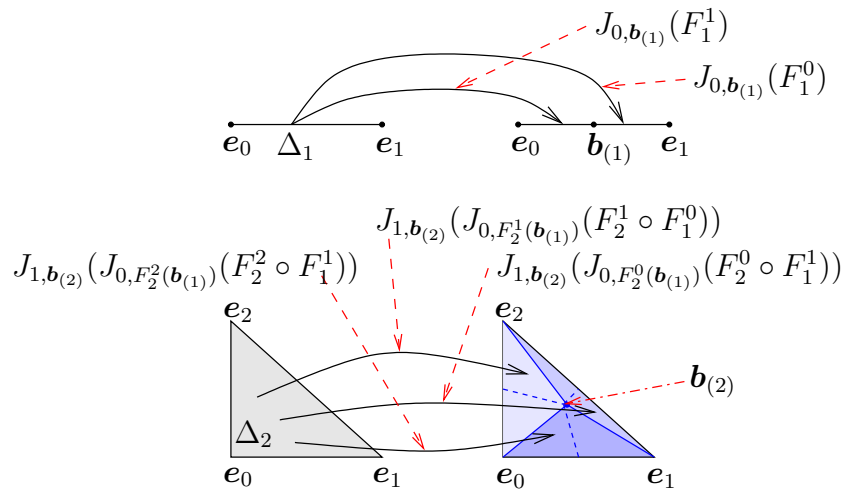


Figure 8.11: We have only sketched the action of three of the terms for $\text{Sd}_2(\delta_2)$.

We may extend the barycentric subdivision operator to $S_k(X; R)$ where X is a topological space and R is an integral domain.

Definition 8.4.70

Let X be a topological space and R be an integral domain. The **barycentric subdivision operator** on $S_k(X; R)$, denoted Sd_k , is the linear operator from $S_k(X; R)$ into itself defined by $\text{Sd}_k(\sigma) = S_k(\sigma)(\text{Sd}_k(\delta_k))$ for all singular k -simplices, where $\delta_k : \Delta_k \rightarrow \Delta_k$ is the identity map. If $c = \sum_{j \in J} a_j \sigma_j$ where $J \subset \mathbb{N}$ is a finite set, σ_j is a singular k -simplex and $a_j \in R$ for all $j \in J$, then $\text{Sd}_k(c)$ is defined as $\text{Sd}_k(c) = \sum_{j \in J} a_j \text{Sd}_k(\sigma_j)$.

Suppose that X and Y are two topological spaces and R is an integral domain. If $f : X \rightarrow Y$ is a continuous function, then we have the following commutative diagram.

$$\begin{array}{ccc} S_k(X; R) & \xrightarrow{S_k(f)} & S_k(Y; R) \\ \downarrow \text{Sd}_k & & \downarrow \text{Sd}_k \\ S_k(X; R) & \xrightarrow{S_k(f)} & S_k(Y; R) \end{array} \quad (8.4.23)$$

It is effectively easy to verify this property. If σ is a singular k -simplex in X , then

$$\begin{aligned} S_k(f)(\text{Sd}_k(\sigma)) &= S_k(f)(S_k(\sigma)(\text{Sd}_k(\delta_k))) = (S_k(f) \circ S_k(\sigma))(\text{Sd}_k(\delta_k)) \\ &= S_k(f \circ \sigma)(\text{Sd}_k(\delta_k)) = \text{Sd}_k(f \circ \sigma) = \text{Sd}_k(S_k(f)(\sigma)) . \end{aligned}$$

The following result is a little harder to proof.

Proposition 8.4.71

Let X be a topological space and R be an integral domain. Then $\text{Sd}_{k-1}(\partial_k(c)) = \partial_k(\text{Sd}_k(c))$ for all $c \in S_k(X; R)$ and $k > 0$.

Proof.

i) We first prove by induction that $\text{Sd}_{k-1}(\partial_k(\delta)) = \partial_k(\text{Sd}_k(\delta))$ is true for affine functions $\delta : \Delta_k \rightarrow \Delta_k$.

For $k = 1$, we have

$$\begin{aligned} \partial_1(\text{Sd}_1(\delta)) &= \partial_1(J_{0,\delta(\mathbf{b}_{(1)})}(\text{Sd}_0(\partial_1(\delta)))) = \partial_1(J_{0,\delta(\mathbf{b}_{(1)})}(\text{Sd}_0(\delta \circ F_1^0 - \delta \circ F_1^1))) \\ &= \partial_1(J_{0,\delta(\mathbf{b}_{(1)})}(\delta \circ F_1^0 - \delta \circ F_1^1)) = (\delta \circ F_1^0 - e_{\delta(\mathbf{b}_{(1)})}) - (\delta \circ F_1^1 - e_{\delta(\mathbf{b}_{(1)})}) \\ &= \delta \circ F_1^0 - \delta \circ F_1^1 = \partial_1(\delta) = \text{Sd}_0(\partial_1(\delta)) \end{aligned}$$

where the third equality comes from the definition of Sd_0 and the fourth equality comes from (8.4.21).

Suppose that the result is true for $k \geq 1$ and consider affine functions $\delta : \Delta_{k+1} \rightarrow \Delta_{k+1}$. Then

$$\begin{aligned} \partial_{k+1}(\text{Sd}_{k+1}(\delta)) &= \partial_{k+1}(J_{k,\delta}(\mathbf{b}_{(k+1)})(\text{Sd}_k(\partial_{k+1}(\delta)))) \\ &= \text{Sd}_k(\partial_{k+1}(\delta)) - J_{k-1,\delta}(\mathbf{b}_{(k+1)})(\partial_k(\text{Sd}_k(\partial_{k+1}(\delta)))) \\ &= \text{Sd}_k(\partial_{k+1}(\delta)) - J_{k-1,\delta}(\mathbf{b}_{(k+1)})(\text{Sd}_{k-1}(\partial_k(\partial_{k+1}(\delta)))) \\ &= \text{Sd}_k(\partial_{k+1}(\delta)) - J_{k-1,\delta}(\mathbf{b}_{(k+1)})(\text{Sd}_{k-1}(0)) = \text{Sd}_k(\partial_{k+1}(\delta)) \end{aligned}$$

where we have used (8.4.22) to get the second equality and our hypothesis of induction to get the third equality.

ii) We can now prove that the statement of the proposition is true for a singular k -simplex σ .

$$\begin{aligned} \partial_k(\text{Sd}_k(\sigma)) &= \partial_k(S_k(\sigma)(\text{Sd}_k(\delta_k))) = S_{k-1}(\sigma)(\partial_k(\text{Sd}_k(\delta_k))) = S_{k-1}(\sigma)(\text{Sd}_{k-1}(\partial_k(\delta_k))) \\ &= S_{k-1}(\sigma) \left(\text{Sd}_{k-1} \left(\sum_{j=0}^k (-1)^j \delta_k \circ F_k^j \right) \right) = S_{k-1}(\sigma) \left(\sum_{j=0}^k (-1)^j \text{Sd}_{k-1}(F_k^j) \right) \\ &= S_{k-1}(\sigma) \left(\sum_{j=0}^k (-1)^j S_{k-1}(F_k^j)(\text{Sd}_{k-1}(\delta_{k-1})) \right) = \sum_{j=0}^k (-1)^j S_{k-1}(\sigma \circ F_k^j)(\text{Sd}_{k-1}(\delta_{k-1})) \\ &= \sum_{j=0}^k (-1)^j \text{Sd}_{k-1}(\sigma \circ F_k^j) = \text{Sd}_{k-1} \left(\sum_{j=0}^k (-1)^j \sigma \circ F_k^j \right) = \text{Sd}_{k-1}(\partial_k \sigma) \end{aligned}$$

where we have used (i) to get the third equality and the fact that $\delta_k : \Delta_k \rightarrow \Delta_k$ is the identity map to get the fifth equality. ■

Let X be a topological space and R be an integral domain. To obtain a theory of singular approximation, we need to introduce a map $T_k : S_k(X; R) \rightarrow S_{k+1}(X; R)$.

Definition 8.4.72

We define a linear function, denoted T_k , that maps an affine function $\delta : \Delta_k \rightarrow \Delta_k$ to an affine function $T_k(\delta) : \Delta_{k+1} \rightarrow \Delta_{k+1}$ as it follows.

$$T_k(\delta) = \begin{cases} 0 & \text{if } k = 0 \\ J_{k,\delta}(\mathbf{b}_{(k)})(\delta - \text{Sd}_k(\delta) - T_{k-1}(\partial_k(\delta))) & \text{if } k > 0 \end{cases}$$

for all affine functions $\delta : \Delta_k \rightarrow \Delta_k$, where $\mathbf{b}_{(k)} = (1+k)^{-1} \sum_{j=0}^k \mathbf{e}_k$ is the barycentre of

Δ_k . If $c = \sum_{j \in J} a_j \delta_j$ where $J \subset \mathbb{N}$ is a finite set, $\delta_j : \Delta_k \rightarrow \Delta_k$ is an affine function and

$a_j \in \mathbb{R}$ for $j \in J$, then $T_k(c)$ is defined as $T_k(c) = \sum_{j \in J} a_j T_k(\delta_j)$.

We extend the function T_k to $S_k(X; R)$ where X is a topological space and R is an integral domain.

Definition 8.4.73

Let X be a topological space and R be an integral domain. We define a linear function $T_k : S_k(X; R) \rightarrow S_{k+1}(X; R)$ by $T_k(\sigma) = S_{k+1}(\sigma)(T_k(\delta_k))$ for all singular k -simplices, where $\delta_k : \Delta_k \rightarrow \Delta_k$ is the identity map. If $c = \sum_{j \in J} a_j \sigma_j$ where $J \subset \mathbb{N}$ is a finite set, σ_j is a singular k -simplex and $a_j \in R$ for all $j \in J$, then $T_k(c)$ is defined as $T_k(c) = \sum_{j \in J} a_j T_k(\sigma_j)$.

We note that $S_{k+1}(\sigma)(T_k(\delta_k))$ is well defined because $T_k(\delta_k)(\Delta_{k+1}) \subset \Delta_k$ and $\sigma : \Delta_k \rightarrow X$. In fact, we may extend σ to Δ_{k+1} by setting $\sigma(e_{k+1}) = 0$.

Suppose that X and Y are two topological spaces and R is an integral domain. If $f : X \rightarrow Y$ is a continuous function, then the reader can proceed as we did for Sd_k to verify that we have the following commutative diagram.

$$\begin{array}{ccc} S_k(X; R) & \xrightarrow{S_k(f)} & S_k(Y; R) \\ \downarrow T_k & & \downarrow T_k \\ S_{k+1}(X; R) & \xrightarrow{S_{k+1}(f)} & S_{k+1}(Y; R) \end{array} \quad (8.4.24)$$

Proposition 8.4.74

Let X be a topological space and R be an integral domain. Then $\partial_{k+1}(T_k(c)) + T_{k-1}(\partial_k(c)) = c - \text{Sd}_k(c)$ for all $c \in S_k(X; R)$ and $k > 0$.

Proof.

i) We first prove by induction the $\partial_{k+1}(T_k(\delta)) + T_{k-1}(\partial_k(\delta)) = \delta - \text{Sd}_k(\delta)$ is true for affine functions $\delta : \Delta_k \rightarrow \Delta_k$.

For $k = 1$, we have

$$\begin{aligned} \partial_2(T_1(\delta)) &= \partial_2 \left(J_{1, \delta(\mathbf{b}_{(1)})}(\delta - \text{Sd}_1(\delta) - T_0(\partial_1(\delta))) \right) = \partial_2 \left(J_{1, \delta(\mathbf{b}_{(1)})}(\delta - \text{Sd}_1(\delta)) \right) \\ &= \delta - \text{Sd}_1(\delta) - J_{0, \delta(\mathbf{b}_{(1)})}(\partial_1(\delta - \text{Sd}_1(\delta))) = \delta - \text{Sd}_1(\delta) - J_{0, \delta(\mathbf{b}_{(1)})}(\partial_1(\delta) - \partial_1(\text{Sd}_1(\delta))) \\ &= \delta - \text{Sd}_1(\delta) - J_{0, \delta(\mathbf{b}_{(1)})}(\partial_1(\delta) - \text{Sd}_0(\partial_1(\delta))) = \delta - \text{Sd}_1(\delta) - J_{0, \delta(\mathbf{b}_{(1)})}(\partial_1(\delta) - \partial_1(\delta)) \\ &= \delta - \text{Sd}_1(\delta) = \delta - \text{Sd}_1(\delta) - T_0(\partial_1(\delta)) \end{aligned}$$

for all affine functions $\delta : \Delta_1 \rightarrow \Delta_1$, where we have used (8.4.22) to get the third equality and Proposition 8.4.71 to get the fifth equality.

Suppose that the result is true for $k > 0$ and consider affine functions $\delta : \Delta_{k+1} \rightarrow \Delta_{k+1}$. Then

$$\partial_{k+2}(T_{k+1}(\delta)) = \partial_{k+2} \left(J_{k+1, \delta(\mathbf{b}_{(k+1)})}(\delta - \text{Sd}_{k+1}(\delta) - T_k(\partial_{k+1}(\delta))) \right)$$

$$\begin{aligned}
&= (\delta - \text{Sd}_{k+1}(\delta) - T_k(\partial_{k+1}(\delta))) - J_{k,\delta(\mathbf{b}_{(k+1)})}(\partial_{k+1}(\delta - \text{Sd}_{k+1}(\delta) - T_k(\partial_{k+1}(\delta)))) \\
&= (\delta - \text{Sd}_{k+1}(\delta) - T_k(\partial_{k+1}(\delta))) - J_{k,\delta(\mathbf{b}_{(k+1)})}(\partial_{k+1}(\delta) - \partial_{k+1}(\text{Sd}_{k+1}(\delta)) \\
&\quad - \partial_{k+1}(T_k(\partial_{k+1}(\delta)))) \\
&= (\delta - \text{Sd}_{k+1}(\delta) - T_k(\partial_{k+1}(\delta))) - J_{k,\delta(\mathbf{b}_{(k+1)})}(\partial_{k+1}(\delta) - \text{Sd}_k(\partial_{k+1}(\delta)) \\
&\quad - (\partial_{k+1}(\delta) - \text{Sd}_k(\partial_{k+1}(\delta)) - T_{k-1}(\partial_k(\partial_{k+1}(\delta)))) \\
&= \delta - \text{Sd}_{k+1}(\delta) - T_k(\partial_{k+1}(\delta))
\end{aligned}$$

where we have used (8.4.22) to get the second equality, and Proposition 8.4.71 and the hypothesis of induction to get the fourth equation.

ii) We can now prove that the statement of the proposition is true for a singular k -simplex σ .

$$\begin{aligned}
\partial_{k+1}(T_k(\sigma)) &= \partial_{k+1}(S_{k+1}(\sigma)(T_k(\delta_k))) = S_k(\sigma)(\partial_{k+1}(T_k(\delta_k))) \\
&= S_k(\sigma)(\delta_k - \text{Sd}_k(\delta_k) - T_{k-1}(\partial_k(\delta_k))) = S_k(\sigma)\left(\delta_k - \text{Sd}_k(\delta_k) - T_{k-1}\left(\sum_{j=0}^k (-1)^j \delta_k \circ F_k^j\right)\right) \\
&= S_k(\sigma)\left(\delta_k - \text{Sd}_k(\delta_k) - \sum_{j=0}^k (-1)^j T_{k-1}(F_k^j)\right) \\
&= S_k(\sigma)\left(\delta_k - \text{Sd}_k(\delta_k) - \sum_{j=0}^k (-1)^j S_k(F_k^j)(T_{k-1}(\delta_{k-1}))\right) \\
&= S_k(\sigma)(\delta_k) - S_k(\sigma)(\text{Sd}_k(\delta_k)) - S_k\left(\sum_{j=0}^k (-1)^j \sigma \circ F_k^j\right)(T_{k-1}(\delta_{k-1})) \\
&= \sigma - \text{Sd}_k(\sigma) - S_k(\partial_k \sigma)(T_{k-1}(\delta_{k-1})) = \sigma - \text{Sd}_k(\sigma) - T_{k-1}(\partial_k(\sigma))
\end{aligned}$$

where we have used (i) to get the third equality. We have also used several times the fact that δ_q is the identity map on Δ_q for $q = k$ and $k - 1$. ■

Since the topological spaces X that we are considering for the singular k -chains may not have a metric, we have to formulate the distance between singular k -chains in terms of open sets.

Definition 8.4.75

Let $\mathcal{U} = \{U_i\}_{i \in I}$ be an open cover of a topological space X . We say that a singular k -simplex σ in X is **small of order $\mathcal{U}$** if there is an $i \in I$ such that $\sigma(\delta_k) \subset U_i$.

We associate to every linear combination $c = \sum_{j \in J} a_j \delta_j$ where $J \subset \mathbb{N}$ is a finite set, $a_j \in \mathbb{R}$ and $\delta_j : \Delta_k \rightarrow \Delta_k$ is an affine function for all $j \in J$, the value $d(c) = \max_{j \in J} \text{diam}(\delta_j(\Delta_k))$. We have a result similar to Proposition 8.2.17. The proof is very similar.

Proposition 8.4.76

Let $\delta : \Delta_k \rightarrow \Delta_k$ be an affine function. Then $d(\text{Sd}_k(\delta)) \leq \frac{k}{k+1} d(\delta)$.

Proof (Sketch).

There is a result equivalent to Lemma 8.2.15 for affine functions $\delta : \Delta_k \rightarrow \Delta_k$. Namely, $\text{diam}(\delta(\Delta_k)) = \|\delta(\mathbf{e}_{i_1}) - \delta(\mathbf{e}_{i_2})\|$ for some $i_1, i_2 \in \{0, 1, 2, \dots, k\}$. The proof is basically identical because $\delta(\mathbf{x}) = t\delta(\mathbf{x}) + (1-t)\delta(\mathbf{x})$ for all $\mathbf{x} \in \Delta_k$.

By definitions of $\text{Sd}_k(\delta)$ and the previous paragraph, it follows that $d(c) = \|\delta(\mathbf{b}_{(p)}) - \delta(\mathbf{b}_{(q)})\|$ for some $p, q \in \{0, 1, \dots, k\}$ where $\mathbf{b}_{(p)}$ and $\mathbf{b}_{(q)}$ are vertices on the image of the same function

$$\text{Sd}_k(\delta_k) = J_{k-1, \mathbf{b}_{(k)}}(J_{k-2, F_k^{s_k}(\mathbf{b}_{(k-1)})}(\dots(J_{0, F_k^{s_k} \circ \dots \circ F_2^{s_2}(\mathbf{b}_{(1)})}(F_k^{s_k} \circ F_{k-1}^{s_{k-1}} \circ \dots \circ F_1^{s_1})))) : \Delta_k \rightarrow \Delta_k$$

for some $0 \leq s_j \leq j$ with $1 \leq j \leq k$, where δ_k is the identity map. We may assume that $p < q$, $\mathbf{b}_{(p)} = \frac{1}{1+p} \sum_{j=0}^p \mathbf{e}_{i_j}$ and $\mathbf{b}_{(q)} = \frac{1}{1+q} \sum_{j=0}^q \mathbf{e}_{i_j}$. Note that the image of $\text{Sd}_k(\delta_k)$ is nothing else than $K^{[1]}$ defined in Subsection 8.2.2 where K is the k -simplex $[e_0, e_1, \dots, e_k]$. We can then proceed as in the proof of Proposition 8.2.17 to show that

$$\begin{aligned} d(c) &= \|\delta(\mathbf{b}_{(p)}) - \delta(\mathbf{b}_{(q)})\| = \left\| \frac{1}{p+1} \sum_{j=0}^p \delta(\mathbf{e}_{i_j}) - \frac{1}{q+1} \sum_{j=0}^q \delta(\mathbf{e}_{i_j}) \right\| \\ &\leq \frac{1}{(q+1)(p+1)} \sum_{j=0}^q \sum_{j=0}^p \|\delta(\mathbf{e}_{i_j}) - \delta(\mathbf{e}_{i_j})\| \leq \frac{q(p+1)}{(q+1)(p+1)} d(\delta) \\ &= \frac{q}{q+1} d(\delta) \leq \frac{k}{k+1} d(\delta) \end{aligned}$$

because $\|\delta(\mathbf{e}_{i_j}) - \delta(\mathbf{e}_{i_j})\| \leq d(\delta)$ for all i and j , and $q \leq k$. ■

Given an affine function $\delta : \Delta_k \rightarrow \Delta_k$, it follows from the previous proposition that $d((\text{Sd}_k)^q(\delta)) \leq \left(\frac{k}{k+1}\right)^q d(\delta) \rightarrow 0$ as $q \rightarrow \infty$, where

$$(\text{Sd}_k)^q(\delta) = \underbrace{\text{Sd}_k(\text{Sd}_k(\dots(\text{Sd}_k(\delta))))}_{q \text{ times}}.$$

for $q \in \mathbb{N}$.

The theorem in singular homology that corresponds to Theorem 8.2.24 is the following theorem.

Theorem 8.4.77

Let $\mathcal{U} = \{U_i\}_{i \in I}$ be an open cover of a topological space X and R be an integral domain. Given a singular k -simplex σ in X , there exists $q \in \mathbb{N}$ such that $(\text{Sd}_k)^q(\sigma)$ is a linear combination of singular k -simplices small of order $\mathcal{U}$.

Proof.

Since $\{\sigma^{-1}(U_i)\}_{i \in I}$ is an open cover of the compact set Δ_k , there exists a Lebesgue number ϵ such that, for every $\mathbf{x} \in \Delta_k$, we have that $B_\epsilon(\mathbf{x}) \cap \Delta_k \subset \sigma^{-1}(U_{i_{\mathbf{x}}})$ for some $i_{\mathbf{x}} \in I$; namely, $\sigma(B_\epsilon(\mathbf{x}) \cap \Delta_k) \subset U_{i_{\mathbf{x}}}$ for some $i_{\mathbf{x}} \in I$.

Choose q large enough to have $d((\text{Sd}_k)^q(\delta_k)) < \epsilon$ where δ_k is the identity map on Δ_k . It follows from the commutative diagram in (8.4.23) that $(\text{Sd}_k)^q(\sigma) = S_k(\sigma)((\text{Sd}_k)^q(\delta_k))$. So $(\text{Sd}_k)^q(\sigma)$ is the linear combinations of singular k -simplices of the form $\sigma \circ \delta$, where $\delta : \Delta_k \rightarrow \Delta_k$ is an affine function such that $\text{diam}(\delta(\Delta_k)) < \epsilon$. Since $\delta(\Delta_k) \subset B_\epsilon(\mathbf{x}) \cap \Delta_k$ for some $\mathbf{x} \in \Delta_k$, we have that $(\sigma \circ \delta)(\Delta_k) \subset U_{i_{\mathbf{x}}}$ for some $i_{\mathbf{x}} \in I$. ■

Proposition 8.4.78

Let A be a subset of a topological space X and R be an integral domain. Moreover, let $\mathcal{U} = \{U_i\}_{i \in I}$ be an open cover of X . Every homology class in $H_k(X, A; R)$ can be represented by a relative k -cycle of $X \bmod A$ which is a linear combination of singular k -simplices small of order $\mathcal{U}$.

Proof.

Given $[c]_{X,A} \in H_k(X, A; R)$, we have that c is a relative k -cycle of $X \bmod A$. Recall that this means that $\partial_k(c) \in S_{k-1}(A; R)$.

It follows from Proposition 8.4.74 that

$$\partial_{k+1}(T_k(c)) + T_{k-1}(\partial_k(c)) = c - \text{Sd}_k(c). \quad (8.4.25)$$

Since $\partial_k(c) \in S_{k-1}(A; R)$, we have that $T_{k-1}(\partial_k(c)) \in S_k(A; R)$ follows from the commutative diagram in (8.4.24) with f replaced by the inclusion map $\iota : A \rightarrow X$. Moreover, $\text{Sd}_k(c)$ is a relative k -cycle of $X \bmod A$ because $\partial_k(\text{Sd}_k(c)) = \text{Sd}_{k-1}(\partial_k(c))$ according to Proposition 8.4.71 and $\partial_k(c) \in S_{k-1}(A; R)$ implies that $\text{Sd}_{k-1}(\partial_k(c)) \in S_{k-1}(A; R)$ due to the commutative diagram in (8.4.23) with f replaced by the inclusion map $\iota : A \rightarrow X$. It therefore follows from (8.4.25) and the definition of homologous in $X \bmod A$ that $c \stackrel{h}{\sim} \text{Sd}_k(c)$ in $X \bmod A$. We get that $[c]_{X,A} = [\text{Sd}_k(c)]_{X,A}$ in $H_k(X, A; R)$.

It follows by induction that $[c]_{X,A} = [(\text{Sd}_k)^q(c)]_{X,A}$ for all $q \in \mathbb{N}$. If we apply Theorem 8.4.77 to each singular k -simplex of the linear combination of c , then we can find q large enough to have that $(\text{Sd}_k)^q(c)$ a linear combination of singular k -simplices small of order $\mathcal{U}$. ■

We can finally prove the excision theorem.

Proof (of Theorem 8.4.66).

Let $\iota : A \rightarrow X$ be the inclusion map. We prove that $H_k(\iota) : H_k(X \setminus U, A \setminus U; R) \rightarrow H_k(X, A; R)$ is an isomorphism.

i) We first prove that $H_k(\iota)$ is onto. Suppose that $[c]_{X,A} \in H_k(X, A; R)$. It follows from Proposition 8.4.78 that we may assume that $c = \sum_{j \in J} a_j \sigma_j$ is a relative k -cycle where each σ_j is small of order $\mathcal{U} = \{X \setminus \overline{U}, A^\circ\}$.

If $\sigma_j(\Delta_k) \not\subset X \setminus U$, then we must have that $\sigma_j(\Delta_k) \subset A^\circ$ because $X \setminus \overline{U} \subset X \setminus U$. It follows that $\sigma_j \in S_k(A; R)$. In other words, we have that $\sigma_j \in B_k(X, A; R)$ and so $[\sigma_j]_{X,A} = [0]_{X,A}$. We can then remove all this terms from the summation and still get a summation homologous in $X \bmod A$ to c . We may therefore assume that $\sigma_j(\Delta_k) \subset X \setminus U$ for all j .

Since c is a relative k -cycle and $c \in S_k(X \setminus U; R)$, we get that $\partial_k(c) \in S_k(A; R) \cap S_k(X \setminus U; R)$. Thus $\partial_k(c) \in S_k(A \setminus U; R)$. Hence c is a relative k -cycle in $X \setminus U \bmod A \setminus U$. We have that $[c]_{X \setminus U, A \setminus U} \in H_k(X \setminus U; A \setminus U; R)$ and $H_k(\iota)([c]_{X \setminus U, A \setminus U}) = [c]_{X,A} \in H_k(X, A; R)$.

ii) We now prove that $H_k(\iota)$ is one-to-one. Suppose that $[c]_{X \setminus U, A \setminus U} \in H_k(X \setminus U, A \setminus U; R)$ satisfies $H_k(\iota)([c]_{X \setminus U, A \setminus U}) = [c]_{X,A} = [0]_{X,A} \in H_k(X, A; R)$. Since $[[c]]_{X,A} \in \mathcal{R}(\partial_k) = B_k(X, A; R)/S_k(A; R)$, we have that c is a relative k -boundary in $X \bmod A$. Therefore, there exists $b \in S_{k+1}(X; R)$ such that $\tilde{c} = c + \partial_{k+1}(b) \in S_k(A; R)$.

We get from Proposition 8.4.71 that

$$(\text{Sd}_k)^q(\tilde{c}) = (\text{Sd}_k)^q(c) + \partial_{k+1}((\text{Sd}_{k+1})^q(b)) \quad (8.4.26)$$

for $q \geq 0$. According to Proposition 8.4.78, we may choose q large enough such that $(\text{Sd}_k)^q(b)$ is small of order $\mathcal{U} = \{X \setminus \overline{U}, A^\circ\}$. Thus, we may write $(\text{Sd}_k)^q(b) = b_1 + b_2$ for some $b_1 \in S_{k+1}(X \setminus \overline{U}; R) \subset S_{k+1}(X \setminus U; R)$ and $b_2 \in S_{k+1}(A^\circ; R) \subset S_{k+1}(A; R)$. We get from (8.4.26) that

$$(\text{Sd}_k)^q(\tilde{c}) - \partial_{k+1}(b_2) = (\text{Sd}_k)^q(c) + \partial_{k+1}(b_1),$$

where the right hand side is in $S_k(X \setminus U; R)$ and the left hand side is in $S_k(A; R)$. Therefore, both sides of the previous equality must be in $S_k(A \setminus U; R)$. We conclude that $(\text{Sd}_k)^q(c) \stackrel{h}{\sim} (\text{Sd}_k)^q(\tilde{c}) - \partial_{k+1}(b_2) \in S_k(A \setminus U; R)$. Thus $(\text{Sd}_k)^q(c)$ is a relative k -boundary in $X \setminus U \bmod A \setminus U$ because $(\text{Sd}_k)^q(c) \in S_k(X \setminus U; R)$ since $c \in S_k(X \setminus U; R)$. It follows that $[c]_{X \setminus U, A \setminus U} = [(\text{Sd}_k)^q(c)]_{X \setminus U, A \setminus U} = [0]_{X \setminus U, A \setminus U} \in H_k(X \setminus U, A \setminus U; R)$. Note that $[c]_{X \setminus U, A \setminus U} = [(\text{Sd}_k)^q(c)]_{X \setminus U, A \setminus U}$ as we saw in the context of the proof of Proposition 8.4.78. ■

Before closing this subsection, we should mention some results to illustrate the power of the technique of excision. First, we have to strengthen the concept of retraction introduced in definition 8.1.29 to the context of relative homology.

Definition 8.4.79

Let A and Y be two subsets of a topological space X and let $B = Y \cap A$ (Figure 8.12). We say that (Y, B) is a **deformation retract** of (X, A) if there exists a function

$r : X \rightarrow Y$ such that $r(A) \subset B$, $r \circ \iota = \text{Id}_Y$ where $\iota : Y \rightarrow X$ is the inclusion map (i.e. Y is a retract of X), and $\text{Id}_X \stackrel{\text{rel}}{\sim} \iota \circ r$. The homotopy $H : X \times [0, 1] \rightarrow X$ from Id_X to $\iota \circ r$ is called **deformation retraction**.

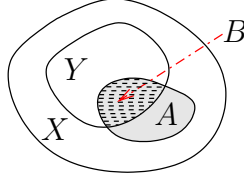


Figure 8.12: (Y, B) is a deformation retract of (X, A) .

In the previous definition, to apply the definition of homotopic functions given in Definition 8.4.57, we consider $\text{Id}_X : (X, A) \rightarrow (X, A)$ and $\iota \circ r : (X, A) \rightarrow (X, A)$. To get $(\iota \circ r)(A) \subset B$, we have to require that $r(A) \subset B$.

Proposition 8.4.80

Let X be a topological space and R be an integral domain. Assume that $V \subset U \subset A \subset X$ (Figure 8.13). If V can be excised from (X, A) and $(X \setminus U, A \setminus U)$ is a deformation retract of $(X \setminus V, A \setminus V)$, then U can be excised from (X, A) .

Proof.

We have that $X \setminus U \subset X \setminus V$, $A \setminus U \subset A \setminus V$ and $(A \setminus V) \cap (X \setminus U) = A \setminus U$ (Figure 8.13). As usual, let $\iota : Y \rightarrow X$ be the inclusion map. Let $r : X \setminus V \rightarrow X \setminus U$ be the function satisfying $r(x) = x$ for all $x \in X \setminus U$, $r(A \setminus V) \subset A \setminus U$, and $\text{Id}_{X \setminus V} \stackrel{\text{rel}}{\sim} \iota \circ r$.

It follows from $r \circ \iota = \text{Id}_{X \setminus U}$, $\text{Id}_{X \setminus V} \stackrel{\text{rel}}{\sim} \iota \circ r$, and Propositions 8.4.53 and 8.4.58, that $H_k(r) \circ H_k(\iota) = H_k(\text{Id}_{X \setminus U}) = \text{Id}_{H_k(X \setminus U, A \setminus U; R)}$ and $H_k(\iota) \circ H_k(r) = H_k(\text{Id}_{X \setminus V}) = \text{Id}_{H_k(X \setminus V, A \setminus V; R)}$. Thus $H_k(\iota) : H_k(X \setminus U, A \setminus U; R) \rightarrow H_k(X \setminus V, A \setminus V; R)$ is an isomorphism.

Since V can be excised from (X, A) , we have that $H_k(\iota) : H_k(X \setminus V, A \setminus V; R) \rightarrow H_k(X, A; R)$ is an isomorphism. It follows that $H_k(\iota) : H_k(X \setminus U, A \setminus U; R) \cong H_k(X, A; R)$. ■

Example 8.4.81

Let $S_+^n = \{\mathbf{x} \in S^n : x_{n+1} \geq 0\}$ and $S_-^n = \{\mathbf{x} \in S^n : x_{n+1} \leq 0\}$. We use the previous proposition to prove below that $(S_-^n)^\circ = \{\mathbf{x} \in S^n : x_{n+1} < 0\}$ can be excised from (S^n, S_-^n) . Hence $H_k(S_+^n, S_-^{n-1}; R) \cong H_k(S^n, S_-^n; R)$ because $S^n \setminus (S_-^n)^\circ = S_+^n$ and $S_-^n \setminus (S_-^n)^\circ = S_-^{n-1}$.

We cannot directly use the excision theorem, Theorem 8.4.66, because we are in the situation where $U = (S_-^n)^\circ$ and $A = S_-^n$ in the statement of the theorem. Therefore $\overline{U} \not\subset A^\circ$. We have to use the previous proposition with $V = \{\mathbf{x} \in S^n : x_{n+1} \leq -\epsilon\}$ and $0 \leq \epsilon < 1$ (Figure 8.14). We have from Theorem 8.4.66 that V can be excised from (S^n, S_-^n) because $\overline{V} \subset (S_-^n)^\circ$. Moreover, $(S_+^n, S_-^{n-1}) = (S^n \setminus (S_-^n)^\circ, S_-^n \setminus (S_-^n)^\circ)$ is a deformation retract of

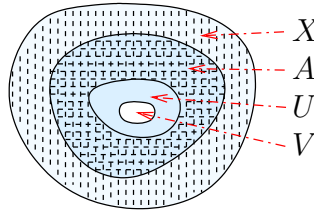


Figure 8.13: Representation of the sets involved in the statement of Proposition 8.4.80. $X \setminus V$ is the union of the pale and darker blue regions. $A \setminus V$ is the darker blue region. $X \setminus U$ is the union of the regions with horizontal dashed lines and the region with horizontal and vertical dashed lines. $A \setminus U$ is the region with horizontal and vertical dashed lines.

$(S^n \setminus V, S_-^n \setminus V)$. A possible function $r : S^n \setminus V \rightarrow S_+^n$ for the deformation retract is represented by the arrows along the great circles in Figure 8.14. Only the $\mathbf{x} \in (S_-^n)^\circ$ are moving along the great circles toward S^{n-1} . ♣

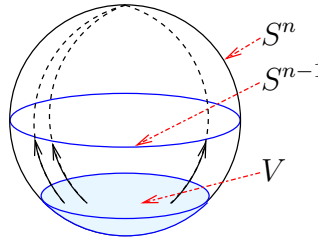


Figure 8.14: Representation of the set V involved in Example 8.4.81. The arrows illustrate the map r use to show that $(S^n \setminus (S_-^n)^\circ, S_-^n \setminus (S_-^n)^\circ)$ is a deformation retract of $(S^n \setminus V, S_-^n \setminus V)$. Only the $\mathbf{x} \in (S_-^n)^\circ$ are moving along the great circles toward S^{n-1} .

Example 8.4.82

We may use the tools that we have introduced so far to compute $H_k(S^n)$ for $k, n > 0$.

i) Let $B^n = \overline{B_1(\mathbf{0})} \subset \mathbb{R}^n$. So $\partial B^n = S^{n-1}$. We get from Proposition 8.4.60 that

$$\xrightarrow{C_{k+1}} H_k(S^{n-1}; R) \xrightarrow{H_k(\iota)} H_k(B^n; R) \xrightarrow{H_k(\text{Id}_{B^n})} H_k(B^n, S^{n-1}; R) \quad (8.4.27)$$

$$\xrightarrow{C_k} H_{k-1}(S^{n-1}; R) \xrightarrow{H_{k-1}(\iota)} H_{k-1}(B^n; R) \xrightarrow{H_{k-1}(\text{Id}_{B^n})} \rightarrow$$

k > 1) Since B^n is contractible, we get from Corollary 8.4.41 that $H_k(B^n; R) = 0$ and $H_{k-1}(B^n; R) = 0$. We deduce from (8.4.27) that

$$0 \xrightarrow{H_k(\text{Id}_{B^n})} H_k(B^n, S^{n-1}; R) \xrightarrow{C_k} H_{k-1}(S^{n-1}; R) \xrightarrow{H_{k-1}(\iota)} 0$$

Since $\mathcal{N}(C_k) = \mathcal{R}(H_k(\text{Id}_{B^n})) = \{[0]_{B^n, S^{n-1}}\}$ and $\mathcal{R}(C_k) = \mathcal{N}(H_{k-1}(\iota)) = H_{k-1}(S^{n-1}; R)$, we find that C_k is a isomorphism between $H_k(B^n, S^{n-1}; R)$ and $H_{k-1}(S^{n-1}; R)$.

k = 1) We consider two cases: $n > 1$ and $n = 1$.

For $n > 1$, we still have that $H_1(B^n; R) = 0$ because B^n is contractible. Moreover, we have from Proposition 8.4.56 that $H_0(B^n, S^{n-1}; R) = 0$ because B^n is path-connected. Hence, we get from (8.4.27) that

$$0 \xrightarrow{H_1(\text{Id}_{B^n})} H_1(B^n, S^{n-1}; R) \xrightarrow{C_1} H_0(S^{n-1}; R) \xrightarrow{H_0(\iota)} H_0(B^n; R) \xrightarrow{H_0(\text{Id}_{B^n})} 0$$

Since $\mathcal{N}(C_1) = \mathcal{R}(H_1(\text{Id}_{B^n})) = \{[0]_{B^n, S^{n-1}}\}$, we still have that C_1 is one-to-one. We now have that $H_0(B^n; R) \cong R$ and $H_0(S^{n-1}; R) \cong R$ according to Proposition 8.4.11 because B^n and S^{n-1} are path-connected. Hence, since $\mathcal{R}(H_0(\iota)) = \mathcal{N}(H_0(\text{Id}_{B^n})) = H_0(B^n; R) \cong R$, we get that $\mathcal{N}(H_0(\iota)) = \{[0]\} \in H_0(S^{n-1}; R) \cong R$. It follows that $\mathcal{R}(C_1) = \mathcal{N}(H_0(\iota)) = \{[0]_{S^{n-1}}\}$. Since C_1 is one-to-one, we must have that $H_1(B^n, S^{n-1}; R) = 0$.

For $n = 1$, we still have that $H_1(B^1; R) = 0$ because $B^1 = [-1, 1]$ is contractible and $H_0(B^1, S^0; R) = 0$ because B^1 is path-connected. Hence, we get from (8.4.27) that

$$0 \xrightarrow{H_1(\text{Id}_{B^1})} H_1(B^1, S^0; R) \xrightarrow{C_1} H_0(S^0; R) \xrightarrow{H_0(\iota)} H_0(B^1; R) \xrightarrow{H_0(\text{Id}_{B^1})} 0$$

Since $\mathcal{N}(C_1) = \mathcal{R}(H_1(\text{Id}_{B^1})) = \{[0]_{B^1, S^0}\}$, we still have that C_1 is one-to-one. We now have that $H_0(B^1; R) \cong R$ and $H_0(S^0; R) \cong R^2$ according to Propositions 8.4.11 and 8.4.10. Hence, since $\mathcal{R}(H_0(\iota)) = \mathcal{N}(H_0(\text{Id}_{B^1})) = H_0(B^1; R) \cong R$, we get that $\mathcal{N}(H_0(\iota)) \cong R$. It follows that $\mathcal{R}(C_1) = \mathcal{N}(H_0(\iota)) \cong R$. Since C_1 is one-to-one, we have that $H_1(B^1, S^0; R) \cong R$.

In summary, we have found so far that $H_k(B^n, S^{n-1}; R) \cong H_{k-1}(S^{n-1}; R)$ for $k > 1$ and $n > 0$, $H_1(B^n, S^{n-1}; R) = 0$ for $n > 1$ and $H_1(B^1, S^0; R) \cong R$.

ii) The projection $\pi : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$ defined by $\pi(\mathbf{x}) = (x_1, x_2, \dots, x_n)$ for $\mathbf{x} \in \mathbb{R}^{n+1}$ and $n > 0$ provides a homeomorphism $\pi : (S_+^n, S^{n-1}) \rightarrow (B^n, S^{n-1})$. Hence $H_k(S_+^n, S^{n-1}; R) \cong H_k(B^n, S^{n-1}; R)$ for $n > 0$.

Moreover, we get from Example 8.4.81 that $H_k(S_+^n, S^{n-1}; R) \cong H_k(S^n, S_-^n; R)$ for $n > 0$. Thus $H_k(B^n, S^{n-1}; R) \cong H_k(S^n, S_-^n; R)$ for $n > 0$.

It follows from Proposition 8.4.63 that $H_k(S^n, S_-^n; R) \cong H_k(S^n; R)$ for $k > 0$ and $n > 0$ because S^n is path connected and S_-^n is contractible. Therefore $H_k(B^n, S^{n-1}; R) \cong H_k(S^n; R)$ for $k > 0$

In conclusion, we have found that

1. $H_0(S^0; R) \cong R^2$ comes from Propositions 8.4.11 and 8.4.10 because $S^0 = \{-1, 1\}$ has two components,
2. $H_0(S^n; R) \cong R$ for $n > 0$ comes from Propositions 8.4.11 because S^n is path-connected,
3. $H_1(S^0; R) = 0$ comes from Example 8.4.9 and Proposition 8.4.10 because $S^0 = \{-1, 1\}$,
4. $H_1(S^1; R) \cong H_1(B^1, S^0; R) \cong R$,

5. $H_1(S^n; R) \cong H_1(B^n, S^{n-1}; R) = 0$ for $n > 1$, and
6. $H_k(S^n; R) \cong H_k(B^n, S^{n-1}; R) \cong H_{k-1}(S^{n-1}; R)$ for $k > 1$ and $n > 0$.



8.4.8 Mayer-Vietoris Sequence

The Mayer-Vietoris sequence alluded to in Subsection 7.2.3 is an import topic. To prove Mayer-Vietoris theorem below, we need along the way a very famous result. The proof of this result is amusing.

Theorem 8.4.83 (Five Lemma)

Let R be an integral domain. Suppose that we have the following commutative diagram

$$\begin{array}{ccccccccc}
 A_1 & \xrightarrow{f_1} & A_2 & \xrightarrow{f_2} & A_3 & \xrightarrow{f_3} & A_4 & \xrightarrow{f_4} & A_5 \\
 \downarrow h_1 & & \downarrow h_2 & & \downarrow h_3 & & \downarrow h_4 & & \downarrow h_5 \\
 B_1 & \xrightarrow{g_1} & B_2 & \xrightarrow{g_2} & B_3 & \xrightarrow{g_3} & B_4 & \xrightarrow{g_4} & B_5
 \end{array}$$

where the A_j and B_j are R -module, and the f_j , g_j and h_j are R -module homomorphism. Moreover, suppose that each row represents an exact sequence and that the h_j for $j \neq 3$ are isomorphism. Then h_3 is also an isomorphism.

Proof.

i) h_3 is one-to-one.

Suppose that $h_3(a_3) = 0$. Using the diagram above, we get that $g_3(h(a_3)) = 0 = h_4(f_3(a_3))$. Thus $f_3(a_3) = 0$ because h_4 is an isomorphism. Therefore, there exists $a_2 \in A_2$ such that $f_2(a_2) = a_3$ because $\mathcal{R}(f_2) = \mathcal{N}(f_3)$. Using the diagram above, we get that $h_3(a_3) = 0 = h_3(f_2(a_2)) = g_2(h_2(a_2))$. Since $\mathcal{R}(g_1) = \mathcal{N}(g_2)$, there exists $b_1 \in B_1$ such that $g_1(b_1) = h_2(a_2)$. Since h_1 is an isomorphism, there exists $a_1 \in A_1$ such that $h_1(a_1) = b_1$. Hence, using the diagram above, we get that $h_2(a_2) = g_1(h_1(a_1)) = h_2(f_1(a_1))$. Since h_2 is an isomorphism, we get that $a_2 = f_1(a_1)$. Finally, $a_3 = f_2(a_2) = f_2(f_1(a_1)) = 0$ because $\mathcal{R}(f_1) = \mathcal{N}(f_2)$. This proves that h_3 is one-to-one.

ii) h_3 is onto.

Given $b_3 \in B_3$, there exists $a_4 \in A_4$ such that $h_4(a_4) = g_3(b_3)$ because h_4 is an isomorphism. Using the diagram above and $\mathcal{R}(g_3) = \mathcal{N}(g_4)$, we then get that $h_5(f_4(a_4)) = g_4(h_4(a_4)) = g_4(g_3(b_3)) = 0$. Since h_5 is an isomorphism, we get that $f_4(a_4) = 0$. Since $\mathcal{R}(f_3) = \mathcal{N}(f_4)$, there exists $a_3 \in A_3$ such that $f_3(a_3) = a_4$. Hence, using the diagram above, we get that $g_3(h_3(a_3)) = h_4(f_3(a_3)) = h_4(a_4) = g_3(b_3)$. Thus $h_3(a_3) - b_3 \in \mathcal{N}(g_3)$. Since $\mathcal{R}(g_2) = \mathcal{N}(g_3)$, there exists $b_2 \in B_2$ such that $g_2(b_2) = h_3(a_3) - b_3$. Since h_2 is an isomorphism, there exists $a_2 \in A_2$ such that $h_2(a_2) = b_2$. Using the diagram above for the last time,

we get that $h_3(f_2(a_2)) = g_2(h_2(a_2)) = g_2(b_2) = h_3(a_3) - b_3$. Therefore $h_3(a_3 - f_2(a_2)) = b_3$. Since b_3 is arbitrary, we have shown that $\mathcal{R}(h_3) = B_3$. ■

Definition 8.4.84

Suppose that X_1 and X_2 are two topological spaces. The triad $(X_1 \cup X_2, X_1, X_2)$ is said to be **Exact** if $X_1 \setminus X_2$ can be excised from $(X_1 \cup X_2, X_1)$ and $X_2 \setminus X_1$ can be excised from $(X_1 \cup X_2, X_2)$.

In other words, The triad $(X_1 \cup X_2, X_1, X_2)$ is exact if the inclusion maps $\iota_2 : (X_2, X_1 \cap X_2) \rightarrow (X_1 \cup X_2, X_1)$ and $\iota_1 : (X_1, X_1 \cap X_2) \rightarrow (X_1 \cup X_2, X_2)$ are excisions. For instance, using the notation introduced in Example 8.4.81, we have that (S^n, S_n^+, S_n^-) is an exact triad.

Lemma 8.4.85

Suppose that X_1 and X_2 are two topological spaces. Let

$$\tilde{H}_k = \mathcal{N}(\partial_k|_{S_k(X_1;R)+S_k(X_2;R)}) / \mathcal{N}(\partial_{k+1}|_{S_{k+1}(X_1;R)+S_{k+1}(X_2;R)}) .$$

Then $(X_1 \cup X_2, X_1, X_2)$ is an exact triad if and only if $h_k : \tilde{H}_k \rightarrow H_k(X_1 \cup X_2; R)$ defined by $h_k([c_1 + c_2]) = [c_1 + c_2]_X$ for all $c_1 \in S_k(X_1; R)$ and $c_2 \in S_k(X_2; R)$ is an isomorphism.

Proof.

Let $X = X_1 \cup X_2$ and $B = X_1 \cap X_2$.

We first note that h_k is well defined. Suppose that $[c_1 + c_2] = [\tilde{c}_1 + \tilde{c}_2]$ with $c_1, \tilde{c}_1 \in S_k(X_1; R)$ and $c_2, \tilde{c}_2 \in S_k(X_2; R)$. Then $(c_1 + c_2) - (\tilde{c}_1 + \tilde{c}_2) = \partial_{k+1}(b_1 + b_2)$ for some $b_1 \in S_{k+1}(X_1; R)$ and $b_2 \in S_{k+1}(X_2; R)$. Since $b_1 + b_2 \in S_{k+1}(X; R)$, we get that $[c_1 + c_2]_X = [\tilde{c}_1 + \tilde{c}_2]_X$.

i) We prove that the sequence

$$0 \longrightarrow S_k(X_2; R) \xrightarrow{\iota} S_k(X_1; R) + S_k(X_2; R) \xrightarrow{\pi_1} S_k(X_1; R)/S_k(B; R) \longrightarrow 0 \quad (8.4.28)$$

is exact where $\iota(c) = c$ for all $c \in S_k(X_2; R)$ and $\pi_1(c_1 + c_2) = [[c_1]]_{X_1, B}$ for all $c_1 \in S_k(X_1; R)$ and $c_2 \in S_k(X_2; R)$.

We first have to prove that π_1 is well defined. Suppose $c_1, \tilde{c}_1 \in S_k(X_1; R)$ and $c_2, \tilde{c}_2 \in S_k(X_2; R)$ are such that $c_1 + c_2 = \tilde{c}_1 + \tilde{c}_2$. Then $c_1 - \tilde{c}_1 = \tilde{c}_2 - c_2$ where the right side is in $S_k(X_1; R)$ and the left side in $S_k(X_2; R)$. Hence $c_1 - \tilde{c}_1 \in S_k(B; R)$. Thus $[[c_1]]_{X_1, B} = [[\tilde{c}_1]]_{X_1, B}$.

The function ι is clearly one-to-one and the function π_1 is clearly onto. We also have that $\mathcal{R}(\iota) = S_q(X_2; R) = \mathcal{N}(\pi_1)$. Thus the sequence (8.4.28) is exact.

ii) Using the Zig-Zag lemma, Lemma 8.4.27, we get from (8.4.28) the long exact sequence

$$\xrightarrow{\tilde{\partial}_{k+1}} H_k(X_2; R) \xrightarrow{H_k(\iota)} \tilde{H}_k \xrightarrow{H_k(\pi_1)} H_k(X_1, B; R) \xrightarrow{\tilde{\partial}_k} H_{k-1}(X_2; R) \xrightarrow{H_{k-1}(\iota)} \quad (8.4.29)$$

Moreover, using again the Zig-Zag lemma with the exact sequence

$$0 \longrightarrow S_k(X_2; R) \xrightarrow{\iota} S_k(X; R) \xrightarrow{\pi_2} S_k(X; R)/S_k(X_2; R) \longrightarrow 0$$

we get the long exact sequence

$$\begin{aligned} \xrightarrow{\tilde{\partial}_{k+1}} H_k(X_2; R) &\xrightarrow{H_k(\iota)} H_k(X; R) \xrightarrow{H_k(\pi_2)} H_k(X, X_2; R) \xrightarrow{\tilde{\partial}_k} \\ &H_{k-1}(X_2; R) \xrightarrow{H_{k-1}(\iota)} \end{aligned} \quad (8.4.30)$$

Note that we could also have referred to Proposition 8.4.60 to get (8.4.30). In fact, this is an equivalent proof of Proposition 8.4.60.

iii) Suppose that (X, X_1, X_2) is an exact triad. Thus $H_k(\iota_1) : H_k(X_1, B; R) \rightarrow H_k(X, X_2; R)$ is an isomorphism. We join the long exact sequences in (8.4.29) and (8.4.30) to get the following commutative diagram.

$$\begin{array}{ccccccccccc} \xrightarrow{H_{k+1}(\pi_1)} & H_{k+1}(X_1, B; R) & \xrightarrow{\tilde{\partial}_{k+1}} & H_k(X_2; R) & \xrightarrow{H_k(\iota)} & \tilde{H}_k & \xrightarrow{H_k(\pi_1)} & H_k(X_1, B; R) & \xrightarrow{\tilde{\partial}_k} & \\ & \downarrow H_{k+1}(\iota_1) & & \downarrow \text{Id} & & \downarrow h_k & & \downarrow H_k(\iota_1) & & \\ \xrightarrow{H_{k+1}(\pi_2)} & H_{k+1}(X, X_2; R) & \xrightarrow{\tilde{\partial}_{k+1}} & H_k(X_2; R) & \xrightarrow{H_k(\iota)} & H_k(X; R) & \xrightarrow{H_k(\pi_2)} & H_k(X, X_2; R) & \xrightarrow{\tilde{\partial}_k} & \\ & & & & & & & & & \\ & & & H_{k-1}(X_2; R) & \xrightarrow{H_{k-1}(\iota)} & & & & & \\ & & & \downarrow \text{Id} & & & & & & \\ & & & H_{k-1}(X_2; R) & \xrightarrow{H_{k-1}(\iota)} & & & & & \end{array}$$

where $\iota_1 : (X_1, B) \rightarrow (X, X_2)$ is the inclusion map. The only condition that may require a little bit attention in order to prove that we really have a commutative diagram is $H_k(\iota_1) \circ H_k(\pi_1) = H_k(\pi_2) \circ h_k$. Given $[c_1 + c_2] \in \tilde{H}_k$, we have that $H_k(\pi_2)(h_k([c_1 + c_2])) = H_k(\pi_2)([c_1 + c_2]_X) = [c_1 + c_2]_{X, X_2} = [c_1]_{X, X_2}$ because $c_2 \in S_k(X_2; R)$. Moreover, $H_k(\iota_1)(H_k(\pi_1)([c_1 + c_2])) = H_k(\iota_1)([c_1]_{X_1, B}) = [c_1]_{X, X_2}$.

Since $H_k(\iota_1)$, $H_{k+1}(\iota_1)$ and all the identity maps are isomorphism, we get from the Five Lemma that h_k is an isomorphism.

iv) Suppose $h_k : \tilde{H}_k \rightarrow H_k(X_1 \cup X_2; R)$ defined by $h_k([c_1 + c_2]) = [c_1 + c_2]_X$ for all $c_1 \in S_k(X_1; R)$ and $c_2 \in S_k(X_2; R)$ is an isomorphism. We join the long exact sequences

(8.4.29) and (8.4.30) to get the following commutative diagram.

$$\begin{array}{ccccccccccc}
 \xrightarrow{\tilde{\partial}_{k+1}} & H_k(X_2; R) & \xrightarrow{H_k(\iota)} & \tilde{H}_k & \xrightarrow{H_k(\pi_1)} & H_k(X_1, B; R) & \xrightarrow{\tilde{\partial}_k} & H_{k-1}(X_2; R) & \xrightarrow{H_{k-1}(\iota)} & \\
 & \downarrow \text{Id} & & \downarrow h_k & & \downarrow H_k(\iota_1) & & \downarrow \text{Id} & & \\
 \xrightarrow{\tilde{\partial}_{k+1}} & H_k(X_2; R) & \xrightarrow{H_k(\iota)} & H_k(X; R) & \xrightarrow{H_k(\pi_2)} & H_k(X, X_2; R) & \xrightarrow{\tilde{\partial}_k} & H_{k-1}(X_2; R) & \xrightarrow{H_{k-1}(\iota)} & \\
 & & & & & & & & & \\
 & & & \tilde{H}_{k-1} & \xrightarrow{H_{k-1}(\pi_1)} & & & & & \\
 & & & \downarrow h_{k-1} & & & & & & \\
 & & & \tilde{H}_{k-1} & \xrightarrow{H_{k-1}(\pi_1)} & & & & &
 \end{array}$$

Since h_k , h_{k-1} and all the identity mappings are isomorphism, we get from the Five Lemma that $H_k(\iota_1)$ is an isomorphism. $\blacksquare$

Proposition 8.4.86 (Mayer-Vietoris)

Suppose that $(X_1 \cup X_2, X_1, X_2)$ is an exact triad. Then the sequence

$$\begin{array}{c}
 \xrightarrow{H_{k+1}} H_k(X_1 \cap X_2; R) \xrightarrow{F_k} H_k(X_1; R) \oplus H_k(X_2; R) \xrightarrow{G_k} \\
 \\
 H_k(X_1 \cup X_2; R) \xrightarrow{H_k} H_{k-1}(X_1 \cap X_2; R) \xrightarrow{F_{k-1}}
 \end{array}$$

is a long exact sequence where $F_k([c]_{X_1 \cap X_2}) = ([c]_{X_2}, -[c]_{X_1})$, $G_k([c_2]_{X_2}, [c_1]_{X_1}) = [c_1]_{X_1 \cup X_2} + [c_2]_{X_1 \cup X_2}$ and $H_k : H_k(X_1 \cup X_2; R) \rightarrow H_k(X_1 \cap X_2; R)$ is defined in the proof below. This sequence is called a **Mayer-Vietoris sequence**

Proof.

Let $B = X_1 \cap X_2$ and $X = X_1 \cup X_2$.

We consider the following three chain complexes. $\mathcal{A} = \{(S_k(B; R), \partial_k)\}_{k \in \mathbb{Z}}$, $\mathcal{B} = \{(S_k(X_1; R) \oplus S_k(X_2; R), \partial_k)\}_{k \in \mathbb{Z}}$ and $\mathcal{C} = \{(S_k(X_1; R) + S_k(X_2; R), \partial_k)\}_{k \in \mathbb{Z}}$. Let $\mathcal{F} = \{f_k\}_{k \in \mathbb{Z}}$ be the chain map between $\mathcal{A}$ and $\mathcal{B}$ defined by $f_k(c) = (c, -c)$ for all $c \in S_k(B; R)$, and $\mathcal{G} = \{g_k\}_{k \in \mathbb{Z}}$ is a chain map between $\mathcal{B}$ and $\mathcal{C}$ defined by $g_k(c_1, c_2) = c_1 + c_2$ for all $c_1 \in S_k(X_1; R)$ and $c_2 \in S_k(X_2; R)$. Note that ∂_k is defined on $S_k(X_1; R) \oplus S_k(X_2; R)$ by $\partial_k(c_1, c_2) = (\partial_k(c_1), \partial_k(c_2))$ for all $c_1 \in S_k(X_1; R)$ and $c_2 \in S_k(X_2; R)$.

The sequence

$$0 \longrightarrow S_k(B; R) \xrightarrow{f_k} S_k(X_1; R) \oplus S_k(X_2; R) \xrightarrow{g_k} S_k(X_1; R) + S_k(X_2; R) \longrightarrow 0$$

is an exact sequence for all k because f_k is one-to-one, g_k is onto and $\mathcal{R}(f_k) = \mathcal{N}(g_k) = \{(c, -c) : c \in S_k(B; R)\}$. It follows from the Zig-Zag lemma, Lemma 8.4.27, that

$$\xrightarrow{H_{k+1}} H_k(\mathcal{A}) \xrightarrow{F_k} H_k(\mathcal{B}) \xrightarrow{G_k} H_k(\mathcal{C}) \xrightarrow{H_k} H_{k-1}(\mathcal{A}) \xrightarrow{F_{k-1}}$$

is a long exact sequence where $H_k(\mathcal{A}) = H_k(B; R)$, $H_k(\mathcal{C}) \cong H_k(X; R)$ according to the previous lemma, and

$$\begin{aligned} H_k(\mathcal{B}) &= \mathcal{N}(\partial_k|_{S_k(X_1; R) \oplus S_k(X_2; R)}) / \mathcal{R}(\partial_{k+1}|_{S_{k+1}(X_1; R) \oplus S_{k+1}(X_2; R)}) \\ &\cong \mathcal{N}(\partial_k|_{S_k(X_1; R)}) / \mathcal{R}(\partial_{k+1}|_{S_{k+1}(X_1; R)}) \oplus \mathcal{N}(\partial_k|_{S_k(X_2; R)}) / \mathcal{R}(\partial_{k+1}|_{S_{k+1}(X_2; R)}) \\ &\cong H_k(X_1; R) \oplus H_k(X_2; R) \end{aligned}$$

because $\mathcal{N}(\partial_k|_{S_k(X_1; R) \oplus S_k(X_2; R)}) = \mathcal{N}(\partial_k|_{S_k(X_1; R)}) \oplus \mathcal{N}(\partial_k|_{S_k(X_2; R)})$ and $\mathcal{R}(\partial_{k+1}|_{S_{k+1}(X_1; R) \oplus S_{k+1}(X_2; R)}) = \mathcal{R}(\partial_{k+1}|_{S_{k+1}(X_1; R)}) \oplus \mathcal{R}(\partial_{k+1}|_{S_{k+1}(X_2; R)})$. Note that $F_k = H_k(f_k)$ and $G_k = H_k(g_k)$ in the notation of the Zig-Zag lemma.

It follows from the Zig-Zag lemma that $H_k : H_k(X_1 \cup X_2; R) \rightarrow H_{k-1}(X_1 \cap X_2; R)$ is defined by $H_k([c]_{X_1 \cup X_2}) = [a]_{X_1 \cap X_2}$ where $a \in Z_{k-1}(X_1 \cap X_2; R)$ is such that $f_{k-1}(a) = (a, -a) = (\partial_k(b_1), \partial_k(b_2))$ for $b_1 \in Z_k(X_1; R)$ and $b_2 \in Z_k(X_2; R)$ with $g_k(b_1, b_2) = b_1 + b_2 = c$. ■

Example 8.4.87

The torus $T^2 \subset \mathbb{R}^3$ is the union of two annulus, X_1 and X_2 , as in Figure 8.15. We have that $X_1 \cap X_2$ is the union of two disjoint circles.

We prove that (T^2, X_1, X_2) is an exact triad. We first prove that $U = X_2^\circ = X_2 \setminus X_1$ can be excised from (T^2, X_2) . We cannot use Theorem 8.4.66 to justify this operation because $\bar{U} \not\subset X_2^\circ$. To justify the excision of U from (T^2, X_2) , we have to use Proposition 8.4.80 with V as in Figure 8.15. We have that $\bar{V} \subset X_2^\circ$. Therefore, V can be excised from (T^2, X_2) according to Theorem 8.4.66. Moreover, we have that $(T^2 \setminus U, X_2 \setminus U) = (X_1, \partial X_1)$ is a deformation retract of $(T^2 \setminus V, X_2 \setminus V)$. It then follows from Proposition 8.4.80 that X_2° can be excised from (T^2, X_2) . Similarly, $X_1^\circ = X_1 \setminus X_2$ can be excised from (T^2, X_1) . This proves that (T^2, X_1, X_2) is a triad.

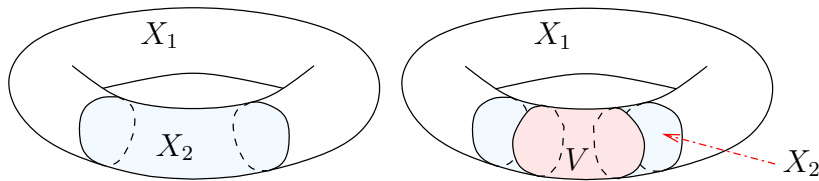


Figure 8.15: The torus is the union of two annulus, X_1 and X_2 , whose intersection is the union of two transversal sections of the torus homeomorphic to a circle.

We get the following Mayer-Vietoris sequence from the Mayer-Vietoris theorem.

$$\xrightarrow{H_{k+1}} H_k(\partial X_1; R) \xrightarrow{F_k} H_k(X_2; R) \oplus H_k(X_1; R) \xrightarrow{G_k} H_k(T^2; R) \xrightarrow{H_k} H_{k-1}(\partial X_1; R) \xrightarrow{F_{k-1}}$$

In particular, for $k = 1$ and using reduced homology, we get the exact sequence

$$\begin{aligned} 0 \xrightarrow{G_2} H_2(\mathbb{T}^2; R) \xrightarrow{H_2} H_1(\partial X_1; R) \xrightarrow{F_1} H_1(X_2; R) \oplus H_1(X_1; R) \xrightarrow{G_1} \\ H_1(\mathbb{T}^2; R) \xrightarrow{H_1} H_0^\sharp(\partial X_1; R) \xrightarrow{F_0} 0 \end{aligned}$$

because $H_0^\sharp(X_2; R) \cong H_0^\sharp(S^1; R) = 0$, $H_0^\sharp(X_1; R) \cong H_0^\sharp(S^1; R) = 0$, $H_2(X_2; R) \cong H_2(S^1; R) = 0$ and $H_2(X_1; R) \cong H_2(S^1; R) = 0$ according to the information collected at the end of Example 8.4.82. We also have that $H_1(X_2; R) \cong H_1(S^1; R) \cong R$ and $H_1(X_1; R) \cong H_1(S^1; R) \cong R$. Moreover, we get from Proposition 8.4.10 that $H_1(\partial X_1; R) \cong H_1(S^1; R) \oplus H_1(S^1; R) \cong R \oplus R$. Finally, we have that $H_0^\sharp(\partial X_1; R) \cong R$ by direct computation.

In summary, the sequence above is equivalent to

$$0 \xrightarrow{G_2} H_2(\mathbb{T}^2; R) \xrightarrow{H_2} R \oplus R \xrightarrow{F_1} R \oplus R \xrightarrow{G_1} H_1(\mathbb{T}^2; R) \xrightarrow{H_1} R \xrightarrow{F_0} 0 \quad (8.4.31)$$

If we determine the kernel and range of F_1 , then we will be able to determine the singular 1st and 2nd homology module of the torus $\mathbb{T}^2$. Recall that $F_1 = (H_1(\iota) \times H_1(\iota)) \circ \eta_1$ where the first inclusion map is from ∂X_1 into X_2 , the second inclusion map is from ∂X_1 into X_1 , and $\eta_1 : H_1(\partial X_1; R) \rightarrow H_1(\partial X_1; R) \times H_1(\partial X_1; R)$ is defined by $\eta_1([c]_{\partial X_1}) = ([c]_{\partial X_1}, [c]_{\partial X_1})$ for all $[c]_{\partial X_1} \in H_1(\partial X_1; R)$. Since $\partial X_1 = S^1$, we have from Example 8.4.82 that $H_1(\partial X_1; R) \cong R$. Thus $\mathcal{R}(F_1) \cong R$. This implies that $\mathcal{N}(F_1) \cong R$.

We may now combine this information with the sequence in (8.4.31). Since H_2 is one-to-one and $\mathcal{R}(H_2) = \mathcal{N}(F_1) \cong R$, we get that $H_2(\mathbb{T}^2; R) \cong R$. Since $\mathcal{N}(G_1) = \mathcal{R}(F_1) \cong R$, we have that $\mathcal{R}(G_1) \cong R$. Since $\mathcal{N}(H_1) = \mathcal{R}(G_1) \cong R$ and $\mathcal{R}(H_1) \cong R$, we get that $H_1(\mathbb{T}^2; R) \cong R \oplus R$. ♣

Example 8.4.88

To give another example of how Mayer-Vietoris sequences can be used, suppose that S and T are two topological spaces, and that $s_0 \in S$ and $t_0 \in T$. Let $X_1 = \{(s, t_0) : s \in S\}$ and $X_2 = \{(s_0, t) : t \in T\}$. They are both subsets of $S \times T$. If S_0 is an open neighbourhood of s_0 such that $S \setminus S_0$ is a deformation retract of $S \setminus \{s_0\}$ and T_0 is an open neighbourhood of t_0 such that $T \setminus T_0$ is a deformation retract of $T \setminus \{t_0\}$, then it follows from Proposition 8.4.80 that $(X_1 \cup X_2, X_1, X_2)$ is an exact triad. To be more precise, it follows from Proposition 8.4.80 with $A = X_2$, $X = X_1 \cup X_2$, $U = X_2 \setminus \{(s_0, t_0)\} = X \setminus X_1$ and $V = X_2 \setminus (\{s_0\} \times T_0)$ that the inclusion map $\iota_1 : (X \setminus U, A \setminus U) = (X_1, X_1 \cap X_2) \rightarrow (X, A) = (X_1 \cup X_2, X_2)$ is an excision. Note that $(X \setminus U, A \setminus U) = (X_1, \{(s_0, t_0)\})$ is a deformation retract of $(X \setminus V, A \setminus V) = (X_1 \cup (\{s_0\} \times T_0), \{s_0\} \times T_0)$ because $T \setminus T_0$ is a deformation retract of $T \setminus \{t_0\}$. Similarly, we have that the inclusion map $\iota_2 : (X_2, X_1 \cap X_2) \rightarrow (X_1 \cup X_2, X_1)$ is an excision.

Since $(X_1 \cup X_2, X_1, X_2)$ is an exact triad, we get the Mayer-Vietoris sequence

$$\begin{aligned} \xrightarrow{H_{k+1}} H_k(\{(s_0, t_0)\}; R) \xrightarrow{F_k} H_k(X_2; R) \oplus H_k(X_1; R) \xrightarrow{G_k} \\ H_k(X_1 \cup X_2; R) \xrightarrow{H_k} H_{k-1}(\{(s_0, t_0)\}; R) \xrightarrow{F_{k-1}} \end{aligned}$$

because $X_1 \cap X_2 = \{(s_0, t_0)\}$. All the previous theory is also valid for reduced homology. Since $H_k^\sharp(\{(s_0, t_0)\}; R) = 0$ for all $k \geq 0$, we get the exact sequence

$$0 \xrightarrow{F_k} H_k^\sharp(X_2; R) \oplus H_k^\sharp(X_1; R) \xrightarrow{G_k} H_k^\sharp(X_1 \cup X_2; R) \xrightarrow{H_k} 0$$

Thus $H_k^\sharp(X_1 \cup X_2; R) \cong H_k^\sharp(X_2; R) \oplus H_k^\sharp(X_1; R)$.

Suppose that S and T are two subsets of $\mathbb{R}^n$ that are homeomorphic to S^1 and intersect at a single point $\mathbf{x}_0$ that we may assume to be associated to (s_0, t_0) . The set $G_2 = X_1 \cup X_2$ is represented in Figure 8.16. We get from the previous discussion that $H_k^\sharp(G_2; R) \cong H_k^\sharp(S^1; R) \oplus H_k^\sharp(S^1; R)$.

We can generalize this construction by induction. Suppose that G_{q+1} is the union of G_q with another “leave” L which is homeomorphic to S^1 and such that $L \cap G_q = \{\mathbf{x}_0\}$ as in Figure 8.16 for $q = 4$. We can show that (G_{q+1}, G_q, L) is an exact triad and use the Mayer-Vietoris sequence to show that $H_k^\sharp(G_{q+1}; R) = H_k^\sharp(S^1; R) \oplus H_k^\sharp(G_q; R)$.

Using the information collected at the end of Example 8.4.82, we conclude that

$$H_k(G_q) \cong \begin{cases} 0 & \text{if } k > 1 \\ R^q & \text{if } k = 1 \end{cases}$$

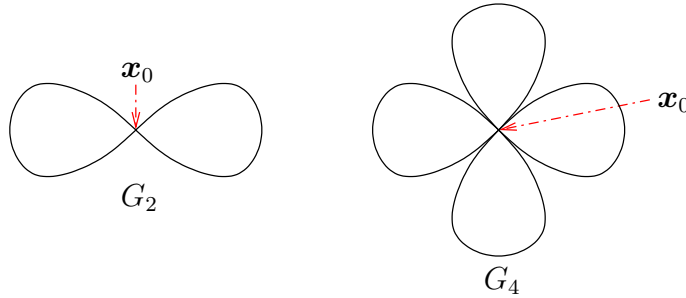


Figure 8.16: The topological spaces G_2 and G_4 referred to in Example 8.4.88

8.4.9 Further Studies

There are a lot more results about singular homology that could be presented in a first course. We list a few of them below.

Vector Fields

Singular homology theory can be used to give another proof of Proposition 7.3.4 stating that there is not a nowhere null vector field on S^{k-1} if k is odd.

Jordan-Brouwer Separation Theorem

The “obvious” but non-trivial Jordan-Brouwer separation theorem can be proved using singular homology. Let $s_{n-1} \subset S^n$ be a set homeomorphic to S^{n-1} . The Jordan-Brouwer separation theorem states that $S^n \setminus s_{n-1}$ is the union of two components having s_{n-1} as boundary. In particular, suppose that $\sigma : [0, 1] \rightarrow \mathbb{R}^2$ is a closed path such that $\sigma(t_1) \neq \sigma(t_2)$ for $t_1 \neq t_2$. Since the sphere S^2 is homeomorphic to the extended space $\mathbb{R}^2$ through stereographic projection (Figure 8.17), we get that $\sigma([0, 1])$ splits $\mathbb{R}^2$ into two components whose boundary is $\sigma([0, 1])$ because $\sigma([0, 1])$ lifts to a closed path s_1 on S^2 . In fact, the result is also true for a curve $\sigma :]0, 1[\rightarrow \mathbb{R}^2$ that does not intersect itself and such that $\sigma(t) \rightarrow \infty$ as $t \rightarrow 0$ and $t \rightarrow 1$ because this curve can also be lifted to a closed path on S^2 .

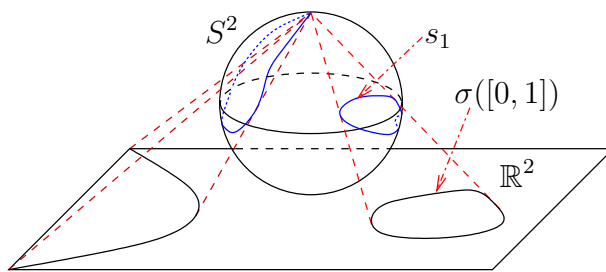


Figure 8.17: The stereographic projection of S^2 on $\mathbb{R}^2$.

8.5 Singular Cohomology

8.5.1 Cohomotopy Modules

Definition 8.5.1

Let X be a topological space and R be an integral domain. The dual space $S_k(X; R)^*$ of $S_k(X; R)$ is denoted $S^k(X; R)$. Namely, $S^k(X; R)$ is the R -module of all R -linear functionals on $S_k(X; R)$. The elements of $S^k(X; R)$ are called **cochains**.

Definition 8.5.2

Let X be a topological space and R be an integral domain. The adjoint of the boundary map $\partial_k : S_k(X; R) \rightarrow S_{k-1}(X; R)$ is the map $d_{k-1} : S^{k-1}(X; R) \rightarrow S^k(X; R)$ defined by

$$(d_{k-1}(\phi))(c) = \phi(\partial_k(c)) \quad (8.5.1)$$

for all $c \in S_k(X; R)$ and $\phi \in S^{k-1}(X; R)$. The map d_k is called the **coboundary operator**.

It is clear that d_k is well and uniquely defined by (8.5.1). Since $\partial_{k+1} \circ \partial_{k+2} = 0$, we have that $d_{k+1} \circ d_k = 0$.

Let X and Y be two topological spaces and R be an integral domain. Given $f : X \rightarrow Y$, we define an homomorphism. $S^k(f) : S^k(Y; R) \rightarrow S^k(X; R)$ by $(S^k(f)(\phi))(c) = \phi(S(f)(c))$ for all $c \in S^k(X; R)$ and $\phi \in S^k(Y; R)$. It is easy to verify that $d_k \circ S^k(f) = S^{k+1}(f) \circ d_k$.

Definition 8.5.3

Let X be a topological space and R be an integral domain. The R -modules of **cocycles** and **coboundaries** are respectively defined by $Z^k(X; R) = \{\phi \in S^k(X; R) : d_k(\phi) = 0\}$ and $B^k(X; R) = \{d_{k-1}(\psi) : \psi \in S^{k-1}(X; R)\}$. The **k^{th} cohomology module** of X is the quotient space $H^k(X; R) = Z^k(X; R)/B^k(X; R)$. The equivalence class of $H^k(X; R)$ associated to $\phi \in Z^k(X; R)$ is denoted $[\phi]_X$.

Remark 8.5.4

Let X be a topological space and R be an integral domain. Consider $[\phi] \in H^k(X; R)$. If $\tilde{\phi} \in Z^k(X; R)$ is another member of the equivalence class $[\phi] \in H^k(X; R)$, then there exists $\psi \in S^{k-1}(X; R)$ such that $\phi = \tilde{\phi} + d_{k-1}(\psi)$. Despite the fact that ϕ and $\tilde{\phi}$ may not be equal on $S_k(X; R)$, we have that $\phi(c) = \tilde{\phi}(c) + (d_{k-1}(\psi))(c) = \tilde{\phi}(c) + \psi(\partial_k(c)) = \tilde{\phi}(c)$ for all $c \in Z_k(X; R)$. ♠

Many of the results in singular homology have equivalent results in singular cohomology. For instance Proposition 8.4.10 becomes the following proposition. The proof of this proposition is almost identical to the proof of Proposition 8.4.10.

Proposition 8.5.5

If the topological space X is the union of distinct (connected) components X_j for $j \in J$, then $H^k(X; R) = \bigoplus_{j \in J} H^k(X_j; R)$.

We have a result similar to Theorem 8.3.4 for simplicial cohomology.

Theorem 8.5.6

Let X be a topological space and R be an integral domain. Then $H^k(X; R)$ is isomorphic to $(H_k(X; R))^*$.

Proof.

The proof is basically identical to the proof of Theorem 8.3.4. We repeat it here for the sake of those who skipped the sections on simplicial homology and cohomology.

The map $h : H^k(X; R) \rightarrow (H_k(X; R))^*$ defined by $h([\phi]_X) = \phi$ for $\phi \in Z^k(X; R)$ is well defined because

$$\begin{aligned} \phi \in Z^k(X; R) &\Rightarrow d_k(\phi) = 0 \Rightarrow \phi(\partial_{k+1}(c)) = (d_k\phi)(c) = 0 \text{ for all } c \in S_{k+1}(X; R) \\ &\Rightarrow \phi = 0 \text{ on } B_k(X; R) \Rightarrow \phi \in (H_k(X; R))^* . \end{aligned}$$

We have that h is one-to-one. Suppose that $[\phi_1]_X = [\phi_2]_X$. Then $\phi_1 - \phi_2 = \psi \in B^k(X; R)$. Hence $h([\phi_1]_X) = h([\phi_2]_X) + h([\psi]_X) = h([\phi_2]_X)$ because

$$\begin{aligned} \psi \in B^k(X; R) &\Rightarrow \psi = d_{k-1}\eta \text{ for some } \eta \in S^{k-1}(X; R) \\ &\Rightarrow \psi(c) = (d_{k-1}\eta)(c) = \eta(\partial_k(c)) = 0 \text{ for all } c \in Z_k(X; R) \\ &\Rightarrow \psi = 0 \text{ on } Z_k(X; R) \Rightarrow h([\psi]_X) = 0. \end{aligned}$$

We have that h is onto. Given $\phi \in (H_k(X; R))^*$, then $[\phi]_X \in H^k(X; R)$ because

$$\begin{aligned} \phi(c) = 0 \text{ for all } c \in B_k(X; R) &\Rightarrow (d_k\phi)(c) = \phi(\partial_{k+1}(c)) = 0 \text{ for all } c \in S_{k+1}(X; R) \text{ since} \\ &\quad \partial_{k+1}(c) \in B_k(X; R) \text{ for all } c \in S_{k+1}(X; R) \\ &\Rightarrow d_k\phi = 0 \Rightarrow \phi \in Z^k(K; R). \end{aligned}$$

Thus $h([\phi]_K) = \phi$.

Hence, the map h is an isomorphism between $H^k(X; R)$ and $(H_k(X; R))^*$. ■

Corollary 8.5.7

Let X be a topological space and R be an integral domain. Then $H^k(X; R)$ is isomorphic to $H_k(X; R)$.

Proof.

This follows from the previous theorem and the fact that $H_k(X; R) \cong (H_k(X; R))^*$. ■

Let X and Y be two topological spaces and R be an integral domain. Given $f : X \rightarrow Y$, we define an homomorphism. $H^k(f) : H^k(Y; R) \rightarrow H^k(X; R)$ by $H^k(f)([\phi]_Y) = [S^k(f)(\phi)]_X$ for all $[\phi]_Y \in H^k(Y; R)$. This map is well defined because $S^k(f)$ maps $Z^k(Y; R)$ to $Z^k(X; R)$ and $B^k(Y; R)$ to $B^k(X; R)$.

8.5.2 Cochain Complexes

It should not be a surprise that there is a version of the Zig-Zag lemma for cohomology. But before stating it, we need a couple of definitions.

Definition 8.5.8

Let R be an integral domain. A **cochain complex** is a sequence $\mathcal{C} = \{(C^k, d_k)\}_{k \in \mathbb{Z}}$ where the sets C^k are R -modules and the maps $d_k : C_k \rightarrow C_{k+1}$ are homomorphism such that $d_{k+1} \circ d_k = 0$ for $k \in \mathbb{Z}$.

Definition 8.5.9

A **cochain map** between two chain complexes given by $\mathcal{C} = \{(C^k, d_k)\}_{k \in \mathbb{Z}}$ and $\tilde{\mathcal{C}} = \{(\tilde{C}^k, \tilde{d}_k)\}_{k \in \mathbb{Z}}$ is a sequence $\mathcal{F} = \{f_k\}_{k \in \mathbb{Z}}$ where the maps $f_k : C^k \rightarrow \tilde{C}^k$ are homomorphism such that $\tilde{d}_k \circ f_k = f_{k+1} \circ d_k$ for $k \in \mathbb{Z}$.

Definition 8.5.10

Consider a cochain complex given by $\mathcal{C} = \{(C^k, d_k)\}_{k \in \mathbb{Z}}$. A **k-cocycle** in $\mathcal{C}$ is an element $c \in C^k$ such that $d_k(c) = 0$. The set of all k -cocycles in C^k is a R -module denoted $Z^k(\mathcal{C})$.

A **k-coboundary** in $\mathcal{C}$ is an element $c \in C^k$ with the property that there exists an element $\tilde{c} \in C^{k-1}$ such that $d_{k-1}(\tilde{c}) = c$. The set of all k -coboundaries in C^k is a R -module denoted $B^k(\mathcal{C})$.

The **singular kth cohomology module** of $\mathcal{C}$ is the R -module defined as $H^k(\mathcal{C}) = Z^k(\mathcal{C})/B^k(\mathcal{C})$. The equivalence class of $H^k(\mathcal{C})$ associated to $c \in Z^k(\mathcal{C})$ is denoted $[c]$.

It follows from the definition that a cochain map $\mathcal{F} = \{f_k\}_{k \in \mathbb{Z}}$ between two chain complexes $\mathcal{C} = \{(C^k, d_k)\}_{k \in \mathbb{Z}}$ and $\tilde{\mathcal{C}} = \{(\tilde{C}^k, \tilde{d}_k)\}_{k \in \mathbb{Z}}$ sends k -cocycles in $\mathcal{C}$ to k -cocycles in $\tilde{\mathcal{C}}$, and k -coboundaries in $\mathcal{C}$ to k -coboundaries in $\tilde{\mathcal{C}}$. Hence, we may define $H^k(\mathcal{F}) : H^k(\mathcal{C}) \rightarrow H^k(\tilde{\mathcal{C}})$ by $H^k(\mathcal{F})([c]) = [f_k(c)]$ for all $c \in Z^k(\mathcal{C})$.

We can now state the version of the Zig-Zag lemma for cohomology.

Lemma 8.5.11 (Zig-Zag Lemma)

Suppose that $\mathcal{F} = \{f_k\}_{k \in \mathbb{Z}}$ is a cochain map between the cochain complexes $\mathcal{A} = \{(A^k, d_k^{\mathcal{A}})\}_{k \in \mathbb{Z}}$ and $\mathcal{B} = \{(B^k, d_k^{\mathcal{B}})\}_{k \in \mathbb{Z}}$, and $\mathcal{G} = \{g_k\}_{k \in \mathbb{Z}}$ is a cochain map between the cochain complexes $\mathcal{B} = \{(B^k, d_k^{\mathcal{B}})\}_{k \in \mathbb{Z}}$ and $\mathcal{C} = \{(C^k, d_k^{\mathcal{C}})\}_{k \in \mathbb{Z}}$. If

$$0 \longrightarrow A^k \xrightarrow{f_k} B^k \xrightarrow{g_k} C^k \longrightarrow 0$$

is an exact sequence for all k , then there exists a long exact sequence

$$\xrightarrow{\tilde{d}_{k-1}} H^k(\mathcal{A}) \xrightarrow{H^k(\mathcal{F})} H^k(\mathcal{B}) \xrightarrow{H^k(\mathcal{G})} H^k(\mathcal{C}) \xrightarrow{\tilde{d}_k} H^{k+1}(\mathcal{A}) \xrightarrow{H^k(\mathcal{F})} \dots$$

where $\tilde{d}_k$ is induced by $d_k^{\mathcal{B}}$.

The proof of this lemma is very similar (dual) to the proof of Lemma 8.4.27 and is left as exercise to the reader.

There is also a dual to Proposition 8.4.86 which could be proved in a similar fashion or using a duality argument (see [15, 20])⁹.

⁹Duality and category theories are necessary tools in the study of cohomology and would have simplified the proofs of some results. Since we are not planning to systematically study cohomology theory, the inclusion

Proposition 8.5.12 (Mayer-Vietoris)

Suppose that $(X_1 \cup X_2, X_1, X_2)$ is an exact triad. Then the sequence

$$\begin{aligned} \xrightarrow{H_{k-1}^*} H^k(X_1 \cup X_2; R) &\xrightarrow{F_k^*} H^k(X_2; R) \oplus H^k(X_1; R) \xrightarrow{G_k^*} \\ &H^k(X_1 \cap X_2; R) \xrightarrow{H_k^*} H^{k+1}(X_1 \cup X_2; R) \xrightarrow{F_{k+1}^*} \end{aligned}$$

is an exact sequence where $F_k^*([\phi]_{X_1 \cup X_2}) = ([\phi]_{X_2}, [\phi]_{X_1})$, $G_k([\phi_2]_{X_2}, [\phi_1]_{X_1}) = [\phi_1]_{X_1 \cap X_2} - [\phi_2]_{X_1 \cap X_2}$ and $H_k^* : H^k(X_1 \cap X_2; R) \rightarrow H^k(X_1 \cup X_2; R)$ is the dual of $H_k : H_k(X_1 \cup X_2; R) \rightarrow H_k(X_1 \cap X_2; R)$; namely, $H_k^*([\phi]_{X_1 \cap X_2})([c]_{X_1 \cup X_2}) = [\phi]_{X_1 \cap X_2}(H_{k+1}([c]_{X_1 \cup X_2}))$ for all $[c]_{X_1 \cup X_2} \in H_{k+1}(X_1 \cup X_2; R)$ and $[\phi]_{X_1 \cap X_2} \in H^k(X_1 \cap X_2; R)$.

8.5.3 Relative Cohomology

It is also possible to define relative k^{th} cohomology modules.

Definition 8.5.13

Let A be a subset of a topological space X and R be an integral domain. The adjoint of the boundary operator $\bar{\partial}_k : S_k(X; R)/S_k(A; R) \rightarrow S_{k-1}(X; R)/S_{k-1}(A; R)$ is the operator $\bar{d}_{k-1} : (S_{k-1}(X; R)/S_{k-1}(A; R))^* \rightarrow (S_k(X; R)/S_k(A; R))^*$ defined

$$(\bar{d}_{k-1}(\phi))([c]_{X,A}) = \phi(\bar{\partial}_k([c]_{X,A})) \quad (8.5.2)$$

for all $[c]_{X,A} \in S_k(X; R)/S_k(A; R)$ and $\phi \in (S_{k-1}(X; R)/S_{k-1}(A; R))^*$.

It is clear that $\bar{d}_k$ is well and uniquely defined by (8.5.2).

Definition 8.5.14

Let A be a subset of a topological space X and R be an integral domain. The **relative k^{th} cohomology module** of X mod A is the quotient space $H^k(X, A; R) = \mathcal{N}(\bar{d}_k)/\mathcal{R}(\bar{d}_{k-1})$. The equivalence class of $H^k(X, A; R)$ associated to $\phi \in \mathcal{N}(\bar{d}_k)$ is denoted $[\phi]_{X,A}$.

of these subjects in these lecture notes would have added too much extra material to these lecture notes; in particular, to this chapter which is already longer than initially intended.

Proposition 8.5.15

Let A be a subset of a topological space X and R be an integral domain. The short exact sequence

$$0 \longrightarrow S_k(A; R) \xrightarrow{S_k(\iota)} S_k(X; R) \xrightarrow{\pi_k} S_k(X; R)/S_k(A; R) \longrightarrow 0$$

yields the short exact sequence

$$0 \longrightarrow (S_k(X; R)/S_k(A; R))^* \xrightarrow{\pi_k^*} S^k(X; R) \xrightarrow{S_k(\iota)^*} S^k(A; R) \longrightarrow 0$$

where $S_k(\iota)^* : S^k(X; R) \rightarrow S^k(A; R)$ is defined by $(S_k(\iota)^*(\phi))(c) = \phi(S_k(\iota)(c))$ for all $c \in S_k(A; R)$ and $\phi \in S^k(X; R)$, and $\pi_k^* : (S_k(X; R)/S_k(A; R))^* \rightarrow S^k(X; R)$ is defined by $(\pi_k^*(\phi))(c) = \phi([c]_{X,A})$ for all $c \in S_k(X; R)$ and $\phi \in (S_k(X; R)/S_k(A; R))^*$.

Proof.

i) $S_k(\iota)^*$ is onto. If $\phi \in S^k(A; R)$, then we can extend it to a linear functional $\tilde{\phi}$ on $S_k(X; R)$ by setting $\tilde{\phi}(\sigma) = \phi(\sigma)$ for all singular k -simplices with image in A and $\tilde{\phi}(\sigma) = 0$ for all singular k -simplices whose image is not in A . Hence, $(S_k(\iota)^*(\tilde{\phi}))(c) = \tilde{\phi}(S_k(\iota)(c)) = \phi(c)$ for all $c \in S_k(A; R)$.

ii) π_k^* is obviously one-to-one because $(\pi_k^*(\phi))(c) = 0$ for all $c \in S_k(X; R)$ means that $\phi([c]_{X,A}) = 0$ for all $[c]_{X,A} \in S_k(X; R)/S_k(A; R)$. Thus $\phi = 0$.

iii) If $\phi \in \mathcal{N}(S_k(\iota)^*)$, then $0 = (S_k(\iota)^*(\phi))(c) = \phi(S_k(\iota)(c))$ for all $c \in S_k(A; R)$. Therefore $\phi(c) = 0$ for all $c \in S_k(A; R)$. Hence, if we set $\tilde{\phi}([c]_{X,A}) = \phi(c)$ for all $[c]_{X,A} \in S_k(X; R)/S_k(A; R)$, then we get a well defined linear functional on $S_k(X; R)/S_k(A; R)$. Since $(\pi_k^*(\tilde{\phi}))(c) = \tilde{\phi}([c]_{X,A}) = \phi(c)$ for all $c \in S_k(X; R)$, we get that $(\pi_k^*(\tilde{\phi})) = \phi$ and so $\phi \in \mathcal{R}(\pi_k^*)$.

We have shown that $\mathcal{N}(S_k(\iota)^*) \subset \mathcal{R}(\pi_k^*)$. To prove the reverse inequality, suppose that $\phi = \pi_k^*(\tilde{\phi})$ for $\tilde{\phi} \in (S_k(X; R)/S_k(A; R))^*$. Then

$$\begin{aligned} ((S_k(\iota)^*(\phi))(c) &= ((S_k(\iota)^*(\pi_k^*(\tilde{\phi}))(c) = (\pi_k^*(\tilde{\phi}))(S_k(\iota)(c)) = \tilde{\phi}(\pi_k(S_k(\iota)(c))) \\ &= \tilde{\phi}([c]_{X,A}) = 0 \end{aligned}$$

for all $c \in S_k(A; R)$. Thus $\phi \in \mathcal{N}(S_k(\iota)^*)$. Hence $\mathcal{R}(\pi_k^*) \subset \mathcal{N}(S_k(\iota)^*)$. ■

It follows from the previous proposition that π_k^* is an isomorphism between $(S_k(X; R)/S_k(A; R))^*$ and the set $\{\phi \in S^k(X; R) : \phi(c) = 0 \text{ for all } c \in S_k(A; R)\}$.

Definition 8.5.16

The set $S^k(X, A; R) = \{\phi \in S^k(X; R) : \phi(c) = 0 \text{ for all } c \in S_k(A; R)\}$ is the **annihilator** of $S_k(A; R)$. The set of **relative cocycles** is the set $Z^k(X, A; R) = \{\phi \in S^k(X, A; R) : \phi(c) = 0 \text{ for all } c \in B_k(X, A; R)\}$ and the set of **relative coboundaries**

is the set $B^k(X, A; R) = \{d_{k-1}(\phi) : \phi \in S^{k-1}(X, A; R)\}$.

As for the relative homology, we have that $H^k(X, A; R) \cong Z^k(X, A; R)/B^k(X, A; R)$.

We can use $\pi_k^* : (S_k(X; R)/S_k(A; R))^* \rightarrow S^k(X; R)$ in Proposition 8.5.15 to obtain an isomorphism between $(S_k(X; R)/S_k(A; R))^*$ and $S^k(X, A; R)$. Using this isomorphism, then $\bar{d}_k$ is a restriction of d_k .

We have that $B^k(X, A; R) \subset Z^k(X, A; R)$ because $(d_{k-1}(\phi))(c) = \phi(\partial_k(c)) = 0$ for all $c \in B_k(X, A; R)$ if $\phi \in B^k(X, A; R)$. To be more precise, $c \in B_k(X, A; R)$ implies that $c = b + \partial_{k+1}(a)$ with $b \in S_k(B; R)$ and $a \in S_{k+1}(X; R)$. Thus $\phi(\partial_k(c)) = \phi(\partial_k(b)) = 0$ because $\phi(g) = 0$ for all $g \in S_{k-1}(A; R)$.

Unfortunately, we do not have a result as strong as in Theorem 8.5.6 for relative homology.

Theorem 8.5.17

Let A be a subset of a topological space X and R be a principal ideal domain. Then $\mathcal{H} : H^k(X, A; R) \rightarrow (H_k(X, A; R))^*$ defined by $\mathcal{H}([\phi]_{X,A})([c]_{X,A}) = \phi(c)$ for all $[c]_{X,A} \in H_k(X, A; R)$ and $[\phi]_{X,A} \in H^k(X, A; R)$ is an epimorphism; namely, a surjective homomorphism.

Instead of giving the definition of principal ideal domain, let us just say that $\mathbb{Z}$ and $\mathbb{R}$ are principal ideal domain. The proof of the previous theorem can be found in ([9]).

It is probably clear to the reader that we could go on and develop a full theory of cohomology and relative cohomology as we have done for singular homology and relative singular homology. We will not do so and refer the interested reader to [9, 21]. We have already gone much further than our initial intention at the beginning of the chapter to provide a short survey of algebraic topology.

8.5.4 Cup Product

There is however another topic that needs to be mentioned. There is a feature of singular cohomology theory that does not have any equivalent in singular homology theory. This feature is present in our study of differential forms on manifolds. It is the wedge product that combines a differential k_1 -form with a differential k_2 -form to produce a differential $(k_1 + k_2)$ -form. In cohomology theory, this new feature is called the “cup product.” We have already seen it in our study of simplicial cohomology.

Let $\delta_1 : \Delta_{k_1} \rightarrow \Delta_{k_1+k_2}$ and $\delta_2 : \Delta_{k_2} \rightarrow \Delta_{k_1+k_2}$ be the affine maps defined by $\delta_1(e_j) = e_j$ for $0 \leq j \leq k_1$ and $\delta_2(e_j) = e_{k_1+j}$ for $0 \leq j \leq k_2$.

Definition 8.5.18

Let X be a topological space and R an integral domain, and let $S^*(X; R) =$

$\bigoplus_{k \geq 0} S^k(X; R)$. The **cup product** of $\phi_1 \in S^{k_1}(X; R)$ with $\phi_2 \in S^{k_2}(X; R)$ is the linear functional $\phi_1 \cup \phi_2 \in S^{k_1+k_2}(X; R)$ defined by $(\phi_1 \cup \phi_2)(\sigma) = \phi_1(\sigma \circ \delta_1) \phi_2(\sigma \circ \delta_2)$ for all singular $(k_1 + k_2)$ -simplices σ and extended linearly to $S_{k_1+k_2}(S; R)$.

Given $\phi = \sum_{j \in \mathbb{N}} a_j \phi_j$ and $\psi = \sum_{j \in \mathbb{N}} b_j \psi_j$, where all but a finite number of $a_j \in \mathbb{R}$ and $b_j \in \mathbb{R}$ are null, and $\phi_j, \psi_j \in S^j(X; R)$ for all j . We have by definition that

$$\phi \cup \psi = \sum_{i,j \in \mathbb{N}} a_j b_i \phi_j \cup \psi_i .$$

It is not hard to prove that the cup product on $S^*(X; R)$ is bilinear, associative and that the identity element in $S^*(X; R)$ is the linear functional $\phi \in S^0(X; R)$ defined by $\phi(\sigma) = 1$ for all singular 0-simplices σ . The cup product also satisfies the relation

$$d_{k_1+k_1}(\phi_1 \cup \phi_2) = d_{k_1}(\phi_1) \cup \phi_2 + (-1)^{k_1} \phi_1 \cup d_{k_2}(\phi_2) \quad (8.5.3)$$

for all $\phi_i \in S^{k_i}(X; R)$ and $\phi_2 \in S^{k_2}(X; R)$. This is similar to item (2) of Theorem 4.5.7 about differential forms.

Let $Z^*(X; R) = \bigoplus_{k \geq 0} Z^k(X; R)$ and $B^*(X; R) = \bigoplus_{k \geq 0} B^k(X; R)$. We have that $B^*(X; R)$ is a (two-sided) ideal of $Z^*(X; R)$. Note that the cup product of two elements in $Z^*(X; R)$ yields an element in $Z^*(X; R)$. Hence $H^*(X; R) = \bigoplus_{k \geq 0} H^k(X; R)$ is a graded R -algebra. It is graded because it is the direct sums of R -modules and it is an algebra because of the cup product.

Let X and Y be two topological spaces and R be an integral domain. Given $f : X \rightarrow Y$, we define an homomorphism $S^*(f) : S^*(Y; R) \rightarrow S^*(X; R)$ as it follows. If $\phi = \sum_{j \in \mathbb{N}} a_j \phi_j$

where all but a finite number of $a_j \in \mathbb{R}$ are null and $\phi_j \in S^j(X; R)$ for all j , then we set $S^*(f)(\phi) = \sum_{j \in \mathbb{N}} a_j S^j(f)(\phi_j)$. We may use $S^*(f)$ to define $H^*(f) : H^*(Y; R) \rightarrow H^*(X; R)$.

Namely, we set $H^*(f)([\phi]_Y) = [S^*(f)(\phi)]_X$ for all $[\phi]_Y \in H^*(Y; R)$.

We define the cup product of $[\phi_1]_X, [\phi_2]_X \in H^*(X; R)$ as $[\phi_1]_X \cup [\phi_2]_X = [\phi_1 \cup \phi_2]_X$. To prove that this cup product is well defined, it suffices to consider $[\phi_i]_X \in H^{k_i}(X; R)$ with $k_i \in \mathbb{N}$ for $i = 1, 2$. Suppose that $\phi_i = \psi_i + d_k(\tau_i)$ for some $\psi_i \in Z^{k_i}(X; R)$ and $\tau_i \in S^{k_i-1}(S; R)$ for $i = 1, 2$. It follows from (8.5.3) that

$$\begin{aligned} (\psi_1 + d_{k_1}(\tau_1)) \cup \phi_2 &= \psi_1 \cup \phi_2 + d_{k_1}(\tau_1) \cup \phi_2 \\ &= \psi_1 \cup \phi_2 + (d_{k_1+k_1}(\tau_1 \cup \phi_2) + (-1)^{k_1+1} \tau_1 \cup d_{k_2}(\phi_2)) \\ &= \psi_1 \cup \phi_2 + d_{k_1+k_1}(\tau_1 \cup \phi_2) \end{aligned}$$

because $\phi_2 \in Z^{k_2}(X; R)$. Similarly,

$$\psi_1 \cup (\psi_2 + d_{k_2}(\tau_2)) = \psi_1 \cup \psi_2 + \psi_1 \cup d_{k_2}(\tau_2)$$

$$\begin{aligned}
&= \psi_1 \cup \psi_2 + (-1)^{k_2} (d_{k_1+k_1}(\psi_1 \cup \tau_2) - d_{k_1}(\psi_1) \cup \tau_2) \\
&= \psi_1 \cup \psi_2 + (-1)^{k_2} d_{k_1+k_1}(\psi_1 \cup \tau_2)
\end{aligned}$$

because $\psi_1 \in Z^{k_1}(X; R)$. Hence, by transitivity, we have

$$[\phi_1 \cup \phi_2]_X = [(\psi_1 + d_{k_1}(\tau_1)) \cup \phi_2]_X = [\psi_1 \cup \phi_2]_X = [\psi_1 \cup (\psi_2 + d_{k_2}(\tau_2))]_X = [\psi_1 \cup \psi_2]_X.$$

The relation in singular cohomology that is equivalent to item (4) in Proposition 4.1.12 is

$$[\phi_1]_X \cup [\phi_2]_X = (-1)^{k_2} [\phi_2]_X \cup [\phi_1]_X$$

for $[\phi_i]_X \in H^{k_i}(X; R)$ with $k_i \in \mathbb{N}$. To quote [9], “The proof of this theorem is surprisingly complicated.” The proof can be found in [9].

8.5.5 Relation Between Singular Cohomology and de Rham Cohomology

For this section, we follow the approach in [15].

Suppose that $S \subset \mathbb{R}^n$ is a k -dimensional smooth manifold. We are going to prove that $H^q(S)$ is (linearly) isomorphic to $H^q(S; \mathbb{R})$ for all $q \in \mathbb{N}$. From now on, $H^q(S)$ without any mention to the integral domain $\mathbb{R}$ will always refer to the q^{th} (de Rham) Cohomology Module of S defined in Definition 7.2.14.

We define a linear map $\mathcal{J} : H^q(S) \rightarrow H^q(S; \mathbb{R})$ for $q \in \mathbb{N}$ as it follows. Given $[\omega] \in H^q(S)$, we set

$$\mathcal{J}([\omega])([c]_S) = \int_c \omega \tag{8.5.4}$$

for all $[c]_S \in H_q(S; \mathbb{R})$. This map is well defined because it is independent of the representative of the equivalence class $[c]_S \in H_q(S; \mathbb{R})$ and the representative of the equivalent class $[\omega] \in H^q(S)$ used to define the integral. Suppose that $\tilde{c} \stackrel{h}{\sim} c$. Then $\tilde{c} = c + \partial_{q+1}b$ for some $b \in S_{q+1}(S; \mathbb{R})$. Using Stokes' theorem for chains, we get

$$\int_{\tilde{c}} \omega = \int_c \omega + \int_{\partial_{q+1}b} \omega = \int_c \omega + \int_b d\omega = \int_c \omega$$

because $d\omega = 0$ since $[\omega] \in H^q(S)$. Suppose that $\tilde{\omega} \sim \omega$. Then $\tilde{\omega} = \omega + d\eta$ for some differential $(q-1)$ -form η on S . Again, using Stokes' theorem for chains, we get

$$\int_c \tilde{\omega} = \int_c \omega + \int_c d\eta = \int_c \omega + \int_{\partial_q c} \eta = \int_c \omega$$

because $\partial_q c = 0$ since $[c]_S \in H_q(S; \mathbb{R})$. It is easy to verify that $\mathcal{J}(\omega) \in H^q(S; \mathbb{R})$.

There is a version of Proposition 8.5.12 for cohomology on manifolds. Suppose that S is a k -dimensional smooth manifold. To reformulate our definition of the cohomology module $H^q(S)$ given in Definition 7.2.14 of Chapter 7, we set $Z^k(\Omega^q(S), d) = \{\omega \in \Omega^q(S) : d\omega = 0\}$ and $B^k(\Omega^q(S), d) = \{d\omega : \omega \in \Omega^{q-1}(S)\}$. Then $H^q(S) = Z^q(S)/B^q(S)$.

Proposition 8.5.19 (Mayer-Vietoris)

Suppose that S is a k -dimensional smooth manifold and that $S = U \cup V$ where U and V are two open subsets of S . Then the sequence

$$\xrightarrow{\tilde{H}_{q-1}^*} H^q(S) \xrightarrow{\tilde{F}_q^*} H^q(U) \oplus H^q(V) \xrightarrow{\tilde{G}_q^*} H^q(U \cap V) \xrightarrow{\tilde{H}_q^*} H^{q+1}(S) \xrightarrow{\tilde{F}_{q+1}^*}$$

is an exact sequence where $\tilde{F}_q^*([\omega]) = ([\omega|_U], [\omega|_V])$, $\tilde{G}_q([\omega_1], [\omega_2]) = [\omega_1|_{U \cap V}] - [\omega_2|_{U \cap V}]$ and $\tilde{H}_q^* : H^q(U \cap V) \rightarrow H^{q+1}(S)$ is defined in proof below.

Proof.

Let $B = U \cap V$. We consider the following three cochain complexes. $\mathcal{A} = \{(\Omega^q(S), d)\}_{k \in \mathbb{Z}}$, $\mathcal{B} = \{(\Omega^q(U) \oplus \Omega^q(V), d)\}_{k \in \mathbb{Z}}$ and $\mathcal{C} = \{(\Omega^q(B), d)\}_{k \in \mathbb{Z}}$. Let $\mathcal{F} = \{\tilde{f}_k\}_{k \in \mathbb{Z}}$ be the cochain map between $\mathcal{A}$ and $\mathcal{B}$ defined by $\tilde{f}_k(\eta) = (\eta|_U, \eta|_V)$ for all $\eta \in \Omega^q(S)$, and $\mathcal{G} = \{\tilde{g}_k\}_{k \in \mathbb{Z}}$ be the cochain map between $\mathcal{B}$ and $\mathcal{C}$ defined by $\tilde{g}_k(\eta_1, \eta_2) = \eta_1|_{U \cap V} - \eta_2|_{U \cap V}$ for all $\eta_1 \in \Omega^q(U_1)$ and $\eta_2 \in \Omega^q(U_2)$. Note that the differential operator d is defined on $\Omega^q(U) \oplus \Omega^q(V)$ by $d(\eta_1, \eta_2) = (d(\eta_1), d(\eta_2))$ for all $\eta_1 \in \Omega^q(U_1)$ and $\eta_2 \in \Omega^q(U_2)$.

Our goal is to prove that the sequence

$$0 \longrightarrow \Omega^q(S) \xrightarrow{\tilde{f}_q} \Omega^q(U) \oplus \Omega^q(V) \xrightarrow{\tilde{g}_q} \Omega^q(B) \longrightarrow 0$$

is an exact sequence. The conclusion of the proposition will then follow from the Zig-Zag lemma, Lemma 8.5.11.

It follows from the Zig-Zag lemma that $\tilde{H}_q^* : H^q(U \cap V) \rightarrow H^{q+1}(S)$ is defined by $\tilde{H}_q^*([\gamma]) = [\alpha]$ where $\alpha \in Z^{q+1}(S)$ is such that $\tilde{f}_{k+1}(\alpha) = (\alpha|_U, \alpha|_V) = (d\beta_1, d\beta_2)$ for $\beta_1 \in Z^q(U)$ and $\beta_2 \in Z^q(V)$ with $\tilde{g}_k(\beta_1, \beta_2) = \beta_1|_{U \cap V} - \beta_2|_{U \cap V} = \gamma$.

i) We have that $\tilde{f}_q$ is one-to-one because $\tilde{f}_q(\eta) = (\eta|_U, \eta|_V) = (0, 0)$ implies that $\eta|_S = 0$ since $S = U \cup V$.

ii) To prove that $\mathcal{R}(\tilde{f}_q) = \mathcal{N}(\tilde{g}_k)$, we note that

$$\tilde{g}_q(\tilde{f}_q(\eta)) = \tilde{g}_q(\eta|_U, \eta|_V) = \eta|_B - \eta|_B = 0$$

for all $\eta \in \Omega^q(S)$. Thus $\mathcal{R}(\tilde{f}_q) \subset \mathcal{N}(\tilde{g}_k)$.

Suppose that $(\eta_1, \eta_2) \in \mathcal{N}(\tilde{g}_k)$. Then $\eta_1|_B - \eta_2|_B = 0|_B$. Let

$$\eta = \begin{cases} \eta_1 & \text{on } U \\ \eta_2 & \text{on } U \setminus V \end{cases}$$

We have that $\eta \in \Omega^q$ because $\eta_1 = \eta_2$ on $B = U \cap V$. Since $\tilde{f}_k(\eta) = (\eta|_U, \eta|_V) = (\eta_1, \eta_2)$, we have that $(\eta_1, \eta_2) \in \mathcal{R}(\tilde{f}_k)$. This proves that $\mathcal{N}(\tilde{g}_k) \subset \mathcal{R}(\tilde{f}_q)$.

iii) We now prove that $\tilde{g}_q$ is onto. Given $\eta \in \Omega^q(B)$, we need to find $\eta_1 \in \Omega^q(U)$ and $\eta_2 \in \Omega^q(V)$ such that $\tilde{g}_q(\eta_1, \eta_2) = \eta_1|_B - \eta_2|_B = \eta$.

Since $\{U, V\}$ is an open cover of S , we may use Theorem 3.4.1 to get a partition of unity $\{\psi_j\}_{j \in \mathbb{N}^+}$ subordinate to the open cover $\{U, V\}$ of $\text{supp } \eta$. Let $\phi_1 = \sum_{\text{supp } \psi_j \subset U} \psi_j$ and $\phi_2 = \sum_{\text{supp } \psi_j \subset V} \psi_j$. The function ϕ_1 and ϕ_2 are well defined and of class C^∞ because the sums are finite on compact sets according to Proposition 3.4.4. We have that $\text{supp } \phi_1 \subset U$ and $\text{supp } \phi_2 \subset V$. Let

$$\eta_1 = \begin{cases} \phi_2 \eta & \text{on } U \cap \text{supp } \phi_2 \\ 0 & \text{on } U \setminus \text{supp } \phi_2 \end{cases} \quad \text{and} \quad \eta_2 = \begin{cases} -\phi_1 \eta & \text{on } V \cap \text{supp } \phi_1 \\ 0 & \text{on } V \setminus \text{supp } \phi_1 \end{cases}$$

Since $U \cap \text{supp } \phi_2 \subset U \cap V$ and $V \cap \text{supp } \phi_1 \subset V \cap U$ where η is defined, we get that η_1 and η_2 are well defined, $\eta_1 \in \Omega^q(U)$ and $\eta_2 \in \Omega^q(V)$. We also have that $\tilde{g}_k(\eta_1, \eta_2) = \eta_1|_B - \eta_2|_B = \phi_2 \eta + \phi_1 \eta = (\phi_2 + \phi_1) \eta = \eta$ on B . ■

The operator $\mathcal{J}$ has some natural properties.

Proposition 8.5.20

Suppose that S_1 and S_2 are k -dimensional smooth manifolds. If $f : S_1 \rightarrow S_2$ is a smooth function, then

$$\begin{array}{ccc} H^q(S_2) & \xrightarrow{f^\#} & H_k^q(S_1) \\ \downarrow \mathcal{J} & & \downarrow \mathcal{J} \\ H^q(S_2; \mathbb{R}) & \xrightarrow{H^q(f)} & H^q(S_1; \mathbb{R}) \end{array}$$

is a commutative diagram for all $q \in \mathbb{N}$, where $f^\#$ is defined in Subsection 7.2.2.

Proof.

If $[\omega] \in H^q(S_2)$ and σ is a q -cycle in S_1 , then $f^\#([\omega]) = [f^*(\omega)]$ and

$$\begin{aligned} \mathcal{J}(f^\#([\omega]))([\sigma]_{S_1}) &= \int_{\sigma} f^*(\omega) = \int_{\Delta_q} \sigma^*(f^*(\omega)) = \int_{\Delta_q} (f \circ \sigma)^*(\omega) = \int_{f \circ \sigma} \omega \\ &= \mathcal{J}([\omega])([f \circ \sigma]_{S_2}) = \mathcal{J}([\omega])([S_q(f)(\sigma)]_{S_2}) = H^q(f)(\mathcal{J}([\omega]))([\sigma]_{S_1}). \end{aligned}$$

Hence, by linearity, $\mathcal{J}(f^\#([\omega]))([c]_{S_1}) = H^q(f)(\mathcal{J}([\omega]))([c]_{S_1})$ for all $[c]_{S_1} \in H_q(S_1; \mathbb{R})$. ■

Lemma 8.5.21

Suppose that S is a k -dimensional smooth manifold and that $S = U \cap V$ where U and V are open subsets of S . Then (S, U, V) is an exact triad.

Proof.

We have that $S \setminus U = V \setminus U$ is a closed subset (in the topology induced on S) of the open set V . It follows from Excision theorem, Theorem 8.4.66, that $V \setminus U$ can be excised from

(S, V) ; namely, $H_k(\iota) : H_k(U, U \cap V; \mathbb{R}) \rightarrow H_k(S, V; \mathbb{R})$ is an isomorphism. Similarly, $U \setminus V$ can be excised from (S, U) ; namely, $H_k(\iota) : H_k(V, U \cap V; \mathbb{R}) \rightarrow H_k(S, V; \mathbb{R})$. ■

Proposition 8.5.22

Suppose that S is k -dimensional smooth manifold and that $S = U \cup V$ where U and V are open subsets of S . Then

$$\begin{array}{ccc} H^{q-1}(U \cap V) & \xrightarrow{\tilde{H}_{q-1}^*} & H^q(S) \\ \downarrow \mathcal{J} & & \downarrow \mathcal{J} \\ H^{q-1}(U \cap V; \mathbb{R}) & \xrightarrow{H_{q-1}^*} & H^q(S; \mathbb{R}) \end{array}$$

is a commutative diagram for all $q \in \mathbb{N}$ where H_q^* and $\tilde{H}_q^*$ are defined in Propositions 8.5.12 and 8.5.19 respectively.

Proof.

Note that we can refer to Propositions 8.5.12 in the statement of the proposition because (S, U, V) is an exact triad according to Lemma 8.5.21.

The diagram states that

$$\mathcal{J}(\tilde{H}_{q-1}^*([\omega]))([c]_S) = H_{q-1}^*(\mathcal{J}([\omega]))([c]_S) \quad (8.5.5)$$

for all $[c]_S \in H_q(S; \mathbb{R})$ and $[\omega] \in H^{q-1}(U \cap V)$. From the duality mentioned in Proposition 8.5.12, we get that (8.5.5) is equivalent to

$$\mathcal{J}(\tilde{H}_{q-1}^*([\omega]))([c]_S) = \mathcal{J}([\omega])(H_q([c]_S)) \quad (8.5.6)$$

for all $[c]_S \in H_q(S; \mathbb{R})$ and $[\omega] \in H^{q-1}(U \cap V)$. We can restate (8.5.6) as

$$\int_c \eta = \int_a \omega \quad (8.5.7)$$

where $[\eta] = \tilde{H}_{q-1}^*([\omega]) \in H^q(S)$ and $[a]_{U \cap V} = H_q([c]_S) \in H_q(U \cap V)$ for all $[c]_S \in H_q(S; \mathbb{R})$ and $[\omega] \in H^{q-1}(U \cap V)$. We prove (8.5.7).

If we refer to the definition of H_q given in the proof of Proposition 8.4.86, then we have that $[a]_{U \cap V} = H_q([c]_S)$ is given by $a \in Z_{q-1}(U \cap V; \mathbb{R})$ such that $f_{k-1}(a) = (a, -a) = (\partial_k(b_1), \partial_k(b_2))$ for $b_1 \in Z_k(U; \mathbb{R})$ and $b_2 \in Z_k(V; \mathbb{R})$ with $g_k(b_1, b_2) = b_1 + b_2 = c$ in S . In particular, we have that $a = \partial_k(b_1) = \partial_k(b_2)$ with $\partial_k(b_1)(\Delta_k) \subset U \cap V$ and $\partial_k(b_2)(\Delta_k) \subset U \cap V$.

If we now refer to the definition of $\tilde{H}_{q-1}^*$ given in the proof of Proposition 8.4.86, then we have that $[\eta] = \tilde{H}_{q-1}^*([\omega])$ is given by $\eta \in Z^{q+1}(S)$ is such that $\tilde{f}_{k+1}(\eta) = (\eta|_U, \eta|_V) = (d\beta_1, d\beta_2)$ for $\beta_1 \in Z^q(U)$ and $\beta_2 \in Z^q(V)$ with $\tilde{g}_k(\beta_1, \beta_2) = \beta_1|_{U \cap V} - \beta_2|_{U \cap V} = \omega$. In particular, we have that $\eta|_U = d\beta_1$ and $\eta|_V = d\beta_2$.

Since $\partial_q(b_1) + \partial_q(b_2) = \partial_q(c) = 0$ because $c \in Z_q(S; \mathbb{R})$ and $d\beta_1|_{U \cap V} - d\beta_2|_{U \cap V} = d\omega = 0$ because $\omega \in Z^{q-1}(U \cap V)$, we get from Stokes' theorem that

$$\begin{aligned} \int_a \omega &= \int_{\partial_k(b_1)} \omega = \int_{\partial_k(b_1)} (\beta_1 - \beta_2) = \int_{\partial_k(b_1)} \beta_1 - \int_{\partial_k(b_1)} \beta_2 = \int_{\partial_k(b_1)} \beta_1 + \int_{\partial_k(b_2)} \beta_2 \\ &= \int_{b_1} d\beta_1 + \int_{b_2} d\beta_2 = \int_{b_1} \eta + \int_{b_2} \eta = \int_c \eta. \end{aligned}$$

This proves (8.5.7). ■

Theorem 8.5.23 (de Rham)

Let $S \subset \mathbb{R}^n$ be a k -dimensional smooth manifold and $q \in \mathbb{N}$. The map $\mathcal{J} : H^q(S) \rightarrow H^q(S; \mathbb{R})$ defined in (8.5.4) is an isomorphism.

Proof.

For convenience, we will say that a smooth manifold M is a **de Rham** manifold if $\mathcal{J} : H^q(M) \rightarrow H^q(M; \mathbb{R})$ is an isomorphism.

Claim 1) If U is a convex open subset of $\mathbb{R}^k$, then U is a de Rham manifold.

It follows from Corollary 7.2.22 (see comments after the corollary) and Corollary 8.4.41 that $H^q(U) = 0$ and $H^q(U; \mathbb{R}) \cong H_q(U; \mathbb{R}) = 0$ for $q > 0$, where Corollary 8.5.7 was used for the last relation. The condition for being a de Rham manifold is therefore trivially satisfied.

If $q = 0$, then we get from Corollaries 7.2.21 and 8.4.41 that $H^0(U) \cong \mathbb{R}$ and $H^0(U; \mathbb{R}) \cong H_0(U; \mathbb{R}) \cong \mathbb{R}$. Hence $H^0(U)$ is the set of constant functions on U and $H^0(U; \mathbb{R})$ is the set of constant functions on $\{e_u : u \in U\}$ where $e_u : \Delta_0 \rightarrow U$ is defined by $e_u(e_0) = u$ for all $u \in U$. If we identify e_u to u , then we have the set of constant functions f on U in both cases. Hence

$$\mathcal{J}(f)(\sigma) = \int_{\sigma} f = \int_{\Delta_0} \sigma^*(f) = f(\sigma(e_0)) = f(u) \neq 0$$

for all singular 0-simplices $\sigma = e_u$ in U unless $f = 0$. Thus $\mathcal{J} : H^0(U) \rightarrow H^0(U; \mathbb{R})$ is an isomorphism.

Claim 2) If $\{S_j\}_{j=0}^{\infty}$ is a collections of disjoint de Rham manifolds where each S_j is connected, then $\bigcup_{j=0}^{\infty} S_j$ is a de Rham manifold.

This follows from the fact that the inclusion $\iota : S_i \rightarrow \bigcup_{j=0}^{\infty} S_j$ induces an isomorphism

$\tilde{\iota} : \bigoplus_{j=0}^{\infty} H_q(S_j; \mathbb{R}) \rightarrow H_q\left(\bigcup_{j=0}^{\infty} S_j; \mathbb{R}\right)$. Hence, using Propositions 8.4.10 and 8.5.5, we get that

$H_q(\tilde{\iota}) : \bigoplus_{j=0}^{\infty} H_q(S_j; \mathbb{R}) \rightarrow H_q\left(\bigcup_{j=0}^{\infty} S_j; \mathbb{R}\right)$ and $\tilde{\iota}^{\#} : \bigoplus_{j=0}^{\infty} H^q(S_j) \rightarrow H^q\left(\bigcup_{j=0}^{\infty} S_j\right)$ are isomorphism

(see comment after Proposition 7.2.23). We have from Proposition 8.5.20 that

$$\begin{array}{ccc} H^q\left(\bigcup_{j=0}^{\infty} S_j\right) & \xrightarrow{\tilde{\iota}^\#} & \bigoplus_{j=0}^{\infty} H^q(S_j) \\ \downarrow \mathcal{J} & & \downarrow \mathcal{J} \\ H^q\left(\bigcup_{j=0}^{\infty} S_j; \mathbb{R}\right) & \xrightarrow{H_q(\tilde{\iota})} & \bigoplus_{j=0}^{\infty} H^q(S_j; \mathbb{R}) \end{array}$$

is a commutative diagram. Since $H_q(\tilde{\iota})$ and $\tilde{\iota}^\#$ are isomorphism and $\mathcal{J} : H^q(S_j) \rightarrow H^q(S_j; \mathbb{R})$ is an isomorphism for all j , we get that $\mathcal{J} : H^q\left(\bigcup_{j=0}^{\infty} S_j\right) \rightarrow H^q\left(\bigcup_{j=0}^{\infty} S_j; \mathbb{R}\right)$ is an isomorphism.

Claim 3) If $U = U_1 \cup U_2$ where U_1 and U_2 are two open subsets of S , and the sets U_1 , U_2 and $U_1 \cap U_2$ are de Rham manifolds, then U is a de Rham manifold.

According to Lemma 8.5.21, $(U_1 \cup U_2, U_1, U_2)$ is an exact triad. Hence, we get from Proposition 8.5.12 and Proposition 8.5.19 that each row of the following diagram is exact.

$$\begin{array}{ccccccc} \xrightarrow{\tilde{F}_{q-1}^*} & H^{q-1}(U_1) \oplus H^{q-1}(U_2) & \xrightarrow{\tilde{G}_{q-1}^*} & H^{q-1}(U_1 \cap U_2) & \xrightarrow{\tilde{H}_{q-1}^*} & H^q(U) & \xrightarrow{\tilde{F}_q^*} \\ & \downarrow \mathcal{J} & & \downarrow \mathcal{J} & & \downarrow \mathcal{J} & \\ \xrightarrow{F_{q-1}^*} & H^{q-1}(U_2; \mathbb{R}) \oplus H^{q-1}(U_1; \mathbb{R}) & \xrightarrow{G_{q-1}^*} & H^{q-1}(U_1 \cap U_2; \mathbb{R}) & \xrightarrow{H_{q-1}^*} & H^k(U; \mathbb{R}) & \xrightarrow{F_k^*} \\ & & & & & & \\ & H^q(U_1) \oplus H^q(U_2) & \xrightarrow{\tilde{G}_q^*} & H^q(U_1 \cap U_2) & \xrightarrow{\tilde{H}_q^*} & & \\ & \downarrow \mathcal{J} & & \downarrow \mathcal{J} & & & \\ & H^k(U_2; \mathbb{R}) \oplus H^k(U_1; \mathbb{R}) & \xrightarrow{G_k^*} & H^k(U_1 \cap U_2; \mathbb{R}) & \xrightarrow{H_k^*} & & \end{array}$$

It follows from Propositions 8.5.20 and 8.5.22 that this diagram is commutative. Since the first, second, fourth and fifth maps $\mathcal{J}$ are isomorphism by assumption, we get from the Five Lemma that $\mathcal{J} : H^k(U) \rightarrow H^k(U; \mathbb{R})$ is also an isomorphism.

Claim 4) If $U = \bigcup_{j=1}^J U_j$ where each U_j is an open subset of S and a de Rham manifold, and any finite intersection of U_j with $1 \leq j \leq J$ is a de Rham manifold, then U is a de Rham manifold.

The proof is by induction on J . The result is trivial if $J = 1$. Suppose that the result is true for J and that $U = \bigcup_{j=1}^{J+1} U_j$ where each U_j is an open subset of S and a de Rham manifold, and any finite intersection of U_j for $1 \leq j \leq J+1$ is a de Rham manifold.

By induction, we have that $\tilde{U} = \bigcup_{j=1}^J U_j$ is a de Rham manifold. If we apply Claim 3 to $\tilde{U}$ and U_{J+1} , then we get that $U = \tilde{U} \cup U_{J+1}$ is a de Rham manifold. Note that

$\tilde{U} \cap U_{J+1} = \bigcup_{j=1}^J (U_j \cap U_{J+1})$ is a de Rham manifold because each $U_j \cap U_{J+1}$ is a de Rham manifold by assumption and any finite intersection of $U_j \cap U_{J+1}$ is a finite intersection of U_j with $1 \leq j \leq J+1$ which is also a de Rham manifold by assumption. Thus, by induction hypothesis with $U_j \cap U_{J+1}$ with $1 \leq j \leq J$ instead of U_j with $1 \leq j \leq J$, we get that $U = \tilde{U} \cup U_{J+1}$ is a de Rham manifold.

Claim 5) If S has a basis of open sets $\{U_i\}_{i \in I}$ such that each U_i is a connected, bounded and open set which is also a de Rham manifold, and finite intersections of sets from $\{U_i\}_{i \in I}$ are de Rham manifolds, then S is a de Rham manifold.

Let $\{\phi_j\}_{j \in \mathbb{N}^+}$ be a partition of unity subordinated to $\{U_i\}_{i \in I}$. Consider the function $g = \sum_{j=1}^{\infty} j\phi_j$ on S . Note that the sum $\sum_{j=1}^{\infty} j\phi_j$ is finite on compact subsets of S according

to Proposition 3.4.4. Thus g is well defined. We also have that $g(\mathbf{u}) = \sum_{j=1}^{\infty} j\phi_j(\mathbf{u}) \geq \sum_{j=1}^{\infty} \phi_j(\mathbf{u}) = 1$ for all $\mathbf{u} \in S$.

Moreover $g^{-1}([0, a]) \subset \bigcup_{j=1}^{J_a} \overline{U}_{i_j}$ for some $i_j \in I$ such that $\text{supp } \phi_j \subset U_{i_j}$ and J_a is a positive

integer greater or equal to a . To prove this statement, we prove that, if $\mathbf{u} \notin \bigcup_{j=1}^{J_a} \overline{U}_{i_j}$, then $d(\mathbf{u}) > 1$. Since $\phi_j(\mathbf{u}) = 0$ for $1 \leq j \leq J_a$ because $\mathbf{u} \notin U_{i_j} \supset \text{supp } \phi_j$ for $1 \leq j \leq J_a$, we get that $g(\mathbf{u}) = \sum_{j=1}^{\infty} j\phi_j(\mathbf{u}) = \sum_{j=J_a+1}^{\infty} j\phi_j(\mathbf{u}) \geq (J_a + 1) \sum_{j=J_a+1}^{\infty} \phi_j(\mathbf{u}) = J_a + 1 > c$.

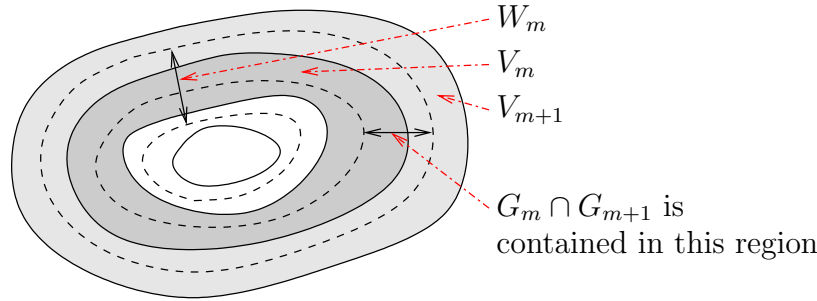
Note that $g^{-1}([0, a])$ is a closed subset of the compact set $\bigcup_{j=1}^{J_a} \overline{U}_{i_j}$. Thus $g^{-1}([0, a])$ is a compact subset of S for all $a > 0$.

Let $V_m = g^{-1}([m, m+1])$ and $W_m = g^{-1}(]m-1/2, m+3/2[)$ for $m \in \mathbb{N}$. Since W_m is an open subset of S , it is the union of some open sets from $\{U_i\}_{i \in I}$. Since $V_m \subset W_m$, these subsets form an open cover of the compact set V_m . Thus, there exists a finite subcover $\{U_{i_{m,r}}\}_{r=1}^{R_m}$ of V_m . We have from Claim 4 that $G_m = \bigcup_{r=1}^{R_m} U_{i_{m,r}}$ is a de Rham manifold. Note

that G_m is the union of more than one distinct connected de Rham manifolds if $\bigcap_{r=1}^{R_m} U_{i_{m,r}} = \emptyset$.

By construction, we have that $V_m \subset G_m \subset W_m$. Thus $G_{m_1} \cap G_{m_2} = \emptyset$ for $|m_2 - m_1| > 1$. Hence $G_e = \bigcup_{m \text{ even}} G_m$ and $G_o = \bigcup_{m \text{ odd}} G_m$ are both the union of distinct connected de Rham manifolds. It follows from Claim 2 that G_e and G_o are de Rham manifolds.

Since $G_m \cap G_{m+1} = \bigcup_{r=1}^{R_m} \bigcup_{\tilde{r}=1}^{R_{m+1}} U_{i_{m,r}} \cap U_{i_{m+1,\tilde{r}}}$, we get that $G_m \cap G_{m+1}$ is the finite union of open and connected de Rham manifolds because any finite intersection of U_i is de Rham manifolds by assumption. Thus $G_m \cap G_{m+1}$ is the disjoint union of open and connected de Rham Manifolds according to Claim 4. Since $(G_{m_1} \cap G_{m_1+1}) \cap (G_{m_2} \cap G_{m_2+1}) = \emptyset$ for $m_1 \neq m_2$, it follows that $G_e \cap G_o = \bigcup_{m=1}^{\infty} (G_m \cap G_{m+1})$ is the disjoint union of open and connected de Rham Manifolds. We get from Claim 2 that $G_e \cap G_o$ is a de Rham manifold. Finally, it follows from Claim 3 that $S = G_e \cup G_o$ is a de Rham manifold.



Claim 6) If $U \subset \mathbb{R}^k$ is an open set, then U is a de Rham manifold.

We have that $U = \bigcup_{i \in I} U_i$ where each U_i is an open ball in $\mathbb{R}^k$. It follows from Claim 1 that the U_i are de Rham manifold since open balls are convex. The U_i are also connected and bounded. Since the finite intersection of open balls is also a convex set, we have that finite intersections of sets from $\{U_i\}_{i \in I}$ are de Rham manifolds. Therefore the collection $\{U_i\}_{i \in I}$ meets the requirements of Claim 5. Hence U is de Rham.

Claim 7) Finally, if $S \subset \mathbb{R}^n$ is a k -dimensional smooth manifold, then S is a de Rham manifold.

The manifold S has a basis of open set $\{U_i\}_{i \in I}$ where each U_i is associated to a local chart (W_i, U_i, ϕ_i) of S . We may select the U_i to be bounded and connected¹⁰. Since each U_i is diffeomorphic to the open set $W_i \subset \mathbb{R}^k$ and W_i is a de Rham manifold according to Claim 6, we get from Proposition 8.5.20 that U_i is a de Rham manifold. This is also true for the finite intersection of set from $\{U_i\}_{i \in I}$ be definition of local charts. It follows from Claim 5 that S is a de Rham manifold. ■

8.6 Relation Between Simplicial and Singular Theory

This section is not needed for the next chapter and, in a sense, is not related to the goal of these lecture notes. However, after having introduced both the simplicial and singular homology theory, we feel that it is important to know how they are related and thus how

¹⁰For instance, for each $\mathbf{u}$, we may use $U_{\mathbf{u}} = S \cap B_r(\mathbf{u})$ for r small enough, and use $I = S$ as set of indices.

they may interact. After having read the sections about the simplicial and singular theories, the reader has probably realized that there are a lot of similarity between the theoretical results and techniques used in both theories. As we demonstrate in this section, this is not really surprising.

The style of this section is a little less detailed than in the rest of these lecture notes.

8.6.1 Relation Between Simplicial Homology and Singular Homology

We follow the presentation in [21].

Through out this section, L is a subcomplex of a simplicial complex K and R is an integral domain. We assume that an indexing of the vertices of K has been selected. More specifically, we assume that $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots$ are the vertices of K . In particular, we assume that there could be countably infinitely many open simplices in K . This is more general than in Sections 8.2 and 8.3 where we assumed that there were only a finite number of open simplices in K . Most of what we said in Sections 8.2 and 8.3 is true in the case where K contains countably infinitely many open simplices. However, there are a few exceptions. Barycentric subdivisions and simplicial approximation are two concepts that need to be generalized to address the issue that K may contain countably infinitely many open simplices. The reader may find more information about this subject in [21].

Adding the possibility that the simplicial complex K could contain countably infinitely many open simplices does not really increase the level of complexity in this section. The only extra work is in part (II) of the proof of Proposition 8.6.7.

Ordered Simplicial Homology

We need to enlarge $C_k(K; R)$ to be able to define a mapping onto $S_k([K]; R)$. For instance, we need a set large enough to be able to map an element of this set to a singular k -simplex $\sigma : \Delta_k \rightarrow [K]$ representing a constant function into $[K]$.

In addition to the k -simplices $(s) \in K$, we also consider elements of the form $(\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_k)$ where $\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_k$ are vertices of a k -simplex $(s) \in K$, and the $\mathbf{y}_j$ may not be all distinct.

Definition 8.6.1

Elements of the form $(\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_k)$ as described above are called **ordered k -simplices** of K .

Definition 8.6.2

Let $\tilde{C}_k(K; R)$ be the free abelian group generated by the elements $(\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_k)$ as described above. Elements of $\tilde{C}_k(K; R)$ are called **ordered k -chains** of K .

We define the boundary operator $\tilde{\partial}_k$ on $\tilde{C}_K(K; R)$ as we did for the boundary operator ∂_k on $C_K(K; R)$; namely,

$$\tilde{\partial}_k(\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_k) = \sum_{i=0}^k (-1)^i (\mathbf{y}_0, \mathbf{y}_1, \dots, \widehat{\mathbf{y}}_i, \dots, \mathbf{y}_k) .$$

In analogy to $H_k(K; R) = \mathcal{N}(\partial_k)/\mathcal{R}(\partial_{k-1})$, we can define $\tilde{H}_k(K; R) = \mathcal{N}(\tilde{\partial}_k)/\mathcal{R}(\tilde{\partial}_{k-1})$. Likewise, we define the quotient space $\tilde{C}_k(K, L; R) = \tilde{C}_k(K; R)/\tilde{C}_k(L; R)$. Since $\tilde{\partial}_k : \tilde{C}_k(L; R) \rightarrow \tilde{C}_{k-1}(L; R)$, we may define the operator $\tilde{\partial}_k : \tilde{C}_k(K, L; R) \rightarrow \tilde{C}_{k-1}(K, L; R)$ by $\tilde{\partial}_k([\![c]\!]_{K,L}) = [\![\tilde{\partial}_k(c)]\!]_{K,L}$ for all $c \in \tilde{C}_k(K; R)$, where $[\![\cdot]\!]_{K,L}$ represents an equivalence class in $\tilde{C}_k(K, L; R)$. We also define $\tilde{H}_k(K, L; R) = \mathcal{N}(\tilde{\partial}_k)/\mathcal{R}(\tilde{\partial}_{k-1})$.

Remark 8.6.3

We will use the same notation $[\cdot]_K$ to denote equivalence classes in $H_k(K; R)$, $\tilde{H}_k(K; R)$ and their reduced form. Similarly, we will use the same notation $[\cdot]_{K,L}$ to denote equivalence classes in $H_k(K, L; R)$, $\tilde{H}_k(K, L; R)$ and their reduced form. This will simplify the notation. The context is generally sufficient to determine in which space the equivalence class belongs to. ♠

Definition 8.6.4

Let K_1 and K_2 be two simplicial complexes, R be an integral domain, and f be a simplicial map between K_1 and K_2 . The function $\tilde{C}_k(f) : \tilde{C}_k(K_1; R) \rightarrow \tilde{C}_k(K_2; R)$ is the linear function defined by $\tilde{C}_k(f)((\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_k)) = (f(\mathbf{y}_0), f(\mathbf{y}_1), \dots, f(\mathbf{y}_k))$ for all ordered k -simplices $(\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_k)$ and extended linearly to all of $\tilde{C}_k(K_1; R)$.

Since $\tilde{C}_k(f)$ commutes with $\tilde{\partial}_k$, we get the following proposition.

Proposition 8.6.5

In the context of the previous definition, the map $\tilde{H}_k(f) : \tilde{H}_k(K_1; R) \rightarrow \tilde{H}_k(K_2; R)$ defined by $\tilde{H}_k(f)([c]_{K_1}) = [\tilde{C}_k(f)(c)]_{K_2}$ for all $[c]_{K_1} \in \tilde{H}_k(K_1; R)$ is an homomorphism.

The following result is easy to proof.

Proposition 8.6.6

Suppose that L_i is a subcomplex of a simplicial complex K_i for $i = 1, 2$, and that R is an integral domain. If $f : (K_1, L_1) \rightarrow (K_2, L_2)$, then the map $\tilde{H}_k(f) : \tilde{H}_k(K_1, L_1; R) \rightarrow \tilde{H}_k(K_2, L_2; R)$ defined by $\tilde{H}_k(f)([\![c]\!]_{K_1, L_1}) = [\![\tilde{C}_k(f)(c)]\!]_{K_2, L_2}$ for all $[\![c]\!]_{K_1, L_1} \in \tilde{H}_k(K_1, L_1; R)$ is an homomorphism.

The theory of reduced and relative simplicial homology presented in Section 8.2.5 can be replicated to obtain the same theoretical results for reduced and relative ordered simplicial homology.

The free abelian group $\tilde{C}_k(K; R)$ is a useful tool to proof the relation between simplicial and singular homology but it is not useful for computations because it is much bigger than $C_k(K; R)$.

We first construct an augmentation-preserving chain map $\mathcal{F} = \{f_k\}_{k \in \mathbb{N}}$ between the chain complex $\tilde{\mathcal{C}} = \{\tilde{C}_k(K; R), \tilde{\partial}_k\}_{k \in \mathbb{N}}$ and the chain complex $\mathcal{S} = \{S_k([K]; R), \partial_k\}_{k \in \mathbb{N}}$. Recall that $[K] = \bigcup_{(s) \in K} (s)$. Given an ordered k -simplex $(s) = (\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_k)$, we set $F((\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_k)) = \delta_{(s)}$, where $\delta_{(s)} : \Delta_k \rightarrow [K]$ is the affine map defined by the conditions $\delta_{(s)}(\mathbf{e}_i) = \mathbf{y}_i$ for $0 \leq i \leq k$. We extend F to $\tilde{C}_k(K; R)$ by linearity; namely,

$$F\left(\sum_{\substack{(s) \text{ an ordered} \\ k\text{-simplex}}} a_{(s)}(s)\right) = \sum_{\substack{(s) \text{ an ordered} \\ k\text{-simplex}}} a_{(s)} F((s)) ,$$

where $a_{(s)} \in R$ and the sum is finite. The chain map $\mathcal{F} = \{f_k\}_{k \in \mathbb{Z}}$ is defined by $f_k = F$ for all k . We also have that $\mathcal{F} = \{f_k\}_{k \in \mathbb{N}}$ is a chain map between the chain complex $\{\tilde{C}_k(L; R), \tilde{\partial}_k\}_{k \in \mathbb{N}}$ and the chain complex $\{S_k([L]; R), \partial_k\}_{k \in \mathbb{N}}$. Since the chain map $\mathcal{F}$ is augmentation preserving, $\mathcal{F}$ is also a chain map between the augmented chain complex $\tilde{\mathcal{C}}^\# = \{\tilde{C}_k^\#(K; R), \tilde{\partial}_k^\#\}_{k \in \mathbb{N}}$ and the augmented chain complex $\mathcal{S}^\# = \{S_k^\#([K]; R), \partial_k^\#\}_{k \in \mathbb{N}}$.

Using F , we may define the following homomorphisms:

- $\tilde{H}_k(F) : \tilde{H}_k(K; R) \rightarrow H_k([K]; R)$ by $\tilde{H}_k(F)([c]_K) = [F(c)]_{[K]}$ for all $[c]_K \in \tilde{H}_k(K; R)$,
- $\tilde{H}_k(F) : \tilde{H}_k(K, L; R) \rightarrow H_k([K], [L]; R)$ by $\tilde{H}_k(F)([c]_{K,L}) = [F(c)]_{[K],[L]}$ for all $[c]_{K,L} \in \tilde{H}_k(K, L; R)$, and
- $\tilde{H}_k^\#(F) : \tilde{H}_k^\#(K; R) \rightarrow H_k^\#([K]; R)$ by $\tilde{H}_k^\#(F)([c]_K) = [F(c)]_{[K]}$ for all $[c]_K \in \tilde{H}_k^\#(K; R)$.

Proposition 8.6.7

Let L be a subcomplex of a simplicial complex K and R be an integral domain. Then

1. $\tilde{H}_k(F) : \tilde{H}_k(K; R) \rightarrow H_k([K]; R)$ is an isomorphism,
2. $\tilde{H}_k^\#(F) : \tilde{H}_k^\#(K; R) \rightarrow H_k^\#([K]; R)$ is an isomorphism, and
3. $\tilde{H}_k(F) : \tilde{H}_k(K, L; R) \rightarrow H_k([K], [L]; R)$ is an isomorphism.

Proof.

It follows from Proposition 8.4.33 that (1) is equivalent to (2). So, we only need to prove (2) and (3).

I) We first assume that K is a finite set. The proof is by induction on the number of elements in K .

I.i) If K has only one element, then $K = \{(\mathbf{x}_i)\}$ for some $\mathbf{x}_i \in \mathbb{R}^n$. We have that $\tilde{C}_k(K; R)$ is generated by the ordered k -simplex $(s) = (\mathbf{x}_i, \mathbf{x}_i, \dots, \mathbf{x}_i)$ and $\delta_{(s)} : \Delta_k \rightarrow [K] = \{\mathbf{x}_i\}$ is

the affine map defined by $\delta_{(s)}(\mathbf{x}) = \mathbf{x}_i$ for all $\mathbf{x} \in \Delta_k$. Therefore $F : \tilde{C}_k(K; R) \rightarrow S_k([K]; R)$ is given by $F(r(s)) = r\delta_s$ for all $r \in R$. Hence F is an isomorphism. It follows that (2) is true. The statement (3) is obviously true because the only non-trivial relative homology modules are $\tilde{H}_k(K, \emptyset; R) = \tilde{H}_k(K; R)$ and $H_k([K], \emptyset; R) = H_k([K]; R)$.

I.ii) We assume that (2) and (3) are true if K has less than N elements and prove that it is also true if K has N elements.

It suffices to prove that $\tilde{H}_k^\sharp(F) : \tilde{H}_k^\sharp(K; R) \rightarrow H_k^\sharp([K]; R)$ is an isomorphism. Suppose that this is true. We get the following commuting diagram from Proposition 8.4.60 ¹¹

$$\begin{array}{ccccccc}
 & \xrightarrow{C_{k+1}} & \tilde{H}_k(L; R) & \xrightarrow{\tilde{H}_k(\iota)} & \tilde{H}_k(K; R) & \xrightarrow{\tilde{H}_k(\text{Id}_X)} & \tilde{H}_k(K, L; R) \\
 & & \downarrow \tilde{H}_k(F) & & \downarrow \tilde{H}_k(F) & & \downarrow \tilde{H}_k(F) \\
 & \xrightarrow{C_{k+1}} & H_k([L]; R) & \xrightarrow{H_k(\iota)} & H_k([K]; R) & \xrightarrow{H_k(\text{Id}_X)} & H_k([K], [L]; R) \\
 \\
 & \xrightarrow{C_k} & \tilde{H}_{k-1}(L; R) & \xrightarrow{\tilde{H}_{k-1}(\iota)} & \tilde{H}_{k-1}(K; R) & \xrightarrow{\tilde{H}_{k-1}(\text{Id}_X)} & \\
 & & \downarrow \tilde{H}_{k-1}(F) & & \downarrow \tilde{H}_{k-1}(F) & & \\
 & \xrightarrow{C_k} & H_{k-1}([L]; R) & \xrightarrow{H_{k-1}(\iota)} & H_{k-1}([K]; R) & \xrightarrow{H_{k-1}(\text{Id}_X)} &
 \end{array}$$

where each row is an exact sequence. Since $\tilde{H}_k^\sharp(F) : \tilde{H}_k^\sharp(L; R) \rightarrow H_k^\sharp([L]; R)$ is an isomorphism because L has less than N elements, and $\tilde{H}_k^\sharp(F) : \tilde{H}_k^\sharp(K; R) \rightarrow H_k^\sharp([K]; R)$ is an isomorphism by assumption, we get from Proposition 8.4.33 that $\tilde{H}_k(F) : \tilde{H}_k(L; R) \rightarrow H_k([L]; R)$ and $\tilde{H}_k(F) : \tilde{H}_k(K; R) \rightarrow H_k([K]; R)$ are isomorphism. Hence, we get from the Five Lemma that $\tilde{H}_k(F) : \tilde{H}_k(K, L; R) \rightarrow H_k([K], [L]; R)$ is an isomorphism.

We now prove that $\tilde{H}_k^\sharp(F) : \tilde{H}_k^\sharp(K; R) \rightarrow H_k^\sharp([K]; R)$ is an isomorphism. Suppose that $(s) \in K$ is a simplex which is not the face of any other simplex but itself. Then $K_1 = K \setminus \{s\}$ is simplicial subcomplex of K and K_1 has $N - 1$ elements. Let K_2 be the simplicial complex which is the collection of (s) and all its faces, and $K_3 = K_2 \setminus \{(s)\}$ (Figure 8.18).

We get the following commutative diagram.

$$\begin{array}{ccc}
 \tilde{H}_k^\sharp(K; R) & \xrightarrow{\tilde{H}_k^\sharp(F)} & H_k^\sharp([K]; R) \\
 \tilde{H}_k(\text{Id}) \downarrow & & \downarrow H_k(\text{Id}) \\
 \tilde{H}_k(K, K_2; R) & \xrightarrow{\tilde{H}_k(F)} & H_k([K], [K_2]; R) \\
 \tilde{H}_k(\iota) \uparrow & & \uparrow H_k(\iota) \\
 \tilde{H}_k(K_1; K_3; R) & \xrightarrow{\tilde{H}_k(F)} & H_k([K_1], [K_3]; R)
 \end{array}$$

where ι denotes an inclusion map. We have by induction that $\tilde{H}_k(F) : \tilde{H}_k(K_1; K_3; R) \rightarrow H_k([K_1], [K_3]; R)$ is an isomorphism. To prove that $\tilde{H}_k^\sharp(F) : \tilde{H}_k^\sharp(K; R) \rightarrow H_k^\sharp([K]; R)$ is an isomorphism, we prove that all the maps $\tilde{H}_k(\iota)$, $H_k(\iota)$, $\tilde{H}_k(\text{Id})$ and $H_k(\text{Id})$ are isomorphism.

¹¹There are versions of Proposition 8.4.60 for the reduced homology and the ordered simplicial homology.

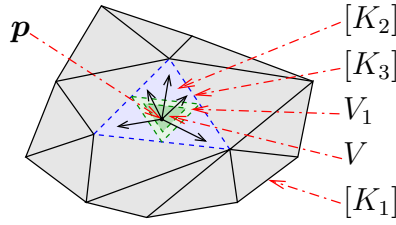


Figure 8.18: Figure associated to the proof of Proposition 8.6.7. The set $[K_1]$ is the region in grey including the continuous black lines. $[K_2]$ is the region in blue including the dashed blue lines forming the boundary. $[K_3]$ is the union of the dashed blue lines. V_1 is the region in pale green excluding the dashed green lines forming the boundary. Finally, V is the region in darker green excluding the dashed green lines forming the boundary. The arrows in black describe the retraction r of $[K_2] \setminus V$ onto $[K_3]$.

Since $[K_2] = [s]$ is contractible, we get from Proposition 8.4.64 that $H_k(\text{Id}) : H_k^\sharp([K]; R) \rightarrow H_k^\sharp([K], [K_2]; R)$ is an isomorphism for $k \geq 0$. Since $H_k^\sharp([K], [K_2]; R) = H_k([K], [K_2]; R)$ for $k > 0$, we get that $H_k(\text{Id}) : H_k^\sharp([K]; R) \rightarrow H_k([K], [K_2]; R)$ is an isomorphism for $k > 0$. For $k = 0$, we assume that $[K]$ is the disjoint union of path-connected components X_j for $j \in J$. There exists $j_0 \in J$ such that $(s) \subset X_{j_0}$. We have from Propositions 8.4.55 and 8.4.56 that

$$\begin{aligned} H_0([K], [K_2]; R) &= \left(\bigoplus_{j \in J \setminus \{j_0\}} H_0(X_j; R) \right) \oplus H_0(X_{j_0}, [K_2]; R) \\ &= \left(\bigoplus_{j \in J \setminus \{j_0\}} H_0(X_j; R) \right) \oplus 0. \end{aligned}$$

We get from Propositions 8.4.10 and 8.4.45 that

$$\left(\bigoplus_{j \in J \setminus \{j_0\}} H_0(X_j; R) \right) \oplus H_0(X_{j_0}; R) = H_0([K]; R) \cong H_0^\sharp([K]; R) \oplus R.$$

Since $H_0(X_j; R) \cong R$ for all j according to Corollary 8.4.41, we get that $H_0([K], [K_2]; R) = \bigoplus_{j \in J \setminus \{j_0\}} H_0(X_j; R) \cong H_0^\sharp([K]; R)$. Since the previous isomorphism is deduced from the identity

map, we again have that $H_0(\text{Id}) : H_0^\sharp([K]; R) \rightarrow H_0([K], [K_2]; R)$ is an isomorphism. The same reasoning shows that $\tilde{H}_k(\text{Id}) : \tilde{H}_k^\sharp(K; R) \rightarrow \tilde{H}_k(K, K_2; R)$ is an isomorphism for $k \geq 0$ ¹².

Consider the inclusion map $\iota : \tilde{C}_k(K_1; R)/\tilde{C}_k(K_3; R) \rightarrow \tilde{C}_k(K; R)/\tilde{C}_k(K_2; R)$. The map ι is an isomorphism because $K_1 \setminus K_3 = K \setminus K_2$. More precisely, if $[[c]]_{K, K_2} \in \tilde{C}_k(K; R)/\tilde{C}_k(K_2; R)$, then $c - b \in \tilde{C}_k(K_2; R)$ for some $b \in \tilde{C}_k(K_1; R)$ because $(s) \in K_2$. Hence $\iota([b]_{K_1, K_3}) = [[c]]_{K, K_2}$ and ι is onto. If $[[c_1]]_{K, K_2} = \iota([c_1]_{K_1, K_3}) = \iota([c_2]_{K_1, K_3}) =$

¹²The courageous reader is invited to repeat for relative and reduced simplicial homology what we have done for relative and reduced singular homology.

$[[c_2]]_{K,K_2}$ with $c_1, c_2 \in \tilde{C}_k(K_1; R)$, then $c_1 - c_2 = b \in \tilde{C}_k(K_2; R) \cap \tilde{C}_k(K_1; R)$. Thus $c_1 - c_2 = b \in \tilde{C}_k(K_3; R)$ and therefore $[[c_1]]_{K_1, K_3} = [[c_2]]_{K_1, K_3}$. This implies that $\tilde{H}_k(\iota) : \tilde{H}_k(K_1, K_3; R) \rightarrow \tilde{H}_k(K, K_2; R)$ is also an isomorphism.

To prove that $H_k(\iota) : H_k([K_1], [K_3]; R) \rightarrow H_k([K], [K_2]; R)$ is an isomorphism is not as simple as the proof that $\tilde{H}_k(\iota)$ is an isomorphism given above. We plan to use excision to prove that $H_k(\iota) : H_k([K_1], [K_3]; R) \rightarrow H_k([K], [K_2]; R)$ is an isomorphism. Unfortunately, the excision theorem, Theorem 8.4.66, cannot be used directly because we have the situation where the set that we want to excise, namely (s) , is such that $\overline{(s)} = [s]$ is not a subset of $[K_2]^\circ = (s)$, where closure and interior of sets are relative to the induced topology on $[K]$ from $\mathbb{R}^n$. We therefore need to use Proposition 8.4.80.

Suppose that (s) is a k -simplex and that $(s_0), (s_1), \dots, (s_k)$ are the faces of (s) which are $(k-1)$ -simplices. Let V_1 be the k -simplex $(\mathbf{b}_{(s_0)}, \mathbf{b}_{(s_1)}, \dots, \mathbf{b}_{(s_k)})$ where $\mathbf{b}_{(s_j)}$ is the barycentre of (s_j) for $0 \leq j \leq k$. Let $\mathbf{p}$ be the barycentre of V_1 . Referring to the statement of Proposition 8.4.80, the set $V \subset (s)$ that we consider is given by

$$V = \mathbf{p} + \epsilon(V_1 - \mathbf{p}) = \{\mathbf{p} + \epsilon(\mathbf{x} - \mathbf{p}) : \mathbf{x} \in V_1\}$$

for $0 < \epsilon < 1$ fixed (Figure 8.18).

We have that $\overline{V} \subset (s)^\circ = (s)$. There is a retraction r of $[K_2] \setminus V$ onto $[K_3]$ (Figure 8.18). By defining r to be the identity on $[K_1] \setminus [K_2]$, we get a deformation retraction of $([K] \setminus V, [K_2] \setminus V)$ onto $([K_1], [K_3]) = ([K] \setminus (s), [K_2] \setminus (s))$. We may therefore use Proposition 8.4.80 to conclude that $H_k(\iota) : H_k([K_1], [K_3]; R) \rightarrow H_k([K] \setminus V, [K_2] \setminus V; R)$ is an isomorphism.

Moreover, since $\overline{V} \subset (s)^\circ = (s)$, we may use the excision theorem, Theorem 8.4.66, to conclude that $H_k(\iota) : H_k([K] \setminus V, [K_2] \setminus V; R) \rightarrow H_k([K], [K_2]; R)$ is an isomorphism. It follows from the previous paragraph that $H_k(\iota) : H_k([K_1], [K_3]; R) \rightarrow H_k([K], [K_2]; R)$ is an isomorphism.

II) The fact that we have proved the proposition when K has any finite number of elements does not automatically imply that the proposition is true when K has infinitely many elements. We do not have a concept of limit for the spaces $C_k(K; R)$, $S_k(K; R)$, and so on, as the size of K increases. We now assume that K has countably infinitely many elements.

It still follows from Proposition 8.4.33 that (1) is equivalent to (2). Moreover, if we prove that $\tilde{H}_k^\sharp(F) : \tilde{H}_k^\sharp(K; R) \rightarrow H_k^\sharp([K]; R)$ is an isomorphism for all k , then we can use long exact sequences and the Five Lemma exactly as we have done in (I.ii) to show that this implies that (3) is true for all k .

Thus, we only have to prove that $\tilde{H}_k^\sharp(F) : \tilde{H}_k^\sharp(K; R) \rightarrow H_k^\sharp([K]; R)$ where K has infinitely many elements.

II.i) Suppose that $c \in S_k([K]; R)$. We may express c as a finite sum $c = \sum_{j \in J} a_j \sigma_j$, where

$a_j \in R$ and σ_j is a singular k -simplex in $[K]$ for all j . We prove that there exists a simplicial subcomplex L_c of K such that L_c has a finite number of elements and $Q = \bigcup_{j \in J} \sigma_j(\Delta_k) \subset [L_c]$.

Since the σ_j are continuous functions and Δ_k is compact, we have that $\sigma_j(\Delta_k)$ is a compact

set. Thus Q is a compact set because it is the finite union of compact sets. Suppose that there is no finite number of simplices in K that cover Q . Then there exists an infinite collection $\{(s_i)\}_{i \in \mathbb{N}}$ of distinct simplices in K such that $(s_i) \cap Q \neq \emptyset$ for all $i \in \mathbb{N}$ ¹³. Choose $z_i \in (s_i) \cap Q$ for each $i \in \mathbb{N}$. Let $U_m = Q \setminus \bigcup_{i \neq m} \{z_i\}$ for all $m \in \mathbb{N}$. The sets U_m are open sets

(see Remark 8.6.8 below) and $Q = \bigcup_{m \in \mathbb{N}} U_m$. Thus $\{U_m\}_{m \in \mathbb{N}}$ is an open cover of the compact

set Q . Therefore, there exists a finite subcover $\{U_{m_t}\}_{0 \leq t \leq T}$ of Q . But $Q \not\supseteq \bigcup_{0 \leq t \leq T} U_{m_t}$ because

$z_i \notin \bigcup_{0 \leq t \leq T} U_{m_t}$ if $i \neq m_t$ for $0 \leq t \leq T$. This is a contradiction. Hence, there are a finite number of simplices in K whose union cover Q . If we add all the faces of these simplices, then we get a finite simplicial subcomplex L_c of K with $Q \subset [L_c]$.

II.ii) We first prove that $\tilde{H}_k^\sharp(F) : \tilde{H}_k^\sharp(K; R) \rightarrow H_k^\sharp([K]; R)$ is onto. Suppose that $[c]_K \in H_k^\sharp([K]; R)$ and that $c = \sum_{j \in J} a_j \sigma_j$, where $a_j \in R$ and σ_j is a singular k -simplex in $[K]$ for all

$j \in J$. Using (II.i), we get a finite simplicial subcomplex L_c of K such that $c \in S_k^\sharp([L_c]; R)$ and thus $[c]_{L_c} \in H_k^\sharp([L_c]; R)$. we consider the commutative diagram

$$\begin{array}{ccc} \tilde{H}_k^\sharp(L_c; R) & \xrightarrow{\tilde{H}_k^\sharp(F)} & H_k^\sharp([L_c]; R) \\ \tilde{H}_k(\iota) \downarrow & & \downarrow H_k(\iota) \\ \tilde{H}_k^\sharp(K; R) & \xrightarrow{\tilde{H}_k^\sharp(F)} & H_k^\sharp([K]; R) \end{array}$$

We have from (I) that $\tilde{H}_k^\sharp(F) : \tilde{H}_k^\sharp(L_c; R) \rightarrow H_k^\sharp([L_c]; R)$ is an isomorphism. Thus, there exists $[b]_{L_c} \in \tilde{H}_k^\sharp(L_c; R)$ such that $\tilde{H}_k^\sharp(F)([b]_{L_c}) = [c]_{[L_c]}$. we therefore get that $\tilde{H}_k^\sharp(F)(\tilde{H}_k(\iota)([b]_{L_c})) = H_k(\iota)(\tilde{H}_k^\sharp(F)([b]_{L_c})) = [c]_{[K]}$. Hence $[c]_{[K]}$ is in the image of $\tilde{H}_k^\sharp(F) : \tilde{H}_k^\sharp(K; R) \rightarrow H_k^\sharp([K]; R)$.

To prove that $\tilde{H}_k^\sharp(F) : \tilde{H}_k^\sharp(K; R) \rightarrow H_k^\sharp([K]; R)$ is one-to-one, suppose that $\tilde{H}_k^\sharp(F)([c]_K) = [F(c)]_{[K]} = [0]_{[K]}$ for some $[c]_K \in \tilde{H}_k^\sharp(K; R)$. This means that $F(c) = \partial^\sharp(b)$ for some $b \in S_{k+1}^\sharp([K]; R)$. Using (II.i), we get a finite simplicial subcomplex L_c of K such that $F(c) \in S_k^\sharp([L_c]; R)$ and a finite simplicial subcomplex L_b of K such that $b \in S_{k+1}^\sharp([L_b]; R)$. Let $L = L_c \cup L_b$. We have that L is a finite simplicial subcomplex of K with $F(c) \in S_k^\sharp([L]; R)$ and $b \in S_{k+1}^\sharp([L]; R)$. We also have that $c \in \tilde{C}_k^\sharp(L; R)$ because c is a finite linear combinations of ordered k -simplices which are included in L by construction¹⁴. We consider the commutative

¹³These simplices are obtained inductively using the fact that $Q \subset [K] = \bigcup_{(s) \in K} (s)$.

¹⁴It follows from the definition of F that if (s) is one of the ordered k -simplices involved in the linear combination for c , then (s) must comes from one of the initial simplices that cover Q in (II.i). Recall that if two open simplices intersect, then they must be equal.

diagram

$$\begin{array}{ccc} \tilde{H}_k^\sharp(L; R) & \xrightarrow{\tilde{H}_k^\sharp(F)} & H_k^\sharp([L]; R) \\ \tilde{H}_k(\iota) \downarrow & & \downarrow H_k(\iota) \\ \tilde{H}_k^\sharp(K; R) & \xrightarrow{\tilde{H}_k^\sharp(F)} & H_k^\sharp([K]; R) \end{array}$$

Since $F(c) = \partial^\sharp(b) \in S_k^\sharp([L]; R)$ with $b \in S_{k+1}^\sharp([L]; R)$, we have that $[F(c)]_{[L]} = [0]_{[L]} \in \tilde{H}_k^\sharp([L]; R)$. Since $\tilde{H}_k^\sharp(F) : \tilde{H}_k^\sharp(L; R) \rightarrow H_k^\sharp([L]; R)$ is an isomorphism according to (I), we have that $[c]_L = [0]_L \in \tilde{H}_k^\sharp(L; R)$. Thus $[c]_K = \tilde{H}(\iota)([c]_L) = [0]_K \in \tilde{H}_k^\sharp(K; R)$. ■

Remark 8.6.8

We prove that the sets U_m in part (II.i) of the previous proof are open sets. By adding the missing faces if needed to $\{(s_i)\}_{i \in \mathbb{N}}$, we get a simplicial subcomplex L of K . Consider $\mathbf{z} \in U_m$. If $\mathbf{z} \in U_m \setminus [L]$, then $V = Q \setminus [L]$ can be used as an open neighbourhood of $\mathbf{z}$ such that $V \subset U_m$. If $\mathbf{z} \in [L] = \bigcup_{(s) \in L} (s)$, then there exists a unique $(s) \in L$ such that $\mathbf{z} \in (s)$. The finite union $W = \bigcup_{\substack{(t) \in K \\ [t] \cap [s] \neq \emptyset}} (t)$ is an open neighbourhood of $\mathbf{z}$ in the induce

topology on $[K]$. Moreover, W contains only a finite number of $\mathbf{z}_i$ with $i \neq m$ because the open simplices are distinct and we have only one $\mathbf{z}_i$ per (s_i) ; We are referring to the $\mathbf{z}_i$ with $i \neq m$ that may be in one of the (t) listed in the union. We do not want to reject $\mathbf{z}_m$ if it is in W ; in particular if $\mathbf{z} = \mathbf{z}_m$. If $\mathbf{z}_{m_s}$ for $m_s \neq m$ and $0 \leq s \leq S$ are the $\mathbf{z}_i \in W$, then $Q \cap (W \setminus \{\mathbf{z}_{m_s} : 0 \leq s \leq S\}) \subset U_m$ is an open neighbourhood of $\mathbf{z}$. ♠

Definition of the Isomorphism

We really want a version of the previous proposition for oriented simplicial homology. Instead of ordered simplices, we want to work with oriented simplices. Instead of the map F defined previously, we would like to use the following map. Given an oriented k -simplex $[[s]] = [[\mathbf{x}_{i_0}, \mathbf{x}_{i_1}, \dots, \mathbf{x}_{i_k}]]$, we set $G([s]) = \delta_{[s]}$ where as before $\delta_{[s]} : \Delta_k \rightarrow [K]$ is the affine map defined by $\delta_{[s]}(\mathbf{e}_j) = \mathbf{x}_{i_j}$ for $0 \leq j \leq k$. We extend G to $C_k(K; R)$ by linearity; namely, $G\left(\sum_{\substack{(s) \in K \\ \dim(s)=k}} a_{(s)} [[s]]\right) = \sum_{\substack{(s) \in K \\ \dim(s)=k}} a_{(s)} G([s])$ where $a_{(s)} \in R$ and the sum is finite. The rest of this subsection is devoted to justify how we may “replace” F by G .

Let $\phi : C_k(K; R) \rightarrow \tilde{C}_k(K; R)$ be the function defined by $\phi([[\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_k}]]]) = (\mathbf{x}_{j_0}, \mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_k})$ for all oriented k -simplices with $j_0 < j_1 < \dots < j_k$ and extended to $C_k(K; R)$ by linearly, and $\psi : \tilde{C}_k(K; R) \rightarrow C_k(K; R)$ be the function defined by

$$\psi((\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_k)) = \begin{cases} [[\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_k]] & \text{if } (\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_k) \in K \\ 0 & \text{otherwise} \end{cases}$$

for all ordered k -simplices and extended to $\tilde{C}_k(K; R)$ by linearity. We note that $\psi((s)) = 0$ when (s) has at least two equal vertices. It is clear that $\tilde{\partial}_k \circ \phi = \phi \circ \partial_k$. It is also true that $\psi \circ \tilde{\partial}_k = \partial_k \circ \psi$ but the proof is not completely obvious. If $(s) \in \tilde{C}_k(K; R)$ does not have any equal vertices, then the result is obvious. If $(s) = (\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_k)$ has two equal vertices, say $\mathbf{y}_m = \mathbf{y}_{m+1}$ for some $0 \leq m < k$, then $\partial_k(\psi((s))) = \partial_K(0) = 0$ and

$$\begin{aligned} \psi(\tilde{\partial}_k(s)) &= \sum_{i=0}^k (-1)^i \psi((\mathbf{y}_0, \mathbf{y}_1, \dots, \widehat{\mathbf{y}}_i, \dots, \mathbf{y}_k)) \\ &= (-1)^m \psi((\mathbf{y}_0, \mathbf{y}_1, \dots, \widehat{\mathbf{y}}_m, \dots, \mathbf{y}_k)) + (-1)^{m+1} \psi((\mathbf{y}_0, \mathbf{y}_1, \dots, \widehat{\mathbf{y}}_{m+1}, \dots, \mathbf{y}_k)) = 0 \end{aligned}$$

because $(\mathbf{y}_0, \mathbf{y}_1, \dots, \widehat{\mathbf{y}}_m, \dots, \mathbf{y}_k) = (\mathbf{y}_0, \mathbf{y}_1, \dots, \widehat{\mathbf{y}}_{m+1}, \dots, \mathbf{y}_k)$. If $(s) = (\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_k)$ has more than two equal vertices, then $\partial_k(\psi((s))) = \partial_K(0) = 0$ and

$$\psi(\tilde{\partial}_k(s)) = \sum_{i=0}^k (-1)^i \psi((\mathbf{y}_0, \mathbf{y}_1, \dots, \widehat{\mathbf{y}}_i, \dots, \mathbf{y}_k)) = 0$$

because each $(\mathbf{y}_0, \mathbf{y}_1, \dots, \widehat{\mathbf{y}}_i, \dots, \mathbf{y}_k)$ has at last two equal vertices.

We have the following commutative diagram

$$\begin{array}{ccccccc} \xrightarrow{\partial_{k+2}} C_{k+1}(K; R) & \xrightarrow{\partial_{k+1}} & C_k(K; R) & \xrightarrow{\partial_k} & C_{k-1}(K; R) & \xrightarrow{\partial_{k-1}} & \\ & \psi \uparrow & \downarrow \phi & & \psi \uparrow & \downarrow \phi & \\ & \psi \uparrow & \downarrow \phi & & \psi \uparrow & \downarrow \phi & \\ \xrightarrow{\tilde{\partial}_{k+2}} \tilde{C}_{k+1}(K; R) & \xrightarrow{\tilde{\partial}_{k+1}} & \tilde{C}_k(K; R) & \xrightarrow{\tilde{\partial}_k} & \tilde{C}_{k-1}(K; R) & \xrightarrow{\tilde{\partial}_{k-1}} & \end{array}$$

Namely, $\{\phi_k\}_{k \in \mathbb{Z}}$ with $\phi_k = \phi$ for all k is a chain map from the chain complex $\mathcal{C} = \{(C_k(K; R), \partial_k)\}_{k \in \mathbb{Z}}$ to the chain complex $\tilde{\mathcal{C}} = \{(\tilde{C}_k(K; R), \tilde{\partial}_k)\}_{k \in \mathbb{Z}}$ and $\{\psi_k\}_{k \in \mathbb{Z}}$ with $\psi_k = \psi$ for all k is a chain map from $\tilde{\mathcal{C}}$ to $\mathcal{C}$.

Since $\phi : \mathcal{N}(\partial_k) \rightarrow \mathcal{N}(\tilde{\partial}_k)$ and $\phi : \mathcal{R}(\partial_{k-1}) \rightarrow \mathcal{R}(\tilde{\partial}_{k-1})$, we can define the homomorphism $\tilde{H}_k(\phi) : H_k(K; R) \rightarrow \tilde{H}_k(K; R)$ by $\tilde{H}_k(\phi)([c]_K) = [\phi(c)]_K$ for all $c \in C_k(K; R)$, where as usual $[\cdot]_K$ represents equivalence classes in both $H_k(K; R)$ and $\tilde{H}_k(K; R)$. Similarly, since $\psi : \mathcal{N}(\tilde{\partial}_k) \rightarrow \mathcal{N}(\partial_k)$ and $\psi : \mathcal{R}(\tilde{\partial}_{k-1}) \rightarrow \mathcal{R}(\partial_{k-1})$, we may define the homomorphism $\tilde{H}_k(\psi) : \tilde{H}_k(K; R) \rightarrow H_k(K; R)$ by $\tilde{H}_k(\psi)([c]_K) = [\psi(c)]_K$ for all $c \in \tilde{C}_k(K; R)$.

We can repeat all the previous discussion for the reduced homology (i.e. with ∂_k replaced by $\partial_k^\#$ and $\tilde{\partial}_k$ replaced by $\tilde{\partial}_k^\#$) and define $\tilde{H}_k^\#(\phi) : H_k^\#(K; R) \rightarrow \tilde{H}_k^\#(K; R)$ and $\tilde{H}_k^\#(\psi) : \tilde{H}_k^\#(K; R) \rightarrow H_k^\#(K; R)$. We assume that $\tilde{C}_{-1}^\#(K; R) = R$ and $C_{-1}^\#(K; R) = R$, and that $\phi : C_{-1}^\#(K; R) \rightarrow \tilde{C}_{-1}^\#(K; R)$ and $\psi : \tilde{C}_{-1}^\#(K; R) \rightarrow C_{-1}^\#(K; R)$ are the identity maps.

According to Proposition 8.4.33, if we prove that $\tilde{H}_k^\#(\phi)$ is an isomorphism, then we will have that $\tilde{H}_k(\phi)$ is an isomorphism. To prove that $\tilde{H}_k^\#(\phi)$ is an isomorphism, we prove that $\tilde{H}_k^\#(\psi)$ is the inverse of $\tilde{H}_k^\#(\phi)$.

It is clear that $\psi \circ \phi = \text{Id}$ on $C_k^\#(K; R)$. Thus $\tilde{H}_k^\#(\psi) \circ \tilde{H}_k^\#(\phi) = \text{Id}$ on $H_k^\#(K; R)$. To prove that $\tilde{H}_k^\#(\phi) \circ \tilde{H}_k^\#(\psi) = \text{Id}$ requires a little bit more work because $\phi \circ \psi$ is not the

identity. We need to prove that there exists a chain homotopy between $\mathcal{F} = \{f_k\}_{k \in \mathbb{Z}}$ with $f_k = \text{Id} : \tilde{C}_k^\sharp(K; R) \rightarrow \tilde{C}_k^\sharp(K; R)$ and $\mathcal{G} = \{g_k\}_{k \in \mathbb{Z}}$ with $g_k = \phi \circ \psi : \tilde{C}_k^\sharp(K; R) \rightarrow \tilde{C}_k^\sharp(K; R)$. Namely, we need to find $\mathcal{D} = \{D_k\}_{k \in \mathbb{Z}}$ with $D_k : \tilde{C}_k^\sharp(K; R) \rightarrow \tilde{C}_{k+1}^\sharp(K; R)$ such that $\tilde{\partial}_{k+1}^\sharp \circ D_k + D_{k-1} \circ \tilde{\partial}_k^\sharp = \phi \circ \psi - \text{Id}$ on $\tilde{C}_k^\sharp(K; R)$. After this is done, we will be able to use Proposition 8.4.29¹⁵ to conclude that $\tilde{H}_k^\sharp(\text{Id}) = \tilde{H}_k^\sharp(\phi \circ \psi)$ for all k ; namely, $\tilde{H}_k^\sharp(\phi) \circ \tilde{H}_k^\sharp(\psi) = \text{Id} : \tilde{H}_k^\sharp(K; R) \rightarrow \tilde{H}_k^\sharp(K; R)$.

Proposition 8.6.9

There exists $\mathcal{D} = \{D_k\}_{k \in \mathbb{Z}}$ with $D_k : \tilde{C}_k^\sharp(K; R) \rightarrow \tilde{C}_{k+1}^\sharp(K; R)$ such that $\tilde{\partial}_{k+1}^\sharp \circ D_k + D_{k-1} \circ \tilde{\partial}_k^\sharp = \phi \circ \psi - \text{Id}$ on $\tilde{C}_k^\sharp(K; R)$.

Proof.

i) Suppose that $(\mathbf{p}, [K])$ is an general position. We consider the cone $\mathbf{p} * K$. Recall that it is a simplicial complex consisting of all the simplices of the form $(\mathbf{p}, \mathbf{x}_{i_0}, \mathbf{x}_{i_1}, \dots, \mathbf{x}_{i_k})$ and all their faces where $(\mathbf{x}_{i_0}, \mathbf{x}_{i_1}, \dots, \mathbf{x}_{i_k}) \in K$. Hence $K \subset \mathbf{p} * K$.

We prove that the chain complex $\{(\tilde{C}_k^\sharp(\mathbf{p} * K; R), \tilde{\partial}_k^\sharp)\}_{k \in \mathbb{Z}}$ is acyclic. Let $Q_k : \tilde{C}_k^\sharp(\mathbf{p} * K; R) \rightarrow \tilde{C}_{k+1}^\sharp(\mathbf{p} * K; R)$ be the mapping define by $Q_k((\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_k)) = (\mathbf{p}, \mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_k)$ for all $(\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_k) \in \mathbf{p} * K$, and extended by linearity to $\tilde{C}_k^\sharp(\mathbf{p} * K; R)$.

Proceeding as we did in Section 8.4.7, we find that $\tilde{\partial}_1^\sharp(Q_0(c)) = c - \tilde{\partial}_0^\sharp(c)(\mathbf{p})$ for all $c \in \tilde{C}_0^\sharp(\mathbf{p} * K; R)$ and $\tilde{\partial}_{k+1}^\sharp(Q_k(c)) = c - Q_{k-1}(\tilde{\partial}_k^\sharp(c))$ for all $c \in \tilde{C}_k^\sharp(\mathbf{p} * K; R)$ if $k > 0$.

If $c \in \tilde{Z}_0^\sharp(\mathbf{p} * K; R)$, then $\tilde{\partial}_0^\sharp(c) = 0$. Therefore $c = \tilde{\partial}_1^\sharp(Q_0(c))$. Hence $c \in \tilde{B}_0^\sharp(\mathbf{p} * K; R)$. Thus $[c]_{\mathbf{p} * K} = [0]_{\mathbf{p} * K}$. We get that $\tilde{H}_0^\sharp(\mathbf{p} * K; R) = 0$.

If $c \in \tilde{Z}_k^\sharp(\mathbf{p} * K; R)$ with $k > 0$, then $\tilde{\partial}_k^\sharp(c) = 0$. Therefore $c = \tilde{\partial}_{k+1}^\sharp(Q_k(c))$. Hence $c \in \tilde{B}_k^\sharp(\mathbf{p} * K; R)$. Thus $[c]_{\mathbf{p} * K} = [0]_{\mathbf{p} * K}$. We get that $\tilde{H}_k^\sharp(\mathbf{p} * K; R) = 0$.

ii) Let $\eta = \phi \circ \psi$. Given $\sigma = (\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_k) \in \mathbf{p} * K$, Let L_σ be the smallest subcomplex of $\mathbf{p} * K$ containing the vertices $\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_k, \eta(\mathbf{y}_0), \eta(\mathbf{y}_1), \dots, \eta(\mathbf{y}_{k-1})$ and $\eta(\mathbf{y}_k)$. We have that $\tilde{\partial}_k^\sharp(\sigma) \in \tilde{C}_k^\sharp(L_\sigma; R)$. Moreover, if $\sigma \in \tilde{C}_k^\sharp(K; R)$, then $L_\sigma \subset K$ because $\eta : \tilde{C}_k^\sharp(K; R) \rightarrow \tilde{C}_k^\sharp(K; R)$.

We use a proof by induction to construct $D_k : \tilde{C}_k^\sharp(\mathbf{p} * K; R) \rightarrow \tilde{C}_{k+1}^\sharp(\mathbf{p} * K; R)$ such that $\tilde{\partial}_{k+1}^\sharp \circ D_k + D_{k-1} \circ \tilde{\partial}_k^\sharp = \eta - \text{Id}$ on $\tilde{C}_k^\sharp(\mathbf{p} * K; R)$ and such that $D_k(\tilde{C}_k^\sharp(Q; R)) \subset \tilde{C}_{k+1}^\sharp(Q; R)$ for all simplicial subcomplexes of $\mathbf{p} * K$; in particular, for $Q = K$.

k = -1) We define $D_{-1} : \tilde{C}_{-1}^\sharp(\mathbf{p} * K; R) = R \rightarrow \tilde{C}_0^\sharp(\mathbf{p} * K; R)$ by $D_{-1}(r) = 0$ for $r \in R$.

k = 0) Consider $\sigma = (\mathbf{y}) \in \tilde{C}_1^\sharp(\mathbf{p} * K; R)$. Thus $\mathbf{y}$ is a vertex of $\mathbf{p} * K$. Since

$$\tilde{\partial}_0^\sharp(\eta(\sigma) - \text{Id}(\sigma)) = \eta(\tilde{\partial}_0^\sharp(\sigma)) - \text{Id}(\tilde{\partial}_0^\sharp(\sigma)) = \eta(1) - \text{Id}(1) = 1 - 1 = 0,$$

we get that $\eta(\sigma) - \text{Id}(\sigma) \in \tilde{Z}_0^\sharp(\mathbf{p} * K; R)$. Since $\tilde{H}_0^\sharp(\mathbf{p} * K; R) = 0$ by (i), we have that

¹⁵To be precise, a version of this proposition for ordered simplicial homology.

$\eta(\sigma) - \text{Id}(\sigma) \in \tilde{B}_0^\sharp(\mathbf{p} * K; R)$. Thus, there exists $b_\sigma \in \tilde{C}_1^\sharp(\mathbf{p} * K; R)$ such that $\tilde{\partial}_1^\sharp(b_\sigma) = \eta(\sigma) - \text{Id}(\sigma)$. We may assume that $b_\sigma \in \tilde{C}_1^\sharp(L_\sigma; R)$ because $\eta(\sigma) - \text{Id}(\sigma) \in \tilde{C}_0^\sharp(L_\sigma; R)$.

We set $D_0(\sigma) = b_\sigma$ for all ordered 0-simplices σ of $\mathbf{p} * K$, and extend D_0 by linearity to $\tilde{C}_0^\sharp(\mathbf{p} * K; R)$.

By linearity, we have

$$\tilde{\partial}_1^\sharp(D_0(c)) + D_{-1}(\tilde{\partial}_0^\sharp(c)) = \tilde{\partial}_1^\sharp(D_0(c)) = \eta(c) - \text{Id}(c)$$

for all $c \in \tilde{C}_0^\sharp(\mathbf{p} * K; R)$. If $c \in \tilde{C}_0^\sharp(Q; R)$ for some simplicial subcomplex Q of $\mathbf{p} * K$, then c is the sum (after simplification) of oriented 0-simplices σ in Q . Since $D_0(\sigma) \in \tilde{C}_1^\sharp(L_\sigma; R) \subset \tilde{C}_1^\sharp(Q, R)$ for all of these oriented 0-simplices because $L_\sigma \subset Q$ for each of them, we have that $D_0(c) \in \tilde{C}_1^\sharp(Q, R)$. In particular, $D_0(\tilde{C}_0^\sharp(K; R)) \subset \tilde{C}_1^\sharp(K; R)$.

k > 0 We assume for $q < k$ that $\tilde{\partial}_{q+1}^\sharp \circ D_q + D_{q-1} \circ \tilde{\partial}_q^\sharp = \eta - \text{Id}$ on $\tilde{C}_q^\sharp(\mathbf{p} * K; R)$ and $D_q(\tilde{C}_q^\sharp(Q; R)) \subset \tilde{C}_{q+1}^\sharp(Q; R)$ for all simplicial subcomplexes Q of $\mathbf{p} * K$.

Suppose that σ is an ordered k -simplex of $\mathbf{p} * K$. Let $c_\sigma = \eta(\sigma) - \text{Id}(\sigma) - D_{k-1}(\tilde{\partial}_k^\sharp(\sigma))$. We have that $c_\sigma \in \tilde{C}_k^\sharp(L_\sigma; R)$ because $D_{k-1}(\tilde{C}_{k-1}^\sharp(L_\sigma; R)) \subset \tilde{C}_k^\sharp(L_\sigma; R)$ by hypothesis of induction.

From our hypothesis of induction, we also have

$$\begin{aligned} \tilde{\partial}_k^\sharp(c_\sigma) &= \tilde{\partial}_k^\sharp(\eta(\sigma)) - \tilde{\partial}_k^\sharp(\text{Id}(\sigma)) - \tilde{\partial}_k^\sharp(D_{k-1}(\tilde{\partial}_k^\sharp(\sigma))) \\ &= \tilde{\partial}_k^\sharp(\eta(\sigma)) - \tilde{\partial}_k^\sharp(\text{Id}(\sigma)) - (-D_{k-2}(\tilde{\partial}_{k-1}^\sharp(\tilde{\partial}_k^\sharp(\sigma))) + \eta(\tilde{\partial}_k^\sharp(\sigma)) - \text{Id}(\tilde{\partial}_k^\sharp(\sigma))) \\ &= \eta(\tilde{\partial}_k^\sharp(\sigma)) - (\text{Id})(\tilde{\partial}_k^\sharp(\sigma)) + D_{k-2}(\tilde{\partial}_{k-1}^\sharp(\tilde{\partial}_k^\sharp(\sigma))) - \eta(\tilde{\partial}_k^\sharp(\sigma)) + \text{Id}(\tilde{\partial}_k^\sharp(\sigma)) \\ &= 0 \end{aligned}$$

because $\tilde{\partial}_{k-1}^\sharp(\tilde{\partial}_k^\sharp(\sigma)) = 0$. Thus $c_\sigma \in \tilde{Z}_k^\sharp(\mathbf{p} * K; R)$. Since $\tilde{H}_{k+1}^\sharp(\mathbf{p} * K; R) = 0$, we have that $c_\sigma \in \tilde{B}_k^\sharp(\mathbf{p} * K; R)$. Therefore, there exists $b_\sigma \in \tilde{C}_{k+1}^\sharp(\mathbf{p} * K; R)$ such that $\tilde{\partial}_{k+1}^\sharp(b_\sigma) = c_\sigma$. We may assume that $b_\sigma \in \tilde{C}_{k+1}^\sharp(L_\sigma; R)$ because $c_\sigma \in \tilde{C}_k^\sharp(L_\sigma; R)$.

As expected, we set $D_k(\sigma) = b_\sigma$ for all ordered k -simplices σ of $\mathbf{p} * K$, and extend D_k by linearity to $\tilde{C}_k^\sharp(\mathbf{p} * K; R)$.

By linearity, we have

$$\begin{aligned} \tilde{\partial}_{k+1}^\sharp(D_k(c)) + D_{k-1}(\tilde{\partial}_k^\sharp(c)) &= (\eta(c) - \text{Id}(c) - D_{k-1}(\tilde{\partial}_k^\sharp(c))) + D_{k-1}(\tilde{\partial}_k^\sharp(c)) \\ &= \eta(c) - \text{Id}(c) \end{aligned}$$

for all $c \in \tilde{C}_k^\sharp(\mathbf{p} * K; R)$. If $c \in \tilde{C}_k^\sharp(Q; R)$ for some simplicial subcomplex Q of $\mathbf{p} * K$, then c is the sum (after simplification) of oriented k -simplices σ in Q . Since, $D_k(\sigma) \in \tilde{C}_{k+1}^\sharp(L_\sigma; R) \subset \tilde{C}_{k+1}^\sharp(Q, R)$ for all of these oriented k -simplices because $L_\sigma \subset Q$ for each of them, we have that $D_k(c) \in \tilde{C}_{k+1}^\sharp(Q, R)$. In particular, $D_k(\tilde{C}_k^\sharp(K; R)) \subset \tilde{C}_{k+1}^\sharp(K; R)$.

The operator D_k that we are looking for in the statement of the proposition is D_K defined above restricted to $\tilde{C}_k^\sharp(K; R)$. ■

We have that $G = F \circ \phi$. Therefore, $H_k^\sharp(G) = \tilde{H}_k^\sharp(F) \circ \tilde{H}_k^\sharp(\phi) : H_k^\sharp(K; R) \rightarrow H_k^\sharp([K]; R)$ is an homomorphism. Since $\tilde{H}_k^\sharp(\phi) : H_k^\sharp(K; R) \rightarrow \tilde{H}_k^\sharp(K; R)$ is an isomorphism and $\tilde{H}_k^\sharp(F) : \tilde{H}_k^\sharp(K; R) \rightarrow H_k^\sharp([K]; R)$ is an isomorphism according to the Proposition 8.6.7, we have that $H_k^\sharp(G) : H_k^\sharp(K; R) \rightarrow H_k^\sharp([K]; R)$ is an isomorphism. For similar reasons, $H_k(G) : H_k(K; R) \rightarrow H_k([K]; R)$ is an isomorphism.

We can repeat the previous discussion in the context of relative homology. Since $\phi : C_k(L; R) \rightarrow \tilde{C}_k(L; R)$, we may define the homomorphism $\phi : C_k(K, L; R) \rightarrow \tilde{C}_k(K, L; R)$ by $\phi([c]_{K,L}) = [\phi(c)]_{K,L}$ for all $[c]_{K,L} \in C_k(K, L; R)$, where as usual $[\cdot]_{K,L}$ represents the equivalence classes in both $C_k(K, L; R)$, and $\tilde{C}_k(K, L; R)$. Similarly, Since $\psi : \tilde{C}_k(L; R) \rightarrow C_k(L; R)$, we may define the homomorphism $\psi : \tilde{C}_k(K, L; R) \rightarrow C_k(K, L; R)$ by $\psi([c]_{K,L}) = [\psi(c)]_{K,L}$ for all $[c]_{K,L} \in \tilde{C}_k(K, L; R)$. We have the following commutative diagram

$$\begin{array}{ccccccc}
 \xrightarrow{\bar{\partial}_{k+2}} & C_{k+1}(K, L; R) & \xrightarrow{\bar{\partial}_{k+1}} & C_k(K, L; R) & \xrightarrow{\bar{\partial}_k} & C_{k-1}(K, L; R) & \xrightarrow{\bar{\partial}_{k-1}} \\
 & \psi \uparrow \downarrow \phi & & \psi \uparrow \downarrow \phi & & \psi \uparrow \downarrow \phi & \\
 \xrightarrow{\bar{\partial}_{k+2}} & \tilde{C}_{k+1}(K, L; R) & \xrightarrow{\bar{\partial}_{k+1}} & \tilde{C}_k(K, L; R) & \xrightarrow{\bar{\partial}_k} & \tilde{C}_{k-1}(K, L; R) & \xrightarrow{\bar{\partial}_{k-1}}
 \end{array}$$

Namely, $\{\phi_k\}_{k \in \mathbb{Z}}$ with $\phi_k = \phi$ for all k is a chain map from the chain complex $\mathcal{C} = \{(C_k(K, L; R), \bar{\partial}_k)\}_{k \in \mathbb{Z}}$ to the chain complex $\tilde{\mathcal{C}} = \{(\tilde{C}_k(K, L; R), \bar{\partial}_k)\}_{k \in \mathbb{Z}}$ and $\{\psi_k\}_{k \in \mathbb{Z}}$ with $\psi_k = \psi$ for all k is a chain map from $\tilde{\mathcal{C}}$ to $\mathcal{C}$.

We may define the homomorphism $\tilde{H}_k(\phi) : H_k(K, L; R) \rightarrow \tilde{H}_k(K, L; R)$ by $\tilde{H}_k(\phi)([c]_{K,L}) = [\phi(c)]_{K,L}$ for all $[c]_{K,L} \in H_k(K, L; R)$, where as usual $[\cdot]_{K,L}$ represents the equivalence classes in both $H_k(K, L; R)$, and $\tilde{H}_k(K, L; R)$. As well, we may define the homomorphism $\tilde{H}_k(\psi) : \tilde{H}_k(K, L; R) \rightarrow H_k(K, L; R)$ by $\tilde{H}_k(\psi)([c]_{K,L}) = [\psi(c)]_{K,L}$ for all $[c]_{K,L} \in \tilde{H}_k(K, L; R)$.

We can repeat all the previous discussion for the reduced homology and define $\tilde{H}_k^\sharp(\phi) : H_k^\sharp(K, L; R) \rightarrow \tilde{H}_k^\sharp(K, L; R)$ and $\tilde{H}_k^\sharp(\psi) : \tilde{H}_k^\sharp(K, L; R) \rightarrow H_k^\sharp(K, L; R)$.

As we did for the regular homology, we can use Proposition 8.4.33 to conclude that $\tilde{H}_k^\sharp(\phi)$ is an isomorphism if and only if $\tilde{H}_k(\phi)$ is an isomorphism. To prove that $\tilde{H}_k^\sharp(\phi)$ is an isomorphism, we prove that $\tilde{H}_k^\sharp(\psi)$ is the inverse of $\tilde{H}_k^\sharp(\phi)$. The proof is almost identical to the proof that we gave above for the reduced homology. The only difference is that we use the following proposition instead of Proposition 8.6.9.

Proposition 8.6.10

There exists $\mathcal{D} = \{D_k\}_{k \in \mathbb{Z}}$ with $D_k : \tilde{C}_k^\sharp(K, L; R) \rightarrow \tilde{C}_{k+1}^\sharp(K, L; R)$ such that $\tilde{\partial}_{k+1}^\sharp \circ D_k + D_{k-1} \circ \tilde{\partial}_k^\sharp = \phi \circ \psi - \text{Id}$ on $\tilde{C}_k^\sharp(K, L; R)$.

Proof (Sketch).

We can use the chain homotopy $\mathcal{D} = \{D_k\}_{k \in \mathbb{Z}}$ obtained in Proposition 8.6.9 to induce a chain homotopy between $\mathcal{F} = \{f_k\}_{k \in \mathbb{Z}}$ with $f_k = \text{Id} : \tilde{C}_k^\sharp(K, L; R) \rightarrow \tilde{C}_k^\sharp(K, L; R)$ and $\mathcal{G} = \{g_k\}_{k \in \mathbb{Z}}$ with $g_k = \phi \circ \psi : \tilde{C}_k^\sharp(K, L; R) \rightarrow \tilde{C}_k^\sharp(K, L; R)$.

It suffices to note that $D_k : \tilde{C}_k^\sharp(K; R) \rightarrow \tilde{C}_{k+1}^\sharp(K; R)$ was defined in such a way to have that $D_k : \tilde{C}_k^\sharp(L; R) \rightarrow \tilde{C}_{k+1}^\sharp(L; R)$. Thus we can define $D_k : \tilde{C}_k^\sharp(K, L; R) \rightarrow \tilde{C}_{k+1}^\sharp(K, L; R)$ in a natural way as we have done with $\tilde{\partial}_k^\sharp$ for instance to get $\tilde{\partial}_k^\sharp$. ■

We get from $G = F \circ \phi$ that $H_k(G) = \tilde{H}_k(F) \circ \tilde{H}_k(\phi) : H_k(K, L; R) \rightarrow H_k([K], [L]; R)$ is an homomorphism. Since $\tilde{H}_k(\phi) : H_k(K, L; R) \rightarrow \tilde{H}_k(K, L; R)$ is an isomorphism and $\tilde{H}_k(F) : \tilde{H}_k(K, L; R) \rightarrow H_k([K], [L]; R)$ is an isomorphism according to the Proposition 8.6.7, we have that $H_k(G) : H_k(K, L; R) \rightarrow H_k([K], [L]; R)$ is an isomorphism.

In summary, we get the following result.

Proposition 8.6.11

Let L be a subcomplex of a simplicial complex K and R is an integral domain. Then $H_k(G) : H_k(K; R) \rightarrow H_k([K]; R)$ and $H_k(G) : H_k(K, L; R) \rightarrow H_k([K], [L]; R)$ are isomorphisms. We have a similar result for the reduced homology.

To conclude this subsection, we should mention the following result. Suppose that $f : ([K_1], [L_1]) \rightarrow ([K_2], [L_2])$ is a continuous maps where L_i is a simplicial subcomplex of the simplicial complex K_i for $i = 1, 2$. Then we have the following commutative diagram.

$$\begin{array}{ccc} H_k(K_1, L_1; R) & \xrightarrow{H_k(f)} & H_k(K_2, L_2; R) \\ H_k(G) \downarrow & & \downarrow H_k(G) \\ H_k([K_1], [L_1]; R) & \xrightarrow{H_k(f)} & H_k([K_2], [L_2]; R) \end{array}$$

The proof of this result is interesting. It makes use of simplicial approximation. This is a naturality result for those who may know about category theory.

8.6.2 Relation Between Simplicial Cohomology and Singular Cohomology

Suppose that L is a simplicial subcomplex of a simplicial complex K , and that R is an integral domain. We have defined in the previous subsection an homomorphism $G : C_k(K; R) \rightarrow S_k([K]; R)$ and shown that it induces isomorphism $H_k(G) : H_k(K; R) \rightarrow H_k([K]; R)$ and $H_k(G) : H_k(K, L; R) \rightarrow H_k([K], [L]; R)$.

By duality, we get the following result.

Proposition 8.6.12

Let L be a subcomplex of a simplicial complex K and R is an integral domain. Then $H^k(G) : H^k([K]; R) \rightarrow H^k(K; R)$ and $H^k(G) : H^k([K], [L]; R) \rightarrow H^k(K, L; R)$ are isomorphisms.

As for $H_k(G)$, suppose that $f : ([K_1], [L_1]) \rightarrow ([K_2], [L_2])$ is a continuous maps where L_i is a simplicial subcomplex of the simplicial complex K_i for $i = 1, 2$. Then we have the following commutative diagram.

$$\begin{array}{ccc}
 H^k([K_2], [L_2]; R) & \xrightarrow{H^k(f)} & H^k([K_1], [L_1]; R) \\
 \downarrow H^k(G) & & \downarrow H^k(G) \\
 H^k(K_2, L_2; R) & \xrightarrow{H^k(f)} & H^k(K_1, L_1; R)
 \end{array}$$

Chapter 9

Riemannian Geometry

To be precise, this chapter is an introduction to the Riemannian geometry of surfaces (i.e. 2-dimensional manifolds). For a more complete study, the reader should consult [25, 27]. This chapter is inspired in large part on [25].

One of the major results of this chapter is Gauss-Bonnet theorem, Theorem 9.6.1, that relates integration of the curvature over a manifold to the Euler characteristic of the simplicial complex used to triangulate the manifold. We will conclude this chapter and these lecture notes by relating the sign of the curvature of a manifold with constant curvature to the type of geometry on the manifold. The study of the possible types of geometry is a really fascinating subject with deep historical background.

9.1 Riemannian Manifold

Definition 9.1.1

A **Riemannian manifold** is a manifold S of class C^∞ for which there exists a map $\tau : S \rightarrow \bigcup_{\mathbf{u} \in S} \mathcal{T}^2(T_{\mathbf{u}}S)$ of class C^∞ such that $\langle \cdot, \cdot \rangle_{\mathbf{u}} = \tau(\mathbf{u})$ is an inner product on $T_{\mathbf{u}}S$ (See Definition 5.7.2). The map τ is called a **Riemannian metric** on S .

The map τ in the previous definition is said to be smooth. An equivalent way to define smoothness is as it follows. We say that τ is smooth if the map defined by $\mathbf{u} \mapsto \langle F_1(\mathbf{u}), F_2(\mathbf{u}) \rangle_{\mathbf{u}}$ for $\mathbf{u} \in S$ is of class C^∞ for all smooth vector fields F_1 and F_2 on S .

It follows from Proposition 5.7.3 that it is always possible to define a Riemannian metric on a manifold of $\mathbb{R}^n$. If M is a submanifold of a Riemannian manifold S , then M inherits a Riemannian metric by restricting $\tau : S \rightarrow \bigcup_{\mathbf{u} \in S} \mathcal{T}^2(T_{\mathbf{u}}S)$ to M .

Definition 9.1.2

Let S_1 and S_2 be two Riemannian manifolds. An **isometry** between S_1 and S_2 is an isomorphism $f : S_1 \rightarrow S_2$ such that f and $f^{-1} : S_2 \rightarrow S_1$ are of class C^∞ , and $\langle f_*(\mathbf{u}, \mathbf{x}_1), f_*(\mathbf{u}, \mathbf{x}_2) \rangle_{f(\mathbf{u})} = \langle (\mathbf{u}, \mathbf{x}_1), (\mathbf{u}, \mathbf{x}_2) \rangle_{\mathbf{u}}$ for all $(\mathbf{u}, \mathbf{x}_1), (\mathbf{u}, \mathbf{x}_2) \in T_{\mathbf{u}}S_1$ and all $\mathbf{u} \in S_1$.

Example 9.1.3

There may be more than one Riemannian metric on a manifold S . Moreover, S equipped with one of these Riemannian metric may not be isometric to S equipped with another Riemannian metric.

For instance, the torus T^2 seen as a submanifold of $\mathbb{R}^3$ inherits the Riemannian metric of $\mathbb{R}^3$. Namely, $\langle (\mathbf{u}, \mathbf{x}_1), (\mathbf{u}, \mathbf{x}_2) \rangle_{\mathbf{u}} = \langle \mathbf{x}_1, \mathbf{x}_2 \rangle$ for all $(\mathbf{u}, \mathbf{x}_1), (\mathbf{u}, \mathbf{x}_2) \in T_{\mathbf{u}}T^2$ and $\mathbf{u} \in T^2$ where the second inner product is the standard inner product in $\mathbb{R}^3$.

We also have that $T^2 \cong S^1 \times S^1 \subset \mathbb{R}^4$. Thus we may define a Riemannian metric on T^2 based on the Riemannian metric on S^1 . If $\tau^\circ : S^1 \rightarrow \bigcup_{\mathbf{u} \in S^1} \mathcal{T}^2(T_{\mathbf{u}}S^1)$ is a Riemannian metric on S^1 and $\langle \cdot, \cdot \rangle_{\mathbf{u}}^\circ = \tau^\circ(\mathbf{u})$ for all $\mathbf{u} \in S^1$, then we can define the Riemannian metric $\tau : S^1 \times S^1 \rightarrow \bigcup_{(\mathbf{u}_1, \mathbf{u}_2) \in S^1 \times S^1} \mathcal{T}^2(T_{(\mathbf{u}_1, \mathbf{u}_2)}(S^1 \times S^1))$ on $T^2 \cong S^1 \times S^1$ as it follows. We set $\tau(\mathbf{u}_1, \mathbf{u}_2) = \langle \cdot, \cdot \rangle_{(\mathbf{u}_1, \mathbf{u}_2)}$ for all $(\mathbf{u}_1, \mathbf{u}_2) \in S^1 \times S^1$ where

$$\begin{aligned} & \langle ((\mathbf{u}_1, \mathbf{u}_2), (\mathbf{x}_{1,1}, \mathbf{x}_{2,1})), ((\mathbf{u}_1, \mathbf{u}_2), (\mathbf{x}_{1,2}, \mathbf{x}_{2,2})) \rangle_{(\mathbf{u}_1, \mathbf{u}_2)} \\ &= \langle (\mathbf{u}_1, \mathbf{x}_{1,1}), (\mathbf{u}_1, \mathbf{x}_{1,2}) \rangle_{\mathbf{u}_1}^\circ + \langle (\mathbf{u}_2, \mathbf{x}_{2,1}), (\mathbf{u}_2, \mathbf{x}_{2,2}) \rangle_{\mathbf{u}_2}^\circ \end{aligned}$$

for $(\mathbf{u}_i, \mathbf{x}_{i,j}) \in T_{\mathbf{u}_i}S^1$ with $1 \leq i, j \leq 2$. These two metrics are not equivalent because $S^1 \times S^1$ with its Riemannian metric cannot be imbedded as a submanifold of $\mathbb{R}^3$. The proof is based on the fact that the “curvature” of $S^1 \times S^1$ is everywhere null. Consult [25] for more information. 

9.2 Preliminaries

The first section of this chapter is technical and tedious but clear the path for the study of Riemannian geometry later by giving us some of the basic elements that we will play with later. So, it is worth the pain.

9.2.1 Group of Rotations

Let S be a 2-dimension Riemannian manifold with a Riemannian metric $\tau : S \rightarrow \bigcup_{\mathbf{u} \in S} \mathcal{T}^2(T_{\mathbf{u}}S)$. As usual, we let $\langle \cdot, \cdot \rangle_{\mathbf{u}} = \tau(\mathbf{u})$ and $\| \cdot \|_{\mathbf{u}} = \langle \cdot, \cdot \rangle_{\mathbf{u}}^{1/2}$.

For $\mathbf{u} \in S$, we define the action of the group of rotations S^1 on $T_{\mathbf{u}}S$ as it follows. Given $\eta = e^{i\theta} \in S^1 = \{z \in \mathbb{C} : |z| = 1\}$ and $(\mathbf{u}, \mathbf{x}) \in T_{\mathbf{u}}S$, we set $\eta(\mathbf{u}, \mathbf{x}) = (\mathbf{u}, \theta.\mathbf{x})$ where $\theta.\mathbf{x}$ is the rotation by θ radians about the origin of the vector $\mathbf{x}$ in the 2-dimension space $T_{\mathbf{u}}S$.

We first prove that the function

$$\begin{aligned}\Phi : TS \times S^1 &\rightarrow TS \\ ((\mathbf{u}, \mathbf{x}), \eta) &\rightarrow \eta(\mathbf{u}, \mathbf{x}) = (\mathbf{u}, \theta.\mathbf{x})\end{aligned}$$

is a smooth function. We have seen that if (W, U, ϕ) is a local chart of S , then $(\tilde{W}, \tilde{U}, \tilde{\phi})$ is a local chart of TS where $\tilde{W} = TW = W \times \mathbb{R}^2$, $\tilde{U} = TU = \bigcup_{\mathbf{u} \in U} T_{\mathbf{u}}S$, and $\tilde{\phi} : \tilde{W} \rightarrow \tilde{U}$ is defined by $\tilde{\phi}(\mathbf{w}, \mathbf{y}) = \phi_*(\mathbf{w}, \mathbf{y}) = (\phi(\mathbf{w}), D\phi(\mathbf{w})\mathbf{y})$ for all $(\mathbf{w}, \mathbf{y}) \in \tilde{W}$. We consider a local chart of $TS \times S^1$ of the form $(\tilde{W} \times I, \tilde{U} \times C, (\tilde{\phi}, \psi))$ where $I \subset \mathbb{R}$ is an open interval of length less than 2π , the function $\psi : \mathbb{R} \rightarrow S^1$ is defined by $\psi(\theta) = e^{i\theta}$ for $\theta \in \mathbb{R}$, and $C = \psi(I)$. The condition that I be of length less than 2π is to ensure that $\tilde{\psi}$ is one-to-one.

Since we assume that local chart (W, U, ϕ) of S is orientation preserving, we may assume that the orientation on $T_{\phi(\mathbf{w})}S$ is given by $\mu_{\phi(\mathbf{w})} = [(\phi(\mathbf{w}), D\phi(\mathbf{w})\mathbf{e}_1), (\phi(\mathbf{w}), D\phi(\mathbf{w})\mathbf{e}_2)]$ while the orientation on $T_{\mathbf{w}}W$ is given by $\mu_{\mathbf{w}} = [(\mathbf{w}, \mathbf{e}_1), (\mathbf{w}, \mathbf{e}_2)]$. Do not forget that $D\phi(\mathbf{w}) : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ is of rank 2 for all $\mathbf{w} \in W$.

We use Gram-Schmidt orthogonalization and change of bases to express Φ in local coordinates. Let

$$\mathbf{p}_1(\mathbf{w}) = D\phi(\mathbf{w})\mathbf{e}_1$$

and

$$\mathbf{p}_2(\mathbf{w}) = D\phi(\mathbf{w})\mathbf{e}_2 - \frac{\langle D\phi(\mathbf{w})\mathbf{e}_1, D\phi(\mathbf{w})\mathbf{e}_2 \rangle_{\phi(\mathbf{w})}}{\langle D\phi(\mathbf{w})\mathbf{e}_1, D\phi(\mathbf{w})\mathbf{e}_1 \rangle_{\phi(\mathbf{w})}} D\phi(\mathbf{w})\mathbf{e}_1.$$

Let $\mathcal{A} = \{(\phi(\mathbf{w}), D\phi(\mathbf{w})\mathbf{e}_1), (\phi(\mathbf{w}), D\phi(\mathbf{w})\mathbf{e}_2)\}$ and $\mathcal{B} = \{(\phi(\mathbf{w}), \mathbf{q}_1(\mathbf{w})), (\phi(\mathbf{w}), \mathbf{q}_2(\mathbf{w}))\}$ where $\mathbf{q}_i(\mathbf{w}) = \|\mathbf{p}_i(\mathbf{w})\|_{\phi(\mathbf{w})}^{-1} \mathbf{p}_i(\mathbf{w})$ for $i = 1, 2$. We have that $\mathcal{B}$ is orthonormal basis for $T_{\phi(\mathbf{w})}S$. If $Q(\mathbf{w}) = [\text{Id}]_{\mathcal{B}}^{\mathcal{A}}$ is the matrix of change of bases from $\mathcal{A}$ to $\mathcal{B}$, then

$$[\theta.(D\phi(\mathbf{w})\mathbf{y})]_{\mathcal{A}} = (Q(\mathbf{w}))^{-1} \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} Q(\mathbf{w}) \underbrace{[D\phi(\mathbf{w})\mathbf{y}]_{\mathcal{A}}}_{=\mathbf{y}} \quad (9.2.1)$$

for $\mathbf{y} \in \mathbb{R}^2$. Hence, the local representation $\tilde{\Phi}$ of Φ satisfies

$$\tilde{\phi}(\tilde{\Phi}((\mathbf{w}, \mathbf{y}), \theta)) = \Phi((\tilde{\phi}, \psi)((\mathbf{w}, \mathbf{y}), \theta)) = \Phi((\phi(\mathbf{w}), D\phi(\mathbf{w})\mathbf{y}), \theta) = (\phi(\mathbf{w}), \theta.(D\phi(\mathbf{w})\mathbf{y}))$$

for $((\mathbf{w}, \mathbf{y}), \theta) \in \tilde{W} \times I$. Since the right hand side in (9.2.1) is a smooth function, we deduce that $\tilde{\Phi}$ is a smooth function.

Now that we have dealt with the smoothness of Φ , we should focus on the geometric aspect of the action of $\eta \in S^1$ on $T_{\mathbf{u}}S$ for all $\mathbf{u}$. This is the essential element for the rest of this chapter. Suppose that $\mathcal{B} = \{(\mathbf{u}, \mathbf{v}_1), (\mathbf{u}, \mathbf{v}_2)\}$ is an oriented and orthonormal basis of $T_{\mathbf{u}}S$. Every element $(\mathbf{u}, \mathbf{x}) \in T_{\mathbf{u}}S$ can be expressed as $(\mathbf{u}, \mathbf{x}) = (\mathbf{u}, a_1\mathbf{v}_1 + a_2\mathbf{v}_2)$ for

$a_1, a_2 \in \mathbb{R}$. Then $[\mathbf{x}]_{\mathcal{B}} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$. As we saw above, a rotation by $\theta \in \mathbb{R}$ of $[\mathbf{x}]_{\mathcal{B}}$ is represented in the basis $\mathcal{B}$ by

$$[\theta.\mathbf{x}]_{\mathcal{B}} = \theta.[\mathbf{x}]_{\mathcal{B}} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} a_1 \cos(\theta) - a_2 \sin(\theta) \\ a_1 \sin(\theta) + a_2 \cos(\theta) \end{pmatrix}. \quad (9.2.2)$$

Thus $\theta.\mathbf{x} = (a_1 \cos(\theta) - a_2 \sin(\theta))\mathbf{v}_1 + (a_1 \sin(\theta) + a_2 \cos(\theta))\mathbf{v}_2$.

The norm on $T_{\mathbf{u}}S$ is defined by $\|(\mathbf{u}, \mathbf{x})\|_{\mathbf{u}} = (\langle (\mathbf{u}, \mathbf{x}), (\mathbf{u}, \mathbf{x}) \rangle_{\mathbf{u}})^{1/2}$ for all $(\mathbf{u}, \mathbf{x}) \in T_{\mathbf{u}}S$. Since $\{(\mathbf{u}, \mathbf{v}_1), (\mathbf{u}, \mathbf{v}_2)\}$ is an orthonormal basis, we have that $\|(\mathbf{u}, \mathbf{x})\|_{\mathbf{u}} = \|(\mathbf{u}, a_1\mathbf{v}_1 + a_2\mathbf{v}_2)\|_{\mathbf{u}} = \sqrt{a_1^2 + a_2^2}$. We also have that $\|(\mathbf{u}, \theta.\mathbf{x})\|_{\mathbf{u}} = \sqrt{a_1^2 + a_2^2}$. In particular, if $\|(\mathbf{u}, \mathbf{x})\|_{\mathbf{u}} = 1$, then $\|(\mathbf{u}, \theta.\mathbf{x})\|_{\mathbf{u}} = 1$.

For $(\mathbf{u}, \mathbf{x})$ fixed, we may use (9.2.2) to compute the derivative of the map from $R : \mathbb{R} \rightarrow T_{\mathbf{u}}S$ defined by $R(\theta) = \eta(\mathbf{u}, \mathbf{x}) = (\mathbf{u}, \theta.\mathbf{x})$. We find that

$$R'(\theta) = (\mathbf{u}, (\pi/2 + \theta).\mathbf{x}) = v(\eta(\mathbf{u}, \mathbf{x})) \quad (9.2.3)$$

where $v = e^{i\pi/2} = i$. Note that this relation can easily be deduced if we express (9.2.2) in $\mathbb{C} \cong \mathbb{R}^2$.

9.2.2 The Tangent Bundle $T(TS)$

Let S be a 2-dimensional Riemannian manifold. Recall that TS is a 4-dimensional manifold. In particular, if (W, U, ϕ) is a local chart of S , then $(\tilde{W}, \tilde{U}, \tilde{\phi})$ is a local chart of TS where $\tilde{W} = W \times \mathbb{R}^2$, $\tilde{U} = TU = \bigcup_{\mathbf{u} \in U} T_{\mathbf{u}}S$, and $\tilde{\phi} : \tilde{W} \rightarrow \tilde{U}$ is defined by $\tilde{\phi}(\mathbf{w}, \mathbf{y}) = \phi_*(\mathbf{w}, \mathbf{y}) = (\phi(\mathbf{w}), D\phi(\mathbf{w})\mathbf{y})$ for all $(\mathbf{w}, \mathbf{y}) \in \tilde{W}$.

Using the procedure used to obtain the local chart $(\tilde{W}, \tilde{U}, \tilde{\phi})$ above, we find that a local chart of $T(TS)$ is given by $(\check{W}, \check{U}, \check{\phi})$ where $\check{W} = (W \times \mathbb{R}^2) \times (\mathbb{R}^2 \times \mathbb{R}^2)$, $\check{U} = T\tilde{U} = \bigcup_{(\mathbf{u}, \mathbf{x}) \in \tilde{U}} T_{(\mathbf{u}, \mathbf{x})}(TS)$, and $\check{\phi} : \check{W} \rightarrow \check{U}$ is defined by

$$\begin{aligned} \check{\phi}((\mathbf{w}, \mathbf{y}), (\mathbf{r}, \mathbf{s})) &= \check{\phi}_*((\mathbf{w}, \mathbf{y}), (\mathbf{r}, \mathbf{s})) = \left(\tilde{\phi}(\mathbf{w}, \mathbf{y}), D_{\mathbf{w}, \mathbf{y}} \tilde{\phi}(\mathbf{w}, \mathbf{y}) \begin{pmatrix} \mathbf{r} \\ \mathbf{s} \end{pmatrix} \right) \\ &= ((\phi(\mathbf{w}), D\phi(\mathbf{w})\mathbf{y}), (D\phi(\mathbf{w})\mathbf{r}, D_{\mathbf{w}}(D\phi(\mathbf{w})\mathbf{y})\mathbf{r} + D\phi(\mathbf{w})\mathbf{s})) \end{aligned}$$

for all $((\mathbf{w}, \mathbf{y}), (\mathbf{r}, \mathbf{s})) \in \check{W}$. Note that $(\mathbf{y}, \mathbf{r}) \mapsto D_{\mathbf{w}}(D\phi(\mathbf{w})\mathbf{y})\mathbf{r}$ is a bilinear map.

The important information to remember is that $T(TS)$ is a 8-dimensional manifold, that $T_{\mathbf{u}, \mathbf{x}}(TS)$ is a 4-dimensional vector space for all $(\mathbf{u}, \mathbf{x}) \in TS$ and that the elements of $T_{\mathbf{u}, \mathbf{x}}(TS)$ are of the form $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q}))$ where $(\mathbf{u}, \mathbf{p}) \in T_{\mathbf{u}}S$.

Let $\pi_S : TS \rightarrow S$ be the projection defined by $\pi_S(\mathbf{u}, \mathbf{x}) = \mathbf{u}$ for all $(\mathbf{u}, \mathbf{x}) \in T_{\mathbf{u}}S$. This is obviously a smooth map since the representation of π_S with respect to a local charts (W, U, ϕ) and $(\tilde{W}, \tilde{U}, \tilde{\phi})$ defined above is $\pi(\mathbf{w}, \mathbf{y}) = \mathbf{w}$ for all $(\mathbf{w}, \mathbf{y}) \in W \times \mathbb{R}^2$.

Lemma 9.2.1

We have that $(\pi_S)_* : T(TS) \rightarrow TS$ is given by $(\pi_S)_*((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) = (\mathbf{u}, \mathbf{p})$ for all $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \in T_{(\mathbf{u}, \mathbf{x})}(TS)$ and $(\mathbf{u}, \mathbf{x}) \in TS$.

Proof.

Suppose that $(\tilde{W}, \tilde{U}, \tilde{\phi})$ and $(\check{W}, \check{U}, \check{\phi})$ are the local charts of TS and $T(TS)$ respectively defined above. Moreover, suppose that $((\mathbf{w}, \mathbf{y}), (\mathbf{r}, \mathbf{s})) \in \check{W}$ is such that $\check{\phi}((\mathbf{w}, \mathbf{y}), (\mathbf{r}, \mathbf{s})) = ((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \in \tilde{U}$ with $\mathbf{u} = \phi(\mathbf{w})$ and $\mathbf{p} = D\phi(\mathbf{w})\mathbf{r}$.

If $\pi : W \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is the projection defined by $\pi(\mathbf{w}, \mathbf{y}) = \mathbf{w}$ for all $(\mathbf{w}, \mathbf{y}) \in W \times \mathbb{R}^2$, then $\pi_*((\mathbf{w}, \mathbf{y}), (\mathbf{r}, \mathbf{s})) = (\pi(\mathbf{w}, \mathbf{y}), D\pi(\mathbf{w})(\mathbf{r}, \mathbf{s})) = (\mathbf{w}, \pi(\mathbf{r}, \mathbf{s})) = (\mathbf{w}, \mathbf{r})$ for all $((\mathbf{w}, \mathbf{y}), (\mathbf{r}, \mathbf{s})) \in \check{W}$.

Since $\phi \circ \pi = \pi_S \circ \tilde{\phi}$ on W , we get that $\tilde{\phi} \circ \pi_* = \phi_* \circ \pi_* = (\pi_S)_* \circ \tilde{\phi}_* = (\pi_S)_* \circ \tilde{\phi}$ on $\check{W} = TW$. Hence

$$\begin{aligned} (\pi_S)_*((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) &= (\pi_S)_*(\check{\phi}((\mathbf{w}, \mathbf{y}), (\mathbf{r}, \mathbf{s}))) = \tilde{\phi}(\pi_*((\mathbf{w}, \mathbf{y}), (\mathbf{r}, \mathbf{s}))) = \tilde{\phi}(\mathbf{w}, \mathbf{r}) \\ &= (\phi(\mathbf{w}), D\phi(\mathbf{w})\mathbf{r}) = (\mathbf{u}, \mathbf{p}) . \end{aligned}$$

Proposition 9.2.2

Suppose that S_1 and S_2 are two 2-dimensional Riemannian manifolds and that $f : S_1 \rightarrow S_2$ is an isometry. Let $\tilde{f} = f_* : TS_1 \rightarrow TS_2$. Then

1. $\pi_{S_2} \circ \tilde{f} = f \circ \pi_{S_1}$ on TS_1 , and
2. $(\pi_{S_2})_* \circ \tilde{f}_* = \tilde{f} \circ (\pi_{S_1})_*$ on $T(TS_1)$.

Moreover, if S_1 and S_2 are oriented and f preserves the orientation, then

3. $\eta \circ \tilde{f} = \tilde{f} \circ \eta$ for all $\eta \in S^1$ acting on TS_1 and TS_2 .

Proof.

Let (W_i, U_i, ϕ_i) , $(\tilde{W}_i, \tilde{U}_i, \tilde{\phi}_i)$ and $(\check{W}_i, \check{U}_i, \check{\phi}_i)$ be the previously defined local charts of S_i , TS_i and $T(TS_i)$ respectively for $i = 1, 2$.

Let $g : W_1 \rightarrow W_2$ be the local representation of f ; namely, $f \circ \phi_1 = \phi_2 \circ g$ on W_1 . We then have that $\tilde{f} \circ (\phi_1)_* = f_* \circ (\phi_1)_* = (\phi_2)_* \circ g_*$ on $TW_1 = \tilde{W}_1$.

1) Given $(\mathbf{u}, \mathbf{x}) \in \tilde{U}_1 \subset T_{\mathbf{u}}S_1$, we have that $(\mathbf{u}, \mathbf{x}) = (\phi_1)_*(\mathbf{w}, \mathbf{y})$ for some $(\mathbf{w}, \mathbf{y}) \in \tilde{W}_1$. Hence

$$\begin{aligned} (\pi_{S_2} \circ \tilde{f})(\mathbf{u}, \mathbf{x}) &= \pi_{S_2}((\tilde{f} \circ (\phi_1)_*)(\mathbf{w}, \mathbf{y})) = \pi_{S_2}((\phi_2)_* \circ g_*)(\mathbf{w}, \mathbf{y})) \\ &= \pi_{S_2}((\phi_2 \circ g)_*(\mathbf{w}, \mathbf{y})) = (\phi_2 \circ g)(\mathbf{w}) = (\phi_2 \circ g \circ \phi_1^{-1})(\phi_1(\mathbf{w})) \\ &= f(\mathbf{x}) = (f \circ \pi_{S_1})(\mathbf{u}, \mathbf{x}) . \end{aligned}$$

2) Since $\pi_{S_2} \circ \tilde{f} = f \circ \pi_{S_1}$, we get

$$(\pi_{S_2})_* \circ \tilde{f}_* = (\pi_{S_2} \circ \tilde{f})_* = (f \circ \pi_{S_1})_* = f_* \circ (\pi_{S_1})_* = \tilde{f} \circ (\pi_{S_1})_* .$$

3) We need to show that $\eta(\tilde{f}(\mathbf{u}, \mathbf{x})) = \tilde{f}(\eta(\mathbf{u}, \mathbf{x}))$ for all $(\mathbf{u}, \mathbf{x}) \in T_{\mathbf{u}}S_1$ and $\eta \in S^1$.

Since

$$\left\langle \tilde{f}(\eta(\mathbf{u}, \mathbf{x})), \tilde{f}(\mathbf{u}, \mathbf{x}) \right\rangle_{f(\mathbf{u})} = \langle \eta(\mathbf{u}, \mathbf{x}), (\mathbf{u}, \mathbf{x}) \rangle_{\mathbf{u}} ,$$

the angle between $\tilde{f}(\eta(\mathbf{u}, \mathbf{x}))$ and $\tilde{f}(\mathbf{u}, \mathbf{x})$ is equal modulo the direction of the rotation to the angle between $\eta(\mathbf{u}, \mathbf{x})$ and $(\mathbf{u}, \mathbf{x})$; namely, it is equal to η modulo the direction of the rotation. But we also assume that f is orientation preserving, thus $\tilde{f}(\eta(\mathbf{u}, \mathbf{x})) = \eta(\tilde{f}(\mathbf{u}, \mathbf{x}))$ for all $(\mathbf{u}, \mathbf{x}) \in T_{\mathbf{u}}S_1$ and $\eta \in S^1$. ■

Proposition 9.2.3

Let S be a 2-dimensional oriented and Riemannian manifold, and $f : S \rightarrow S$ be an orientation preserving isometry such that $f(\mathbf{u}) = \mathbf{u}$ for some $\mathbf{u} \in S$, then $f_* : T_{\mathbf{u}}S \rightarrow T_{\mathbf{u}}S$ is a simple rotation by an angle $\eta \in S^1$.

Proof.

Choose $(\mathbf{u}, \mathbf{x}) \in T_{\mathbf{u}}S$. Since f_* preserve the length, we have that $\|f_*(\mathbf{u}, \mathbf{x})\|_{\mathbf{u}} = \|(\mathbf{u}, \mathbf{x})\|_{\mathbf{u}}$. Let $\eta \in S^1$ be the angle such that $f_*(\mathbf{u}, \mathbf{x}) = \eta(\mathbf{u}, \mathbf{x})$.

Given $(\mathbf{u}, \mathbf{y}) \in T_{\mathbf{u}}S$, we have that $(\mathbf{u}, \mathbf{y}) = a(\mathbf{u}, \tilde{\mathbf{y}}) = (\mathbf{u}, a\tilde{\mathbf{y}})$ where $a \in \mathbb{R}$ and $\|(\mathbf{u}, \tilde{\mathbf{y}})\|_{\mathbf{u}} = \|(\mathbf{u}, \mathbf{x})\|_{\mathbf{u}}$. There exists $\mu \in S^1$ such that $(\mathbf{u}, \tilde{\mathbf{y}}) = \mu(\mathbf{u}, \mathbf{x})$. Hence, it follows from the third item of Proposition 9.2.2 that

$$\begin{aligned} f_*(\mathbf{u}, \mathbf{y}) &= f_*(\mathbf{u}, a\tilde{\mathbf{y}}) = af_*(\mathbf{u}, \tilde{\mathbf{y}}) = af_*(\mu(\mathbf{u}, \mathbf{x})) = a\mu(f_*(\mathbf{u}, \mathbf{x})) = a\mu(\eta(\mathbf{u}, \mathbf{x})) \\ &= a\eta(\mu(\mathbf{u}, \mathbf{x})) = a\eta(\mathbf{u}, \tilde{\mathbf{y}}) = \eta(\mathbf{u}, a\tilde{\mathbf{y}}) = \eta(\mathbf{u}, \mathbf{y}) \end{aligned}$$

where we have used the fact that $f_*|_{T_{\mathbf{u}}S}$ is a linear mapping. ■

9.2.3 The Tangent Bundle TS_c

We begin with a definition.

Definition 9.2.4

Let S be a 2-dimensional Riemannian manifold. The **circle bundle** is the set defined by $S_c = \bigcup_{\mathbf{u} \in S} S_c(\mathbf{u})$, where $S_c(\mathbf{u}) = \{(\mathbf{u}, \mathbf{x}) \in T_{\mathbf{u}}S : \langle (\mathbf{u}, \mathbf{x}), (\mathbf{u}, \mathbf{x}) \rangle_{\mathbf{u}} = 1\}$ for all $\mathbf{u} \in S$.

For each $\mathbf{u} \in S$ fixed, the set $S_c(\mathbf{u})$ is the unit circle in $T_{\mathbf{u}}S$ centred at the origin. Therefore, given $\eta = e^{i\theta} \in S^1$, we have that $\eta : S_c \rightarrow S_c$. This defines a smooth mapping.

Proposition 9.2.5

If S be a 2-dimensional Riemannian manifold, then S_c is a 3-dimensional manifold of class C^∞

Proof.

We could prove this proposition by explicitly providing smooth local charts as we will do later. Instead, we consider the map $g : TS \rightarrow \mathbb{R}$ defined by $g((\mathbf{u}, \mathbf{x})) = \langle (\mathbf{u}, \mathbf{x}), (\mathbf{u}, \mathbf{x}) \rangle_{\mathbf{u}}$ for all $(\mathbf{u}, \mathbf{x}) \in T_x S$ and $\mathbf{u} \in S$ and prove that $S_c = g^{-1}(\{1\})$ is a smooth manifold.

Suppose that (W, U, ϕ) is a local chart of the manifold S . The local representation of g is given by

$$\phi^*(g)(\mathbf{w}, \mathbf{y}) = g(\phi_*(\mathbf{w}, \mathbf{y})) = \langle (\phi(\mathbf{w}), D\phi(\mathbf{w})\mathbf{y}), (\phi(\mathbf{w}), D\phi(\mathbf{w})\mathbf{y}) \rangle_{\phi(\mathbf{w})}$$

for all $(\mathbf{w}, \mathbf{y}) \in W \times \mathbb{R}^2 \subset \mathbb{R}^2 \times \mathbb{R}^2$. Since $D_{\mathbf{w}, \mathbf{y}} \phi^*(g)(\mathbf{w}, \mathbf{y})$ is of rank 1 for all $(\mathbf{w}, \mathbf{y}) \in W \times \mathbb{R}^2$ such that $\phi^*(g)(\mathbf{w}, \mathbf{y}) = 1$, We get from Theorem 5.2.15 that $(\phi^*(g))^{-1}(\{1\})$ is a 3-dimensional manifold in $W \times \mathbb{R}^2$.

If $(\tilde{W}, \tilde{U}, \tilde{\phi})$ is a local chart of $(\phi^*(g))^{-1}(\{1\}) \subset W \times \mathbb{R}^2$, then $(\tilde{W}, \phi_*(\tilde{U}), \phi_* \circ \tilde{\phi})$ is local chart of $S_c = g^{-1}(\{1\})$. $\blacksquare$

Let S be a 2-dimensional Riemannian manifold. We first define local charts of S_c . Let (W, U, ϕ) be a local chart of S . We define the smooth function $E : W \rightarrow \bigcup_{\mathbf{w} \in W} D\phi(\mathbf{w})(\mathbb{R}^2)$ by

$$E(\mathbf{w}) = \|D\phi(\mathbf{w})(\mathbf{e}_1)\|_{\phi(\mathbf{w})}^{-1} D\phi(\mathbf{w})(\mathbf{e}_1)$$

for $\mathbf{w} \in W$. Recall that $D\phi(\mathbf{w})$ is of rank 2 for all $\mathbf{w} \in W$ by definition of local charts. Therefore, the norm is never null. A local chart of S_c is defined by $(\tilde{W}, \tilde{U}, \tilde{\phi})$ where $\tilde{W} = W \times I$ with $I \subset \mathbb{R}$ an open interval of length less than 2π , $\tilde{\phi}$ is defined by

$$\tilde{\phi}(\mathbf{w}, \theta) = (\phi(\mathbf{w}), \theta \cdot E(\mathbf{w}))$$

for $(\mathbf{w}, \theta) \in \tilde{W}$, and $\tilde{U} = \tilde{\phi}(\tilde{W}) \subset \bigcup_{\mathbf{u} \in U} S_c(\mathbf{u})$. The condition that I be of length less than 2π

is to ensure that $\tilde{\phi}$ is one-to-one. The proof of the smoothness of the function $\tilde{\phi}$ is similar to the proof of the smoothness of the function Φ in Subsection 9.2.1.

Remark 9.2.6

We have that $S_c(U) = \bigcup_{\mathbf{u} \in U} S_c(\mathbf{u}) \cong W \times S^1$. To prove this statement, we define the mapping $g : W \times S^1 \rightarrow S_c(U)$ by $g(\mathbf{w}, \eta) = \eta(\phi(\mathbf{w}), E(\mathbf{w}))$ for $\mathbf{w} \in W$ and $\eta \in S^1$. The map g is an isomorphism between $W \times S^1$ and $S_c(U)$.

Since $\phi : W \rightarrow U$ is an isomorphism, we have that

$$G(\mathbf{u}, \eta) = g(\phi^{-1}(\mathbf{u}), \eta) = \eta(\mathbf{u}, E(\phi^{-1}(\mathbf{u})))$$

for $\mathbf{u} \in U$ and $\eta \in S^1$ defines an isomorphism between $U \times S^1$ and $S_c(U)$.

However, it is generally false that $S_c \cong S \times S^1$. For instance, we cannot have $S_c \cong S \times S^1$ for $S = S^2$ because the isomorphism implies the existence of a non-null vector field on S and we know from Proposition 7.3.4 that this is not possible on S^2 . ♠

Since S_c is a manifold, then TS_c is well define. Using the procedure to obtain local charts of the previous subsection, a local chart of TS_c is given by $(\check{W}, \check{U}, \check{\phi})$ where $\check{W} = (W \times I) \times (\mathbb{R}^2 \times \mathbb{R})$ with $I \subset \mathbb{R}$ an open interval of length less than 2π , $\check{U} = \check{\phi}_*(\check{W})$, and $\check{\phi} = \check{\phi}_*$. In particular, we have

$$\begin{aligned} \check{\phi}((\mathbf{w}, \theta), (\mathbf{r}, t)) &= \check{\phi}_*((\mathbf{w}, \theta), (\mathbf{r}, t)) = \left(\check{\phi}(\mathbf{w}, \theta), D_{\mathbf{w}, \theta} \check{\phi}(\mathbf{w}, \theta) \begin{pmatrix} \mathbf{r} \\ t \end{pmatrix} \right) \\ &= \left(\check{\phi}(\mathbf{w}, \theta), \left(D\phi(\mathbf{w}) \mathbf{r}, D_{\mathbf{w}, \theta}(\theta.E(\mathbf{w})) \begin{pmatrix} \mathbf{r} \\ t \end{pmatrix} \right) \right) \\ &= \left(\check{\phi}(\mathbf{w}, \theta), \left(D\phi(\mathbf{w}) \mathbf{r}, D_{\mathbf{w}}(\theta.E(\mathbf{w})) \mathbf{r} + t \frac{\partial}{\partial \theta}(\theta.E(\mathbf{w})) \right) \right) \\ &= \left(\check{\phi}(\mathbf{w}, \theta), (D\phi(\mathbf{w}) \mathbf{r}, D_{\mathbf{w}}(\theta.E(\mathbf{w})) \mathbf{r} + t(\pi/2 + \theta).E(\mathbf{w})) \right) \end{aligned} \quad (9.2.4)$$

for all $((\mathbf{w}, \theta), (\mathbf{r}, t)) \in \check{W}$ where we have used (9.2.3) to obtain the last equality.

We define a vector field H on S_c using local charts. Let

$$H(\mathbf{u}, \mathbf{x}) = \check{\phi}_*((\mathbf{w}, \theta), (\mathbf{0}, 1)) \quad (9.2.5)$$

for $(\mathbf{u}, \mathbf{x}) = \check{\phi}(\mathbf{w}, \theta)$. This is a smooth vector field on S_c because its local representation is given by

$$((\check{\phi}_*)^{-1} \circ H \circ \check{\phi})(\mathbf{w}, \theta) = ((\mathbf{w}, \theta), (\mathbf{0}, 1))$$

for $(\mathbf{w}, \theta) \in \check{W} = W \times I$. Note that

$$\begin{aligned} H(\mathbf{u}, \mathbf{x}) &= \check{\phi}_*((\mathbf{w}, \theta), (\mathbf{0}, 1)) = \left((\mathbf{u}, \theta.\tilde{E}(\mathbf{u})), \left(\mathbf{0}, \frac{\partial}{\partial \nu}(\nu.\tilde{E}(\mathbf{u})) \Big|_{\nu=\theta} \right) \right) \\ &= \left((\mathbf{u}, \mathbf{x}), \left(\mathbf{0}, \frac{\partial}{\partial \nu}(\nu.\tilde{E}(\mathbf{u})) \Big|_{\nu=\theta} \right) \right) \end{aligned} \quad (9.2.6)$$

where $\tilde{E}(\mathbf{u}) = E(\phi^{-1}(\mathbf{u})) = E(\mathbf{w})$.

Remark 9.2.7

Using the local charts $(\check{W}, \check{U}, \check{\phi})$ of S_c and the interpretation of the tangent space of S_c has a space of differential linear operators on S_c given in Subsection 5.3.2, we observe that $H(\mathbf{u}, \mathbf{x})$ is represented locally as $\frac{\partial}{\partial \theta} \Big|_{(\mathbf{w}, \theta)}$ where $(\mathbf{u}, \mathbf{x}) = \check{\phi}(\mathbf{w}, \theta)$. ♠

Remark 9.2.8

It is possible to define a basis for $T_{(\mathbf{u}, \mathbf{x})}S_c$. Suppose that $(\check{W}, \check{U}, \check{\phi})$ is a local chart of S_c as defined above with $(\mathbf{u}, \mathbf{x}) \in S_c$. Suppose that $(\mathbf{u}, \mathbf{x}) = \check{\phi}(\mathbf{w}, \psi)$ for $(\mathbf{w}, \psi) \in \check{W} = W \times I$. Let $\rho_i :]-a, a[\rightarrow W$ for $i = 1, 2$ be two curves such that $\rho_1(0) = \rho_2(0) = \mathbf{w}$, and $\rho'_1(0), \rho'_2(0) \in \mathbb{R}^2$ are linearly independent. Consider the curves $\sigma_i :]-a, a[\rightarrow W \times I$ for $i = 1, 2$ defined by

$\sigma_i(t) = (\rho_i(t), \psi)$ for $-a < t < a$. Then $\tilde{\sigma}_i = \tilde{\phi} \circ \sigma_i :]-a, a[\rightarrow S_c$ for $i = 1, 2$ are two curves in S_c such that $\tilde{\sigma}_i(0) = \tilde{\phi}(\rho_i(0), \psi) = \tilde{\phi}(\mathbf{w}, \psi) = (\mathbf{u}, \mathbf{x})$ and

$$\begin{aligned} (\tilde{\sigma}_i)_*(0, 1) &= \left(\tilde{\sigma}_i(0), \frac{d}{dt} \tilde{\sigma}_i(t) \Big|_{t=0} \right) = \left((\mathbf{u}, \mathbf{x}), \left(\frac{d}{dt} \phi(\rho_i(t)) \Big|_{t=0}, \mathbf{q}_i \right) \right) \\ &= \left((\mathbf{u}, \mathbf{x}), (D\phi(\mathbf{w})\rho'_i(0), \mathbf{q}_i) \right) \end{aligned}$$

for some $\mathbf{q}_i \in \mathbb{R}^n$ ¹. Hence $(\tilde{\sigma}_1)_*(0, 1), (\tilde{\sigma}_2)_*(0, 1) \in T_{(\mathbf{u}, \mathbf{x})}S_c$ are linearly independent because $D\phi(\mathbf{w})\rho'_1(0), D\phi(\mathbf{w})\rho'_2(0) \in \mathbb{R}^n$ are linearly independent since $D\phi(\mathbf{w})$ is of rank 2.

Let $\sigma_3 :]-a, a[\rightarrow W \times I$ be the curve defined by $\sigma_3(t) = (\mathbf{w}, \psi + t)$ for $-a < t < a$. We assume that a is small enough to have $[\psi - a, \psi + a] \subset I$. Then $\tilde{\sigma}_3 = \tilde{\phi} \circ \sigma_3 :]-a, a[\rightarrow S_c$ is a curve in S_c such that $\tilde{\sigma}_3(0) = \tilde{\phi}(\mathbf{w}, \psi) = (\mathbf{u}, \mathbf{x})$ and

$$\begin{aligned} (\tilde{\sigma}_3)_*(0, 1) &= \left(\tilde{\sigma}_3(0), \frac{d}{dt} \tilde{\sigma}_3(t) \Big|_{t=0} \right) = \left((\mathbf{u}, \mathbf{x}), \left(\mathbf{0}, \frac{d}{dt} (\psi + t) \cdot E(\mathbf{w}) \Big|_{t=0} \right) \right) \\ &= \left((\mathbf{u}, \mathbf{x}), \left(\mathbf{0}, \frac{d}{d\nu} (\nu \cdot E(\mathbf{w})) \Big|_{\nu=\psi} \right) \right) = H(\mathbf{u}, \mathbf{x}) . \end{aligned}$$

We have that $\{(\tilde{\sigma}_1)_*(0, 1), (\tilde{\sigma}_2)_*(0, 1), (\tilde{\sigma}_3)_*(0, 1)\}$ is a basis of the 3-dimensional space $T_{(\mathbf{u}, \mathbf{x})}S_c$.

It is interesting to note that $(\pi_S)_* : T_{(\mathbf{u}, \mathbf{x})}S_c \rightarrow T_{\mathbf{u}}S$ is onto. This follows from $(\pi_S)_*((\tilde{\sigma}_i)_*(0, 1)) = (\mathbf{u}, D\phi(\mathbf{w})\rho'_i(0))$ for $i = 1, 2$ because $\rho'_1(0)$ and $\rho'_2(0)$ are linearly independent. It is also interesting to note that the vector $\frac{d}{d\nu} (\nu \cdot E(\mathbf{w})) \Big|_{\nu=\psi} \in \mathbb{R}^n$ is of norm 1 because it is the derivative of the parametric representation $\theta \mapsto \theta \cdot E(\mathbf{w})$ of a unit circle in $\mathbb{R}^n$ centred at the origin where θ represents the arc length. ♠

Remark 9.2.9

For the readers who have difficulties to visualize the tangent space $T_{(\mathbf{u}, \mathbf{x})}S_c$, it may help to use the presentation of tangent spaces given in Subsection 5.3.1. Thus $T_{(\mathbf{u}, \mathbf{x})}S_c$ is the set of all equivalent class of differentiable curves $\sigma :]-a, a[\rightarrow S_c \subset TS$ at $(\mathbf{u}, \mathbf{x}) \in S_c$. It is now clear that $T_{(\mathbf{u}, \mathbf{x})}S_c$ is a subspace of $T_{(\mathbf{u}, \mathbf{x})}(TS)$ which is the set of all equivalent class of differentiable curves $\sigma :]-a, a[\rightarrow TS$ at $(\mathbf{u}, \mathbf{x}) \in TS$.

The definition of tangent spaces given in Subsection 5.3.1 is really nice to visualize tangent spaces. However, it is not really convenient to prove smoothness of tangent spaces. ♠

Proposition 9.2.10

Let S be a 2-dimensional Riemannian manifold.

1. $\eta_*(H(\mathbf{u}, \mathbf{x})) = H(\eta(\mathbf{u}, \mathbf{x}))$ for all $(\mathbf{u}, \mathbf{x}) \in S_c$ and $\eta \in S^1$.
2. η_* is invertible and $(\eta_*)^{-1} = (\eta^{-1})_*$ for all $\eta \in S^1$.

¹We do not need to know the complete expression for the $\mathbf{q}_i$.

Proof.

1) Let (W, U, ϕ) , $(\tilde{W}, \tilde{U}, \tilde{\phi})$ and $(\check{W}, \check{U}, \check{\phi})$ be the previously defined local charts of S , S_c and TS_c respectively.

Suppose that $\eta = e^{i\mu}$ for $\mu \in \mathbb{R}$. Since $\eta : S_c \rightarrow S_c$, we get that $\eta_* : TS_c \rightarrow TS_c$.

Given $(\mathbf{u}, \mathbf{x}) \in \tilde{U} \subset S_c$, we have that $(\mathbf{u}, \mathbf{x}) = \tilde{\phi}(\mathbf{w}, \theta)$ for some $(\mathbf{w}, \theta) \in \tilde{W}$; namely, $(\mathbf{u}, \mathbf{x}) = (\phi(\mathbf{w}), \theta.E(\mathbf{w}))$.

Since $\eta \circ \tilde{\phi} = \tilde{\phi} \circ \eta$ where the first η acts on $\tilde{U}$ and the second on $\tilde{W}$, we get that $\eta_* \circ \tilde{\phi}_* = \tilde{\phi}_* \circ \eta_*$ where the first η_* acts on $T\tilde{U}$ and the second one acts on $T\tilde{W}$.

Since the action of η on $\tilde{W}$ gives $\eta(\mathbf{w}, \theta) = (\mathbf{w}, \theta + \mu)$ ², we have

$$\eta_*((\mathbf{w}, \theta), (\mathbf{r}, s)) = \left(\eta(\mathbf{w}, \theta), D_{\mathbf{w}, \theta} \eta(\mathbf{w}, \theta) \begin{pmatrix} \mathbf{r} \\ s \end{pmatrix} \right) = ((\mathbf{w}, \theta + \mu), (\mathbf{r}, s))$$

for all $((\mathbf{w}, \theta), (\mathbf{r}, s)) \in T_{(\mathbf{w}, \theta)}\tilde{W}$.

It then follows from (9.2.5) that

$$\begin{aligned} \eta_*(H(\mathbf{u}, \mathbf{x})) &= \eta_*(\tilde{\phi}_*((\mathbf{w}, \theta), (\mathbf{0}, 1))) = \tilde{\phi}_*(\eta_*((\mathbf{w}, \theta), (\mathbf{0}, 1))) \\ &= \tilde{\phi}_*((\mathbf{w}, \mu + \theta), (\mathbf{0}, 1)) = \tilde{\phi}_*(\eta(\mathbf{w}, \theta), (\mathbf{0}, 1)) = H(\eta(\mathbf{u}, \mathbf{x})) \end{aligned}$$

because $\tilde{\phi}(\eta(\mathbf{w}, \theta)) = \eta(\tilde{\phi}(\mathbf{w}, \theta)) = \eta(\mathbf{u}, \mathbf{x})$.

2) Since $\eta : S_c \rightarrow S_c$ and $\eta^{-1} : S_c \rightarrow S_c$, we have that $\eta_* : TS_c \rightarrow TS_c$ and $(\eta^{-1})_* : TS_c \rightarrow TS_c$. Moreover, since $\eta \circ \eta^{-1} = \text{Id}_{S_c}$ and $\eta^{-1} \circ \eta = \text{Id}_{S_c}$, we have that $\eta_* \circ (\eta^{-1})_* = (\eta \circ \eta^{-1})_* = (\text{Id}_{S_c})_* = \text{Id}_{TS_c}$ and $(\eta^{-1})_* \circ \eta_* = (\eta^{-1} \circ \eta)_* = (\text{Id}_{S_c})_* = \text{Id}_{TS_c}$. Thus η_* is invertible and $(\eta_*)^{-1} = (\eta^{-1})_*$. ■

Remark 9.2.11

Suppose that (W, U, ϕ) , $(\tilde{W}, \tilde{U}, \tilde{\phi})$ and $(\check{W}, \check{U}, \check{\phi})$ are the previously defined local charts of S , S_c and TS_c respectively. Moreover, suppose that $\eta = e^{i\mu}$ for $\mu \in \mathbb{R}$.

If $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) = \check{\phi}((\mathbf{w}, \theta), (\mathbf{r}, t)) = \tilde{\phi}_*((\mathbf{w}, \theta), (\mathbf{r}, t))$ for $((\mathbf{w}, \theta), (\mathbf{r}, t)) \in T_{(\mathbf{w}, \theta)}\tilde{W}$. then

$$\begin{aligned} \eta_*((\mathbf{w}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) &= \eta_*(\check{\phi}((\mathbf{w}, \theta), (\mathbf{r}, t))) = \eta_*(\tilde{\phi}_*((\mathbf{w}, \theta), (\mathbf{r}, t)))) \\ &= \tilde{\phi}_*(\eta_*((\mathbf{w}, \theta), (\mathbf{r}, t))) = \tilde{\phi}_*(\eta(\mathbf{w}, \theta), (\mathbf{r}, t)) \\ &= \left(\tilde{\phi}(\eta(\mathbf{w}, \theta)), D_{\mathbf{v}, \nu} \tilde{\phi}(\mathbf{v}, \nu)|_{(\mathbf{v}, \nu)=\eta(\mathbf{w}, \theta)} \begin{pmatrix} \mathbf{r} \\ t \end{pmatrix} \right) \\ &= \left(\tilde{\phi}(\eta(\mathbf{w}, \theta)), \left(D\phi(\mathbf{w}) \mathbf{r}, D_{\mathbf{w}}((\mu + \theta).E(\mathbf{w})) \mathbf{r} + t(\pi/2 + \mu + \theta).E(\mathbf{w})) \right) \right) \\ &= (\eta(\mathbf{u}, \mathbf{x}), (\mathbf{p}, \tilde{\mathbf{q}})) \end{aligned}$$

for some $\tilde{\mathbf{q}}$. This simple remark will be useful later. ♠

²We may assume that $0 \leq \mu < 2\pi$ and that the open interval I in the definition of $\tilde{W}$ is such that $\theta, \mu \in I$.

Proposition 9.2.12

If S_1 and S_2 are two 2-dimensional Riemannian manifolds and $f : S_1 \rightarrow S_2$ is an isometry, then $f_* : (S_1)_c \rightarrow (S_2)_c$.

Proof.

Since f is an isometry, we have that $\|f_*(\mathbf{u}, \mathbf{x})\|_{f(\mathbf{u})} = \|(\mathbf{u}, \mathbf{x})\|_{\mathbf{u}}$ for all $(\mathbf{u}, \mathbf{x}) \in T_{\mathbf{u}}S_1$. In particular, if $(\mathbf{u}, \mathbf{x}) \in (S_1)_c(\mathbf{u})$, then $f_*(\mathbf{u}, \mathbf{x}) \in (S_2)_c(f(\mathbf{u}))$. Thus $f_* : (S_1)_c(\mathbf{u}) \rightarrow (S_2)_c(f(\mathbf{u}))$ for all $\mathbf{u} \in S_1$. ■

Proposition 9.2.13

Let S_1 and S_2 be two 2-dimensional oriented and Riemannian manifold, and $f : S_1 \rightarrow S_2$ be an orientation preserving isometry. Then $\tilde{f}_* \circ H = H \circ \tilde{f}$ on $(S_1)_c$ where $\tilde{f} = f_*$.

Proof.

Let (W_i, U_i, ϕ_i) , $(\tilde{W}_i, \tilde{U}_i, \tilde{\phi}_i)$ and $(\check{W}_i, \check{U}_i, \check{\phi}_i)$ be the previously defined local charts of S_i , $(S_i)_c$ and $T(S_i)_c$ respectively for $i = 1, 2$. We may assume that $f(U_1) \subset U_2$. Suppose that $g : W_1 \rightarrow W_2$ is the local representation of $f : U_1 \rightarrow U_2$; namely, $f \circ \phi_1 = \phi_2 \circ g$ on W_1 .

Suppose that $(\mathbf{u}, \mathbf{x}) = \tilde{\phi}_1(\mathbf{w}, \theta)$ for some $\theta \in \mathbb{R}$. Since $\tilde{f} \circ \tilde{\phi}_1 = \tilde{\phi}_2 \circ \tilde{g}$ where $\tilde{g} = g_* : \tilde{W}_1 \rightarrow \tilde{W}_2$ is the local representation of $\tilde{f}$, we get that $\tilde{f}_* \circ (\tilde{\phi}_1)_* = (\tilde{\phi}_2)_* \circ \tilde{g}_*$ on $T\tilde{W}_1$.

Since f is an orientation preserving isometry, we have that $\tilde{f} = f_*$ preserves the angle and direction of rotation between vectors. To be more precise, let $\nu \in S^1$ be the angle such that $\tilde{f}(\mathbf{u}, \tilde{E}(\mathbf{u})) = \nu(f(\mathbf{u}), \tilde{E}(f(\mathbf{u})))$ where $\tilde{E}(\mathbf{u}) = E(\phi^{-1}(\mathbf{u}))$ for all $\mathbf{u} \in U_1$. Then, we get from the third item of Proposition 9.2.2 that

$$\tilde{f}(\mu(\mathbf{u}, \tilde{E}(\mathbf{u}))) = \mu(\tilde{f}(\mathbf{u}, \tilde{E}(\mathbf{u}))) = \mu(\nu(f(\mathbf{u}), \tilde{E}(f(\mathbf{u}))))$$

for all $\mu \in S^1$. If $\nu = e^{i\alpha}$ and $\mu = e^{i\beta}$, then the previous statement yields

$$\begin{aligned} \tilde{f}(\tilde{\phi}_1(\mathbf{w}, \beta)) &= \tilde{f}(\mathbf{u}, \beta \cdot \tilde{E}(\mathbf{u})) = (f(\mathbf{u}), (\alpha + \beta) \cdot \tilde{E}(f(\mathbf{u}))) \\ &= (f(\phi_1(\mathbf{w})), (\alpha + \beta) \cdot \tilde{E}(f(\phi_1(\mathbf{w})))) = \tilde{\phi}_2(g(\mathbf{w}), \alpha + \beta) \end{aligned}$$

for all $\beta \in \mathbb{R}$. Thus $\tilde{g}(\mathbf{w}, \beta) = (g(\mathbf{w}), \alpha + \beta)$ for all $(\mathbf{w}, \beta) \in \tilde{W}$. Hence $\frac{\partial \tilde{g}}{\partial \beta}(\mathbf{w}, \beta) = (\mathbf{0}, 1)$ for all $\beta \in \mathbb{R}$. It then follows from (9.2.5) that

$$\begin{aligned} \tilde{f}_*(H(\mathbf{u}, \mathbf{x})) &= \tilde{f}_*((\tilde{\phi}_1)_*((\mathbf{w}, \theta), (\mathbf{0}, 1))) = (\tilde{\phi}_2)_*(\tilde{g}_*((\mathbf{w}, \theta), (\mathbf{0}, 1))) \\ &= (\tilde{\phi}_2)_*\left(\tilde{g}(\mathbf{w}, \theta), D\tilde{g}(\mathbf{w}, \theta) \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix}\right) = (\tilde{\phi}_2)_*\left(\tilde{g}(\mathbf{w}, \theta), \frac{\partial \tilde{g}}{\partial \theta}(\mathbf{w}, \theta)\right) \\ &= (\tilde{\phi}_2)_*(\tilde{g}(\mathbf{w}, \theta), (\mathbf{0}, 1)) = H(\tilde{f}(\mathbf{u}, \mathbf{x})), \end{aligned}$$

because $\tilde{f}(\mathbf{u}, \mathbf{x}) = \tilde{\phi}_2(\tilde{g}(\mathbf{w}, \theta))$. ■

Definition 9.3.1

A curve $\tilde{\sigma}_{(\mathbf{u}, \mathbf{x})} : [a, b] \rightarrow S_c$ that satisfies (2) above is called a **lift** of σ .

Definition 9.3.2

The **parallel translation** of the unit vector $(\mathbf{u}, \mathbf{x}) \in S_c(\mathbf{u})$ along σ to $\sigma(t)$ with $a \leq t \leq b$ is defined as the vector $\tilde{\sigma}_{(\mathbf{u}, \mathbf{x})}(t) \in S_c(\sigma(t))$.

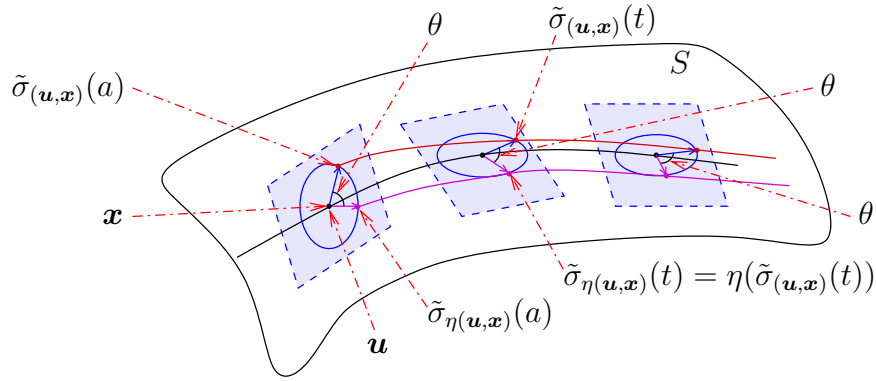


Figure 9.2: This is a schematic representation of parallel translation that aims to illustrate the effect of the translation on the angle between two vectors in S_c . In the figure $\eta = e^{i\theta} \in S^1$. As in the previous figure, elements of the form $\tilde{\sigma}_{(\mathbf{u}, \mathbf{x})}(t) \in S_c(\sigma(t))$ are represented by points on the unit circle $C \subset T_{\sigma(t)}S$ (in blue) centred at the origin which is associated to $\sigma(t)$ in the figure above.

However, the three conditions listed above are not enough to determine uniquely $\tilde{\sigma}_{(\mathbf{u}, \mathbf{x})}$. To justify this statement, we need a little lemma.

Lemma 9.3.3

Given $(\mathbf{u}, \mathbf{x}) \in S_c(\mathbf{u})$, the space $P_{(\mathbf{u}, \mathbf{x})} = ((\pi_S)_*)^{-1}(\{(\mathbf{u}, \mathbf{0})\}) \cap T_{(\mathbf{u}, \mathbf{x})}S_c$ is a 1 dimensional subspace of $T_{(\mathbf{u}, \mathbf{x})}S_c$.

Proof.

We assume that $(\mathbf{u}, \mathbf{x}) = \eta(\mathbf{u}, \tilde{E}(\mathbf{u})) = (\mathbf{u}, \theta, \tilde{E}(\mathbf{u}))$ for some $\eta = e^{i\theta} \in S^1$.

Consider the vector field H on S_c defined in (9.2.6). We have that $H(\mathbf{u}, \mathbf{x}) \in T_{\mathbf{u}, \mathbf{x}}S_c$, $H(\mathbf{u}, \mathbf{x}) = ((\mathbf{u}, \mathbf{x}), (\mathbf{0}, \mathbf{q}))$ with $\mathbf{q} \neq \mathbf{0}$, and $(\pi_S)_*(H(\mathbf{u}, \mathbf{x})) = (\mathbf{u}, \mathbf{0})$. Thus $P_{(\mathbf{u}, \mathbf{x})}$ is of dimension at least 1. The space $P_{(\mathbf{u}, \mathbf{x})}$ is in fact of dimension 1 because

$$(\pi_S)_*((\tilde{\sigma}_i)_*(0, 1)) = (\mathbf{u}, D\phi(\mathbf{w})\rho'_i(0)) \neq (\mathbf{u}, \mathbf{0})$$

for $i = 1, 2$, where $\{(\tilde{\sigma}_i)_*(0, 1)\}_{1 \leq i \leq 3}$ is the basis of $T_{(\mathbf{u}, \mathbf{x})}S_c$ introduced in Remark 9.2.8. The reader may recall that $(\tilde{\sigma}_3)_*(0, 1) = H(\tilde{\sigma}_3(0)) = H(\mathbf{u}, \mathbf{x})$.

Therefore $P_{(\mathbf{u}, \mathbf{x})}$ is the 1-dimensional subspace of $T_{(\mathbf{u}, \mathbf{x})}S_c$ given by

$$P_{(\mathbf{u}, \mathbf{x})} = \{tH(\mathbf{u}, \mathbf{x}) : t \in \mathbb{R}\} = \left\{ \left((\mathbf{u}, \mathbf{x}), t \left(\mathbf{0}, \frac{d}{d\nu}(\nu \cdot \tilde{E}(\mathbf{u})) \Big|_{\nu=\theta} \right) \right) : t \in \mathbb{R} \right\} \quad (9.3.1)$$

(Figure 9.3). ■

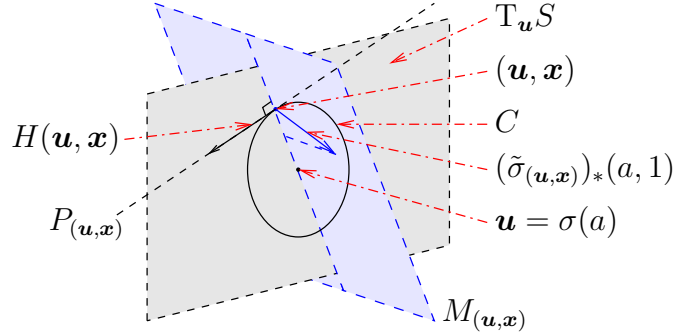


Figure 9.3: This is a schematic representation of a connection. We have projected the subspaces $T_{\mathbf{u}}S$, $M_{(\mathbf{u}, \mathbf{x})}$ and $P_{(\mathbf{u}, \mathbf{x})}$ into $\mathbb{R}^n$. We also have projected $H(\mathbf{u}, \mathbf{x})$ into $\mathbb{R}^n$. Normally, they all live in different spaces. We have added the potential direction $(\tilde{\sigma}_{(\mathbf{u}, \mathbf{x})})_*(a, 1) \in M_{(\mathbf{u}, \mathbf{x})}$ of the lift $\tilde{\sigma}_{(\mathbf{u}, \mathbf{x})}$ of σ at $t = a$. As usual, $(\mathbf{u}, \mathbf{x})$ is represented by the point $\mathbf{u} + \mathbf{x}$ on the unit circle C (in black) in the tangent plane $T_{\mathbf{u}}S$ (in grey) centred at the origin which in the figure above is associated to $\mathbf{u}$. Be aware that $P_{(\mathbf{u}, \mathbf{x})}$ does not generally represent a normal direction to $M_{(\mathbf{u}, \mathbf{x})}$ as it seems to be in the figure. We only can say that $P_{(\mathbf{u}, \mathbf{x})}$ is perpendicular to $M_{(\mathbf{u}, \mathbf{x})} \cap T_{(\mathbf{u}, \mathbf{x})}S$.

Hence, if $(\pi_S)_*((\tilde{\sigma}_{(\mathbf{u}, \mathbf{x})})_*(t, 1)) = \sigma_*(t, 1)$, then $(\pi_S)_*((\tilde{\sigma}_{(\mathbf{u}, \mathbf{x})}(t), (\mathbf{p}, \mathbf{q})) + (\tilde{\sigma}_{(\mathbf{u}, \mathbf{x})})_*(t, 1)) = \sigma_*(t, 1)$ for all $(\tilde{\sigma}_{(\mathbf{u}, \mathbf{x})}(t), (\mathbf{p}, \mathbf{q})) \in P_{\tilde{\sigma}_{(\mathbf{u}, \mathbf{x})}(t)}$ because $(\pi_S)_*(\tilde{\sigma}_{(\mathbf{u}, \mathbf{x})}(t), (\mathbf{p}, \mathbf{q})) = (\sigma(t), 0)$. Therefore, there is a 1-dimensional subspace of the 3-dimensional vector space $T_{\tilde{\sigma}_{(\mathbf{u}, \mathbf{x})}(t)}S_c$ such that the image of this subspace by the projection $(\pi_S)_* : T_{\tilde{\sigma}_{(\mathbf{u}, \mathbf{x})}(t)}S_c \rightarrow T_{\sigma(t)}S$ is $\{\sigma_*(t, 1)\}$. We need to add conditions on the lift associated to a parallel translations to uniquely determine the value of $(\tilde{\sigma}_{(\mathbf{u}, \mathbf{x})})_*(t, 1)$.

Definition 9.3.4

Let S be a 2-dimensional Riemannian manifold. A **connection** on S_c is a collection $\mathcal{M}_c = \{M_{(\mathbf{u}, \mathbf{x})}\}_{(\mathbf{u}, \mathbf{x}) \in S_c}$ such that

1. $M_{(\mathbf{u}, \mathbf{x})}$ is a 2-dimensional subspace of $T_{(\mathbf{u}, \mathbf{x})}S_c$ for all $(\mathbf{u}, \mathbf{x}) \in S_c$,
2. $T_{(\mathbf{u}, \mathbf{x})}S_c = M_{(\mathbf{u}, \mathbf{x})} \oplus P_{(\mathbf{u}, \mathbf{x})}$ for all $(\mathbf{u}, \mathbf{x}) \in S_c$ (Figure 9.3),
3. $\eta_*(M_{(\mathbf{u}, \mathbf{x})}) = M_{\eta(\mathbf{u}, \mathbf{x})}$ for all $(\mathbf{u}, \mathbf{x}) \in S_c$ and $\eta \in S^1$ (Figure 9.4), and
4. S_c can be covered by open sets $\tilde{U} \subset S_c$ such that there exist two smooth vector

fields $F, G : \tilde{U} \rightarrow TS_c$ with the property that $\{F(\mathbf{u}, \mathbf{x}), G(\mathbf{u}, \mathbf{x})\}$ is a basis of $M_{(\mathbf{u}, \mathbf{x})}$ for all $(\mathbf{u}, \mathbf{x}) \in \tilde{U}$.

In the previous definition, we emphasize that $\eta : S_c \rightarrow S_c$ for all fixed $\eta \in S^1$. Thus $\eta_* : TS_c \rightarrow TS_c$.

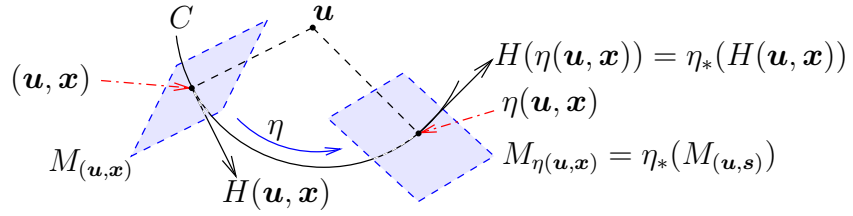


Figure 9.4: Given $(\mathbf{u}, \mathbf{x}) \in S_c$, we give a schematic representation of the rotation of $H(\mathbf{u}, \mathbf{x})$ and $M_{(\mathbf{u}, \mathbf{x})}$ by an angle $\eta \in S^1$.

Example 9.3.5

It is not hard to define a connection locally on S_c .

Suppose that (W, U, ϕ) is local chart of S . Consider the map $G : U \times S^1 \rightarrow S_c(U)$ defined by $G(\mathbf{u}, \eta) = \eta(\mathbf{u}, \tilde{E}(\mathbf{u})) = (\mathbf{u}, \theta \cdot \tilde{E}(\mathbf{u}))$ for $(\mathbf{u}, \theta) \in U \times \mathbb{R}$ and $\eta = e^{i\theta}$, where $\tilde{E}(\mathbf{u}) = E(\phi^{-1}(\mathbf{u}))$ for all $\mathbf{u} \in U$. This is the famous map that we have used to prove that $S_c(U) \cong U \times S^1$ in Remark 9.2.6.

For η fixed, we have that $(G_\eta)_* : T_{\mathbf{u}}S \rightarrow T_{\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))}S_c$. Let $\check{M}_{\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))} = (G_\eta)_*(T_{\mathbf{u}}S)$. Since $(\pi_S)_* \circ (G_\eta)_* = (\pi_S \circ G_\eta)_* = (\text{Id}_S)_*$, we get that $(\pi_S)_*(\check{M}_{\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))}) = T_{\mathbf{u}}S$ for $\mathbf{u} \in U$. Moreover $(\pi_S)_*(H(\eta(\mathbf{u}, \tilde{E}(\mathbf{u})))) = (\mathbf{u}, \mathbf{0})$ for all $\mathbf{u} \in U$. Thus $(\pi_S)_* : \check{M}_{\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))} \rightarrow T_{\mathbf{u}}S$ is a (linear) isomorphism and so $\check{M}_{\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))}$ is of dimension 2. It follows that $T_{\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))} = \check{M}_{\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))} \oplus P_{\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))}$.

We have that $\check{M}_{\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))} = \eta_*(\check{M}_{(\mathbf{u}, \tilde{E}(\mathbf{u}))})$ for $\eta \in S^1$ because $G_\eta = \eta \circ G_1$. We may define $\check{\mathcal{M}}_c = \{\check{M}_{(\mathbf{u}, \mathbf{x})}\}_{(\mathbf{u}, \mathbf{x}) \in S_c(U)}$ because every $(\mathbf{u}, \mathbf{x}) \in S_c(U)$ is of the form $\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))$ for some $\eta \in S^1$.

The existence of two smooth vector fields $P, Q : S_c(U) \rightarrow TS_c$ with the property that $\{P(\mathbf{u}, \mathbf{x}), Q(\mathbf{u}, \mathbf{x})\}$ is a basis of $\check{M}_{(\mathbf{u}, \mathbf{x})}$ for all $(\mathbf{u}, \mathbf{x}) \in S_c(U)$ follows from the isomorphism between $\check{M}_{(\mathbf{u}, \mathbf{x})}$ and $T_{\mathbf{u}}S$ for $(\mathbf{u}, \mathbf{x}) \in S_c(U)$. Two obvious vector fields $\tilde{P}, \tilde{Q} : U \rightarrow TS$ with the property that $\{\tilde{P}(\mathbf{u}), \tilde{Q}(\mathbf{u})\}$ is a basis of $T_{\mathbf{u}}S$ for all $\mathbf{u} \in U$ are given by $\tilde{P}(\mathbf{u}) = (\mathbf{u}, \tilde{E}(\mathbf{u}))$ and $\tilde{Q}(\mathbf{u}) = (\mathbf{u}, (\pi/2) \cdot \tilde{E}(\mathbf{u}))$ for $\mathbf{u} \in U$. We may use $P(\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))) = (G_\eta)_*(\tilde{P}(\mathbf{u}))$ and $Q(\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))) = (G_\eta)_*(\tilde{Q}(\mathbf{u}))$ for $\mathbf{u} \in U$ and $\eta \in S^1$. ♣

The following concept is fundamental in the proof of existence of connections.

Proposition 9.3.6

Let S be a 2-dimensional Riemannian manifold and $\mathcal{M}_c = \{M_{(\mathbf{u}, \mathbf{x})}\}_{(\mathbf{u}, \mathbf{x}) \in S_c}$ be a connection on S_c . There exists a smooth differential 1-form ρ on S_c such that

$$\pi_{P_{(\mathbf{u}, \mathbf{x})}}((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) = \rho(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q}))H(\mathbf{u}, \mathbf{x}) \quad (9.3.2)$$

for all $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \in TS_c$, where $\pi_{P_{(\mathbf{u}, \mathbf{x})}}$ is the projection from $M_{(\mathbf{u}, \mathbf{x})} \oplus P_{(\mathbf{u}, \mathbf{x})}$ onto $P_{(\mathbf{u}, \mathbf{x})}$ with $P_{(\mathbf{u}, \mathbf{x})}$ defined in (9.3.1). The smooth differential 1-form ρ is called the **connection 1-form** associated to the connection $\mathcal{M}_c$.

Proof.

Every element $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \in T_{(\mathbf{u}, \mathbf{x})}S_c$ can be expressed in a unique way as

$$((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) = \lambda((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q}))H(\mathbf{u}, \mathbf{x}) + ((\mathbf{u}, \mathbf{x}), (\tilde{\mathbf{p}}, \tilde{\mathbf{q}}))$$

for some $((\mathbf{u}, \mathbf{x}), (\tilde{\mathbf{p}}, \tilde{\mathbf{q}})) \in M_{(\mathbf{u}, \mathbf{x})}$ and $\lambda((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \in \mathbb{R}$.

We set $\rho(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) = \lambda((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q}))$ for all $(\mathbf{u}, \mathbf{x}) \in S_c$ and $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \in T_{(\mathbf{u}, \mathbf{x})}S_c$.

To prove that ρ is a smooth differential 1-form, we consider an open set $\tilde{U} \subset S_c$ such that there exist two smooth vector fields $F, G : \tilde{U} \rightarrow TS_c$ with the property that $\{F(\mathbf{v}, \mathbf{x}), G(\mathbf{u}, \mathbf{x})\}$ is a basis of $M_{(\mathbf{u}, \mathbf{x})}$ for all $(\mathbf{u}, \mathbf{x}) \in \tilde{U}$. Since $T_{(\mathbf{u}, \mathbf{x})}S_c = M_{(\mathbf{u}, \mathbf{x})} \oplus P_{(\mathbf{u}, \mathbf{x})}$ for all $(\mathbf{u}, \mathbf{x}) \in \tilde{U}$, we have that $\{F(\mathbf{v}, \mathbf{x}), G(\mathbf{u}, \mathbf{x}), H(\mathbf{u}, \mathbf{x})\}$ is a basis of $T_{(\mathbf{u}, \mathbf{x})}S_c$ for all $(\mathbf{u}, \mathbf{x}) \in \tilde{U}$. Let $\{\tilde{F}(\mathbf{v}, \mathbf{x}), \tilde{G}(\mathbf{u}, \mathbf{x}), \tilde{H}(\mathbf{u}, \mathbf{x})\}$ be the dual basis associated to $\{F(\mathbf{v}, \mathbf{x}), G(\mathbf{u}, \mathbf{x}), H(\mathbf{u}, \mathbf{x})\}$ for all $(\mathbf{u}, \mathbf{x}) \in \tilde{U}$. We have by construction that $\tilde{F}$, $\tilde{G}$ and $\tilde{H}$ are smooth differential 1-forms on $\tilde{U}$.

Since $\tilde{H}(\mathbf{u}, \mathbf{x})$ is null on $M_{(\mathbf{u}, \mathbf{x})}$ and $\tilde{H}((\mathbf{u}, \mathbf{x}), (H(\mathbf{u}, \mathbf{x}))) = 1$, we get

$$\tilde{H}(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) = \rho(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q}))$$

for all $(\mathbf{u}, \mathbf{x}) \in \tilde{U}$ and $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \in T_{(\mathbf{u}, \mathbf{x})}\tilde{U}$. Thus $\rho = \tilde{H}$ on $\tilde{U}$. Since $\tilde{U}$ is arbitrary, this proves that ρ is a smooth differential 1-form on S_c . ■

Example 9.3.7 (Example 9.3.5 (Continued))

According to the previous proposition, the connection 1-form $\check{\rho}$ on $S_c(U)$ satisfies (9.3.2) with ρ replaced by $\check{\rho}$ and $(\mathbf{u}, \mathbf{x}) = \eta.(\mathbf{u}, \tilde{E}(\mathbf{u}))$ for $\eta \in S^1$, where $\tilde{E}(\mathbf{u}) = E(\phi^{-1}(\mathbf{u}))$ for all $\mathbf{u} \in U$.

We consider $S^1 \cong \{e^{i\theta} \in \mathbb{C} : \theta \in \mathbb{R}\}$. We have that $T_\eta S^1 = \{(\eta, t i \eta) : t \in \mathbb{R}\}$ for all $\eta = e^{i\theta} \in S^1$. We define a differential 1-form ω on S^1 by $\omega(\eta)(\eta, t i \eta) = t$ for all $\eta \in S^1$ and $t \in \mathbb{R}$.

We now prove that the connection 1-form $\check{\rho}$ associated to $\check{\mathcal{M}}_c$ is given by $\check{\rho} = (G^{-1})^*(\pi_2^*(\omega))$ where $\pi_2 : U \times S^1 \rightarrow S^1$ is the projection defined by $\pi_2(\mathbf{u}, \eta) = \eta$ for all $(\mathbf{u}, \eta) \in U \times S^1$, and G is the function defined in Remark 9.2.6. Since $G : U \times S^1 \rightarrow S_c(U)$ is an isomorphism, we have that $G_* : T(U \times S^1) \rightarrow TS_c(U)$ is an isomorphism. Therefore, given

$(\eta(\mathbf{u}, \tilde{E}(\mathbf{u})), (\mathbf{p}, \mathbf{q})) \in T_{\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))} S_c$ with $\eta = e^{i\theta}$, there exists $((\mathbf{u}, \eta), (\mathbf{r}, t i\eta)) \in T_{(\mathbf{u}, \eta)}(U \times S^1)$ such that

$$\begin{aligned} (\eta(\mathbf{u}, \tilde{E}(\mathbf{u})), (\mathbf{p}, \mathbf{q})) &= G_*((\mathbf{u}, \eta), (\mathbf{r}, t i\eta)) = (G_\eta)_*(\mathbf{u}, \mathbf{r}) + (G_\eta)_*(\eta, t i\eta) \\ &= (G_\eta)_*(\mathbf{u}, \mathbf{r}) + \left(\eta(\mathbf{u}, \tilde{E}(\mathbf{u})), \left(\mathbf{0}, t \frac{d}{d\theta}(\theta \cdot \tilde{E}(\mathbf{u})) \right) \right) \\ &= (G_\eta)_*(\mathbf{u}, \mathbf{r}) + t H(\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))) \end{aligned}$$

where the second to last equality comes from the fact that $\frac{\partial}{\partial \eta} G_\mathbf{u}(\eta) = \frac{\partial}{\partial \eta} G(\mathbf{u}, \eta)$ for $\mathbf{u}$ constant is defined with the help of the local coordinate $\eta = e^{i\theta}$ as

$$\frac{\partial}{\partial \eta} G(\mathbf{u}, \eta) = \frac{\partial}{\partial \theta} G(\mathbf{u}, e^{i\theta}) \left(\frac{d}{d\theta} e^{i\theta} \right)^{-1} = \frac{1}{i\eta} \frac{d}{d\theta} (\mathbf{u}, \theta \cdot \tilde{E}(\mathbf{u})) = \frac{1}{i\eta} \left(\mathbf{0}, \frac{d}{d\theta} \theta \cdot \tilde{E}(\mathbf{u}) \right).$$

Hence

$$\begin{aligned} \check{\rho}(\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))) (\eta(\mathbf{u}, \tilde{E}(\mathbf{u})), (\mathbf{p}, \mathbf{q})) &= (G^{-1})^*(\pi_2^*(\omega)) (\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))) (\eta(\mathbf{u}, \tilde{E}(\mathbf{u})), (\mathbf{p}, \mathbf{q})) \\ &= \pi_2^*(\omega)(\mathbf{u}, \eta) ((\mathbf{u}, \eta), (\mathbf{r}, t i\eta)) = \omega(\pi_2(\mathbf{u}, \eta)) ((\pi_2)_*((\mathbf{u}, \eta), (\mathbf{r}, t i\eta))) \\ &= \omega(\eta)(\eta, t i\eta) = t. \end{aligned}$$

We leave it to the reader to verify that $(\pi_2)_*((\mathbf{u}, \eta), (\mathbf{r}, t i\eta)) = (\eta, t i\eta)$ for all $((\mathbf{u}, \eta), (\mathbf{r}, t i\eta)) \in T_{(\mathbf{u}, \eta)}(U \times S^1)$. Since $(G_\eta)_*(\mathbf{u}, \mathbf{r}) \in \tilde{M}_{\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))}$, we get

$$\begin{aligned} \pi_{P_{\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))}}(\eta(\mathbf{u}, \tilde{E}(\mathbf{u})), (\mathbf{p}, \mathbf{q})) &= \pi_{P_{\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))}}((G_\eta)_*(\mathbf{u}, \mathbf{r})) + t \pi_{P_{\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))}}(H(\eta(\mathbf{u}, \tilde{E}(\mathbf{u})))) \\ &= t H(\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))) = \check{\rho}(\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))) (\eta(\mathbf{u}, \tilde{E}(\mathbf{u})), (\mathbf{p}, \mathbf{q})) H(\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))) \end{aligned}$$

for all $(\eta(\mathbf{u}, \tilde{E}(\mathbf{u})), (\mathbf{p}, \mathbf{q})) \in T S_c(U)$ as required in (9.3.2).

Note that $d\check{\rho} = 0$ because $d(G^{-1})^*(\pi_2^*(\omega)) = (G^{-1})^*(\pi_2^*(d\omega))$ according to Proposition 5.5.6 and $d\omega = 0$ since there is no non-trivial differential 2-form on a 1-dimensional manifold; in the present case, on S^1 . ♣

If ρ is the connection 1-form associated to a connection $\mathcal{M}_c = \{M_{(\mathbf{u}, \mathbf{x})}\}_{(\mathbf{u}, \mathbf{x})}$, then we get from the previous proposition that $\rho(\mathbf{u}, \mathbf{x})$ is null on $M_{(\mathbf{u}, \mathbf{x})}$ for all $(\mathbf{u}, \mathbf{x}) \in S_c$. Moreover, $\rho(\mathbf{u}, \mathbf{x})(H(\mathbf{u}, \mathbf{x})) = 1$. We can also prove the following lemma.

Proposition 9.3.8

Let S be a 2-dimensional Riemannian manifold and $\mathcal{M}_c = \{M_{(\mathbf{u}, \mathbf{x})}\}_{(\mathbf{u}, \mathbf{x}) \in S_c}$ be a connection on S_c where ρ is the connection 1-form associated to $\mathcal{M}_c$. Then $\eta^*(\rho)(\mathbf{u}, \mathbf{x}) = \rho(\mathbf{u}, \mathbf{x})$ for all $\eta \in S^1$ and $(\mathbf{u}, \mathbf{x}) \in S_c$.

Proof.

It follows from Proposition 9.2.10 that $\eta_*(H(\mathbf{u}, \mathbf{x})) = H(\eta(\mathbf{u}, \mathbf{x}))$ for all $\eta \in S^1$. Hence,

since every element of $T S_c$ can be expressed as $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) = \lambda H(\mathbf{u}, \mathbf{x}) + ((\mathbf{u}, \mathbf{x}), (\tilde{\mathbf{p}}, \tilde{\mathbf{q}}))$ with $\lambda = \rho(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q}))$ and $((\mathbf{u}, \mathbf{x}), (\tilde{\mathbf{p}}, \tilde{\mathbf{q}})) \in M_{(\mathbf{u}, \mathbf{x})}$, we get

$$\begin{aligned} \eta^*(\rho)(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) &= \eta^*(\rho)(\mathbf{u}, \mathbf{x})(\lambda H(\mathbf{u}, \mathbf{x})) + \eta^*(\rho)(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\tilde{\mathbf{p}}, \tilde{\mathbf{q}})) \\ &= \rho(\eta(\mathbf{u}, \mathbf{x}))(\lambda \eta_*(H(\mathbf{u}, \mathbf{x}))) = \rho(\eta(\mathbf{u}, \mathbf{x}))(\lambda H(\eta(\mathbf{u}, \mathbf{x}))) \\ &= \lambda \rho(\eta(\mathbf{u}, \mathbf{x}))(H(\eta(\mathbf{u}, \mathbf{x}))) = \lambda = \rho(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \end{aligned}$$

because $\eta_*(M_{(\mathbf{u}, \mathbf{x})}) = M_{\eta(\mathbf{u}, \mathbf{x})}$ and $\rho = 0$ on $M_{(\mathbf{u}, \mathbf{x})}$ for all $(\mathbf{u}, \mathbf{x}) \in S_c$ implies that

$$\eta^*(\rho)(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\tilde{\mathbf{p}}, \tilde{\mathbf{q}})) = \rho(\eta(\mathbf{u}, \mathbf{x}))(\eta_*((\mathbf{u}, \mathbf{x}), (\tilde{\mathbf{p}}, \tilde{\mathbf{q}}))) = 0$$

since $\eta_*((\mathbf{u}, \mathbf{x}), (\tilde{\mathbf{p}}, \tilde{\mathbf{q}})) \in M_{\eta(\mathbf{u}, \mathbf{x})}$. ■

There is kind of a converse to Theorem 9.3.6.

Proposition 9.3.9

Let S be a 2-dimensional Riemannian manifold. Suppose that ρ is a smooth differential 1-form on S_c such that $\rho(\mathbf{u}, \mathbf{x})(H(\mathbf{u}, \mathbf{x})) = 1$ for all $(\mathbf{u}, \mathbf{x}) \in S_c$ and $\eta^*(\rho) = \rho$ for all $\eta \in S^1$. Then $\mathcal{M}_c = \{M_{(\mathbf{u}, \mathbf{x})}\}_{(\mathbf{u}, \mathbf{x}) \in S_c}$ with $M_{(\mathbf{u}, \mathbf{x})} = (\rho(\mathbf{u}, \mathbf{x}))^{-1}(\{((\mathbf{u}, \mathbf{x}), (\mathbf{0}, \mathbf{0}))\})$ for all $(\mathbf{u}, \mathbf{x}) \in S_c$ is a connection on S_c with the property that its connection 1-form is ρ .

Proof.

For each $(\mathbf{u}, \mathbf{y}) \in S_c$, $\rho(\mathbf{u}, \mathbf{x}) : T_{(\mathbf{u}, \mathbf{x})}S_c \rightarrow \mathbb{R}$ is a linear functional. Since $T_{(\mathbf{u}, \mathbf{x})}S_c$ is a 3-dimensional vector space and $\rho(H(\mathbf{u}, \mathbf{x})) \neq 0$, we have that

$$M_{(\mathbf{u}, \mathbf{x})} = (\rho(\mathbf{u}, \mathbf{x}))^{-1}(\{((\mathbf{u}, \mathbf{x}), (\mathbf{0}, \mathbf{0}))\})$$

is a 2-dimensional subspace of $T_{(\mathbf{u}, \mathbf{x})}S_c$ satisfying (1) and (2) of Definition 9.3.4. Since $\eta^*(\rho) = \rho$, we get

$$\rho(\eta(\mathbf{u}, \mathbf{x}))(\eta_*((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q}))) = \eta^*(\rho)(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) = \rho(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) = 0$$

for all $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \in M_{(\mathbf{u}, \mathbf{x})}$. Thus $\eta_*((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \in M_{\eta(\mathbf{u}, \mathbf{x})}$ for all $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \in M_{(\mathbf{u}, \mathbf{x})}$. Hence (3) of Definition 9.3.4 is satisfied.

To prove (4) of Definition 9.3.4, we use the fact that ρ is a smooth differential 1-form on S_c . Suppose that $(\tilde{W}, \tilde{U}, \tilde{\phi})$ is a local chart of S_c as in Subsection 9.2.3. The local representation of ρ is given by $\tilde{\phi}^*(\rho)$. It suffices to construct two smooth vector fields $\tilde{F}, \tilde{G} : \tilde{W} \rightarrow \tilde{W} \times \mathbb{R}^3$ such that $\tilde{\phi}^*(\rho)(\tilde{\mathbf{w}})(\tilde{F}(\tilde{\mathbf{w}})) = \tilde{\phi}^*(\rho)(\tilde{\mathbf{w}})(\tilde{G}(\tilde{\mathbf{w}})) = 0$ and $\tilde{F}(\tilde{\mathbf{w}}), \tilde{G}(\tilde{\mathbf{w}}) \in T_{\tilde{\mathbf{w}}} \tilde{W} = \{\tilde{\mathbf{w}}\} \times \mathbb{R}^3$ are linearly independent for all $\tilde{\mathbf{w}} \in \tilde{W}$. The vector fields $F = \tilde{\phi}_* \circ \tilde{F} \circ \tilde{\phi}^{-1}$ and $G = \tilde{\phi}_* \circ \tilde{G} \circ \tilde{\phi}^{-1}$ on $\tilde{U}$ will then satisfy (4) of Definition 9.3.4. Note that the kernel of $\tilde{\phi}^*(\rho)$ is of dimension 2 for each $\tilde{\mathbf{w}} \in \tilde{W}$ because $M_{\tilde{\phi}(\tilde{\mathbf{w}})}$ is a subspace of dimension 2.

We have that $T_{(\mathbf{u}, \mathbf{x})}S_c = M_{(\mathbf{u}, \mathbf{x})} \oplus P_{(\mathbf{u}, \mathbf{x})}$ because $H(\mathbf{u}, \mathbf{x}) \notin \mathcal{N}(\rho(\mathbf{u}, \mathbf{x}))$. Thus every $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \in T_{(\mathbf{u}, \mathbf{x})}S_c$ can be expressed in a unique way as

$$((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) = \lambda((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q}))H(\mathbf{u}, \mathbf{x}) + ((\mathbf{u}, \mathbf{x}), (\tilde{\mathbf{p}}, \tilde{\mathbf{q}}))$$

with $\lambda((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \in \mathbb{R}$ and $((\mathbf{u}, \mathbf{x}), (\tilde{\mathbf{p}}, \tilde{\mathbf{q}})) \in M_{(\mathbf{u}, \mathbf{x})}$. Hence, $\rho(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) = \lambda((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q}))$ for all $(\mathbf{u}, \mathbf{x}) \in S_c$ and $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \in T_{(\mathbf{u}, \mathbf{x})}S_c$. Therefore,

$$\pi_{P_{(\mathbf{u}, \mathbf{x})}}((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) = \rho(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q}))H(\mathbf{u}, \mathbf{x})$$

for all $(\mathbf{u}, \mathbf{x}) \in S_c$ and $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \in T_{(\mathbf{u}, \mathbf{x})}S_c$. Thus ρ is a connection 1-form associated to $\mathcal{M}_c$. ■

We need a couple of preliminary results to prove that, given a connection $\mathcal{M}_c = \{M_{(\mathbf{u}, \mathbf{x})}\}_{(\mathbf{u}, \mathbf{x}) \in S_c}$ with an associate connection 1-form, there exists a unique lift $\tilde{\sigma}_{(\mathbf{u}, \mathbf{x})} : [a, b] \rightarrow S_c$ such that $(\tilde{\sigma}_{(\mathbf{u}, \mathbf{x})})_*(t, 1) \in M_{\tilde{\sigma}_{(\mathbf{u}, \mathbf{x})}(t)}$ for all $t \in [a, b]$.

Proposition 9.3.10

Let S be a 2-dimensional Riemannian manifold. Suppose that ξ is a differential 1-form on S_c such that $\xi(\mathbf{u}, \mathbf{x})(H(\mathbf{u}, \mathbf{x})) = 0$ for all $(\mathbf{u}, \mathbf{x}) \in S_c$ and $\eta^*(\xi) = \xi$ for all $\eta \in S^1$. Then $\xi = \pi_S^*(\omega)$ for some smooth differential 1-form ω on S .

Proof.

The differential 1-form ω on S is defined as follows. Choose $(\mathbf{u}, \mathbf{p}) \in S_c$ arbitrary but fixed. For $(\mathbf{u}, \mathbf{x}) \in T_{\mathbf{u}}S$, we set

$$\omega(\mathbf{u})(\mathbf{u}, \mathbf{x}) = \xi(\mathbf{u}, \mathbf{p})((\mathbf{u}, \mathbf{p}), (\mathbf{x}, \mathbf{q}))$$

for some $((\mathbf{u}, \mathbf{p}), (\mathbf{x}, \mathbf{q})) \in T_{(\mathbf{u}, \mathbf{p})}S_c$. Such an element exists because $(\pi_S)_* : T_{(\mathbf{u}, \mathbf{p})}S_c \rightarrow T_{\mathbf{u}}S$ is onto. The differential 1-form ω is well defined because the definition is independent of the choice of $((\mathbf{u}, \mathbf{p}), (\mathbf{x}, \mathbf{q})) \in T_{(\mathbf{u}, \mathbf{p})}S_c$. If $((\mathbf{u}, \mathbf{p}), (\mathbf{x}, \tilde{\mathbf{q}}))$ is another element of $T_{(\mathbf{u}, \mathbf{p})}S_c$, then $((\mathbf{u}, \mathbf{p}), (\mathbf{x}, \mathbf{q})) - ((\mathbf{u}, \mathbf{p}), (\mathbf{x}, \tilde{\mathbf{q}})) = \lambda H(\mathbf{u}, \mathbf{p})$ for some $\lambda \in \mathbb{R}$ because $((\mathbf{u}, \mathbf{p}), (\mathbf{x}, \mathbf{q})) - ((\mathbf{u}, \mathbf{p}), (\mathbf{x}, \tilde{\mathbf{q}})) \in P_{(\mathbf{u}, \mathbf{p})}$. Hence

$$\xi(\mathbf{u}, \mathbf{p})((\mathbf{u}, \mathbf{p}), (\mathbf{x}, \mathbf{q})) - \xi(\mathbf{u}, \mathbf{p})((\mathbf{u}, \mathbf{p}), (\mathbf{x}, \tilde{\mathbf{q}})) = \lambda \xi(\mathbf{u}, \mathbf{p})(H(\mathbf{u}, \mathbf{x})) = 0.$$

The differential 1-form ω is also independent of the choice of $(\mathbf{u}, \mathbf{p}) \in S_c$. Any other element of S_c of the form $(\mathbf{u}, \mathbf{z}) \in S_c$ is given by $(\mathbf{u}, \mathbf{z}) = \eta(\mathbf{u}, \mathbf{p}) = (\mathbf{u}, \theta.\mathbf{p})$ for some $\eta = e^{i\theta} \in S^1$. It follows from (2) of Proposition 9.2.10 that η_* is invertible and $(\eta_*)^{-1} = (\eta^{-1})_*$ for all $\eta \in S^1$. Hence, given $((\mathbf{u}, \mathbf{z}), (\mathbf{x}, \mathbf{q})) \in T_{(\mathbf{u}, \mathbf{z})}S_c$, there exists $((\mathbf{u}, \mathbf{p}), (\mathbf{x}, \tilde{\mathbf{q}})) \in T_{(\mathbf{u}, \mathbf{p})}S_c$ such that $\eta_*(\mathbf{u}, \mathbf{p}), (\mathbf{x}, \tilde{\mathbf{q}})) = ((\mathbf{u}, \mathbf{z}), (\mathbf{x}, \mathbf{q}))$ (see Remark 9.2.11). Hence

$$\begin{aligned} \xi(\mathbf{u}, \mathbf{z})((\mathbf{u}, \mathbf{z}), (\mathbf{x}, \mathbf{q})) &= \xi(\eta(\mathbf{u}, \mathbf{p}))(\eta_*((\mathbf{u}, \mathbf{p}), (\mathbf{x}, \tilde{\mathbf{q}}))) = \eta^*(\xi)(\mathbf{u}, \mathbf{p})((\mathbf{u}, \mathbf{p}), (\mathbf{x}, \tilde{\mathbf{q}})) \\ &= \xi(\mathbf{u}, \mathbf{p})((\mathbf{u}, \mathbf{p}), (\mathbf{x}, \tilde{\mathbf{q}})) = \omega(\mathbf{u})(\mathbf{u}, \mathbf{x}), \end{aligned}$$

where the second to last equality comes from the hypothesis that $\eta^*(\xi) = \xi$ and the last equality comes from the independence of the definition of ω on $\tilde{\mathbf{q}}$ proved in the previous paragraph.

To prove that ω is smooth, we consider a local chart (W, U, ϕ) of S and a local chart $(\tilde{W}, \tilde{U}, \tilde{\phi})$ of S_c as given in Subsection 9.2.3. The local representation of ω on W is given by

$$\begin{aligned} \tilde{\phi}^*(\xi)((\mathbf{w}, \psi), (\mathbf{y}, t)) &= \xi(\tilde{\phi}(\mathbf{w}, \psi)) \left(\tilde{\phi}(\mathbf{w}, \psi), D_{\mathbf{w}, \psi} \tilde{\phi}(\mathbf{w}, \psi) \begin{pmatrix} \mathbf{y} \\ t \end{pmatrix} \right) \\ &= \xi(\phi(\mathbf{w}), \psi.E(\mathbf{w})) \left((\phi(\mathbf{w}), \psi.E(\mathbf{w})), \left(D\phi(\mathbf{w})\mathbf{y}, D_{\mathbf{w}, \psi}(\psi.E(\mathbf{w})) \begin{pmatrix} \mathbf{y} \\ t \end{pmatrix} \right) \right) \\ &= \omega(\phi(\mathbf{u}))(\phi(\mathbf{w}), D\phi(\mathbf{w})\mathbf{y}) = \phi^*(\omega)(\mathbf{w})(\mathbf{w}, \mathbf{y}), \end{aligned}$$

where we may assume that $\psi, t \in S^1$ are constant. Thus $\phi^*(\omega)$ is a smooth differential form on W because it is the composition of smooth functions. Note that if $(\mathbf{u}, \mathbf{x}) = \phi_*(\mathbf{w}, \mathbf{y}) \in T_{\mathbf{u}}S$, the equation above states that $\omega(\mathbf{u})(\mathbf{u}, \mathbf{x}) = \xi(\mathbf{u}, \mathbf{p})((\mathbf{u}, \mathbf{p}), (\mathbf{x}, \mathbf{q}))$ with $\tilde{x} = \psi.E(\mathbf{u})$ and $\mathbf{q} = D_{\mathbf{w}, \psi}(\psi.E(\mathbf{w})) \begin{pmatrix} \mathbf{y} \\ t \end{pmatrix}$.

By construction, we have

$$\begin{aligned} \pi_S^*(\omega)(\mathbf{u}, \mathbf{p})((\mathbf{u}, \mathbf{p}), (\mathbf{x}, \mathbf{q})) &= \omega(\pi_S(\mathbf{u}, \mathbf{p}))((\pi_S)_*((\mathbf{u}, \mathbf{p}), (\mathbf{x}, \mathbf{q}))) = \omega(\mathbf{u})(\mathbf{u}, \mathbf{x}) \\ &= \xi(\mathbf{u}, \mathbf{p})((\mathbf{u}, \mathbf{p}), (\mathbf{x}, \mathbf{q})) \end{aligned}$$

for all $((\mathbf{u}, \mathbf{p}), (\mathbf{x}, \mathbf{q})) \in T_{(\mathbf{u}, \mathbf{p})} S_c$. ■

Corollary 9.3.11

Suppose that $\mathcal{M}_c^{[i]} = \{M_{(\mathbf{u}, \mathbf{x})}^{[i]}\}_{(\mathbf{u}, \mathbf{x}) \in S_c}$ for $i = 1, 2$ are two connections on S_c and $\rho^{[i]}$ is a connection 1-form associated to $\mathcal{M}_c^{[i]}$ for $i = 1, 2$. Then $\rho^{[1]} - \rho^{[2]} = \pi_S^*(\omega)$ for some smooth differential 1-form ω on S .

Proof.

It follows from Theorem 9.3.6 and Proposition 9.3.8 that $\rho^{[i]}(\mathbf{u}, \mathbf{x})(H(\mathbf{u}, \mathbf{x})) = 1$ for all $(\mathbf{u}, \mathbf{x}) \in S_c$ and $\eta^*(\rho^{[i]}) = \rho^{[i]}$ for all $\eta \in S^1$ and $i = 1, 2$. Hence, if we set $\xi = \rho^{[1]} - \rho^{[2]}$, then we get that $\xi(\mathbf{u}, \mathbf{x})(H(\mathbf{u}, \mathbf{x})) = 0$ for all $(\mathbf{u}, \mathbf{x}) \in S_c$ and $\eta^*(\xi) = \xi$ for all $\eta \in S^1$. The conclusion of the corollary follows from Proposition 9.3.10. ■

Proposition 9.3.12

Let S be a 2-dimensional Riemannian manifold and $\mathcal{M}_c = \{M_{(\mathbf{u}, \mathbf{x})}\}_{(\mathbf{u}, \mathbf{x}) \in S_c}$ be a connection on S_c where ρ is a connection 1-form associated to $\mathcal{M}_c$. Suppose that

1. $\sigma : [a, b] \rightarrow S$ is a smooth curve,
2. $\tilde{\sigma}^{[i]} : [a, b] \rightarrow S_c$ is smooth curves such that $\pi_S \circ \tilde{\sigma}^{[i]} = \sigma$ for $i = 1, 2$, and
3. $\rho(\tilde{\sigma}^{[1]}(t))((\tilde{\sigma}^{[1]})_*(t, 1)) = 0$ for $a \leq t \leq b$.

Then there exists a smooth function $\theta : [a, b] \rightarrow \mathbb{R}$ such that $\tilde{\sigma}^{[2]}(t) = \eta_t(\tilde{\sigma}^{[1]}(t))$ with $\eta_t = e^{i\theta(t)} \in S^1$ and $\rho(\tilde{\sigma}^{[2]}(t))((\tilde{\sigma}^{[2]})_*(t, 1)) = \theta'(t)$ for $a \leq t \leq b$.

Moreover, if $\tilde{\sigma}^{[1]}(a) = \tilde{\sigma}^{[2]}(a)$, then we can choose θ such that $\theta(a) = 0$.

Proof.

i) Since $\pi_S(\tilde{\sigma}^{[i]}(t)) = \sigma(t)$ for $a \leq t \leq b$ and $i = 1, 2$, we have that $\tilde{\sigma}^{[i]}(t) \in T_{\sigma(t)}S_c$ for $a \leq t \leq b$ and $i = 1, 2$. Thus, for each $t \in [a, b]$, there exists $\eta_t \in S^1$ such that $\tilde{\sigma}^{[2]}(t) = \eta_t(\tilde{\sigma}^{[1]}(t))$. Since $\tilde{\sigma}^{[i]}$ for $i = 1, 2$ are smooth functions, we have that $\eta : [a, b] \rightarrow S^1$ defined by $\eta(t) = \eta_t$ for $a \leq t \leq b$ is also smooth.

Let $\theta : [a, b] \rightarrow \mathbb{R}$ be a function such that $e^{i\theta(t)} = \eta(t)$. Since η is smooth and $[a, b]$ is connected, we get that $\eta([a, b]) \subset S^1$ is connected. We may therefore assume that $\theta([a, b])$ is connected and $\theta : [a, b] \rightarrow \mathbb{R}$ is smooth³. We have that $(\mathbb{R}, e^{i\theta})$ is a covering of S^1 (Figure 8.2).

If $\tilde{\sigma}^{[1]}(a) = \tilde{\sigma}^{[2]}(a)$, then we may take $\theta(a) = 0$. In that case, θ is uniquely defined because $\theta([a, b])$ is connected.

ii) Let (W, U, ϕ) be a local chart of S and $(\tilde{W}, \tilde{U}, \tilde{\phi})$ be the local chart of S_c as defined in Subsection 9.2.3.

Let $\mu : [c, d] \rightarrow W$ be the local representation of σ where $[c, d] \subset [a, b]$; namely, $\sigma = \phi \circ \mu$ on $[c, d]$. Let $\mu^{[i]} : [c, d] \rightarrow \tilde{W}$ be the local representation of $\tilde{\sigma}^{[i]}$ for $i = 1, 2$; namely, $\tilde{\sigma}^{[i]} = \tilde{\phi} \circ \mu^{[i]}$ on $[c, d]$ for $i = 1, 2$. We have for $i = 1, 2$ that $\mu^{[i]}(t) = (\mu(t), \psi^{[i]}(t))$ for some $\psi^{[i]}(t) \in \mathbb{R}$. Moreover,

$$\tilde{\sigma}^{[2]}(t) = \eta_t(\tilde{\sigma}^{[1]}(t)) = \eta_t(\phi(\mu(t)), \psi^{[1]}(t).E(\mu(t))) = (\phi(\mu(t)), (\theta(t) + \psi^{[1]}(t)).E(\mu(t)))$$

for $c \leq t \leq d$. Thus $\psi^{[2]}(t) = \theta(t) + \psi^{[1]}(t)$ for $c \leq t \leq d$.

We have

$$\begin{aligned} \rho(\tilde{\sigma}^{[2]}(t))((\tilde{\sigma}^{[2]})_*(t, 1)) &= \rho((\tilde{\phi} \circ \mu^{[2]})(t))((\tilde{\phi} \circ \mu^{[2]})_*(t, 1)) = \rho((\tilde{\phi}(\mu^{[2]}(t))) (\tilde{\phi}_*(\mu_*^{[2]}(t, 1))) \\ &= \tilde{\phi}^*(\rho)(\mu^{[2]}(t))(\mu_*^{[2]}(t, 1)) = \tilde{\phi}^*(\rho)(\mu^{[2]}(t)) \left(\mu^{[2]}(t), \left(\mu'(t), \frac{d\psi^{[2]}}{dt}(t) \right) \right) \\ &= \tilde{\phi}^*(\rho)(\mu^{[2]}(t)) \left(\mu^{[2]}(t), \left(\mu'(t), \theta'(t) + \frac{d\psi^{[1]}}{dt}(t) \right) \right) \\ &= \tilde{\phi}^*(\rho)(\mu^{[2]}(t)) (\mu^{[2]}(t), (0, \theta'(t))) + \tilde{\phi}^*(\rho)(\mu^{[2]}(t)) \left(\mu^{[2]}(t), \left(\mu'(t), \frac{d\psi^{[1]}}{dt}(t) \right) \right), \end{aligned}$$

where

$$\begin{aligned} \tilde{\phi}^*(\rho)(\mu^{[2]}(t)) (\mu^{[2]}(t), (0, \theta'(t))) &= \rho(\tilde{\sigma}^{[2]}(t)) \left(\tilde{\sigma}^{[2]}(t), D_{\mathbf{w}, \nu} \tilde{\phi}(\mathbf{w}, \nu) \Big|_{(\mathbf{w}, \nu) = \mu^{[2]}(t)} \begin{pmatrix} 0 \\ \theta'(t) \end{pmatrix} \right) \\ &= \rho(\tilde{\sigma}^{[2]}(t)) \left(\tilde{\sigma}^{[2]}(t), \left(0, \theta'(t) \frac{\partial}{\partial \nu} (\nu.E(\mu(t))) \Big|_{\nu = \psi^{[2]}(t)} \right) \right) \end{aligned}$$

³For t in a small enough neighbourhood V of $t_0 \in [a, b]$, we have that $\log(\eta(t)) = \log(|\eta(t)|) + i \arg(\eta(t)) = i\theta(t)$ for $t \in V$. We have used a branch of the logarithm function. Since $\log$ is locally smooth and η is smooth, we have that θ is locally smooth. Since $t_0 \in [a, b]$ is arbitrary, we may conclude that $\theta : [a, b] \rightarrow \mathbb{R}$ is smooth.

$$= \rho(\tilde{\sigma}^{[2]}(t)) (\theta'(t)H(\tilde{\sigma}^{[2]}(t))) = \theta'(t) \rho(\tilde{\sigma}^{[2]}(t)) (H(\tilde{\sigma}^{[2]}(t))) = \theta'(t)$$

and

$$\tilde{\phi}^*(\rho)(\mu^{[2]}(t)) \left(\mu^{[2]}(t), \left(\mu'(t), \frac{d\psi^{[1]}}{dt}(t) \right) \right) = 0$$

as we explain below.

For s fixed, we have that $(\eta_s)_* (\tilde{\sigma}_*^{[1]}(t, 1)) = (\eta_s(\tilde{\sigma}^{[1]}))_* (t, 1)$. Moreover, since

$$\eta_s(\tilde{\sigma}^{[1]}(t)) = \left(\sigma(t), (\theta(s) + \psi^{[1]}(t)).\tilde{E}(\sigma(t)) \right) = \tilde{\phi}(\mu(t), \theta(s) + \psi^{[1]}(t)) ,$$

we get

$$\begin{aligned} (\eta_s(\tilde{\sigma}^{[1]}))_* (t, 1) &= \left(\eta_s(\tilde{\sigma}^{[1]}(t)), \frac{d}{dt} \eta_s(\tilde{\sigma}^{[1]}(t)) \right) = \left(\eta_s(\tilde{\sigma}^{[1]}(t)), \frac{d}{dt} \tilde{\phi}(\mu(t), \theta(s) + \psi^{[1]}(t)) \right) \\ &= \left(\left(\sigma(t), (\theta(s) + \psi^{[1]}(t)).\tilde{E}(\sigma(t)) \right), D_{\mathbf{w}, \nu} \tilde{\phi}(\mathbf{w}, \nu) \Big|_{(\mathbf{w}, \nu) = (\mu(t), \theta(s) + \psi^{[1]}(t))} \left(\frac{\mu'(t)}{\frac{d\psi^{[1]}}{dt}(t)} \right) \right) . \end{aligned}$$

Hence, at $t = s$, we get

$$\begin{aligned} (\eta_t)_* (\tilde{\sigma}_*^{[1]}(t, 1)) &= \left(\left(\sigma(t), \psi^{[2]}(t)).\tilde{E}(\sigma(t)) \right), D_{\mathbf{w}, \nu} \tilde{\phi}(\mathbf{w}, \nu) \Big|_{(\mathbf{w}, \nu) = \mu^{[2]}(t)} \left(\frac{\mu'(t)}{\frac{d\psi^{[1]}}{dt}(t)} \right) \right) \\ &= \left(\tilde{\sigma}^{[2]}(t), D_{\mathbf{w}, \nu} \tilde{\phi}(\mathbf{w}, \nu) \Big|_{(\mathbf{w}, \nu) = \mu^{[2]}(t)} \left(\frac{\mu'(t)}{\frac{d\psi^{[1]}}{dt}(t)} \right) \right) . \end{aligned}$$

It follows that

$$\begin{aligned} \tilde{\phi}^*(\rho)(\mu^{[2]}(t)) \left(\mu^{[2]}(t), \left(\mu'(t), \frac{d\psi^{[1]}}{dt}(t) \right) \right) &= \rho(\tilde{\sigma}^{[2]}(t)) \left(\tilde{\sigma}^{[2]}(t), D_{\mathbf{w}, \nu} \tilde{\phi}(\mathbf{w}, \nu) \Big|_{(\mathbf{w}, \nu) = \mu^{[2]}(t)} \left(\frac{\mu'(t)}{\frac{d\psi^{[1]}}{dt}(t)} \right) \right) \\ &= \rho(\tilde{\sigma}^{[2]}(t)) ((\eta_t)_* (\tilde{\sigma}_*^{[1]}(t, 1))) = \rho(\eta_t(\tilde{\sigma}^{[1]}(t))) ((\eta_t)_* (\tilde{\sigma}_*^{[1]}(t, 1))) \\ &= \eta_t^*(\rho)(\tilde{\sigma}^{[1]}(t)) (\tilde{\sigma}_*^{[1]}(t, 1)) = \rho(\tilde{\sigma}^{[1]}(t)) (\tilde{\sigma}_*^{[1]}(t, 1)) = 0 \end{aligned}$$

where the fifth equality comes from Proposition 9.3.8 and the last equality comes from the condition (3) in the statement of the lemma. ■

Proposition 9.3.13

Let S be a 2-dimensional Riemannian manifold. Moreover, let $\mathcal{M}_c = \{M_{(\mathbf{u}, \mathbf{x})}\}_{(\mathbf{u}, \mathbf{x}) \in S_c}$ be a connection on S_c where ρ is a connection 1-form associated to $\mathcal{M}_c$. Suppose that

$\sigma : [a, b] \rightarrow S$ a piecewise C^∞ curve ⁴ and that $(\sigma(a), \mathbf{x}) \in S_c$. Then there exists a unique piecewise C^∞ curve $\tilde{\sigma} : [a, b] \rightarrow S_c$ such that

1. $\tilde{\sigma}(a) = (\sigma(a), \mathbf{x})$,
2. $\pi_S \circ \tilde{\sigma} = \sigma$, and
3. $\rho(\tilde{\sigma}(t))(\tilde{\sigma}_*(t, 1)) = 0$ for $a \leq t \leq b$.

Thus $\tilde{\sigma}$ is a lift that provides a parallel translation of $(\sigma(a), \mathbf{x})$ along σ . The curve $\tilde{\sigma}$ is called the **horizontal lift** of σ through $(\sigma(a), \mathbf{x})$.

Remark 9.3.14

It follows from Proposition 9.3.6 that $\pi_{P_{\tilde{\sigma}(t)}}(\tilde{\sigma}_*(t, 1)) = 0$ and thus $\tilde{\sigma}_*(t, 1) \in M_{\tilde{\sigma}(t)}$ for $a \leq t \leq b$. We may say that $\tilde{\sigma}_*(t, 1)$ is orthogonal to $P_{\tilde{\sigma}(t)}$.

Moreover, it follows from the uniqueness of $\tilde{\sigma}$ satisfying (1) and Proposition 9.3.8 that the third condition listed before the Definition 9.3.1 is satisfied. ♠

Proof.

It is enough to prove the proposition for a smooth curve $\sigma : [a, b] \rightarrow S_c$ because we can then use the result on each smooth sections of a piecewise C^∞ curve using the end point of the horizontal lift of a smooth subsection as the starting point on S_c of the horizontal lift of the next smooth subsection.

i) Let (W, U, ϕ) be a local charts of S with $\sigma(a) \in U$. We prove that there exists a unique horizontal lift of σ on U .

We have seen in Remark 9.2.6 that $G : U \times S^1 \rightarrow S_c(U) = \bigcup_{\mathbf{u} \in U} S_c(\mathbf{u})$ defined by $G(\mathbf{u}, \eta) = (\mathbf{u}, \theta \cdot \tilde{E}(\mathbf{u}))$ for $\mathbf{u} \in U$ and $\eta = e^{i\theta} \in S^1$ is an isomorphism, where $\tilde{E}(\mathbf{u}) = E(\phi^{-1}(\mathbf{u}))$ for all $\mathbf{u} \in U$. Let $G_1(\mathbf{u}) = G(\mathbf{u}, 1) = (\mathbf{u}, 0 \cdot \tilde{E}(\mathbf{u})) = (\mathbf{u}, \tilde{E}(\mathbf{u}))$ for $\mathbf{u} \in U$. The map $G_1 : U \rightarrow S_c(U)$ is smooth.

We consider the connection $\check{M}_c = \{\check{M}_{(\mathbf{u}, \mathbf{x})}\}_{(\mathbf{u}, \mathbf{x}) \in S_c(U)}$ on $S_c(U)$ with its associated connection 1-form $\check{\rho}$ defined in Examples 9.3.5 and 9.3.7. We prove that there exists an horizontal lift $\check{\sigma}$ of σ on U such that $\check{\sigma}(a) = G_1(\sigma(a))$. To prove the existence, consider $\check{\sigma}(t) = G_1(\sigma(t))$ for $a \leq t \leq c$ and $c - a$ small enough to have $\sigma([a, c]) \subset U$. We have that $\check{\sigma}(a) = G_1(\sigma(a))$ and $\pi_S(\check{\sigma}(t)) = \sigma(t)$ for $a \leq t \leq c$. Moreover, since $\check{M}_{\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))} = (G_\eta)_*(T_{\mathbf{u}}S)$ for $\mathbf{u} \in U$ and $\eta \in S^1$ by definition, we get that $\check{\sigma}_*(t, 1) = (G_1)_*(\sigma_*(t, 1)) \in \check{M}_{\check{\sigma}(t)}$ for $a \leq t \leq c$ because $\sigma_*(t, 1) \in T_{\sigma(t)}S$. Hence $\pi_{P_{\check{\sigma}(t)}}(\check{\sigma}_*(t, 1)) = 0$ for $a \leq t \leq c$. It follows from Proposition 9.3.6 that $\check{\rho}(\check{\sigma}(t))(\check{\sigma}_*(t, 1)) = 0$ for $a \leq t \leq c$ ⁵. Thus $\check{\sigma}$ is a local horizontal lift of σ through $G_1(\sigma(a))$.

⁴Recall that a piecewise C^∞ function $f : [a, b] \rightarrow \mathbb{R}^n$ is a continuous function which is of class C^∞ on all the closed intervals delimited by the finite number of points in $[a, b]$ where it is not C^∞ .

⁵We can prove directly that $\check{\rho}(\check{\sigma}(t))(\check{\sigma}_*(t, 1)) = 0$. Referring to Examples 9.3.5 and 9.3.7, if we define

To prove the uniqueness of the local horizontal lift of σ through $G_1(\sigma(a))$. Suppose that $\tilde{\nu}$ is another local horizontal lift of σ through $G_1(\sigma(a))$ defined on $[a, \tilde{c}]$. By choosing the smallest of c and $\tilde{c}$, we may assume that $c = \tilde{c}$. It follows from Proposition 9.3.12 that there exists $\zeta : [a, c] \rightarrow \mathbb{R}$ with $\zeta(a) = 0$ such that $\tilde{\nu}(t) = \eta_t(\tilde{\sigma}(t))$ with $\eta_t = e^{i\zeta(t)}$ and $\check{\rho}(\tilde{\nu}(t))(\tilde{\nu}_*(t, 1)) = \zeta'(t)$ for $a \leq t \leq c$. Since $\tilde{\nu}$ is a horizontal lift of σ , we have that $\zeta'(t) = \check{\rho}(\tilde{\nu}(t))(\tilde{\nu}_*(t, 1)) = 0$ for $a \leq t \leq c$. Therefore $\zeta(t) = 0$ and so $\tilde{\nu}(t) = \tilde{\sigma}(t)$ for $a \leq t \leq c$.

ii) According to Corollary 9.3.11, there exists a differential 1-form ω on U such that $\check{\rho} - \rho = \pi_S^*(\omega)$. Choose a curve $\tilde{\sigma} : [a, c] \rightarrow S_c(U)$ such that $\pi_S(\tilde{\sigma}(t)) = \sigma(t)$ for $a \leq t \leq c$. It follows from Proposition 9.3.12 that there exists a smooth function $\xi : [a, c] \rightarrow \mathbb{R}$ such that $\tilde{\sigma}(t) = \eta_t(\tilde{\sigma}(t))$ with $\eta_t = e^{i\xi(t)} \in S^1$ and $\check{\rho}(\tilde{\sigma}(t))(\tilde{\sigma}_*(t, 1)) = \xi'(t)$ for $a \leq t \leq c$.

We have from Theorem 9.3.6 that $\tilde{\sigma}$ is a horizontal lift of σ if and only if $\rho(\tilde{\sigma}(t))(\tilde{\sigma}_*(t, 1)) = 0$ for $a \leq t \leq c$. This relation is satisfied if and only if

$$(\check{\rho} - \rho)(\tilde{\sigma}(t))(\tilde{\sigma}_*(t, 1)) = \check{\rho}(\tilde{\sigma}(t))(\tilde{\sigma}_*(t, 1)) = \xi'(t) \quad (9.3.3)$$

for $a \leq t \leq c$. Moreover,

$$\begin{aligned} (\check{\rho} - \rho)(\tilde{\sigma}(t))(\tilde{\sigma}_*(t, 1)) &= \pi_S^*(\omega)(\tilde{\sigma}(t))(\tilde{\sigma}_*(t, 1)) = \omega(\pi_S(\tilde{\sigma}(t)))((\pi_S)_*(\tilde{\sigma}_*(t, 1))) \\ &= \omega(\sigma(t))((\pi_S)_*(\sigma(t), *), (\sigma'(t), *)) \\ &= \omega(\sigma(t))(\sigma(t), \sigma'(t)) = \omega(\sigma(t))(\sigma_*(t, 1)) \end{aligned} \quad (9.3.4)$$

for $a \leq t \leq c$. It follows from (9.3.3) and (9.3.4) that $\tilde{\sigma}$ is a horizontal lift of σ if and only if $\xi'(t) = \omega(\sigma(t))(\sigma_*(t, 1))$ for $a \leq t \leq c$. Namely, $\tilde{\sigma}$ is a horizontal lift of σ if and only if $\xi(t) = \int_a^t \omega(\sigma(s))(\sigma_*(s, 1)) ds + \xi_0$ for $a \leq t \leq c$ and $\xi_0 \in \mathbb{R}$. Hence, each horizontal lift $\tilde{\sigma}$ of σ is of the form $\tilde{\sigma} = \eta_{t, \xi_0}(\tilde{\sigma}(t))$ for $a \leq t \leq c$ and

$$\eta_{t, \xi_0} = e^{i\xi(t)} = e^{i\xi_0} e^{i \int_a^t \omega(\sigma(s))(\sigma_*(s, 1)) ds} \in S^1$$

for some $\xi_0 \in \mathbb{R}$.

Given $(\sigma(a), \mathbf{x}) \in S_c(U)$, there exists a unique $\xi_0 \in [0, 2\pi[$ such that $(\sigma(a), \mathbf{x}) = \eta_{a, \xi_0}(\tilde{\sigma}(a))$. The unique horizontal lift of σ that we are looking for is given by $\tilde{\sigma}(t) = \eta_{t, \xi_0}(\tilde{\sigma}(t))$ for $a \leq t \leq c$ with this value of ξ_0 .

iii) We use a proof by contradiction to demonstrate that $\tilde{\sigma}$ can be extended to the entire interval $[a, b]$. Let

$$t_f = \sup \{t \in [a, b] : \sigma|_{[a, t]} \text{ as a unique horizontal lift } \tilde{\sigma} \text{ with } \tilde{\sigma}(a) = (\sigma(a), \mathbf{x}).\}$$

$\sigma_1(t) = (\sigma(t), 1)$ for $a \leq t \leq c$, then $\check{\sigma}(t) = (G \circ \sigma_1)(t)$ and

$$\begin{aligned} \check{\rho}(\check{\sigma}(t))(\check{\sigma}_*(t, 1)) &= (G^{-1})^*(\pi_2^*(\omega))((G \circ \sigma_1)(t))((G \circ \sigma_1)_*(t, 1)) \\ &= \pi_2^*(\omega)(G^{-1}((G \circ \sigma_1)(t)))((G^{-1})_*((G \circ \sigma_1)_*(t, 1))) \\ &= \pi_2^*(\omega)(\sigma_1(t))((\sigma_1)_*(t, 1)) = \pi_2^*(\omega)((\sigma(t), 1))((\sigma(t), 1), (\sigma'(t), 0)) \\ &= \omega(\pi_2(\sigma(t), 1))((\pi_2)_*((\sigma(t), 1), (\sigma'(t), 0))) = \omega(1)(1, 0) = 0. \end{aligned}$$

It follows from Proposition 9.3.6 that $\pi_{P_{\sigma(t)}}(\check{\sigma}_*(t, 1)) = 0$ and thus $\check{\sigma}_*(t, 1) \in M_{\check{\sigma}(t)}$.

Suppose that $t_f < b$. There exist $\epsilon > 0$ and a local chart (W, U, ϕ) of S such that $\sigma([t_f - \epsilon, t_f + \epsilon]) \subset U$. Using the previous part of the proof, there exists a unique local horizontal lift $\hat{\sigma} : [t_f - \epsilon, t_f + \epsilon] \rightarrow S_c(U)$ of σ such that $\hat{\sigma}(t_f) = (\sigma(t_f), \mathbf{y})$ for $(\sigma(t_f), \mathbf{y}) \in S_c$ arbitrary but fixed.

There exists $\eta \in S^1$ such that $\tilde{\sigma}(t_f - \epsilon) = \eta(\hat{\sigma}(t_f - \epsilon))$. Using Proposition 9.3.8, it is easy to verify that $\eta \circ \tilde{\sigma}$ is a horizontal lift; namely, that (1), (2) and (3) in the statement of the theorem are satisfied for $[a, b]$ replaced by $[t_f - \epsilon, t_f + \epsilon]$, and $(\sigma(a), \mathbf{x})$ replaced by $\tilde{\sigma}(t_f - \epsilon)$. By uniqueness of the horizontal lift of σ through $\tilde{\sigma}(t_f - \epsilon)$, we must have that $\tilde{\sigma}(t) = \eta(\hat{\sigma}(t))$ for $t_f - \epsilon \leq t < t_f$. Therefore

$$\tilde{\sigma}(t) = \begin{cases} \tilde{\sigma}(t) & \text{if } a \leq t < t_f \\ \eta(\hat{\sigma}(t)) & \text{if } t_f \leq t < t_f + \epsilon \end{cases}$$

is a unique horizontal lift of σ with $\tilde{\sigma}(a) = (\sigma(a), \mathbf{x})$. This is a contradiction that t_f is maximal. Note that the uniqueness of the extended map $\tilde{\sigma}$ comes from the uniqueness of $\tilde{\sigma}$ with $\tilde{\sigma}(a) = (\sigma(a), \mathbf{x})$ and the uniqueness of $\eta \circ \hat{\sigma}$ with $\eta(\hat{\sigma}(t_f)) = \eta(\sigma(t_f), \mathbf{y})$. ■

9.4 Structural Equations and Curvature

Definition 9.4.1

Let S be an oriented 2-dimensional Riemannian manifold. We define two differential 1-form ω_1 and ω_2 on S_c as it follows. Given $(\mathbf{u}, \mathbf{x}) \in S_c$, we set

$$\omega_1(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) = \langle (\pi_S)_*((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})), (\mathbf{u}, \mathbf{x}) \rangle_{\mathbf{u}} = \langle (\mathbf{u}, \mathbf{p}), (\mathbf{u}, \mathbf{x}) \rangle_{\mathbf{u}}$$

and

$$\begin{aligned} \omega_2(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) &= \langle (\pi_S)_*((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})), v(\mathbf{u}, \mathbf{x}) \rangle_{\mathbf{u}} \\ &= \langle (\mathbf{u}, \mathbf{p}), v(\mathbf{u}, \mathbf{x}) \rangle_{\mathbf{u}} \end{aligned}$$

for all $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \in T_{(\mathbf{u}, \mathbf{x})}S_c$ where $v = e^{i\pi/2} = i \in S^1$.

Recall from Lemma 9.2.1 that $(\pi_S)_*((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) = (\mathbf{u}, \mathbf{p})$ for all $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \in T_{(\mathbf{u}, \mathbf{x})}(TS)$. Using local charts, it is also easy to prove that ω_1 and ω_2 are smooth because we assume that S is a smooth manifold and the inner product $\langle \cdot, \cdot \rangle_{\mathbf{u}}$ on $T_{\mathbf{u}}S$ is given by a smooth mapping $\tau : S \rightarrow \bigcup_{\mathbf{u} \in S} \mathcal{T}^2(T_{\mathbf{u}}S)$. We will give another proof of this fact later on.

Proposition 9.4.2

Let S be an oriented 2-dimensional Riemannian manifold. Suppose that $\mathcal{M}_c = \{M_{(\mathbf{u}, \mathbf{x})}\}_{(\mathbf{u}, \mathbf{x}) \in S_c}$ is a connection on S_c and that ρ is a connection 1-form associated to $\mathcal{M}_c$. Then $\{\rho(\mathbf{u}, \mathbf{x}), \omega_1(\mathbf{u}, \mathbf{x}), \omega_2(\mathbf{u}, \mathbf{x})\}$ is a basis of $\Omega^1(T_{(\mathbf{u}, \mathbf{x})}S_c)$ for all $(\mathbf{u}, \mathbf{x}) \in S_c$.

Proof.

Since $\Omega^1(T_{(\mathbf{u}, \mathbf{x})}S_c)$ is the dual space of $T_{(\mathbf{u}, \mathbf{x})}S_c$, we have that $\dim \Omega^1(T_{(\mathbf{u}, \mathbf{x})}S_c) = 3$. It therefore suffices to prove that $\rho(\mathbf{u}, \mathbf{x})$, $\omega_1(\mathbf{u}, \mathbf{x})$ and $\omega_2(\mathbf{u}, \mathbf{x})$ are linearly independent. To prove that, it suffices to verify that there is no non-null element $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \in T_{(\mathbf{u}, \mathbf{x})}S_c = M_{(\mathbf{u}, \mathbf{x})} \oplus P_{(\mathbf{u}, \mathbf{x})}$ such that

$$\rho(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) = \omega_1(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) = \omega_2(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) = 0. \quad (9.4.1)$$

Suppose that there is $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \neq ((\mathbf{u}, \mathbf{x}), (\mathbf{0}, \mathbf{0}))$ such that (9.4.1) is satisfied. Then

$$\begin{aligned} (\mathbf{u}, \mathbf{p}) &= (\pi_S)_*((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \\ &= \omega_1(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q}))(\mathbf{u}, \mathbf{x}) + \omega_2(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q}))v(\mathbf{u}, \mathbf{x}) = (\mathbf{u}, \mathbf{0}) \end{aligned}$$

forces $\mathbf{p}$ to be null. Therefore $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) = \lambda H(\mathbf{u}, \mathbf{x})$ for some $\lambda \in \mathbb{R}$. But then $\lambda = \lambda \rho(\mathbf{u}, \mathbf{x})(H(\mathbf{u}, \mathbf{x})) = \rho(\mathbf{u}, \mathbf{x})(\lambda H(\mathbf{u}, \mathbf{x})) = \rho(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) = 0$. This is a contradiction that $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q}))$ is non-null. ■

Proposition 9.4.3

Let S be an oriented 2-dimensional Riemannian manifold. Given $\eta = e^{i\theta} \in S^1$, we have that $\eta^*(\omega_1) = \cos(\theta)\omega_1 + \sin(\theta)\omega_2$ and $\eta^*(\omega_2) = -\sin(\theta)\omega_1 + \cos(\theta)\omega_2$. Moreover, $\eta^*(\omega_1 \wedge \omega_2) = \omega_1 \wedge \omega_2$.

Proof.

Using the results obtained in Subsection 9.2.1 with $(\mathbf{u}, \mathbf{v}_1)$ and $(\mathbf{u}, \mathbf{v}_2)$ replaced by $(\mathbf{u}, \mathbf{x})$ and $v(\mathbf{u}, \mathbf{x})$ respectively, we get that $\eta(\mathbf{u}, \mathbf{x}) = \cos(\theta)(\mathbf{u}, \mathbf{x}) + \sin(\theta)v(\mathbf{u}, \mathbf{x})$ and $v(\mathbf{u}, \mathbf{x}) = -\sin(\theta)(\mathbf{u}, \mathbf{x}) + \cos(\theta)v(\mathbf{u}, \mathbf{x})$. Moreover, it follows from Lemma 9.2.1 and Remark 9.2.11 that $(\pi_S)_*(\eta_*((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q}))) = (\mathbf{u}, \mathbf{p}) = (\pi_S)_*((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q}))$. Therefore,

$$\begin{aligned} \eta^*(\omega_1)(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) &= \omega_1(\eta(\mathbf{u}, \mathbf{x}))(\eta_*((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q}))) \\ &= \langle (\pi_S)_*(\eta_*((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q}))), \eta(\mathbf{u}, \mathbf{x}) \rangle_{\mathbf{u}} \\ &= \langle (\pi_S)_*((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})), \cos(\theta)(\mathbf{u}, \mathbf{x}) + \sin(\theta)v(\mathbf{u}, \mathbf{x}) \rangle \\ &= \cos(\theta) \langle (\pi_S)_*((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})), (\mathbf{u}, \mathbf{x}) \rangle_{\mathbf{u}} + \sin(\theta) \langle (\pi_S)_*((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})), v(\mathbf{u}, \mathbf{x}) \rangle_{\mathbf{u}} \\ &= \cos(\theta)\omega_1(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) + \sin(\theta)\omega_2(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \end{aligned}$$

for all $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \in T_{(\mathbf{u}, \mathbf{x})}S_c$. Similarly, we have

$$\begin{aligned} \eta^*(\omega_2)(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \\ = -\sin(\theta)\omega_1(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) + \cos(\theta)\omega_2(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \end{aligned}$$

for all $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \in T_{(\mathbf{u}, \mathbf{x})}S_c$.

To prove the last statement of the proposition, we use (3) of Proposition 4.1.10 and Proposition 5.5.4 to obtain

$$\begin{aligned} \eta^*(\omega_1 \wedge \omega_2) &= \eta^*(\omega_1) \wedge \eta^*(\omega_2) = (\cos(\theta)\omega_1 + \sin(\theta)\omega_2) \wedge (-\sin(\theta)\omega_1 + \cos(\theta)\omega_2) \\ &= \cos^2(\theta)\omega_1 \wedge \omega_2 - \sin^2(\theta)\omega_2 \wedge \omega_1 = (\cos^2(\theta) + \sin^2(\theta))\omega_1 \wedge \omega_2 \\ &= \omega_1 \wedge \omega_2. \end{aligned}$$

■

Proposition 9.4.4

Let S be an oriented 2-dimensional Riemannian manifold. There exists a differential 2-form ω on S such that $\omega_1 \wedge \omega_2 = (\pi_S)^*(\omega)$.

Proof (Sketch).

The proof is very similar to the proof of Proposition 9.3.10.

We have that $(\omega_1 \wedge \omega_2)(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}_1, \mathbf{q}_1)), ((\mathbf{u}, \mathbf{x}), (\mathbf{p}_2, \mathbf{q}_2))) = 0$ if one of the $((\mathbf{u}, \mathbf{x}), (\mathbf{p}_1, \mathbf{q}_1))$ or $((\mathbf{u}, \mathbf{x}), (\mathbf{p}_2, \mathbf{q}_2))$ is a multiple of $H(\mathbf{u}, \mathbf{x})$ because $\omega_i(\mathbf{u}, \mathbf{x})(H(\mathbf{u}, \mathbf{x})) = 0$ for $i = 1, 2$. Recall that $H(\mathbf{u}, \mathbf{x}) = ((\mathbf{u}, \mathbf{x}), (\mathbf{0}, \mathbf{q}))$ for some $\mathbf{q}$. If we substitute this expression in the definition of ω_i , then we get $\omega_i(\mathbf{u}, \mathbf{x})(H(\mathbf{u}, \mathbf{x})) = 0$ for $i = 1, 2$. Moreover, we have from the previous proposition that $\eta^*(\omega_1 \wedge \omega_2) = \omega_1 \wedge \omega_2$.

The differential 2-form ω on S is defined as follows. Choose $(\mathbf{u}, \mathbf{p}) \in S_c$ arbitrary but fixed. For $(\mathbf{u}, \mathbf{x}_1), (\mathbf{u}, \mathbf{x}_2) \in T_{\mathbf{u}}S$, set

$$\omega(\mathbf{u})((\mathbf{u}, \mathbf{x}_1), (\mathbf{u}, \mathbf{x}_2)) = (\omega_1 \wedge \omega_2)(\mathbf{u}, \mathbf{p})((\mathbf{u}, \mathbf{p}), (\mathbf{x}_1, \mathbf{q}_1)), ((\mathbf{u}, \mathbf{p}), (\mathbf{x}_2, \mathbf{q}_2)))$$

for $((\mathbf{u}, \mathbf{p}), (\mathbf{x}_1, \mathbf{q}_1)), ((\mathbf{u}, \mathbf{p}), (\mathbf{x}_2, \mathbf{q}_2)) \in T_{(\mathbf{u}, \mathbf{p})}S_c$. We can then proceed as in the proof of Proposition 9.3.10 to prove that the definition of ω is independent of the choice of $\mathbf{p}$ and $((\mathbf{u}, \mathbf{p}), (\mathbf{x}_1, \mathbf{q}_1)), ((\mathbf{u}, \mathbf{p}), (\mathbf{x}_2, \mathbf{q}_2)) \in T_{(\mathbf{u}, \mathbf{p})}S_c$.

We also have by construction that $(\pi_S)^*(\omega) = \omega_1 \wedge \omega_2$. ■

Definition 9.4.5

Let S be an oriented 2-dimensional Riemannian manifold. The **volume element** on S , denoted dV , is the differential 2-form on S such that $(\pi_S)^*(dV) = \omega_1 \wedge \omega_2$.

To repeat the warning we have already given in Section 5.7, the notation dV is confusing because it designates the differential 2-form ω defined in the previous proposition which may not be the derivative of any differential 1-form.

Let (W, U, ϕ) be a local chart of S and $(\tilde{W}, \tilde{U}, \tilde{\phi})$ be a local chart of S_c as defined in Subsection 9.2.3. let $g : U \rightarrow S_c(U)$ be the function defined by $g(\mathbf{u}) = (\mathbf{u}, \tilde{E}(\mathbf{u}))$ for $\mathbf{u} \in U$ where $\tilde{E}(\mathbf{u}) = E(\phi^{-1}(\mathbf{u}))$ for all $\mathbf{u} \in U$. The maps $g : U \rightarrow S_c(U)$ and $v \circ g : U \rightarrow S_c(U)$ are smooth functions because their local representations are given by $(\tilde{\phi}^{-1} \circ g \circ \phi)(\mathbf{w}) = (\mathbf{w}, 0)$ and $(\tilde{\phi}^{-1} \circ (v \circ g) \circ \phi)(\mathbf{w}) = (\mathbf{w}, \pi/2)$ for all $\mathbf{w} \in W$ respectively.

Note that $\{g(\mathbf{u}), v(g(\mathbf{u}))\}$ is a basis of $T_{\mathbf{u}}S$ for each $\mathbf{u} \in U$. Let $\{\tilde{\omega}_1(\mathbf{u}), \tilde{\omega}_2(\mathbf{u})\}$ be the dual basis associated to $\{g(\mathbf{u}), v(g(\mathbf{u}))\}$; namely, $\tilde{\omega}_1(\mathbf{u})(g(\mathbf{u})) = 1$, $\tilde{\omega}_1(\mathbf{u})(v(g(\mathbf{u}))) = 0$, $\tilde{\omega}_2(\mathbf{u})(g(\mathbf{u})) = 0$ and $\tilde{\omega}_2(\mathbf{u})(v(g(\mathbf{u}))) = 1$. We have that $\tilde{\omega}_1$ and $\tilde{\omega}_2$ are smooth differential 1-form on U because g and $v \circ g$ are smooth. We get for every $(\mathbf{u}, \mathbf{x}) \in T_{\mathbf{u}}S$ that

$$\begin{aligned} g^*(\omega_1)(\mathbf{u})(\mathbf{u}, \mathbf{x}) &= \omega_1(g(\mathbf{u}))(g_*(\mathbf{u}, \mathbf{x})) = \langle (\pi_S)_*(g_*(\mathbf{u}, \mathbf{x})), g(\mathbf{u}) \rangle \\ &= \langle (\pi_S \circ g)_*(\mathbf{u}, \mathbf{x}), g(\mathbf{u}) \rangle = \langle (\mathbf{u}, \mathbf{x}), g(\mathbf{u}) \rangle = \tilde{\omega}_1(\mathbf{u}, \mathbf{x}) , \end{aligned}$$

where the fourth equality comes from $\pi_S \circ g = \text{Id}_U$. Similarly, $g^*(\omega_2)(\mathbf{u})(\mathbf{u}, \mathbf{x}) = \tilde{\omega}_2(\mathbf{u}, \mathbf{x})$ for all $(\mathbf{u}, \mathbf{x}) \in T_{\mathbf{u}}S$. Thus $\tilde{\omega}_i = g^*(\omega_i)$ for $i = 1, 2$. It follows that

$$\begin{aligned}\tilde{\omega}_1 \wedge \tilde{\omega}_2 &= g^*(\omega_1) \wedge g^*(\omega_1) = g^*(\omega_1 \wedge \omega_2) = g^*((\pi_S)^*(dV)) \\ &= (\pi_S \circ g)^*(dV) = dV\end{aligned}\tag{9.4.2}$$

on U because $\pi_S \circ g = \text{Id}_U$.

If $[g(\mathbf{u}), v(g(\mathbf{u}))] = \mu_{\mathbf{u}}$, the orientation on $T_{\mathbf{u}}S$ for $\mathbf{u} \in S$, then (9.4.2) shows that the definition of volume elements given above is identical to the definition of volume elements given in Definition 5.7.5 of Section 5.7.

Let $\rho_i = (\pi_S)^*(\tilde{\omega}_i)$ for $i = 1, 2$. The ρ_i are smooth differential 1-form on $S_c(U)$ because the $\tilde{\omega}_i$ are smooth differential 1-form on U . We first need to deduce some relations between the ρ_i and ω_i to be able to derive the structural equations later. We have

$$\rho_1 \wedge \rho_2 = (\pi_S)^*(\tilde{\omega}_1) \wedge (\pi_S)^*(\tilde{\omega}_2) = (\pi_S)^*(\tilde{\omega}_1 \wedge \tilde{\omega}_2) = (\pi_S)^*(dV) = \omega_1 \wedge \omega_2 .$$

In fact, we have more than this relation. Given $(g(\mathbf{u}), (\mathbf{p}, \mathbf{q})) \in T_{g(\mathbf{u})}S_c$, we have

$$\begin{aligned}\omega_1(g(\mathbf{u}))(g(\mathbf{u}), (\mathbf{p}, \mathbf{q})) g(\mathbf{u}) + \omega_2(g(\mathbf{u}))(g(\mathbf{u}), (\mathbf{p}, \mathbf{q})) v(g(\mathbf{u})) \\ &= (\mathbf{u}, \mathbf{p}) = (\pi_S)_*(g(\mathbf{u}), (\mathbf{p}, \mathbf{q})) \\ &= \tilde{\omega}_1(\mathbf{u})((\pi_S)_*(g(\mathbf{u}), (\mathbf{p}, \mathbf{q}))) g(\mathbf{u}) + \tilde{\omega}_2(\mathbf{u})((\pi_S)_*(g(\mathbf{u}), (\mathbf{p}, \mathbf{q}))) v(g(\mathbf{u})) \\ &= \tilde{\omega}_1(\pi_S(g(\mathbf{u})))((\pi_S)_*(g(\mathbf{u}), (\mathbf{p}, \mathbf{q}))) g(\mathbf{u}) + \tilde{\omega}_2(\pi_S(g(\mathbf{u})))((\pi_S)_*(g(\mathbf{u}), (\mathbf{p}, \mathbf{q}))) v(g(\mathbf{u})) \\ &= \rho_1(g(\mathbf{u}))(g(\mathbf{u}), (\mathbf{p}, \mathbf{q})) g(\mathbf{u}) + \rho_2(g(\mathbf{u}))(g(\mathbf{u}), (\mathbf{p}, \mathbf{q})) v(g(\mathbf{u}))\end{aligned}$$

where the third identity comes from the fact that $\{\tilde{\omega}_1(\mathbf{u}), \tilde{\omega}_2(\mathbf{u})\}$ is the dual basis of $\{g(\mathbf{u}), v(g(\mathbf{u}))\}$ on $T_{\mathbf{u}}S$ for all $\mathbf{u} \in U$. By matching the coefficients of $g(\mathbf{u})$ and $v(g(\mathbf{u}))$, we get that $\rho_i(g(\mathbf{u})) = \omega_i(g(\mathbf{u}))$ on $T_{g(\mathbf{u})}S_c$ for $i = 1, 2$.

Moreover, given $\eta = e^{i\theta} \in S^1$, we have

$$\eta^*(\rho_i) = \eta^*(\pi_S)^*(\tilde{\omega}_i) = (\pi_S \circ \eta)^*(\tilde{\omega}_i) = (\pi_S)^*(\tilde{\omega}_i) = \rho_i$$

for $i = 1, 2$. Using Proposition 9.4.3, we get

$$\begin{aligned}\omega_1(\eta(g(\mathbf{u}))) (\eta_*(g(\mathbf{u}), (\mathbf{p}, \mathbf{q}))) &= \eta^*(\omega_1)(g(\mathbf{u}))(g(\mathbf{u}), (\mathbf{p}, \mathbf{q})) \\ &= \cos(\theta) \omega_1(g(\mathbf{u}))(g(\mathbf{u}), (\mathbf{p}, \mathbf{q})) + \sin(\theta) \omega_2(g(\mathbf{u}))(g(\mathbf{u}), (\mathbf{p}, \mathbf{q})) \\ &= \cos(\theta) \rho_1(g(\mathbf{u}))(g(\mathbf{u}), (\mathbf{p}, \mathbf{q})) + \sin(\theta) \rho_2(g(\mathbf{u}))(g(\mathbf{u}), (\mathbf{p}, \mathbf{q})) \\ &= \cos(\theta) \eta^*(\rho_1)(g(\mathbf{u}))(g(\mathbf{u}), (\mathbf{p}, \mathbf{q})) + \sin(\theta) \eta^*(\rho_2)(g(\mathbf{u}))(g(\mathbf{u}), (\mathbf{p}, \mathbf{q})) \\ &= \cos(\theta) \rho_1(\eta(g(\mathbf{u}))) (\eta_*(g(\mathbf{u}), (\mathbf{p}, \mathbf{q}))) + \sin(\theta) \rho_2(\eta(g(\mathbf{u}))) (\eta_*(g(\mathbf{u}), (\mathbf{p}, \mathbf{q})))\end{aligned}\tag{9.4.3}$$

for all $(g(\mathbf{u}), (\mathbf{p}, \mathbf{q})) \in T_{g(\mathbf{u})}S_c$.

Let $\tilde{\eta} : S_c \rightarrow S^1$ be the function defined by $\tilde{\eta}(\mathbf{u}, \mathbf{x}) = \eta$ if $(\mathbf{u}, \mathbf{x}) = \eta(g(\mathbf{u}))$. Using local coordinates, it is easy to proof that the function $\tilde{\eta} : S_c \rightarrow S^1$ is a smooth function. Let $\tilde{\theta} : S_c \rightarrow \mathbb{R}$ be the function defined by $\tilde{\theta}(\mathbf{u}, \mathbf{x}) = \theta$ where $\tilde{\eta}(\mathbf{u}, \mathbf{x}) = e^{i\theta}$ with $0 \leq \theta < 2\pi$. Namely, θ is the argument of $\tilde{\eta}(\mathbf{u}, \mathbf{x})$ modulo 2π .

Since all elements of $S_c(\mathbf{u})$ are of the form $\eta(g(\mathbf{u}))$ for some $\eta \in S^1$ and all elements of $T_{\eta(g(\mathbf{u}))}S_c$ are of the form $\eta_*(g(\mathbf{u}), (\mathbf{p}, \mathbf{q}))$ for some $(g(\mathbf{u}), (\mathbf{p}, \mathbf{q})) \in T_{g(\mathbf{u})}$, we get from (9.4.3) that

$$\omega_1(\mathbf{u}, \mathbf{x}) = \cos(\tilde{\theta}(\mathbf{u}, \mathbf{x})) \rho_1(\mathbf{u}, \mathbf{x}) + \sin(\tilde{\theta}(\mathbf{u}, \mathbf{x})) \rho_2(\mathbf{u}, \mathbf{x}) \quad (9.4.4)$$

on $T_{(\mathbf{u}, \mathbf{x})}S_c$. Similarly, we have

$$\omega_2(\mathbf{u}, \mathbf{x}) = -\sin(\tilde{\theta}(\mathbf{u}, \mathbf{x})) \rho_1(\mathbf{u}, \mathbf{x}) + \cos(\tilde{\theta}(\mathbf{u}, \mathbf{x})) \rho_2(\mathbf{u}, \mathbf{x}) \quad (9.4.5)$$

on $T_{(\mathbf{u}, \mathbf{x})}S_c$.

Since ρ_1 and ρ_2 are smooth differential 1-form on $S_c(U)$, and $\tilde{\theta}$ depends smoothly on $(\mathbf{u}, \mathbf{x}) \in \tilde{U}$, we get from the previous two relation that ω_1 and ω_2 are smooth differential 1-form on $S_c(U)$. This is true for all local charts of S and S_c .

Let S be a smooth 2-dimensional manifold and $\mathcal{M}_c = \{M_{(\mathbf{u}, \mathbf{x})}\}_{(\mathbf{u}, \mathbf{x}) \in S_c}$ be a connection on S_c where ρ is a connection 1-form associated to $\mathcal{M}_c$. Since $\{\rho, \omega_1, \omega_2\}$ is a basis of $\Omega^1(T_{(\mathbf{u}, \mathbf{x})}S_c)$ for all $(\mathbf{u}, \mathbf{x}) \in S_c$, we have that all the differential 2-forms on S_c can be expressed uniquely as a linear combination of $\rho \wedge \omega_1$, $\rho \wedge \omega_2$ and $\omega_1 \wedge \omega_2$. In particular, we may express $d\rho$, $d\omega_1$ and $d\omega_2$ as a linear combination of $\rho \wedge \omega_1$, $\rho \wedge \omega_2$ and $\omega_1 \wedge \omega_2$. These linear combinations are called the **Cartan structural formulae**. The rest of this section is devoted to computing these Cartan structural formulae.

9.4.1 First (Cartan) Structural Equation

Let (W, U, ϕ) be a local charts of S . Since $d\rho_i = d((\pi_S)^*(\tilde{\omega}_i)) = (\pi_S)^*(d\tilde{\omega}_i)$ and $d\tilde{\omega}_i = A_i dV$ for some smooth function $A_i : U \rightarrow \mathbb{R}$ because $d\tilde{\omega}_i$ is a smooth differential 2-form on U , we get

$$d\rho_i = (\pi_S)^*(A_i dV) = (A_i \circ \pi_S) (\pi_S)^*(dV) = (A_i \circ \pi_S) \omega_1 \wedge \omega_2$$

for $i = 1, 2$. If we compute the differential of (9.4.4), then we get

$$\begin{aligned} d\omega_1(\mathbf{u}, \mathbf{x}) &= -\sin(\tilde{\theta}) (d\tilde{\theta} \wedge \rho_1)(\mathbf{u}, \mathbf{x}) + \cos(\tilde{\theta}) d\rho_1(\mathbf{u}, \mathbf{x}) \\ &\quad + \cos(\tilde{\theta}) (d\tilde{\theta} \wedge \rho_2)(\mathbf{u}, \mathbf{x}) + \sin(\tilde{\theta}) d\rho_2(\mathbf{u}, \mathbf{x}) \\ &= (d\tilde{\theta} \wedge (-\sin(\tilde{\theta}) \rho_1 + \cos(\tilde{\theta}) \rho_2))(\mathbf{u}, \mathbf{x}) + (\cos(\tilde{\theta}) (A_1 \circ \pi_S)(\mathbf{u}, \mathbf{x}) \\ &\quad + \sin(\tilde{\theta}) (A_2 \circ \pi_S)(\mathbf{u}, \mathbf{x})) (\omega_1 \wedge \omega_2)(\mathbf{u}, \mathbf{x}) \\ &= (d\tilde{\theta} \wedge \omega_2)(\mathbf{u}, \mathbf{x}) + (\cos(\tilde{\theta}) (A_1 \circ \pi_S)(\mathbf{u}, \mathbf{x}) \\ &\quad + \sin(\tilde{\theta}) (A_2 \circ \pi_S)(\mathbf{u}, \mathbf{x})) (\omega_1 \wedge \omega_2)(\mathbf{u}, \mathbf{x}) \end{aligned} \quad (9.4.6)$$

for $(\mathbf{u}, \mathbf{x}) \in S_c(\mathbf{u})$, where we have abbreviated $\tilde{\theta}(\mathbf{u}, \mathbf{x})$ to simply $\tilde{\theta}$.

Using the local chart $(\tilde{W}, \tilde{U}, \tilde{\phi})$ of S_c introduced in Subsection 9.2.3, we get $\tilde{\phi}^*(\tilde{\theta}) : W \times I \rightarrow \mathbb{R}$ is defined by $\tilde{\phi}^*(\tilde{\theta})(\mathbf{w}, \theta) = \theta$. Hence $d(\tilde{\phi}^*(\tilde{\theta}))(\mathbf{w}, \theta)((\mathbf{w}, \theta), (\mathbf{r}, t)) = t$ for all $((\mathbf{w}, \theta), (\mathbf{r}, t)) \in T_{(\mathbf{w}, \theta)}(W \times I) = (W \times I) \times (\mathbb{R} \times \mathbb{R})$. Suppose that $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) =$

$\tilde{\phi}_*((\mathbf{w}, \theta), (\mathbf{r}, t)) \in T_{(\mathbf{u}, \mathbf{x})}S_c$. Using the function G defined in Example 9.3.5 and the differential 1-forms ω and $\check{\rho}$ defined in Example 9.3.7, we get that $(G^{-1} \circ \tilde{\phi})(\mathbf{w}, \theta) = (\mathbf{u}, \eta)$ for $\eta = e^{i\theta}$ and $(G^{-1} \circ \tilde{\phi})_*((\mathbf{w}, \theta), (\mathbf{r}, t)) = ((\mathbf{u}, \eta), (\mathbf{r}, t i \eta))$. Hence

$$\begin{aligned} d\tilde{\theta}(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) &= \tilde{\phi}^*(d\tilde{\theta})(\mathbf{w}, \theta)((\mathbf{w}, \theta), (\mathbf{r}, t)) = d(\tilde{\phi}^*(\tilde{\theta}))(\mathbf{w}, \theta)((\mathbf{w}, \theta), (\mathbf{r}, t)) \\ &= t = \omega(\eta)(\eta, t i \eta) = \pi_2^*(\omega)(\mathbf{u}, \eta)((\mathbf{u}, \eta), (\mathbf{r}, t i \eta)) \\ &= \pi_2^*(\omega)((G^{-1} \circ \tilde{\phi})(\mathbf{w}, \theta))((G^{-1} \circ \tilde{\phi})_*((\mathbf{w}, \theta), (\mathbf{r}, t))) \\ &= \pi_2^*(\omega)(G^{-1}(\tilde{\phi}(\mathbf{w}, \theta)))((G^{-1})_*(\tilde{\phi}_*((\mathbf{w}, \theta), (\mathbf{r}, t)))) \\ &= (G^{-1})^*(\pi_2^*(\omega))(\tilde{\phi}(\mathbf{w}, \theta))(\tilde{\phi}_*((\mathbf{w}, \theta), (\mathbf{r}, t))) \\ &= (G^{-1})^*(\pi_2^*(\omega))(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \\ &= \check{\rho}(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \end{aligned}$$

Thus, it follows from (9.4.6) that

$$\begin{aligned} d\omega_1(\mathbf{u}, \mathbf{x}) &= (\check{\rho} \wedge \omega_2)(\mathbf{u}, \mathbf{x}) + (\cos(\tilde{\theta})(A_1 \circ \pi_S)(\mathbf{u}, \mathbf{x}) \\ &\quad + \sin(\tilde{\theta})(A_2 \circ \pi_S)(\mathbf{u}, \mathbf{x}))(\omega_1 \wedge \omega_2)(\mathbf{u}, \mathbf{x}) \end{aligned}$$

for $(\mathbf{u}, \mathbf{x}) \in S_c(\mathbf{u})$. We may rewrite this latest equation as

$$d\omega_1 = \check{\rho} \wedge \omega_2 + B_1 \omega_1 \wedge \omega_2 \quad (9.4.7)$$

on $S_c(U)$, where $B_1 : S_c(U) \rightarrow \mathbb{R}$ is defined by $B_1(\mathbf{u}, \mathbf{x}) = \cos(\theta)(A_1 \circ \pi_S)(\mathbf{u}, \mathbf{x}) + \sin(\theta)(A_2 \circ \pi_S)(\mathbf{u}, \mathbf{x})$ for $(\mathbf{u}, \mathbf{x}) \in S_c(U)$. The function B_1 is smooth because A_1 , A_2 and $\tilde{\theta}$ are smooth.

Using (9.4.5) and proceeding as we just did above, we get

$$d\omega_2 = -\check{\rho} \wedge \omega_1 + B_2 \omega_1 \wedge \omega_2 \quad (9.4.8)$$

on $S_c(U)$, where $B_2 : S_c(U) \rightarrow \mathbb{R}$ is defined by $B_2(\mathbf{u}, \mathbf{x}) = -\sin(\theta)(A_1 \circ \pi_S)(\mathbf{u}, \mathbf{x}) + \cos(\theta)(A_2 \circ \pi_S)(\mathbf{u}, \mathbf{x})$ for $(\mathbf{u}, \mathbf{x}) \in S_c(U)$. As for the function B_1 , the function B_2 is smooth because A_1 , A_2 and $\tilde{\theta}$ are smooth.

According to Corollary 9.3.11, there exists a differential 1-form ω on U such that $\check{\rho} - \rho = \pi_S^*(\omega)$. Since $\{\tilde{\omega}_1(\mathbf{u}), \tilde{\omega}_2(\mathbf{u})\}$ is a basis of $\Omega^1(T_{\mathbf{u}}S)$ for all $\mathbf{u} \in U$, we have that $\omega = a_1\tilde{\omega}_1 + a_2\tilde{\omega}_2$ for some smooth functions $a_1, a_2 : U \rightarrow \mathbb{R}$. Hence

$$\check{\rho} - \rho = \pi_S^*(\omega) = (a_1 \circ \pi_S) \pi_S^*(\tilde{\omega}_1) + (a_2 \circ \pi_S) \pi_S^*(\tilde{\omega}_2) = (a_1 \circ \pi_S) \rho_1 + (a_2 \circ \pi_S) \rho_2$$

on $S_c(U)$. Since $\{\omega_1(\mathbf{u}, \mathbf{x}), \omega_2(\mathbf{u}, \mathbf{x})\}$ is also a basis of $\Omega^1(T_{(\mathbf{u}, \mathbf{x})}S_c)$ for all $(\mathbf{u}, \mathbf{x}) \in S_c(U)$, we can write

$$\check{\rho} - \rho = b_1 \omega_1 + b_2 \omega_2 \quad (9.4.9)$$

on $S_c(U)$, where $b_1, b_2 : S_c(U) \rightarrow \mathbb{R}$ are two smooth functions. If we substitute the latest formula into (9.4.7) and (9.4.8), then we get

$$\begin{aligned} d\omega_1 &= (\rho + b_1 \omega_1 + b_2 \omega_2) \wedge \omega_2 + B_1 \omega_1 \wedge \omega_2 = \rho \wedge \omega_2 + (b_1 + B_1) \omega_1 \wedge \omega_2 \\ &= \rho \wedge \omega_2 + c_1 \omega_1 \wedge \omega_2 \end{aligned}$$

and

$$\begin{aligned} d\omega_2 &= -(\rho + b_1\omega_1 + b_2\omega_2) \wedge \omega_1 + B_2\omega_1 \wedge \omega_2 = -\rho \wedge \omega_1 + (b_2 + B_2)\omega_1 \wedge \omega_2 \\ &= -\rho \wedge \omega_1 + c_2\omega_1 \wedge \omega_2, \end{aligned}$$

where $c_i = b_i + B_i : S_c(U) \rightarrow \mathbb{R}$ are smooth function for $i = 1, 2$.

The definition of c_1 and c_2 are independent of the local chart since ρ , ω_1 and ω_2 are independent of the local chart. Therefore the previous two relation are true on S_c .

Definition 9.4.6

The **first (Cartan) structural equations** are the equation

$$d\omega_1 = \rho \wedge \omega_2 + c_1\omega_1 \wedge \omega_2 \quad \text{and} \quad d\omega_2 = -\rho \wedge \omega_1 + c_2\omega_1 \wedge \omega_2$$

on S_c .

The first structural equations can be simplified. Obviously, $\rho \wedge \omega_2$ in $d\omega_1$ and $-\rho \wedge \omega_1$ in $d\omega_2$ cannot be removed. However, by selecting the appropriate connection 1-form on S_c , we may get $c_1 = c_2 = 0$ on $T S_c$.

Theorem 9.4.7

Let S be an oriented 2-dimensional Riemannian manifold. There exists a unique connection 1-form ρ on S_c such that

$$d\omega_1 = \rho \wedge \omega_2 \quad \text{and} \quad d\omega_2 = -\rho \wedge \omega_1$$

on S_c . This connection 1-form is called the **Riemannian connection**.

Proof.

Let η_1 be a connection 1-form on S_c . Note that connection 1-forms can be generated using Proposition 9.3.9 for instance. Let η_2 be another connection 1-form on S_c . Proceeding as we did to deduce (9.4.9), we get

$$\eta_1 - \eta_2 = d_1\omega_1 + d_2\omega_2 \tag{9.4.10}$$

on $S_c(U)$ for some smooth functions $d_1, d_2 : S_c(U) \rightarrow \mathbb{R}$. If we substitute $\rho = \eta_1$ in the first structural equations, then we get

$$d\omega_1 = (\eta_2 + d_1\omega_1 + d_2\omega_2) \wedge \omega_2 + c_1\omega_1 \wedge \omega_2 = \eta_2 \wedge \omega_2 + (c_1 + d_1)\omega_1 \wedge \omega_2 \tag{9.4.11}$$

and

$$d\omega_2 = -(\eta_2 + d_1\omega_1 + d_2\omega_2) \wedge \omega_1 + c_2\omega_1 \wedge \omega_2 = -\eta_2 \wedge \omega_1 + (c_2 + d_2)\omega_1 \wedge \omega_2. \tag{9.4.12}$$

Therefore, we get that $d\omega_1 = \eta_2 \wedge \omega_2$ and $d\omega_2 = -\eta_2 \wedge \omega_1$ if we select $\eta_2 = \eta_1 - d_1\omega_1 - d_2\omega_2$ with $d_1 = -c_1$ and $d_2 = -c_2$.

To prove uniqueness, suppose that η_3 is a connection 1-form on S_c such that $d\omega_1 = \eta_3 \wedge \omega_2$ and $d\omega_2 = -\eta_3 \wedge \omega_1$. Instead of (9.4.10), we now have that $\eta_1 - \eta_3 = \tilde{d}_1\omega_1 + \tilde{d}_2\omega_2$ for some smooth functions $\tilde{d}_1, \tilde{d}_2 : S_c(U) \rightarrow \mathbb{R}$. Moreover, (9.4.11) and (9.4.12) become

$$d\omega_1 = \eta_3 \wedge \omega_2 + (c_1 + \tilde{d}_1)\omega_1 \wedge \omega_2 = d\omega_1 + (c_1 + \tilde{d}_1)\omega_1 \wedge \omega_2$$

and

$$d\omega_2 = -\eta_3 \wedge \omega_1 + (c_2 + \tilde{d}_2)\omega_1 \wedge \omega_2 = d\omega_2 + (c_2 + \tilde{d}_2)\omega_1 \wedge \omega_2.$$

Thus $\tilde{d}_1 = -c_1$ and $\tilde{d}_2 = -c_2$. Hence $\eta_3 = \eta_2$. ■

9.4.2 Second (Cartan) Structural Equation

Let (W, U, ϕ) be a local charts of S . As in the proof of Proposition 9.3.13, we consider the connection $\check{\mathcal{M}}_c = \{\check{M}_{(u,x)}\}_{(u,x) \in S_c(U)}$ on $S_c(U)$ with its associated connection 1-form $\check{\rho}$ defined in Examples 9.3.5 and 9.3.7. According to Corollary 9.3.11, there exists a differential 1-form ω on U such that $\check{\rho} - \rho = (\pi_S)^*(\omega)$.

Since $d\check{\rho} = 0$ as we have seen at the end of Example 9.3.7, we have

$$d\rho = d(\rho - \check{\rho}) = -d((\pi_S)^*(\omega)) = -(\pi_S)^*(d\omega)$$

for some differential 1-form ω on U . Since $d\omega$ is a differential 2-form on U , we have that $d\omega = \kappa dV$ for some smooth function $\kappa : U \rightarrow \mathbb{R}$. Hence,

$$d\rho = -(\pi_S)^*(\kappa dV) = -(\pi_S)^*(\kappa) (\pi_S)^*(dV) = -(\pi_S)^*(\kappa) \omega_1 \wedge \omega_2 \quad (9.4.13)$$

on $S_c(U)$. Note that the definition of $\kappa : U \rightarrow \mathbb{R}$ is independent of the local chart since ρ , ω_1 and ω_2 are independent of the local chart. Therefore (9.4.13) is true on S_c .

Definition 9.4.8

The **second (Cartan) structural equation** is the equation

$$d\rho = -(\kappa \circ \pi_S) \omega_1 \wedge \omega_2$$

on S_c . The function κ is called the **curvature** associated to the connection 1-form ρ .

Example 9.4.9

Suppose that S is an open subset of $\mathbb{R}^2$ with the standard Euclidean norm and orientation. This is an oriented 2-dimensional Riemannian manifold with the Riemann metric given by $\langle (u, x), (u, y) \rangle_u = \langle x, y \rangle$ for all $(u, x), (u, y) \in T_u S \cong S \times \mathbb{R}^2$, where the last inner product is the standard inner product in $\mathbb{R}^3$.

We use the construction given in Examples 9.3.5 and 9.3.7. Note that the local construction presented in these examples can be use globally in the present example. Namely, we can use the trivial local chart (W, U, ϕ) with $W = U = S$ and $\phi = \text{Id}$.

The map $G : U \times S^1 \rightarrow S_c(U)$ in Example 9.3.5 is now defined by $G(u, \eta) = \eta(u, e_1) = (u, \theta \cdot e_1)$ where $\eta = e^{i\theta} \in S^1$. We then get a connection on S_c defined by

$\check{\mathcal{M}}_c = \{\check{M}_{\eta(\mathbf{u}, \mathbf{e}_1)}\}_{\eta \in S^1, \mathbf{u} \in S}$ where $\check{M}_{\eta(\mathbf{u}, \mathbf{e}_1)} = (G_\eta)_*(T_{\mathbf{u}}S) = \{(\eta(\mathbf{u}, \mathbf{e}_1), (\mathbf{p}, \mathbf{0})) : \mathbf{p} \in \mathbb{R}^2\}$ for all $\eta \in S^1$ and $\mathbf{u} \in S$.

We have shown in Example 9.3.7 that the connection 1-form associated to $\check{\mathcal{M}}_c$ is $\check{\rho} = (G^{-1})^*(\pi_2^*(\omega))$ where ω is the differential 1-form on S^1 defined by $\omega(\eta)(\eta, t i\eta) = t$ for all $\eta \in S^1$ and $t \in \mathbb{R}$. Note that $T_\eta S^1 = \{(\eta, t i\eta) : t \in \mathbb{R}\}$. Moreover, we have also introduced in Example 9.3.7 the map $\pi_2 : U \times S^1 \rightarrow S^1$ defined by $\pi_2(\mathbf{u}, \eta) = \eta$ for all $(\mathbf{u}, \eta) \in U \times S^1$. Given $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \in T_{(\mathbf{u}, \mathbf{x})}S_c$, there exists $((\mathbf{u}, \eta), (\mathbf{p}, t i\eta)) \in T_{(\mathbf{u}, \eta)}(U \times S^1)$ with $\eta = e^{i\theta}$ such that

$$\begin{aligned} ((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) &= G_*((\mathbf{u}, \eta), (\mathbf{p}, t i\eta)) = \left(\eta(\mathbf{u}, \mathbf{e}_1), \left(\mathbf{p}, t \frac{d}{d\theta}(\theta \cdot \mathbf{e}_1) \right) \right) \\ &= (\eta(\mathbf{u}, \mathbf{e}_1), (\mathbf{p}, t \theta \cdot \mathbf{e}_2)) . \end{aligned}$$

Hence

$$\begin{aligned} \check{\rho}(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) &= (G^{-1})^*(\pi_2^*(\omega))(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \\ &= \pi_2^*(\omega)(\mathbf{u}, \eta)((\mathbf{u}, \eta), (\mathbf{p}, t i\eta)) = \omega(\eta)(\eta, t i\eta) = t \end{aligned}$$

as we have shown in Example 9.3.7.

Consider the map $\tilde{\theta} : S_c \rightarrow \mathbb{R}$ defined at the beginning of Section 9.4. Using the local charts $(\tilde{W}, \tilde{U}, \tilde{\phi})$ of S_c , in particular $\tilde{\phi}(\mathbf{u}, \theta) = e^{i\theta}(\mathbf{u}, \mathbf{e}_1) = (\mathbf{u}, \theta \cdot \mathbf{e}_1)$ because $W = U = S$, we get that $(\tilde{\theta} \circ \tilde{\phi})(\mathbf{u}, \theta) = \theta$ for all $(\mathbf{u}, \theta) \in \tilde{U}$ and

$$\begin{aligned} d\tilde{\theta}(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) &= d\tilde{\theta}(\mathbf{u}, \theta \cdot \mathbf{e}_1)((\mathbf{u}, \theta \cdot \mathbf{e}_1), (\mathbf{p}, t \theta \cdot \mathbf{e}_2)) \\ &= d\tilde{\theta}(\tilde{\phi}(\mathbf{u}, \theta))(\tilde{\phi}_*((\mathbf{u}, \theta), (\mathbf{r}, t))) = \tilde{\phi}^*(d\tilde{\theta})(\mathbf{u}, \theta)((\mathbf{u}, \theta), (\mathbf{r}, t)) \\ &= d(\tilde{\theta} \circ \tilde{\phi})(\mathbf{u}, \theta)((\mathbf{u}, \theta), (\mathbf{r}, t)) = D_{\mathbf{u}, \theta}(\tilde{\theta} \circ \tilde{\phi})(\mathbf{u}, \theta) \begin{pmatrix} \mathbf{r} & t \end{pmatrix}^\top \\ &= \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{r} & t \end{pmatrix}^\top = t \end{aligned}$$

where $\mathbf{r} \in \mathbb{R}^2$ is such that $\mathbf{p} = D\tilde{\phi}(\mathbf{u})\mathbf{r}$. Since $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \in T_{(\mathbf{u}, \mathbf{x})}S_c$ is arbitrary, we therefore have that $d\tilde{\theta} = \check{\rho}$. It is again consistent with the result that we have seen in the previous section.

To show that $\check{\rho}$ is the Riemann connection, we use (9.4.4) and (9.4.5). We have

$$\omega_1(\mathbf{u}, \mathbf{x}) = \cos(\tilde{\theta}) \rho_1(\mathbf{u}, \mathbf{x}) + \sin(\tilde{\theta}) \rho_2(\mathbf{u}, \mathbf{x})$$

and

$$\omega_2(\mathbf{u}, \mathbf{x}) = -\sin(\tilde{\theta}) \rho_1(\mathbf{u}, \mathbf{x}) + \cos(\tilde{\theta}) \rho_2(\mathbf{u}, \mathbf{x})$$

where $\tilde{\theta} = \tilde{\theta}(\mathbf{u}, \mathbf{x})$, $\rho_i = (\pi_S)^*(\tilde{\omega}_i)$ for $i = 1, 2$, and $\{\tilde{\omega}_1, \tilde{\omega}_2\}$ is the dual basis associated to the basis $\{(\mathbf{u}, \mathbf{e}_1), (\mathbf{u}, \mathbf{e}_2)\}$ of $T_{\mathbf{u}}S$. We have for $i = 1, 2$ that $\tilde{\omega}_i = dx_i$ on S because $\tilde{\omega}_i(\mathbf{u})(\mathbf{u}, \mathbf{p}) = \pi_i(\mathbf{p}) = dx_i(\mathbf{u})(\mathbf{u}, \mathbf{p})$ for all $(\mathbf{u}, \mathbf{p}) \in T_{\mathbf{u}}S$. We may therefore write

$$\omega_1(\mathbf{u}, \mathbf{x}) = \cos(\tilde{\theta}) (\pi_S)^*(dx_1)(\mathbf{u}, \mathbf{x}) + \sin(\tilde{\theta}) (\pi_S)^*(dx_2)(\mathbf{u}, \mathbf{x})$$

and

$$\omega_2(\mathbf{u}, \mathbf{x}) = -\sin(\tilde{\theta}) (\pi_S)^*(dx_1)(\mathbf{u}, \mathbf{x}) + \cos(\tilde{\theta}) (\pi_S)^*(dx_2)(\mathbf{u}, \mathbf{x}) .$$

Hence

$$\begin{aligned}
 d\omega_1 &= -\sin(\tilde{\theta}) d\tilde{\theta} \wedge (\pi_S)^*(dx_1) + \cos(\tilde{\theta}) d((\pi_S)^*(dx_1)) \\
 &\quad + \cos(\tilde{\theta}) d\tilde{\theta} \wedge (\pi_S)^*(dx_2) + \sin(\tilde{\theta}) d((\pi_S)^*(dx_2)) \\
 &= d(\tilde{\theta}) \wedge (-\sin(\tilde{\theta}) (\pi_S)^*(dx_1) + \cos(\tilde{\theta}) (\pi_S)^*(dx_2)) \\
 &\quad + \cos(\tilde{\theta}) (\pi_S)^*(d^2x_1) + \sin(\tilde{\theta}) (\pi_S)^*(d^2x_2) = d\tilde{\theta} \wedge \omega_2
 \end{aligned}$$

and similarly

$$d\omega_2 = -d\tilde{\theta} \wedge \omega_1.$$

It then follows from Theorem 9.4.7 that $\check{\rho} = d\tilde{\theta}$ is the Riemann connection on S_c .

We have already seen at the end of Example 9.3.7 that $d\check{\rho} = 0$. Hence, we get from the second structural equation that the curvature κ satisfies $\kappa = 0$ on all of S . ♣

Remark 9.4.10

Using the observation made in Remark 9.2.7 and the associated interpretation of vector field given in Remark 5.4.3, it is clear that $\check{\rho} = d\theta$ because $d\theta\left(\frac{\partial}{\partial\theta}\right) = \frac{\partial\theta}{\partial\theta} = 1$ and $d\theta\left(\frac{\partial}{\partial w_i}\right) = \frac{\partial\theta}{\partial w_i} = 0$ for $i = 0, 1$ since θ does not depend on $\mathbf{w}$ (i.e. on the base point $\mathbf{u}$.) ♠

9.5 Geodesic

Definition 9.5.1

Let S be a 2-dimensional Riemannian manifold and $\sigma : [a, b] \rightarrow S$ be a smooth curve. Let $\underline{\sigma} : [a, b] \rightarrow S_c$ be the smooth curve defined by $\underline{\sigma}(t) = (\sigma(t), \sigma'(t))$ for $a \leq t \leq b$. We say that $\underline{\sigma}$ is a **geodesic** in S if $\underline{\sigma}$ is the horizontal lift $\tilde{\sigma}$ of σ with $\tilde{\sigma}(a) = (\sigma(a), \sigma'(a))$.

If $\underline{\sigma}$ is a geodesic, then $(\sigma(t), \sigma'(t)) \in S_c(\sigma(t))$ and therefore $\|\sigma'(t)\|_{\sigma(t)} = 1$ for all t , where $\|\cdot\|_{\mathbf{u}}$ for $\mathbf{u} \in S$ is given by the Riemannian metric on S . Thus σ is parameterized by arc length.

Definition 9.5.2

Let S be a 2-dimensional Riemannian manifold and $\sigma : [a, b] \rightarrow S$ be a smooth curve parameterized by arc length. Moreover, let ρ be the Riemannian connection on S_c and $\underline{\sigma} : [a, b] \rightarrow S_c$ be the smooth curve defined by $\underline{\sigma}(t) = (\sigma(t), \sigma'(t))$ for $a \leq t \leq b$. The **geodesic curvature** of σ at t , denoted $\kappa_\sigma(t)$, is defined by $\kappa_\sigma(t) = \rho(\underline{\sigma}(t))(\underline{\sigma}_*(t, 1))$ for $a \leq t \leq b$.

We say that $\sigma : [a, b] \rightarrow S$ is a “straight line” in S if $\kappa_\sigma(t) = 0$ for $0 \leq t \leq b$. It then follows from Proposition 9.3.13 that $\underline{\sigma}$ is the horizontal lift $\tilde{\sigma}$ of σ with $\tilde{\sigma}(a) = (\sigma(a), \sigma'(a))$; namely, $\underline{\sigma}$ is a geodesic. In such a situation, the parallel translation of $\underline{\sigma}(a)$ along σ to $\sigma(t)$

is given by $(\sigma(t), \sigma'(t))$ itself. In the literature, geodesic are generally defined as curves with null geodesic curvature.

We have two definitions of curvature for a curve: one given in (6.5.3) and the geodesic curvature defined above. We prove that the curvature defined by (6.5.3) is the geodesic curvature. Without loss of generality, we may assume that $\sigma : [a, b] \rightarrow S \subset \mathbb{R}^3$ is a smooth curve parameterized by arc length.

We have that $\underline{\sigma}_*(t, 1) = ((\sigma(t), \sigma'(t)), (\sigma'(t), \sigma''(t)))$ for $a \leq t \leq b$. We also have for $(\mathbf{u}, \mathbf{x}) \in S_c(\mathbf{u})$ that $\rho(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) = \langle ((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})), H(\mathbf{u}, \mathbf{x}) \rangle_{(\mathbf{u}, \mathbf{x})}$ for $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \in T_{(\mathbf{u}, \mathbf{x})}S_c$ where we are using the metric on $T_{\mathbf{u}, \mathbf{x}}S_c$ defined by

$$\langle ((\mathbf{u}, \mathbf{x}), (\mathbf{p}_1, \mathbf{q}_1)), ((\mathbf{u}, \mathbf{x}), (\mathbf{p}_2, \mathbf{q}_2)) \rangle_{(\mathbf{u}, \mathbf{x})} = \langle \mathbf{p}_1, \mathbf{p}_2 \rangle + \langle \mathbf{q}_1, \mathbf{q}_2 \rangle = \mathbf{p}_1 \cdot \mathbf{p}_2 + \mathbf{q}_1 \cdot \mathbf{q}_2$$

for all $((\mathbf{u}, \mathbf{x}), (\mathbf{p}_i, \mathbf{q}_i)) \in T_{(\mathbf{u}, \mathbf{x})}S_c$ and $i = 1, 2$. If we derive $\|\sigma'(t)\|^2 = 1$ with respect to t , then we get that $\sigma'(t) \cdot \sigma''(t) = 0$ for all t . Thus $\sigma''(t)$ is perpendicular to $\sigma'(t)$ for all t . Hence the fourth component of $H(\sigma(t), \sigma'(t))$ is $\|\sigma''(t)\|^{-1} \sigma''(t)$; namely,

$$H(\sigma(t), \sigma'(t)) = ((\sigma(t), \sigma'(t)), (0, \|\sigma''(t)\|^{-1} \sigma''(t)))$$

(Figure 9.5). We find that the geodesic curvature of σ at t is

$$\begin{aligned} \kappa_\sigma(t) &= \rho(\underline{\sigma}(t))(\underline{\sigma}_*(t, 1)) \\ &= \left\langle ((\sigma(t), \sigma'(t)), (\sigma'(t), \sigma''(t))), ((\sigma(t), \sigma'(t)), (0, \|\sigma''(t)\|^{-1} \sigma''(t))) \right\rangle \\ &= \|\sigma''(t)\|^{-1} \sigma''(t) \cdot \sigma''(t) = \|\sigma''(t)\| \end{aligned}$$

for $a \leq t \leq b$. This is (6.5.3) for a smooth curve σ parameterized by arc length. Note that we have made use of the fact that $\|\sigma''(t) \times \sigma'(t)\| = \|\sigma'(t)\| \|\sigma''(t)\| \sin(\pi/2) = \|\sigma''(t)\|$ because $\sigma'(t)$ is perpendicular to $\sigma''(t)$, so the angle between them is $\pi/2$.

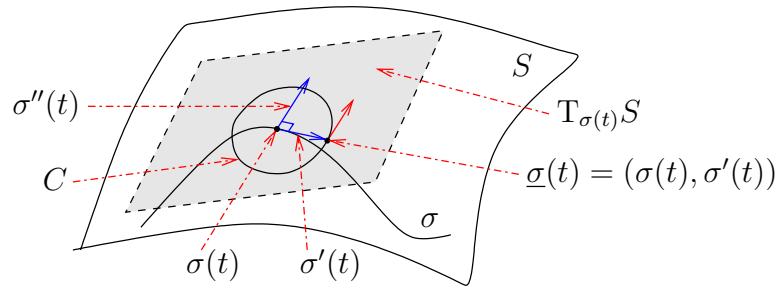


Figure 9.5: Representation of $H(\underline{\sigma}(t))$ along a smooth curve parameterized by arc length. The last component of $H(\underline{\sigma}(t)) \in T_{\underline{\sigma}(t)}S_c$ is represented by the red arrow. C is a circle of radius 1 centred at the origin in $T_{\sigma(t)}S$

9.6 Gauss-Bonnet Theorem

In this section, we use a few concepts from Sections 8.2 and 8.3 on simplicial homology and cohomology. The goal of this section is to prove the following amazing theorem.

Theorem 9.6.1 (Gauss-Bonnet)

Let (S, K, h) be a connected, oriented and smoothly triangulated 2-dimensional Riemannian manifold. Then $\frac{1}{2\pi} \int_S \kappa \, dV = \mathcal{X}(K)$ where $\mathcal{X}(K)$ is the Euler characteristic of the simplicial complex K used to triangulate S .

Recall that we have from Theorem 8.2.36 that $\mathcal{X}(K) = \alpha_0 - \alpha_1 + \alpha_2$ where α_0 is the number of vertices in K (i.e. 0-simplices), α_1 is the number of edges in K (i.e. 1-simplices), and α_2 is the number of faces or triangles in K (i.e. 2-simplices).

Since, for us, simplicial complexes contain only a finite number of elements, it follows from the definition of smoothly triangulated manifold that the manifolds to which the Gauss-Bonnet above can be applied to are compact manifolds like a sphere, a torus, etc.

We need to start with a fairly technical result.

Proposition 9.6.2

Let S be an oriented 2-dimensional Riemannian manifold, and $[[s]] = [[\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_3]]$ be an oriented 2-simplex. Suppose that

1. V is an open set such that $[s] \subset V \subset \mathbb{R}^2$,
2. $h : V \rightarrow S$ is a smooth, non-singular⁶, and orientation preserving function,
3. $\sigma : [a, b] \rightarrow S$ is a parametric representation of the piecewise C^∞ closed curve $h(\partial[s])$ such that the orientation on σ is the orientation on $h(\partial[s])$ induced via h by $\partial_2[[s]]$, and
4. $\tilde{\sigma} : [a, b] \rightarrow S_c$ is a horizontal lift of σ (Figure 9.6).

Then $\eta(\tilde{\sigma}(b)) = \tilde{\sigma}(a)$ for $\eta = e^{i\theta} \in S^1$ with $\theta = - \int_{h([s])} \kappa \, dV$.

Suppose that $[[\mathbf{x}_{i_0}, \mathbf{x}_{i_1}]]$ is the oriented 1-simplex associated to a 1-simplex $(\mathbf{x}_{i_0}, \mathbf{x}_{i_1})$ of a simplicial complex K . For convenience, we adopted the convention that $\int_{[[\mathbf{x}_{i_0}, \mathbf{x}_{i_1}]}} = \int_\alpha$ where α is the singular 1-cube defined by $\alpha(t) = \mathbf{x}_{i_0} + t(\mathbf{x}_{i_1} - \mathbf{x}_{i_0})$ for $0 \leq t \leq 1$ (or any equivalent parametric representation that respect the orientation). More generally, suppose that $[[\mathbf{x}_{i_0}, \mathbf{x}_{i_1}, \mathbf{x}_{i_2}]]$ is the oriented 2-simplex associated to a 2-simplex $(\mathbf{x}_{i_0}, \mathbf{x}_{i_1}, \mathbf{x}_{i_2})$ of a

⁶ h non-singular means that $Dh(\mathbf{u})$ is of rank 2 for all $\mathbf{u} \in V$.

simplicial complex K , then $\int_{\partial_s \llbracket \mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \mathbf{x}_{i_3} \rrbracket} = \int_{\alpha}$ where

$$\sigma(t) = \begin{cases} \mathbf{x}_{i_0} + t(\mathbf{x}_{i_1} - \mathbf{x}_{i_0}) & \text{if } 0 \leq t \leq 1 \\ \mathbf{x}_{i_1} + (t-1)(\mathbf{x}_{i_2} - \mathbf{x}_{i_1}) & \text{if } 1 < t \leq 2 \\ \mathbf{x}_{i_2} + (t-2)(\mathbf{x}_{i_0} - \mathbf{x}_{i_2}) & \text{if } 2 < t \leq 3 \end{cases}$$

(or any equivalent parametric representation that respect the orientation). We have

$\int_{\partial_s \llbracket \mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \mathbf{v}_{i_3} \rrbracket} = \int_{\llbracket \mathbf{v}_{i_0}, \mathbf{v}_{i_1} \rrbracket} - \int_{\llbracket \mathbf{v}_{i_0}, \mathbf{v}_{i_2} \rrbracket} + \int_{\llbracket \mathbf{v}_{i_1}, \mathbf{v}_{i_2} \rrbracket}$. All of these results could have come straight from Section 4.6 if instead of using singular k -cubes we had used triangles as it is sometime done in the literature. The two approaches are equivalent and have their own advantages and disadvantages.

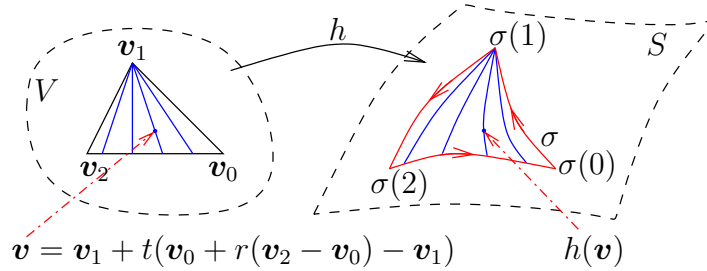


Figure 9.6: This is the first figure associated to the proof of Proposition 9.6.2. The curve σ is in red.

Proof (of Proposition 9.6.2).

Let ρ be the Riemannian connection on S_c given by Theorem 9.4.7.

Without loss of generality, we may use the parametric representation $\sigma : [0, 3] \rightarrow S$ of $h(\partial[s])$ defined by

$$\sigma(t) = \begin{cases} h(\mathbf{v}_0 + t(\mathbf{v}_1 - \mathbf{v}_0)) & \text{if } 0 \leq t \leq 1 \\ h(\mathbf{v}_1 + (t-1)(\mathbf{v}_2 - \mathbf{v}_1)) & \text{if } 1 < t \leq 2 \\ h(\mathbf{v}_2 + (t-2)(\mathbf{v}_0 - \mathbf{v}_2)) & \text{if } 2 < t \leq 3 \end{cases}$$

Suppose that

1. $\tilde{\sigma}(i) = (\sigma(i), \mathbf{x}_i) \in S_c(\sigma(i))$ for $0 \leq i \leq 2$,
2. $\tilde{\nu}_r$ for $0 \leq r \leq 1$ is the horizontal lift of the curve $\nu_r : [0, 1] \rightarrow S$ defined by $\nu_r(t) = h(\mathbf{v}_1 + t(\mathbf{v}_0 + r(\mathbf{v}_2 - \mathbf{v}_0) - \mathbf{v}_1))$ for $0 \leq t \leq 1$ such that $\tilde{\nu}_r(0) = (\sigma(1), \mathbf{x}_1)$, and
3. $\tilde{h} : [s] \rightarrow S_c$ is the map defined by $\tilde{h}(\mathbf{v}) = \tilde{\nu}_r(t)$ for $\mathbf{v} = \mathbf{v}_1 + t(\mathbf{v}_0 + r(\mathbf{v}_2 - \mathbf{v}_0) - \mathbf{v}_1)$ with $0 \leq r, t \leq 1$.

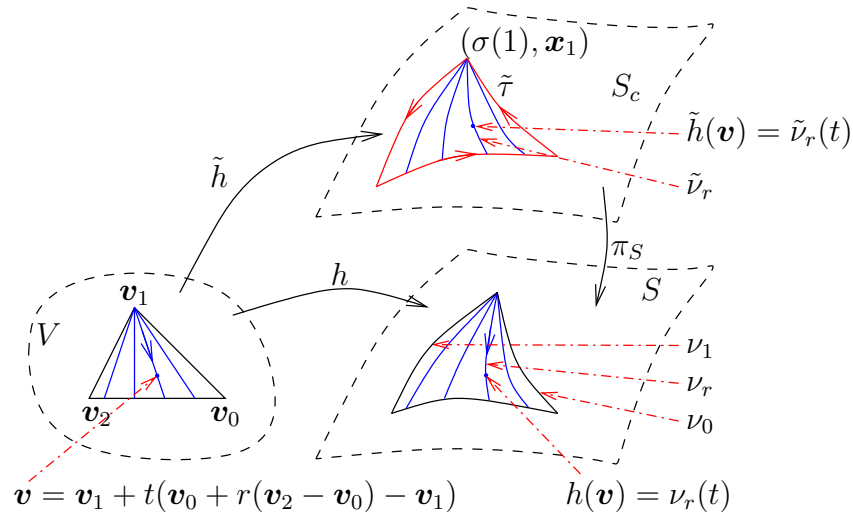


Figure 9.7: This is the second figure associated to the proof of Proposition 9.6.2. We provide a schematic representation of S_c which should not be literally interpreted as what S_c look like. The curve $\tilde{\sigma}$ is in red.

All these assumptions are summarized in Figure 9.7.

We have that $\pi_S \circ \tilde{h} = h$ on $[s]$. Moreover, $\tilde{h}$ is smooth and can be extended to a smooth map from an open set containing $[s]$ to S_c . We only sketch the proof of this claim. To prove that $\tilde{h}$ is a smooth function, a close look at the beginning of the proof of Proposition 9.3.13 shows that $\tilde{\nu}_r(t)$ is given locally by $\tilde{\nu}_r(t) = \eta_{t,r}(\tilde{\sigma}_r(t))$ where $(t, r) \mapsto \eta_{t,r} \in S^1$ is smooth and $\tilde{\sigma}_r(t) = G_1(\nu_r(t))$. It suffices to replace σ in that proof by ν_r . The smooth extension of $\tilde{h}$ to an open neighbourhood of $[s]$ is illustrated in Figure 9.8.

We have

$$\begin{aligned}
 \int_{h([s])} \kappa \, dV &= \int_{[s]} h^*(\kappa \, dV) = \int_{[s]} (\pi_S \circ \tilde{h})^*(\kappa \, dV) = \int_{[s]} \tilde{h}^*((\pi_S)^*(\kappa)(\pi_S)^*(dV)) \\
 &= \int_{[s]} \tilde{h}^*((\kappa \circ \pi_S) \omega_1 \wedge \omega_2) = - \int_{[s]} \tilde{h}^*(d\rho) = - \int_{[s]} d(\tilde{h}^*(\rho)) \\
 &= - \int_{\partial[[s]]} \tilde{h}^*(\rho) = - \int_{[[v_0, v_1]]} \tilde{h}^*(\rho) - \int_{[[v_1, v_2]]} \tilde{h}^*(\rho) - \int_{[[v_2, v_0]]} \tilde{h}^*(\rho) ,
 \end{aligned}$$

where we have use the second structural equation to obtain the fifth equality and Stokes' theorem to obtain the seventh equality.

Let $\tilde{\tau} : [0, 3] \rightarrow S_c$ be the curve defined by

$$\tilde{\tau}(t) = \begin{cases} \tilde{\nu}_0(1-t) & \text{if } 0 \leq t \leq 1 \\ \tilde{\nu}_1(t-1) & \text{if } 1 < t \leq 2 \\ \tilde{\nu}_{3-t}(1) & \text{if } 2 < t \leq 3 \end{cases}$$

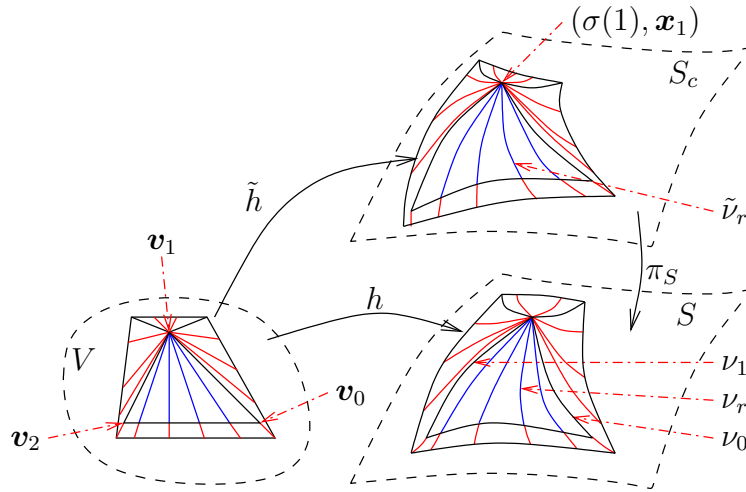


Figure 9.8: This is the third figure associated to the proof of Proposition 9.6.2. The extension of h and $\tilde{h}$ are given by the extension of the curves ν_r and the extension of their horizontal lifts shown in red.

(Figure 9.7). By uniqueness of the horizontal lift as we have seen in Proposition 9.3.13, we have that $\tilde{\tau}(t) = \tilde{\sigma}(t)$ for $0 \leq t \leq 2$. It suffices to compare the horizontal lifts of ν_0 and ν_1 with the horizontal lift of σ with the initial condition $(\sigma(1), \mathbf{x}_1)$ at $t = 0$. We get that $\tilde{\nu}_0(t) = \tilde{\sigma}(1 - t)$ and $\tilde{\nu}_1(t) = \tilde{\sigma}(t + 1)$ for $0 \leq t \leq 1$.

If $\alpha : [0, 1] \rightarrow [\mathbf{v}_0, \mathbf{v}_1]$ is the parametric representation given by $\alpha(t) = \mathbf{v}_0 + t(\mathbf{v}_1 - \mathbf{v}_0)$ for $0 \leq t \leq 1$, then

$$\begin{aligned} \int_{[\mathbf{v}_0, \mathbf{v}_1]} \tilde{h}^*(\rho) &= \int_0^1 \alpha^*(\tilde{h}^*(\rho)) dt = \int_0^1 (\tilde{h} \circ \alpha)^*(\rho) dt = \int_0^1 \tilde{\tau}^*(\rho) dt \\ &= \int_0^1 \rho(\tilde{\tau}(t))(\tilde{\tau}_*(t, 1)) dt = 0 \end{aligned}$$

because $\tilde{\tau}$ on $[0, 1]$ is a horizontal lift of σ and thus $\rho(\tilde{\tau}(t))(\tilde{\tau}_*(t, 1)) = 0$ for $t \in [0, 1]$ (see Proposition 9.3.13). For the same reason,

$$\int_{[\mathbf{v}_1, \mathbf{v}_2]} \tilde{h}^*(\rho) = 0.$$

Since $\tilde{\sigma}$ and $\tilde{\tau}$ are two lifts of σ on $[2, 3]$ such that $\tilde{\tau}(2) = \tilde{\sigma}(2) = (\sigma(2), \mathbf{x}_2)$, we have from Proposition 9.3.12 that there exists a smooth function $\theta : [2, 3] \rightarrow \mathbb{R}$ such that $\theta(2) = 0$, $\tilde{\tau}(t) = \eta_t(\tilde{\sigma}(t))$ with $\eta_t = e^{i\theta(t)} \in S^1$ for $2 \leq t \leq 3$, and $\rho(\tilde{\tau}(t))(\tilde{\tau}_*(t, 1)) = \theta'(t)$ for $2 \leq t \leq 3$. It follows that $\eta_3(\tilde{\sigma}(3)) = \tilde{\tau}(3) = \tilde{\nu}_0(1) = \tilde{\sigma}(0)$.

Moreover, if $\alpha : [2, 3] \rightarrow [\mathbf{v}_0, \mathbf{v}_2]$ is the parametric representation given by $\alpha(t) = \mathbf{v}_2 + (t - 2)(\mathbf{v}_0 - \mathbf{v}_2)$ for $]2, 3[$, then

$$\int_{h([s])} \kappa dV = - \int_{[\mathbf{v}_2, \mathbf{v}_0]} \tilde{h}^*(\rho) = - \int_{[2, 3]} \alpha^*(\tilde{h}^*(\rho)) dt = - \int_{[2, 3]} (\tilde{h} \circ \alpha)^*(\rho) dt$$

$$\begin{aligned}
&= - \int_{[2,3]} \tilde{\tau}^*(\rho) dt = - \int_{[2,3]} \rho(\tilde{\tau}(t)) (\tilde{\tau}_*(t, 1)) dt = - \int_{[2,3]} \theta'(t) dt \\
&= -\theta(3) + \theta(2) = -\theta(3) .
\end{aligned}$$

We get the conclusion of the proposition if we set $\eta = \eta_3$. ■

Remark 9.6.3

We could have chosen any other parametric representation of σ in the previous proposition. We chose the most convenient one.

From the Mean Value Theorem for integrals, we have that $\kappa(\sigma(a))$ is the limit of $\int_{h([s])} \kappa dV / \int_{h([s])} dV$ as $[s]$ shrink to $\mathbf{v}_0$.

There may be no rotation as stated in the previous proposition despite the fact that $\int_{h([s])} \kappa dV$ is not be null. We may have that $\int_{h([s])} \kappa dV$ is a multiple of 2π . ♠

Suppose that S be a 2-dimensional Riemannian manifold and that $\check{\sigma} : [0, L] \rightarrow S$ is a parametric representation by arc length of a piecewise C^∞ curve $\sigma : [a, b] \rightarrow S$ of length L such that $\sigma(t) \neq \mathbf{0}$ for all but a finite number of values of t ⁷. We set

$$\tau(\sigma) = \int_0^L \kappa_{\check{\sigma}}(t) dt ,$$

where $\kappa_{\check{\sigma}}(t)$ is the geodesic curvature of $\check{\sigma}$ at t defined in the previous section. We have that τ defines a linear mapping on the space of all piecewise C^∞ curves on S . If

$$\sigma(t) = \begin{cases} \sigma_1(t) & \text{if } a \leq t \leq c \\ \sigma_2(t) & \text{if } c < t \leq b \end{cases}$$

for some $c \in]a, b[$, and piecewise C^∞ curves $\sigma_1 : [a, c] \rightarrow S$ and $\sigma_2 : [c, b] \rightarrow S$, then $\tau(\sigma) = \tau(\sigma_1) + \tau(\sigma_2)$.

Suppose that:

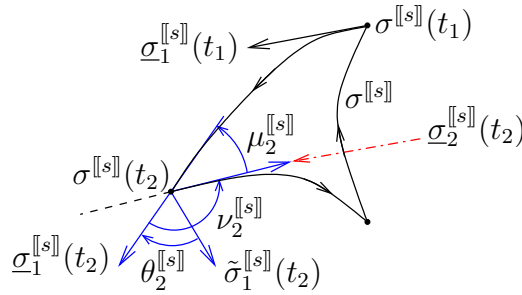
1. $h : V \rightarrow S$ is the smooth orientation preserving function and $[s] = [[\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_3]]$ is an oriented 2-simplex satisfying the hypotheses of Proposition 9.6.2,
2. $\sigma^{[s]} : [a, b] \rightarrow S$ is a parametric representation of the piecewise C^∞ closed curve $h(\partial[s])$ such that the orientation on $\sigma^{[s]}$ is the orientation on $h(\partial[s])$ induces via h by $\partial_2[s]$, and
3. $t_0 = a < t_1 < t_2 < t_3 = b$ are such that $\sigma^{[s]}$ is of class C^∞ on $[t_j, t_{j+1}]$ for $0 \leq j < 3$.

⁷If $\sigma'(t) \neq \mathbf{0}$ for all but a finite number of values of t , then we can define the arc length function $s : [a, b] \rightarrow [0, L]$ by defining it on each of the intervals where $\sigma'(t) \neq \mathbf{0}$ and summing them. As usual, the piecewise C^∞ parametric representation by arc length is given by $\check{\sigma} = \sigma \circ s^{-1} : [0, L] \rightarrow [a, b]$.

Let

1. $\sigma_j^{[s]} = \sigma^{[s]}|_{[t_j, t_{j+1}]}$ for $0 \leq j < 3$.
2. $\underline{\sigma}_j^{[s]} : [t_j, t_{j+1}] \rightarrow TS$ be the curve defined by $\underline{\sigma}_j^{[s]}(t) = (\sigma^{[s]}(t), (\sigma^{[s]})'(t))$ for $t_j \leq t \leq t_{j+1}$ and $0 \leq j < 3$, and
3. $\tilde{\sigma}_j^{[s]} : [t_j, t_{j+1}]$ be the horizontal lift of $\sigma_j^{[s]}$ with $\tilde{\sigma}_j^{[s]}(t_j) = \tilde{\sigma}_{j-1}^{[s]}(t_j)$ for $0 < j < 3$, and $\tilde{\sigma}_0^{[s]}(t_0) = \underline{\sigma}_0^{[s]}(t_0) = (\sigma^{[s]}(t_0), (\sigma^{[s]})'(t_0))$.

The internal angle $\mu_2^{[s]}$ and external angle $\nu_2^{[s]}$ at $\sigma^{[s]}(t_2)$ are defined in the following figure.



We define the internal angle $\mu_1^{[s]}$ and the external angle $\nu_1^{[s]}$ at $\sigma^{[s]}(t_1)$, and the internal angle $\mu_3^{[s]}$ and the external angle $\nu_3^{[s]}$ at $\sigma^{[s]}(t_3) = \sigma^{[s]}(t_0)$, in a similar fashion. An attention particular should be given to the directions of the angles at the vertex $\sigma^{[s]}(t_2)$ indicated by the directed arcs. All the blue vectors drawn from $\sigma^{[s]}(t_2)$ are in fact in $T_{\sigma^{[s]}(t_2)}S$ because all the edges are in S . As well, $\tilde{\sigma}_1^{[s]}(t_2) \in T_{\sigma^{[s]}(t_2)}S$ by construction. We have defined the positive directions of the interior angles as going from one edge to the previous edge on the 1-chain $\partial_2[s]$. The orientation of the 1-chain is illustrated in the figure above.

Lemma 9.6.4

Let S be an oriented 2-dimensional Riemannian manifold, and $[s] = [v_0, v_1, v_3]$ be an oriented 2-simplex. Suppose that

1. V is an open set such that $[s] \subset V \subset \mathbb{R}^2$,
2. $h : V \rightarrow S$ is a smooth, non-singular and orientation preserving function, and
3. $\sigma^{[s]} : [a, b] \rightarrow S$ is a parametric representation of the piecewise C^∞ closed curve $h(\partial[s])$ such that the orientation on $\sigma^{[s]}$ is the orientation on $h(\partial[s])$ induced via h by $\partial_2[s]$.

Then

$$\int_{h([s])} \kappa \, dV = -\pi + \sum_{j=1}^3 (\mu_j^{[s]} - \tau(\sigma_{j-1}^{[s]})) = -\pi + \sum_{j=1}^3 \mu_j^{[s]} - \tau(\sigma^{[s]}) .$$

Proof.

We may assume that $\sigma^{[s]}$ is a parametric representation by arc length because $\tau(\sigma_{j-1}^{[s]})$ is computed using $\sigma^{[s]}$ parameterized by arc length. Let $\tilde{\sigma}^{[s]} : [a, b] \rightarrow S_c$ be the horizontal lift of $\sigma^{[s]}$ with $\tilde{\sigma}^{[s]}(a) = \underline{\sigma}^{[s]}(a) = (\sigma^{[s]}(a), (\sigma^{[s]})'(a))$. We have from Proposition 9.6.2 that $\eta(\tilde{\sigma}^{[s]}(b)) = \tilde{\sigma}^{[s]}(a)$ for $\eta = e^{i\theta} \in S^1$ with $\theta = - \int_{h([s])} \kappa dV$. We also have from Proposition 9.3.12 that $\underline{\sigma}_{j-1}^{[s]}(t_j) = \eta_j(\tilde{\sigma}_{j-1}^{[s]}(t_j))$ with $\eta_j = e^{i\theta_{j-1}^{[s]}}$ where

$$\theta_{j-1}^{[s]} = \int_{t_{j-1}}^{t_j} \rho(\underline{\sigma}_{j-1}^{[s]}(r))((\underline{\sigma}_{j-1}^{[s]})_*(r, 1)) dr = \int_{t_{j-1}}^{t_j} \kappa_{\sigma_{j-1}^{[s]}}(r) dr = \tau(\sigma_{j-1}^{[s]})$$

for $t_{j-1} \leq t \leq t_i$ and $0 < j \leq 3$. Hence

$$\begin{aligned} \theta &= \sum_{j=1}^3 (\nu_j^{[s]} + \theta_j^{[s]}) - 2\pi k = \sum_{i=1}^3 \left(\pi - \mu_j^{[s]} + \tau(\sigma_{j-1}^{[s]}) \right) - 2\pi k \\ &= - \sum_{i=1}^3 \left(\mu_j^{[s]} - \tau(\sigma_{j-1}^{[s]}) \right) + (3 - 2k)\pi \end{aligned} \quad (9.6.1)$$

for some $k \in \mathbb{Z}$.

To prove that $k = 1$, we consider the continuous family of Riemannian metrics on $h(V) \subset S$ defined by $\eta^{[r]} = \eta^{[0]} + r(\eta - \eta^{[0]})$ where η is the Riemannian metric on S from the statement of the lemma and $\eta^{[0]}$ is the Riemannian metric on S induced from $\mathbb{R}^2$ using the function $h : V \rightarrow S$. Basically, we treat $h(V)$ and σ as V and $\partial_2[[s]]$ respectively.

For the Riemannian metric $\eta^{[0]}$, we have that $\kappa = 0$ on $h(V)$ because V is an open subset of $\mathbb{R}^2$ (see Example 9.4.9), and $\kappa_{\sigma_{j-1}} = 0$ for $0 < j \leq 3$ since $\partial_2[[s]]$ is piecewise linear⁸. Hence

(9.6.1) yields $0 = (3 - 2k)\pi - \sum_{j=0}^3 \mu_j^{[s]}$. Since the sum of the internal angles for a triangle is

π , we have that $\sum_{j=0}^3 \mu_j^{[s]} = \pi$. Thus $k = 1$. Since (9.6.1) is true for any Riemannian metric

and varies continuously as r goes from 0 to 1, we must have that k also varies continuously as r goes from 0 to 1. Note that κ and $\tau(\sigma_{j-1})$ are also varying continuously with respect to r ⁹. However, k can only take integer values. Thus $k = 1$ for all $r \in [0, 1]$; in particular for $r = 1$. This gives the conclusion of the lemma. ■

Proof (of Theorem 9.6.1).

Let (S, K, h) be the triple in the definition of a smoothly triangulated manifold, Definition 8.3.9. We assume that the orientation on K is such that h is an orientation preserving map. In particular, $h : (s) \rightarrow h((s))$ is orientation preserving for all 2-simplices $(s) \in K$.

⁸If we use (6.5.3) as we have justified at the end of Section 9.5, then we get that $\kappa_\sigma(t) = 0$ for all t in the interval.

⁹This should be explained more carefully. It is based on the fact that the differential 1-forms ω_1, ω_2 and ρ defined in Definition 9.4.1 and Theorem 9.4.7 vary continuously with respect to r . We leave the details to the reader.

By definition of a smoothly triangulated manifold, for every simplex $(s) = (\mathbf{x}_{i_0}, \mathbf{x}_{i_1}, \mathbf{x}_{i_2})$ of the simplicial complex K , there exist an open set $U_{(s)}$ in the affine space containing $\{\mathbf{x}_{i_0}, \mathbf{x}_{i_1}, \mathbf{x}_{i_2}\}$ and a map $h_{[s]} : U_{[s]} \rightarrow S$ of class C^∞ such that $[s] \subset U_{[s]}$ and $h_{[s]} = h$ on $[s]$. This will justify the use of Lemma 9.6.4 later on.

We first observe that every 1-simplex in K is the edge of exactly two 2-simplices in K . Hence, the image of every 1-simplex by h is the edge of the image of exactly two 2-simplices by h . In non rigorous terms, each edge in the triangulation of S is the edge of two triangles. We obviously have from the definition of simplicial complexes that a 1-simplex cannot be the edge of more than two 2-simplices. We need to determine that it must be two. Suppose that (s) is a 1-simplex in K and that $\mathbf{x} \in (s)$. Then $\mathbf{u} = h(\mathbf{x}) \in h((s))$. Let (W, U, ϕ) be a local chart about $\mathbf{u}$. We may assume that U is connected and small enough to have that $U \cap h((t)) = \emptyset$ for all 1-simplices $(t) \in K$ such that $(s) \neq (t)$. Therefore $U \setminus h((s))$ is the union of two disjoint open sets U_1 and U_2 in U (Figure 9.9). These two open sets are in the image by h of two distinct 2-simplices in K . To convince ourselves that this is true, we note that since U_i is connected and $U_i \cap h((t)) = \emptyset$ for all 1-simplices $(t) \in K$, we have that $U_i \subset h((t_i))$ for some 2-simplex $(t_i) \subset K$. Since the image of an open set by a homeomorphism is an open set, the sets $h((t))$ for (t) a 2-simplex of k are distinct open sets. Therefore, U_i cannot be a subset of more than one set $h((t_i))$ without contradicting the fact that U_i is connected. We must also have that $(t_1) \neq (t_2)$. If $(t_1) = (t_2)$, then $h((s)) \cap U \subset \partial U$ with $h((s)) \cap U \neq \emptyset$. This is a contradiction that U is open.

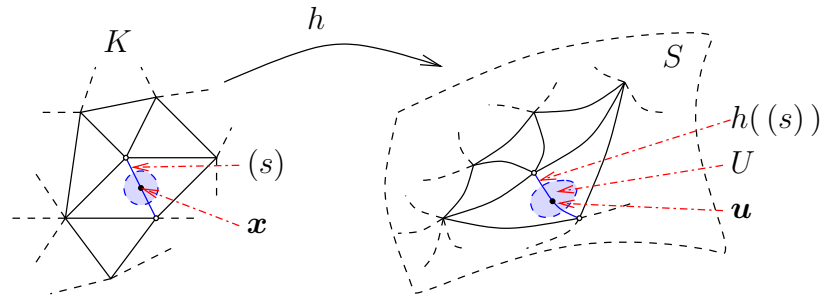


Figure 9.9: This figure is associated to the proof of Theorem 9.6.1. It is used to proof that every 1-simplex in K is the edge of exactly two 2-simplices in K . In non mathematical terms, each edge in the smoothly triangulated manifold S is the edge of two triangles.

From the previous paragraph, we deduce that $\alpha_1 = (3/2)\alpha_2$. Since the Euler characteristic is defined as $\mathcal{X}(K) = \alpha_0 - \alpha_1 + \alpha_2$, we get $\mathcal{X}(K) = \alpha_0 - (3/2)\alpha_2 + \alpha_2 = \alpha_0 - \alpha_2/2$.

It follows from Lemma 9.6.4

$$\frac{1}{2\pi} \int_S \kappa \, dV = \frac{1}{2\pi} \int_{h([K])} \kappa \, dV = \frac{1}{2\pi} \sum_{\substack{(s) \in K \text{ is} \\ \text{a 2-simplex}}} \int_{h([s])} \kappa \, dV$$

$$\begin{aligned}
&= \frac{1}{2\pi} \sum_{\substack{(s) \in K \text{ is} \\ \text{a 2-simplex}}} \left(-\pi + \sum_{j=1}^3 \left(\mu_j^{\llbracket s \rrbracket} - \tau(\sigma_{j-1}^{\llbracket s \rrbracket}) \right) \right) \\
&= -\frac{\alpha_2}{2} + \frac{1}{2\pi} \sum_{\substack{(s) \in K \text{ is} \\ \text{a 2-simplex}}} \left(\sum_{j=1}^3 \mu_j^{\llbracket s \rrbracket} \right) - \frac{1}{2\pi} \sum_{\substack{(s) \in K \text{ is} \\ \text{a 2-simplex}}} \tau(\sigma^{\llbracket s \rrbracket}) .
\end{aligned}$$

The last summation in the relation above is null because each edge of the triangulation of S is associated to exactly two triangles in the triangulation of S ; namely, each 1-simplex in K is the edge of exactly two 2-simplices in K . Suppose that $\sigma_{j_1-1}^{\llbracket s_1 \rrbracket}$ and $\sigma_{j_2-1}^{\llbracket s_2 \rrbracket}$ are the parametric representations of the common edge between the oriented 2-simplices $\llbracket s_1 \rrbracket$ and $\llbracket s_2 \rrbracket$. Then $\tau(\sigma_{j_2-1}^{\llbracket s_2 \rrbracket}) = -\tau(\sigma_{j_1-1}^{\llbracket s_1 \rrbracket})$ because the two parametric representations have opposite directions.

As for the sum in the middle, it is the sum of the interior angles of all the triangles at each vertex of the triangulation of S ; namely, at each 0-simplex in K . The sum at each vertex is 2π (Figure 9.10). Therefore, the sum in the middle is equal to $2\pi\alpha_0$.

Hence we get

$$\frac{1}{2\pi} \int_S \kappa \, dV = -\frac{\alpha_2}{2} + \alpha_0 = \mathcal{X}(K) . \quad \blacksquare$$

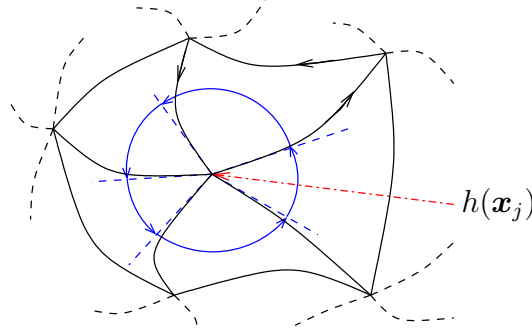


Figure 9.10: The internal angles of all the triangles at a vertex $h(\mathbf{x}_j)$ of the smoothly triangulated manifold S . Recall that all the dashed blue lines delimiting the interior angles are in $T_{h(\mathbf{x}_j)}S$ because all the edges are in S . By construction, we have defined the direction of all the interior angle as going from one edge to the previous edge on the 1-chain $\partial_2\llbracket s \rrbracket$ associated to the triangle, and this for all oriented 2-simplices $s \in K$. The orientation of one of the 1-chain is illustrated in the figure.

Example 9.6.5

We can use the Gauss-Bonnet theorem to prove that $\int_{S^2} \kappa \, dV = 4\pi$ without having to compute the integral.

Suppose that (S^2, K, h) is a smoothly triangulated manifold. We have from Proposition 8.6.11 that $H_j([K]; \mathbb{R}) \cong H_j(K; \mathbb{R})$ for all j , and from Corollary 8.4.16 that $H_j(S^2; \mathbb{R}) \cong H_j([K]; \mathbb{R})$ because $h : [K] \rightarrow S^2$ is a homeomorphism. Therefore, we may use the results in Example 8.4.82 to get the Betti numbers $\beta_j = \dim(H_j(K; \mathbb{R}))$ for $0 \leq j \leq 2$ that we need to compute the Euler characteristic $\mathcal{X}(K) = \sum_{j=0}^2 (-1)^j \beta_j$ as defined in Definition 8.2.35.

We have from Example 8.4.82 that $H_0(S^2; \mathbb{R}) \cong \mathbb{R}$, $H_1(S^2; \mathbb{R}) \cong 0$ and $H_2(S^2; \mathbb{R}) \cong \mathbb{R}$. Hence $\beta_0 = \beta_2 = 1$ and $\beta_1 = 0$. We get that $\mathcal{X}(K) = 2$ and thus $\int_{S^2} \kappa \, dV = 2\pi \mathcal{X}(K) = 4\pi$. ♣

Example 9.6.6

We can use the Gauss-Bonnet theorem to prove that $\int_{T^2} \kappa \, dV = 0$ without having to compute the integral.

As in the previous example, suppose that (T^2, K, h) is a smoothly triangulated manifold. We may use $H_j(T^2; \mathbb{R}) \cong H_j(K; \mathbb{R})$ for $0 \leq j \leq 2$ to get the Betti numbers $\beta_j = \dim(H_j(K; \mathbb{R}))$ for $0 \leq j \leq 2$. Since $H_j(T^2; \mathbb{R}) \cong H_j(S^1; \mathbb{R}) \times H_j(S^1; \mathbb{R})$ for $0 \leq j \leq 2$, we get from Example 8.4.82 that $H_0(T^2; \mathbb{R}) \cong \mathbb{R}^2$, $H_1(T^2; \mathbb{R}) \cong \mathbb{R}^2$ and $H_2(T^2; \mathbb{R}) \cong 0$. Hence $\beta_0 = \beta_1 = 2$ and $\beta_2 = 0$. We get that $\mathcal{X}(K) = 0$ and thus $\int_{S^2} \kappa \, dV = 2\pi \mathcal{X}(K) = 0$. ♣

We have seen in Proposition 7.3.4 that there is no nowhere null vector field on S^{k-1} if k is odd. We can extend this result using Gauss-Bonnet theorem.

Corollary 9.6.7

Let (S, K, h) be a connected, oriented and smoothly triangulated 2-dimensional Riemannian manifold. If S admits a nowhere null smooth vector field, then $\mathcal{X}(K) = 0$.

Proof.

Suppose that $F : S \rightarrow TS$ is a nowhere null smooth vector field. Let $f : S \rightarrow S_c$ be the smooth map defined by $f(\mathbf{u}) = (\|F\mathbf{u}\|_{\mathbf{u}})^{-1} F(\mathbf{u})$ where $\|\cdot\|_{\mathbf{u}}$ is the norm induced by the Riemannian metric $\tau : S \rightarrow \bigcup_{\mathbf{x} \in S} \mathcal{T}^2(T_{\mathbf{x}}S)$.

According to the second structural equation $d\rho = -(\pi_S)^*(\kappa \, dV)$ where V is the volume element on S . Hence

$$d(f^*\rho) = f^*(d\rho) = -f^*((\pi_S)^*(\kappa \, dV)) = -(\pi_S \circ f)^*(\kappa \, dV) = -\kappa \, dV$$

because $\pi_S \circ f = \text{Id}_S$. Thus $\kappa \, dV$ is an exact differential 2-form on S . It follows from Gauss-Bonnet theorem and Stokes' theorem that

$$\mathcal{X}(K) = \frac{1}{2\pi} \int_S \kappa \, dV = -\frac{1}{2\pi} \int_S d(f^*\rho) = -\frac{1}{2\pi} \int_{\partial S} f^*\rho = 0$$

because $\partial S = \emptyset$. ■

It follows from the previous corollary that the connected, oriented and smoothly triangulated 2-dimensional Riemannian manifold (S, K, h) with $\mathcal{X}(K) \neq 0$ do not admit a nowhere null vector field.

9.7 Lie Bracket

As announced at the end of Subsection 5.3.2, we are now going to make use of the definition of tangent spaces as spaces of differential linear operators. In particular, we will use the definition of vector fields given in Remark 5.4.3.

If F is a smooth vector field on a manifold S and $g \in C^\infty(S)$, then $F(g) : \mathbf{u} \mapsto F(\mathbf{u})(g)$ for $\mathbf{u} \in S$ defines a function in $C^\infty(S)$. Therefore, the following definition makes sense.

Definition 9.7.1

Let F_1, F_2 be two smooth vector fields on a manifold S . The **Lie bracket** of F_1 and F_2 , denoted $[F_1, F_2]$, is the smooth vector field on S defined by

$$[F_1, F_2](\mathbf{u})(g) = F_1(\mathbf{u})(F_2(g)) - F_2(\mathbf{u})(F_1(g))$$

for all $\mathbf{u} \in S$ and $g \in C^\infty(S)$.

It is not hard to prove that

1. $[F_1, F_2] = -[F_2, F_1]$,
2. $[F_1 + F_2, F_3] = [F_1, F_3] + [F_2, F_3]$,
3. $[\lambda F_1, F_2] = \lambda[F_1, F_2]$ and
4. $[[F_1, F_2], F_3] + [[F_2, F_3], F_1] + [[F_3, F_1], F_2] = 0$

for all vector field F_i on S and $\lambda \in \mathbb{R}$ with $1 \leq i \leq 3$. The fourth identity above is called the **Jacobi identity**. For those interested in Lie algebra, the four properties above are those required to say that the space of smooth vector fields on S is a Lie algebra. There is a vast literature on this subject.

Note that if ω is a smooth differential 1-form on S and F is a smooth vector field on S , then $\omega(F) : S \rightarrow \mathbb{R}$ defined by $\omega(F)(\mathbf{u}) = \omega(\mathbf{u})(F(\mathbf{u}))$ for all $\mathbf{u} \in S$ is a function of class C^∞ .

Proposition 9.7.2

Suppose that ω is a smooth differential 1-form on a smooth k -dimensional manifold S , and that F_1 and F_2 are two smooth vector fields on S . Then

$$d\omega(\mathbf{u})(F_1(\mathbf{u}), F_2(\mathbf{u})) = F_1(\mathbf{u})(\omega(F_2)) - F_2(\mathbf{u})(\omega(F_1)) - \omega(\mathbf{u})([F_1, F_2](\mathbf{u})) \quad (9.7.1)$$

for all $\mathbf{u} \in S$.

Proof.

As usual, we only need to prove the result for a local chart (W, U, ϕ) of S . We have that $\phi^*(\omega) = \sum_{i=1}^k \omega_i du_i$ where $w_i : W \rightarrow \mathbb{R}$ are smooth functions. Hence ω is locally the sum of terms of the form $h dg$ where $h, g : U \rightarrow \mathbb{R}$ are smooth functions. Since both sides of (9.7.1) are linear relative to ω , we only have to prove (9.7.1) for $h dg$.

Using Definition 4.1.11 and the definition of Alt given in Proposition 4.1.9 and referring to Remark 5.4.3, we have that the left hand side of (9.7.1) is

$$\begin{aligned} d\omega(\mathbf{u})(F_1(\mathbf{u}), F_2(\mathbf{u})) &= (dh \wedge dg)(\mathbf{u})(F_1(\mathbf{u}), F_2(\mathbf{u})) \\ &= dh(\mathbf{u})(F_1(\mathbf{u})) dg(\mathbf{u})(F_2(\mathbf{u})) - dh(\mathbf{u})(F_2(\mathbf{u})) dg(\mathbf{u})(F_1(\mathbf{u})) \\ &= F_1(\mathbf{u})(h) F_2(\mathbf{u})(g) - F_2(\mathbf{u})(h) F_1(\mathbf{u})(g) \end{aligned}$$

for $\mathbf{u} \in U$.

For the right hand side of (9.7.1), we have

$$\begin{aligned} F_1(\mathbf{u})(\omega(F_2)) - F_2(\mathbf{u})(\omega(F_1)) - \omega(\mathbf{u})([F_1, F_2](\mathbf{u})) \\ &= F_1(\mathbf{u})(h dg(F_2)) - F_2(\mathbf{u})(h dg(F_1)) - h(\mathbf{u}) dg(\mathbf{u})([F_1, F_2](\mathbf{u})) \\ &= F_1(\mathbf{u})(h F_2(g)) - F_2(\mathbf{u})(h F_1(g)) - h(\mathbf{u}) [F_1, F_2](\mathbf{u})(g) \\ &= (F_1(\mathbf{u})(h) F_2(\mathbf{u})(g) + h(\mathbf{u}) F_1(\mathbf{u})(F_2(g))) - (F_2(\mathbf{u})(h) F_1(\mathbf{u})(g) + h(\mathbf{u}) F_2(\mathbf{u})(F_1(g))) \\ &\quad - h(\mathbf{u}) (F_1(\mathbf{u})(F_2(g)) - F_2(\mathbf{u})(F_1(g))) = F_1(\mathbf{u})(h) F_2(\mathbf{u})(g) - F_2(\mathbf{u})(h) F_1(\mathbf{u})(g) \end{aligned}$$

for $\mathbf{u} \in U$, where the third equality comes from the product rule applied to the operators $F_1(\mathbf{u})$ and $F_2(\mathbf{u})$, and the definition of $[F_1, F_2]$.

Therefore, the left and right sides of (9.7.1) are equal. ■

Let F_1, F_2 and F_3 be vector fields on S_c such that $\{\omega_1(\mathbf{u}, \mathbf{x}), \omega_2(\mathbf{u}, \mathbf{x}), \rho(\mathbf{u}, \mathbf{x})\}$ is the dual basis associated to $\{F_1(\mathbf{u}, \mathbf{x}), F_2(\mathbf{u}, \mathbf{x}), F_3(\mathbf{u}, \mathbf{x})\}$ for all $(\mathbf{u}, \mathbf{x}) \in S_c$. Recall that $F_3 = H$ defined in Section 9.2.3 because ρ is a connection 1-form. We get from (9.7.1) that

$$\begin{aligned} d\omega_1(\mathbf{u}, \mathbf{x})(F_1(\mathbf{u}, \mathbf{x}), F_2(\mathbf{u}, \mathbf{x})) \\ &= F_1(\mathbf{u}, \mathbf{x})(\omega_1(F_2)) - F_2(\mathbf{u}, \mathbf{x})(\omega_1(F_1)) - \omega_1(\mathbf{u}, \mathbf{x})([F_1, F_2](\mathbf{u}, \mathbf{x})) \\ &= -\omega_1(\mathbf{u}, \mathbf{x})([F_1, F_2](\mathbf{u}, \mathbf{x})) . \end{aligned}$$

Moreover

$$\begin{aligned} (\rho \wedge \omega_2 + c_1 \omega_1 \wedge \omega_2)(\mathbf{u}, \mathbf{x})(F_1(\mathbf{u}, \mathbf{x}), F_2(\mathbf{u}, \mathbf{x})) \\ &= \rho(\mathbf{u}, \mathbf{x})(F_1(\mathbf{u}, \mathbf{x})) \omega_2(\mathbf{u}, \mathbf{x})(F_2(\mathbf{u}, \mathbf{x})) - \rho(\mathbf{u}, \mathbf{x})(F_2(\mathbf{u}, \mathbf{x})) \omega_2(\mathbf{u}, \mathbf{x})(F_1(\mathbf{u}, \mathbf{x})) \\ &\quad + c_1(\mathbf{u}, \mathbf{x}) \omega_1(\mathbf{u}, \mathbf{x})(F_1(\mathbf{u}, \mathbf{x})) \omega_2(\mathbf{u}, \mathbf{x})(F_2(\mathbf{u}, \mathbf{x})) \\ &\quad - c_1(\mathbf{u}, \mathbf{x}) \omega_1(\mathbf{u}, \mathbf{x})(F_2(\mathbf{u}, \mathbf{x})) \omega_2(\mathbf{u}, \mathbf{x})(F_1(\mathbf{u}, \mathbf{x})) = c_1(\mathbf{u}, \mathbf{x}) . \end{aligned}$$

If we substitute these two expressions in the first of the first structural equations in Definition 9.4.6, then we get

$$\omega_1(\mathbf{u}, \mathbf{x})([F_1, F_2](\mathbf{u}, \mathbf{x})) = -c_1(\mathbf{u}, \mathbf{x}) . \quad (9.7.2)$$

Computations similar to those above give

$$d\omega_2(\mathbf{u}, \mathbf{x})(F_1(\mathbf{u}, \mathbf{x}), F_2(\mathbf{u}, \mathbf{x})) = -\omega_2(\mathbf{u}, \mathbf{x})([F_1, F_2](\mathbf{u}, \mathbf{x}))$$

and

$$(-\rho \wedge \omega_1 + c_2 \omega_1 \wedge \omega_2)(\mathbf{u}, \mathbf{x})(F_1(\mathbf{u}, \mathbf{x}), F_2(\mathbf{u}, \mathbf{x})) = c_2(\mathbf{u}, \mathbf{x}) .$$

If we substitute these two expressions in the second of the first structural equations in Definition 9.4.6, then we get

$$\omega_2(\mathbf{u}, \mathbf{x})([F_1, F_2](\mathbf{u}, \mathbf{x})) = -c_2(\mathbf{u}, \mathbf{x}) . \quad (9.7.3)$$

We get from (9.7.1) that

$$\begin{aligned} d\rho(\mathbf{u}, \mathbf{x})(F_1(\mathbf{u}, \mathbf{x}), F_2(\mathbf{u}, \mathbf{x})) \\ &= F_1(\mathbf{u}, \mathbf{x})(\rho(F_2)) - F_2(\mathbf{u}, \mathbf{x})(\rho(F_1)) - \rho(\mathbf{u}, \mathbf{x})([F_1, F_2](\mathbf{u}, \mathbf{x})) \\ &= -\rho(\mathbf{u}, \mathbf{x})([F_1, F_2](\mathbf{u}, \mathbf{x})) . \end{aligned}$$

Moreover

$$\begin{aligned} &(-(\kappa \circ \pi_S) \omega_1 \wedge \omega_2)(\mathbf{u}, \mathbf{x})(F_1(\mathbf{u}, \mathbf{x}), F_2(\mathbf{u}, \mathbf{x})) \\ &= -\kappa(\mathbf{u}) \left(\omega_1(\mathbf{u}, \mathbf{x})(F_1(\mathbf{u}, \mathbf{x})) \omega_2(\mathbf{u}, \mathbf{x})(F_2(\mathbf{u}, \mathbf{x})) \right. \\ &\quad \left. - \omega_1(\mathbf{u}, \mathbf{x})(F_2(\mathbf{u}, \mathbf{x})) \omega_2(\mathbf{u}, \mathbf{x})(F_1(\mathbf{u}, \mathbf{x})) \right) = -\kappa(\mathbf{u}) . \end{aligned}$$

If we substitute these two expressions in the second structural equation in Definition 9.4.8, then we get

$$\rho(\mathbf{u}, \mathbf{x})([F_1, F_2](\mathbf{u}, \mathbf{x})) = \kappa(\mathbf{u}) . \quad (9.7.4)$$

Proceeding as we did above with F_1, F_2 replaced by H, F_1 and H, F_2 , we get the following six equations.

$$\omega_1(\mathbf{u}, \mathbf{x})([H, F_1](\mathbf{u}, \mathbf{x})) = 0 , \quad (9.7.5)$$

$$\omega_2(\mathbf{u}, \mathbf{x})([H, F_1](\mathbf{u}, \mathbf{x})) = 1 , \quad (9.7.6)$$

$$\rho(\mathbf{u}, \mathbf{x})([H, F_1](\mathbf{u}, \mathbf{x})) = 0 , \quad (9.7.7)$$

$$\omega_1(\mathbf{u}, \mathbf{x})([H, F_2](\mathbf{u}, \mathbf{x})) = -1 , \quad (9.7.8)$$

$$\omega_2(\mathbf{u}, \mathbf{x})([H, F_2](\mathbf{u}, \mathbf{x})) = 0 \quad (9.7.9)$$

and

$$\rho(\mathbf{u}, \mathbf{x})([H, F_2](\mathbf{u}, \mathbf{x})) = 0 . \quad (9.7.10)$$

It follows from (9.7.2) to (9.7.4) that

$$[F_1, F_2](\mathbf{u}, \mathbf{x}) = \omega_1(\mathbf{u}, \mathbf{x})([F_1, F_2](\mathbf{u}, \mathbf{x})) F_1(\mathbf{u}, \mathbf{x}) + \omega_2(\mathbf{u}, \mathbf{x})([F_1, F_2](\mathbf{u}, \mathbf{x})) F_2(\mathbf{u}, \mathbf{x})$$

$$\begin{aligned}
& + \rho(\mathbf{u}, \mathbf{x})([F_1, F_2](\mathbf{u}, \mathbf{x})) H(\mathbf{u}, \mathbf{x}) \\
& = -c_1(\mathbf{u}, \mathbf{x}) F_1(\mathbf{u}, \mathbf{x}) - c_2(\mathbf{u}, \mathbf{x}) F_2(\mathbf{u}, \mathbf{x}) + \kappa(\mathbf{u}) H(\mathbf{u}, \mathbf{x})
\end{aligned}$$

for all $(\mathbf{u}, \mathbf{x}) \in S_c$. Similarly, it follows from (9.7.5) to (9.7.7) and from (9.7.8) to (9.7.10) that

$$\begin{aligned}
[H, F_1](\mathbf{u}, \mathbf{x}) &= \omega_1(\mathbf{u}, \mathbf{x})([H, F_1](\mathbf{u}, \mathbf{x})) F_1(\mathbf{u}, \mathbf{x}) + \omega_2(\mathbf{u}, \mathbf{x})([H, F_1](\mathbf{u}, \mathbf{x})) F_2(\mathbf{u}, \mathbf{x}) \\
& + \rho(\mathbf{u}, \mathbf{x})([H, F_1](\mathbf{u}, \mathbf{x})) H(\mathbf{u}, \mathbf{x}) = F_2(\mathbf{u}, \mathbf{x})
\end{aligned}$$

and

$$\begin{aligned}
[H, F_2](\mathbf{u}, \mathbf{x}) &= \omega_1(\mathbf{u}, \mathbf{x})([H, F_2](\mathbf{u}, \mathbf{x})) F_1(\mathbf{u}, \mathbf{x}) + \omega_2(\mathbf{u}, \mathbf{x})([H, F_2](\mathbf{u}, \mathbf{x})) F_2(\mathbf{u}, \mathbf{x}) \\
& + \rho(\mathbf{u}, \mathbf{x})([H, F_2](\mathbf{u}, \mathbf{x})) H(\mathbf{u}, \mathbf{x}) = -F_1(\mathbf{u}, \mathbf{x})
\end{aligned}$$

respectively for all $(\mathbf{u}, \mathbf{x}) \in S_c$. Hence, the first and second structural equations formulated in terms of vector fields F_1 , F_2 and H are

$$[F_1, F_2] = -c_1 F_1 - c_2 F_2 + (\kappa \circ \pi_S) H \quad , \quad [H, F_1] = F_2 \quad \text{and} \quad [H, F_2] = -F_1$$

on S_c . In the special case where ρ is the Riemann connection on S_c , then $[F_1, F_2] = (\kappa \circ \pi_S) H$.

Remark 9.7.3

It follows from the definition of ω_1 and ω_2 in Definition 9.4.1 that

$$(\pi_S)_*(F_1(\mathbf{u}, \mathbf{x})) = \omega_1((\mathbf{u}, \mathbf{x}))(F_1(\mathbf{u}, \mathbf{x}))(\mathbf{u}, \mathbf{x}) + \omega_2((\mathbf{u}, \mathbf{x}))(F_1(\mathbf{u}, \mathbf{x}))v(\mathbf{u}, \mathbf{x}) = (\mathbf{u}, \mathbf{x})$$

and, similarly, $(\pi_S)_*(F_2(\mathbf{u}, \mathbf{x})) = v(\mathbf{u}, \mathbf{x})$ where $v = e^{i\pi/2} \in S^1$. Moreover, $F_1(\mathbf{u}, \mathbf{x})$ and $F_2(\mathbf{u}, \mathbf{x})$ are orthogonal to $H(\mathbf{u}, \mathbf{x})$ because $\rho(\mathbf{u}, \mathbf{x})(F_1(\mathbf{u}, \mathbf{x})) = \rho(\mathbf{u}, \mathbf{x})(F_2(\mathbf{u}, \mathbf{x})) = 0$. These conditions determine uniquely $F_1(\mathbf{u}, \mathbf{x})$ and $F_2(\mathbf{u}, \mathbf{x})$ for all $(\mathbf{u}, \mathbf{x}) \in S_c$. ♠

9.7.1 An Equivalent Definition of the Lie Bracket

There is another way to define the Lie bracket that is often used in the literature. First, we have to reformulate the definition of an **integral curve** at $\mathbf{u} \in S$ for a vector field $F : S \rightarrow TS$. It is a function $\sigma_{\mathbf{u}} : I \rightarrow S$ where I is an open neighbourhood of the origin such that $\sigma_{\mathbf{u}}(0) = \mathbf{u}$ and $\sigma_{\mathbf{u}}^* \left(\frac{d}{dt} \Big|_t \right) = F(\sigma_{\mathbf{u}}(t))$ for all $t \in I$; namely,

$$\sigma_{\mathbf{u}}^* \left(\frac{d}{dt} \Big|_t \right) (g) = \frac{d}{dt} (g \circ \sigma_{\mathbf{u}})(t) = F(\sigma_{\mathbf{u}}(t))(g) \quad (9.7.11)$$

for all $g \in C^\infty(S)$ and $t \in I$. Note that the interval of existence I may vary with $\mathbf{u}$. Let (W, U, ϕ) be a local chart of S . It follows from (9.7.11) that $\frac{d}{dt}(\phi^{-1} \circ \sigma_{\mathbf{u}}) = \tilde{F}(\phi^{-1} \circ \sigma_{\mathbf{u}})$ where

$$\tilde{F}(\mathbf{w}) = \sum_{j=1}^k \tilde{f}_j(\mathbf{w}) \frac{\partial}{\partial w_j} \quad \text{for } \mathbf{w} \in W \text{ is the local representation of } F. \text{ This is expected if (9.7.11)}$$

is equivalent to the condition in our previous definition of integral curves in Definition 5.4.2.

We have from the theorem of existence and uniqueness of solutions for ordinary differential equations that, by shrinking U if necessary, there exists an open interval I containing the origin such that the integral curve $\sigma_{\mathbf{u}}$ at any $\mathbf{u} \in U$ exists on I and $\sigma_{\mathbf{u}}(t) \in S$ for all $t \in I$. Since we are mostly working on local charts, we will assume that this is always satisfied. If we set $X(t, \mathbf{u}) = \sigma_{\mathbf{u}}(t)$ for $\mathbf{u} \in U$ and $t \in I$, then we get what is called a **flow**. It is proved in the theory of ordinary differential equations that $X : I \times U \rightarrow S$ is smooth if the vector field F is smooth (see [1]).

Suppose that $F_1, F_2 : S \rightarrow TS$ are two smooth vector fields. Moreover, suppose that $X : I \times U \rightarrow S$ is the flow associated to the vector field F_1 as explained above. Then $\sigma_{\mathbf{u}}(t) = X_{\mathbf{u}}(t) = X(t, \mathbf{u})$ for $t \in I$ satisfies $\sigma_{\mathbf{u}}^* \left(\frac{d}{dt} \Big|_t \right) = F_1(\sigma_{\mathbf{u}}(t))$ for all $t \in I$ and $\sigma_{\mathbf{u}}(0) = \mathbf{u}$. The **Lie bracket** of F_1 and F_2 is the vector field defined by

$$[F_1, F_2](\mathbf{u}) = \frac{d}{dt} (X_t^*(F_2)(\mathbf{u})) \Big|_{t=0} \quad (9.7.12)$$

where $X_t(\mathbf{u}) = X(t, \mathbf{u})$ for all $\mathbf{u} \in S$. We have from the theory of ordinary differential equations that $X_t : U \rightarrow S$ maps open subsets of U to open subsets of S ; in particular, X_t maps open neighbourhoods of $\mathbf{u} \in U$ to open neighbourhoods of $X_t(\mathbf{u})$ in S . Note that we have use for the first time the pull-back of a vector field.

Let (W, U, ϕ) be a local chart of S . Suppose that the local representation of the vector field F_i is defined by $\tilde{F}_i(\mathbf{w}) = \sum_{j=1}^k (\tilde{f}_i)_j(\mathbf{w}) \frac{\partial}{\partial w_j} \Big|_{\mathbf{w}}$ for $\mathbf{w} \in W$ and $i = 1, 2$, where $(\tilde{f}_i)_j : W \rightarrow \mathbb{R}^k$ for $1 \leq j \leq k$ and $i = 1, 2$ are smooth functions.

The flow $\tilde{X}$ on W associated to the vector field $\tilde{F}_1$ is given by $\phi(\tilde{X}(t, \mathbf{w})) = X(t, \phi(\mathbf{w}))$ for all $\mathbf{w} \in W$ and $|t|$ small enough. This follows from the fact that $X_{\mathbf{u}}^* \left(\frac{d}{dt} \Big|_t \right) = F_1(X_{\mathbf{u}}(t))$ implies that

$$\begin{aligned} \tilde{X}_{\mathbf{w}}^* \left(\frac{d}{dt} \Big|_t \right) (g \circ \phi) &= \frac{d}{dt} g(\phi(\tilde{X}_{\mathbf{w}}(t))) = \frac{d}{dt} g(X_{\phi(\mathbf{w})}(t)) = X_{\phi(\mathbf{w})}^* \left(\frac{d}{dt} \Big|_t \right) (g) \\ &= F_1(X_{\phi(\mathbf{w})}(t))(g) = \sum_{j=1}^k (\tilde{f}_1)_j(\phi^{-1}(X_{\phi(\mathbf{w})}(t))) \frac{\partial}{\partial w_j} (g \circ \phi) \Big|_{\mathbf{w}=\phi^{-1}(X_{\phi(\mathbf{w})}(t))} \\ &= \sum_{j=1}^k (\tilde{f}_1)_j(\tilde{X}_{\mathbf{w}}(t)) \frac{\partial}{\partial w_j} (g \circ \phi) \Big|_{\mathbf{w}=\tilde{X}_{\mathbf{w}}(t)} = \tilde{F}_1(\tilde{X}_{\mathbf{w}}(t))(g \circ \phi) \end{aligned}$$

for all $\phi \in C^\infty(S)$.

We have that $\frac{d}{dt} \tilde{X}(t, \mathbf{w}) = \tilde{f}_1(\tilde{X}(t, \mathbf{w}))$ is equivalent to $\tilde{X}_{\mathbf{w}}^* \left(\frac{d}{dt} \Big|_t \right) (\tilde{g}) = \tilde{F}_1(\tilde{X}_{\mathbf{w}}(t))(\tilde{g})$ for all $\tilde{g} \in C^\infty(W)$ because

$$\tilde{X}_{\mathbf{w}}^* \left(\frac{d}{dt} \Big|_t \right) (\tilde{g}) = \frac{d}{dt} (\tilde{g}(\tilde{X}(t, \mathbf{w}))) = D \tilde{g}(\tilde{X}(t, \mathbf{w})) \frac{d}{dt} \tilde{X}(t, \mathbf{w})$$

and

$$\begin{aligned}\tilde{F}_1(\tilde{X}_{\mathbf{w}}(t))(\tilde{g}) &= \tilde{F}_1(\tilde{X}(t, \mathbf{w}))(\tilde{g}) = \sum_{j=1}^k (\tilde{f}_1)_j(\tilde{X}(t, \mathbf{w})) \frac{\partial \tilde{g}}{\partial w_j}(\tilde{X}(t, \mathbf{w})) \\ &= D \tilde{g}(\tilde{X}(t, \mathbf{w})) \tilde{f}_1(\tilde{X}(t, \mathbf{w}))\end{aligned}$$

for all $\tilde{g} \in C^\infty(W)$ and $|t|$ small enough.

The local representation of the vector field $Q_t = X_t^*(F_2)$ is given by

$$\tilde{Q}_t(\mathbf{w}) = \sum_{j=1}^k (\tilde{q}_t)_j(\mathbf{w}) \frac{\partial}{\partial w_j} \text{ where } \tilde{q}_t(\mathbf{w}) = D \tilde{X}_{-t}(\tilde{X}_t(\mathbf{w})) \tilde{f}_2(\tilde{X}_t(\mathbf{w})) \text{ for } \mathbf{w} \in W \text{ and } t \text{ small}$$

enough to have $\tilde{X}_t(\mathbf{w}) \in W$. To prove this claim, it is more convenient to use our standard definition of vector fields and the isomorphism with the definition in terms of differential linear operators. We have

$$\begin{aligned}X_t^*(F_2)(\mathbf{u}) &= (X_t^{-1})_*(F_2(X_t(\mathbf{u}))) = (X_{-t})_*(F_2(X_t(\mathbf{u}))) = (X_{-t})_*(\phi_*(\tilde{F}_2(\phi^{-1}(X_t(\mathbf{u})))) \\ &= (X_{-t} \circ \phi)_*(\tilde{F}_2(\phi^{-1}(X_t(\mathbf{u})))) = (\mathbf{u}, D(X_{-t} \circ \phi)(\phi^{-1}(X_t(\mathbf{u}))) \tilde{f}_2(\phi^{-1}(X_t(\mathbf{u})))) \\ &= (\mathbf{u}, D \phi(\phi^{-1}(\mathbf{u})) D(\phi^{-1} \circ X_{-t} \circ \phi)(\phi^{-1}(X_t(\mathbf{u}))) \tilde{f}_2(\phi^{-1}(X_t(\mathbf{u})))) \\ &= (\mathbf{u}, D \phi(\phi^{-1}(\mathbf{u})) D \tilde{X}_{-t}(\phi^{-1}(X_t(\mathbf{u}))) \tilde{f}_2(\phi^{-1}(X_t(\mathbf{u})))) \\ &= (\mathbf{u}, D \phi(\phi^{-1}(\mathbf{u})) D \tilde{X}_{-t}(\tilde{X}_t(\phi^{-1}(\mathbf{u}))) \tilde{f}_2(\tilde{X}_t(\phi^{-1}(\mathbf{u}))))\end{aligned}$$

for $\mathbf{u} \in U$ and t small enough. If $\tilde{Q}_t(\mathbf{w}) = (\mathbf{w}, \tilde{q}_t(\mathbf{w}))$ for $\mathbf{w} \in W$ is the local representation of Q , then we get from the previous relation that

$$\tilde{q}_t(\mathbf{w}) = D \tilde{X}_{-t}(\tilde{X}_t(\mathbf{w})) \tilde{f}_2(\tilde{X}_t(\mathbf{w}))$$

for $\mathbf{w} \in W$ and $|t|$ small enough.

Since the local representation of the vector field $Q_t = X_t^*(F_2)$ is given by

$$\begin{aligned}&\sum_{j=1}^k (\tilde{q}_t)_j(\mathbf{w}) \frac{\partial}{\partial w_j} \Big|_{\mathbf{w}=\phi^{-1}(\mathbf{u})}, \text{ the local representation of the vector field } \frac{d}{dt}(X_t^*(F_2)) \Big|_{t=0} \text{ is given} \\ &\text{by } \sum_{i=1}^k \frac{d}{dt}(\tilde{q}_t(\phi^{-1}(\mathbf{u})))_j \Big|_{t=0} \frac{\partial}{\partial w_j} \Big|_{\mathbf{w}=\phi^{-1}(\mathbf{u})}.\end{aligned}$$

We prove that the definition of the Lie bracket in (9.7.12) is equivalent to the definition in Definition 9.7.1. To prove this claim, we first need a little technical result.

Suppose that $g \in C^\infty(S)$. Then

$$\begin{aligned}\frac{d}{dt}X_t^*(g)(\mathbf{u}) &= \frac{d}{dt}g(X(t, \mathbf{u})) = \frac{d}{dt}((g \circ \phi)(\tilde{X}(t, \phi^{-1}(\mathbf{u})))) \\ &= D(g \circ \phi)(\tilde{X}(t, \phi^{-1}(\mathbf{u}))) \frac{d}{dt}\tilde{X}(t, \phi^{-1}(\mathbf{u})) = D(g \circ \phi)(\tilde{X}(t, \phi^{-1}(\mathbf{u}))) \tilde{f}_1(\tilde{X}(t, \phi^{-1}(\mathbf{u}))) \\ &= \sum_{j=1}^k (\tilde{f}_1)_j(\phi^{-1}(X(t, \mathbf{u}))) \frac{\partial}{\partial w_j}(g \circ \phi)(\phi^{-1}(X(t, \mathbf{u}))) = F_1(X(t, \mathbf{u}))(g)\end{aligned} \tag{9.7.13}$$

for all $\mathbf{u} \in U$.

We are now ready to prove that the definition of Lie bracket given in (9.7.12) is equivalent to our Definition 9.7.1.

Given $\mathbf{u} = \phi(\mathbf{w})$ and $g \in C^\infty(S)$, we have

$$\begin{aligned}
 \left. \frac{d}{dt} (X_t^*(F_2)(\mathbf{u})) \right|_{t=0}(g) &= \left. \frac{d}{dt} \left(D(g \circ \phi)(\mathbf{w}) D \tilde{X}_{-t}(\tilde{X}_t(\mathbf{w})) \tilde{f}_2(\tilde{X}_t(\mathbf{w})) \right) \right|_{t=0} \\
 &= \left. \frac{d}{dt} \left(D((g \circ X_{-t} \circ \phi) \circ (\phi^{-1} \circ X_t \circ \phi))(\mathbf{w}) D \tilde{X}_{-t}(\tilde{X}_t(\mathbf{w})) \tilde{f}_2(\tilde{X}_t(\mathbf{w})) \right) \right|_{t=0} \\
 &= \left. \frac{d}{dt} \left(D(g \circ X_{-t} \circ \phi)(\tilde{X}_t(\mathbf{w})) D \tilde{X}_t(\mathbf{w}) D \tilde{X}_{-t}(\tilde{X}_t(\mathbf{w})) \tilde{f}_2(\tilde{X}_t(\mathbf{w})) \right) \right|_{t=0} \\
 &= \left. \frac{d}{dt} \left(D(g \circ X_{-t} \circ \phi)(\tilde{X}_t(\mathbf{w})) \tilde{f}_2(\tilde{X}_t(\mathbf{w})) \right) \right|_{t=0} \\
 &= \left. \frac{d}{dt} \left(D(X_{-t}^*(g) \circ \phi)(\phi^{-1}(X_t(\phi(\mathbf{w})))) \tilde{f}_2(\phi^{-1}(X_t(\phi(\mathbf{w})))) \right) \right|_{t=0} \\
 &= \left. \frac{d}{dt} \left(F_2(X_t(\mathbf{u}))(X_{-t}^*(g)) \right) \right|_{t=0} = \left. \frac{d}{dt} \left(X_t^*(F_2)(\mathbf{u})(X_{-t}^*(g)) \right) \right|_{t=0} \\
 &= \left. \frac{d}{dt} \left(X_t^*(F_2(g))(\mathbf{u}) \right) \right|_{t=0} + F_2(\mathbf{u}) \left(\left. \frac{d}{dt} (X_{-t}^*(g)) \right|_{t=0} \right) \\
 &= F_1(\mathbf{u})(F_2(g)) - F_2(\mathbf{u})(F_1(g))
 \end{aligned}$$

where the fourth equality comes from

$$D \tilde{X}_t(\mathbf{w}) D \tilde{X}_{-t}(\tilde{X}_t(\mathbf{w})) = D(\tilde{X}_t \circ \tilde{X}_{-t})(\tilde{X}_t(\mathbf{w})) = D \text{Id}_W(\tilde{X}_t(\mathbf{w})) = \text{Id}_W,$$

the second to last equality is derived from the Leibniz' rule because

$(G_1, G_2) \mapsto G_1^*(F_2)(\mathbf{u})(G_2^*(g))$ is a bilinear map since $F_2(\mathbf{u})$ is linear, and the last equality comes from (9.7.13).

9.8 Geodesic Coordinate Systems

Throughout this subsection, S is an oriented 2-dimensional Riemannian manifold.

Let F_1 , F_2 and F_3 be the smooth vector fields on S_c such that $\{\omega_1(\mathbf{u}, \mathbf{x}), \omega_2(\mathbf{u}, \mathbf{x}), \rho(\mathbf{u}, \mathbf{x})\}$ is the dual basis associated to $\{F_1(\mathbf{u}, \mathbf{x}), F_2(\mathbf{u}, \mathbf{x}), F_3(\mathbf{u}, \mathbf{x})\}$ for all $(\mathbf{u}, \mathbf{x}) \in S_c$; namely, $\omega_i(\mathbf{u}, \mathbf{x})(F_j(\mathbf{u}, \mathbf{x})) = \delta_{i,j}$ for $1 \leq i \leq 2$ and $1 \leq j \leq 3$, and $\rho(\mathbf{u}, \mathbf{x})(F_j(\mathbf{u}, \mathbf{x})) = \delta_{3,j}$ for $1 \leq j \leq 3$. We have already noted that $F_3 = H$ where H is defined in Section 9.2.3.

Given $(\mathbf{u}, \mathbf{x}) \in S_c$, as we stated in the previous subsection, there exists a local smooth flow $X :]-\epsilon_{(\mathbf{u}, \mathbf{x})}, \epsilon_{(\mathbf{u}, \mathbf{x})}[\times U_{(\mathbf{u}, \mathbf{x})} \rightarrow S_c$ associated to the smooth vector field F_1 where $U_{(\mathbf{u}, \mathbf{x})} \subset S_c$ is an open neighbourhood of $(\mathbf{u}, \mathbf{x})$. In particular, for each $(\tilde{\mathbf{u}}, \tilde{\mathbf{x}}) \in U_{(\mathbf{u}, \mathbf{x})}$, the function $\tilde{\sigma}_{(\mathbf{u}, \mathbf{x})} :]-\epsilon_{(\mathbf{u}, \mathbf{x})}, \epsilon_{(\mathbf{u}, \mathbf{x})}[\rightarrow S_c$ defined by $\tilde{\sigma}_{(\mathbf{u}, \mathbf{x})}(t) = X(t, (\tilde{\mathbf{u}}, \tilde{\mathbf{x}}))$ for $t \in]-\epsilon_{(\mathbf{u}, \mathbf{x})}, \epsilon_{(\mathbf{u}, \mathbf{x})}[$ satisfies $(\tilde{\sigma}_{(\mathbf{u}, \mathbf{x})})_*(t, 1) = F_1(\tilde{\sigma}_{(\mathbf{u}, \mathbf{x})}(t))$ for all $t \in]-\epsilon_{(\mathbf{u}, \mathbf{x})}, \epsilon_{(\mathbf{u}, \mathbf{x})}[$ and $\tilde{\sigma}_{(\mathbf{u}, \mathbf{x})}(0) = (\tilde{\mathbf{u}}, \tilde{\mathbf{x}})$ by uniqueness of solutions.

We have that $\{U_{(\mathbf{u}, \mathbf{x})}\}_{(\mathbf{u}, \mathbf{x}) \in S_c(\mathbf{u})}$ is an open cover of the compact set $S_c(\mathbf{u})$. Therefore, there exists a finite subcover $\{U_{(\mathbf{u}, \mathbf{x}_j)}\}_{j \in J}$ of $S_c(\mathbf{u})$. Let $\epsilon_{\mathbf{u}} = \min_{j \in J} \epsilon_{(\mathbf{u}, \mathbf{x}_j)}$. We have that $X :] - \epsilon_{\mathbf{u}}, \epsilon_{\mathbf{u}}[\times S_c(\mathbf{u}) \rightarrow S_c$ is well defined.

Let $Q : [0, \epsilon_{\mathbf{u}}[\times S^1 \rightarrow S_c$ be the function defined by $Q(t, \eta) = X(t, \eta(\mathbf{u}, \tilde{E}(\mathbf{u})))$ for $t \in [0, \epsilon_{\mathbf{u}}[$ and $\eta \in S^1$, where $\tilde{E}(\mathbf{u}) = E(\phi^{-1}(\mathbf{u}))$ and (W, U, ϕ) is a local chart of S about $\mathbf{u}$. The local chart (W, U, ϕ) of S with $\mathbf{u} \in U$ chosen to define $\tilde{E}(\mathbf{u})$ is irrelevant¹⁰. The map Q is smooth because it is the composition of smooth maps.

For η fixed, we consider the curves $\sigma_{\eta} : [0, \epsilon_{\mathbf{u}}[\rightarrow S$ and $\tilde{\sigma}_{\eta} : [0, \epsilon_{\mathbf{u}}[\rightarrow S_c$ defined respectively by $\sigma_{\eta}(t) = \pi_S(Q(t, \eta))$ and $\tilde{\sigma}_{\eta}(t) = Q(t, \eta)$ for $t \in [0, \epsilon_{\mathbf{u}}[$.

Proposition 9.8.1

The curve $\tilde{\sigma}_{\eta} : [0, \epsilon_{\mathbf{u}}[\rightarrow S_c$ is the horizontal lift of $\sigma_{\eta} : [0, \epsilon_{\mathbf{u}}[\rightarrow S$ through $\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))$

Proof.

We have that

1. $\tilde{\sigma}_{\eta}(0) = Q(0, \eta) = X(0, \eta(\mathbf{u}, \tilde{E}(\mathbf{u}))) = \eta(\mathbf{u}, \tilde{E}(\mathbf{u}))$,
2. $\pi_S \circ \tilde{\sigma}_{\eta} = \sigma_{\eta}$ and
3. $\rho(\tilde{\sigma}_{\eta}(t))((\tilde{\sigma}_{\eta})_*(t, 1)) = \rho(\tilde{\sigma}_{\eta}(t))(F_1(\tilde{\sigma}_{\eta}(t))) = 0$ for $0 \leq t < \epsilon_{\mathbf{u}}$ by definition of F_1 .

Thus, all the conditions for an horizontal lift are satisfied. ■

Moreover, we have the following result.

Proposition 9.8.2

$\sigma_{\eta} : [0, \epsilon_{\mathbf{u}}[\rightarrow S$ is a geodesic.

Proof.

Since $\pi_S \circ \tilde{\sigma}_{\eta} = \sigma_{\eta}$, we have that $(\pi_S)_* \circ (\tilde{\sigma}_{\eta})_* = (\pi_S \circ \tilde{\sigma}_{\eta})_* = (\sigma_{\eta})_*$. Thus

$$\begin{aligned} (\sigma_{\eta}(t), \sigma'_{\eta}(t)) &= (\sigma_{\eta})_*(t, 1) = (\pi_S)_*((\tilde{\sigma}_{\eta})_*(t, 1)) \\ &= \omega_1(\tilde{\sigma}_{\eta}(t))((\tilde{\sigma}_{\eta})_*(t, 1)) \tilde{\sigma}_{\eta}(t) + \omega_2(\tilde{\sigma}_{\eta}(t))((\tilde{\sigma}_{\eta})_*(t, 1)) v(\tilde{\sigma}_{\eta}(t)) \\ &= \omega_1(\tilde{\sigma}_{\eta}(t))(F_1(\tilde{\sigma}_{\eta}(t))) \tilde{\sigma}_{\eta}(t) + \omega_2(\tilde{\sigma}_{\eta}(t))(F_1(\tilde{\sigma}_{\eta}(t))) v(\tilde{\sigma}_{\eta}(t)) = \tilde{\sigma}_{\eta}(t) \end{aligned}$$

for $0 \leq t < \epsilon_{\mathbf{u}}$. ■

¹⁰Any $(\mathbf{u}, \mathbf{x}) \in S_c(\mathbf{u})$ with $\mathbf{x}$ fixed could have been used instead of $(\mathbf{u}, \tilde{E}(\mathbf{u}))$.

Let

$$D_{\epsilon_u} = \{(\mathbf{u}, \mathbf{x}) \in T_{\mathbf{u}}S : \|(\mathbf{u}, \mathbf{x})\|_{\mathbf{u}} < \epsilon_u\} \cong B_{\epsilon_u}(\mathbf{0}) \subset \mathbb{R}^2$$

where the norm is given by the Riemannian metric on S . The isomorphism between $B_{\epsilon_u}(\mathbf{0})$ and D_{ϵ_u} can be defined by $t(\theta, \mathbf{e}_1) \mapsto (\mathbf{u}, t(\theta, \tilde{E}(\mathbf{u})))$ for $0 \leq t < \epsilon_u$ and $\theta \in \mathbb{R}$.

Let $P : [0, \epsilon_u[\times S^1 \rightarrow D_{\epsilon_u} \cong B_{\epsilon_u}(\mathbf{0})$ be the function defined by $P(t, \eta) = t(\theta, \mathbf{e}_1)$ for $(t, \eta) \in [0, \epsilon_u[\times S^1$ and $\eta = e^{i\theta}$. Since $P|_{[0, \epsilon_u[\times S^1}$ is an isomorphism, we may define $R : D_{\epsilon_u} \rightarrow [0, \epsilon_u[\times S^1$ by

$$R(\mathbf{x}) = \begin{cases} (t, e^{i\theta}) & \text{if } \mathbf{x} = t(\theta, \mathbf{e}_1) \text{ with } t > 0 \\ (0, 0) & \text{if } \mathbf{x} = \mathbf{0} \end{cases}$$

We have that $(P|_{[0, \epsilon_u[\times S^1})^{-1} = R|_{D_{\epsilon_u} \setminus \{\mathbf{0}\}}$.

Let $G : D_{\epsilon_u} \rightarrow S$ be the map defined by $G = \pi_S \circ Q \circ R$. Note that G is well defined because $\pi_S(Q(0, \eta)) = \mathbf{u}$ for all $\eta \in S^1$.

By taking ϵ_u smaller if necessary, we now prove that $G : D_{\epsilon_u} \rightarrow G(D_{\epsilon_u})$ is a diffeomorphism. This implies that $G(D_{\epsilon_u})$ is an open neighbourhood of $\mathbf{u}$ in S , and G_* maps $T_{\mathbf{0}}D_{\epsilon_u}$ onto $T_{\mathbf{u}}S$. We use $D_{\epsilon_u} = B_{\epsilon_u}(\mathbf{0}) \subset \mathbb{R}^2$ and show that $DG(\mathbf{0}) : \mathbb{R}^2 \rightarrow T_{\mathbf{u}}S$ is of rank 2. It will then follow from the Inverse Function Theorem that $G : D_{\epsilon_u} \rightarrow S$ is a diffeomorphism for ϵ_u small enough. Consider $\mathbf{x} = \theta, \mathbf{e}_1 \in \mathbb{R}^2$ with $0 \leq \theta < 2\pi$ fixed and let $\eta = e^{i\theta} \in S^1$. We get the directional derivative

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{1}{t} (G(t\mathbf{x}) - G(\mathbf{0})) &= \lim_{t \rightarrow 0} \frac{1}{t} (\pi_S(Q(t, \eta)) - \pi_S(Q(0, \eta))) = \lim_{t \rightarrow 0} \frac{1}{t} (\sigma_\eta(t) - \sigma_\eta(0)) \\ &= \sigma'_\eta(0) = \theta, \tilde{E}(\mathbf{u}) \end{aligned}$$

because $(\sigma_\eta(0), \sigma'_\eta(0)) = (\sigma_\eta)_*(0, 1) = \tilde{\sigma}_\eta(0) = \eta(\mathbf{u}, \tilde{E}(\mathbf{u}))$. Since $\eta \in S^1$ is arbitrary, all the directions from the origin in $T_{\mathbf{u}}S$ are included in the range of $DG(\mathbf{0})$. Thus $DG(\mathbf{0})$ is of rank 2.

The following diagram summarizes the construction of G .

$$\begin{array}{ccc} [0, \epsilon_u[\times S^1 & \xrightarrow{Q} & S_c \\ R \uparrow & & \downarrow \pi_S \\ D_{\epsilon_u} & \xrightarrow{G} & S \end{array}$$

Definition 9.8.3

The map $G : D_{\epsilon_u} \rightarrow S$ is called a **polar or geodesic coordinate system** for S about $\mathbf{u}$ if ϵ_u is small enough to ensure that G is a diffeomorphism onto its image. (Figure 9.11).

For convenience, the lines $t \mapsto t(\theta, \mathbf{e}_1)$ in $D_{\epsilon_u} \cong B_{\epsilon_u}(\mathbf{0})$ for $0 \leq t < \epsilon_u$ and $\eta = e^{i\theta} \in S^1$ fixed are called **radial geodesic**.

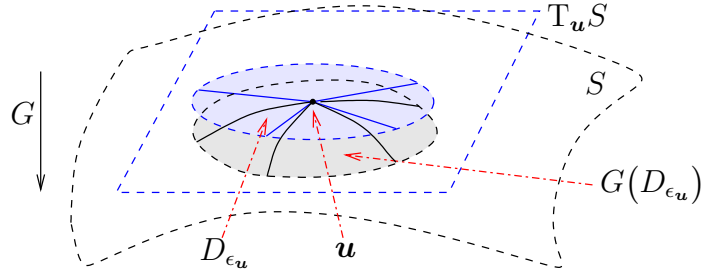


Figure 9.11: Schematic representation of a geodesic coordinate system.

The following technical results will be quite useful for our study of geodesic coordinate systems.

In the rest of the section, we will regularly refer to local charts $(\check{W}, \check{U}, \check{\phi})$ of $[0, \epsilon_u[\times S^1$ where $\check{W} = [0, \epsilon_u[\times I$ for I a connected and open interval of length less than 2π , $\check{\phi}(r, \theta) = (r, e^{i\theta})$ for $(r, \theta) \in \check{W}$, and $\check{U} = \check{\phi}(\check{W})$.

There are two natural vector fields V_1 and V_2 on $]0, \epsilon_u[\times S^1$ that we now define.

Let $\check{V}_1$ and $\check{V}_2$ be the vector fields defined on $\check{W}$ by $\check{V}_1(r, \theta) = ((r, \theta), (1, 0))$ and $\check{V}_2(r, \theta) = ((r, \theta), (0, 1))$ respectively for $(r, \theta) \in \check{W}$. Are there any simpler non-trivial vector fields on W ? If $(t, \eta) = \check{\phi}(r, \theta)$, then

$$\begin{aligned} V_1(t, \eta) &= \check{\phi}_*(\check{V}_1(\check{\phi}^{-1}(t, \eta))) = \check{\phi}_*(\check{V}_1(r, \theta)) = \check{\phi}_*((r, \theta), (1, 0)) \\ &= (\check{\phi}(r, \theta), (D\check{\phi})(r, \theta) \begin{pmatrix} 1 & 0 \end{pmatrix}^\top) = \left(\check{\phi}(r, \theta), \frac{\partial \check{\phi}}{\partial r}(r, \theta)\right) = ((t, \eta), (1, 0)) \end{aligned}$$

and

$$\begin{aligned} V_2(t, \eta) &= \check{\phi}_*(\check{V}_2(\check{\phi}^{-1}(t, \eta))) = \check{\phi}_*(\check{V}_2(r, \theta)) = \check{\phi}_*((r, \theta), (0, 1)) \\ &= (\check{\phi}(r, \theta), (D\check{\phi})(r, \theta) \begin{pmatrix} 0 & 1 \end{pmatrix}^\top) = \left(\check{\phi}(r, \theta), \frac{\partial \check{\phi}}{\partial \theta}(r, \theta)\right) \\ &= ((t, \eta), (0, i\eta)) = ((t, \eta), (0, v\eta)) \end{aligned}$$

where $v = e^{i(\pi/2)}$ as usual.

Since $(P \circ \check{\phi})(r, \theta) = r(\theta \cdot \mathbf{e}_1)$ for all $(r, \theta) \in \check{W}$, we get

$$\begin{aligned} P_*(V_1(t, \eta)) &= (P_* \circ \check{\phi}_*)((\check{\phi}^{-1} \circ V_1 \circ \check{\phi})(t, \theta)) = (P_* \circ \check{\phi}_*)(\check{V}_1(t, \theta)) \\ &= (P \circ \check{\phi})_*((t, \theta), (1, 0)) = (t(\theta \cdot \mathbf{e}_1), \theta \cdot \mathbf{e}_1) \end{aligned}$$

where $V_1(t, \eta)$ represents a vector field in D_{ϵ_u} that point in the radial direction $\eta = e^{i\theta}$, and

$$\begin{aligned} P_*(V_2(t, \eta)) &= (P_* \circ \check{\phi}_*)((\check{\phi}^{-1} \circ V_2 \circ \check{\phi})(t, \theta)) = (P_* \circ \check{\phi}_*)(\check{V}_2(t, \theta)) \\ &= (P \circ \check{\phi})_*((t, \theta), (0, 1)) = (t(\theta \cdot \mathbf{e}), t((\pi/2 + \theta) \cdot \mathbf{e})) \end{aligned}$$

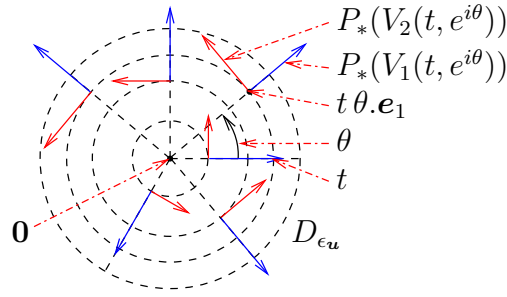


Figure 9.12: Representation of the vector field $P_* \circ V_1$ on D_{ϵ_u} in blue and the vector field $P_* \circ V_2$ on D_{ϵ_u} in red.

represents a vector field in D_{ϵ_u} that point in the rotational direction as illustrated in Figure 9.12.

It will often be more convenient when performing some algebraic manipulations to use the definition of vector fields given in Remark 5.4.3. The isomorphism between $\check{T}_{\mathbf{u}}S$ and $T_{\mathbf{u}}S$ introduced in Section 5.3.2 shows that the vector fields V_1 and V_2 on $]0, \epsilon_{\mathbf{u}}[\times S^1$ can also be respectively represented locally as

$$\check{V}_1 = \frac{\partial}{\partial r} \Big|_{(r,\theta)=\check{\phi}^{-1}(t,\eta)} \quad \text{and} \quad \check{V}_2 = \frac{\partial}{\partial \theta} \Big|_{(r,\theta)=\check{\phi}^{-1}(t,\eta)} \quad (9.8.1)$$

for $(t, \eta) \in]0, \epsilon_{\mathbf{u}}[\times S^1$.

Lemma 9.8.4

1. $Q_*(V_1(t, \eta)) = F_1(\tilde{\sigma}_\eta(t))$ for $(t, \eta) \in]0, \epsilon_{\mathbf{u}}[\times S^1$.
2. $Q^*(\rho)(V_1) = Q^*(\omega_2)(V_1) = 0$ and $Q^*(\omega_1)(V_1) = 1$ on $]0, \epsilon_{\mathbf{u}}[\times S^1$.
3. $[V_1, V_2](t, \eta) = ((t, \eta), (0, 0))$ for all $(t, \eta) \in]0, \epsilon_{\mathbf{u}}[\times S^1$.
4. $Q_*(V_2(t, \eta))|_{t=0} = H(\tilde{\sigma}_\eta(0))$.

Proof.

1) Using a local chart $(\check{W}, \check{U}, \check{\phi})$ of $]0, \epsilon_{\mathbf{u}}[\times S^1$ as defined above, we get for $(t, \eta) = \check{\phi}(r, \theta)$ that

$$\begin{aligned} Q_*(V_1(t, \eta)) &= Q_*(\check{\phi}_*(\check{V}_1(r, \theta))) = (Q \circ \check{\phi})_*(\check{V}_1(r, \theta)) = (Q \circ \check{\phi})_*((t, \theta), (1, 0)) \\ &= \left(Q(\check{\phi}(t, \theta)), D_{t,\theta}(Q \circ \check{\phi})(t, \theta) \begin{pmatrix} 1 & 0 \end{pmatrix}^\top \right) = \left(Q(t, \eta), \frac{\partial}{\partial t}(Q \circ \check{\phi})(t, \theta) \right) \\ &= \left(Q(t, \eta), \frac{\partial}{\partial t} Q(t, \eta) \right) = (\tilde{\sigma}_\eta(t), \tilde{\sigma}'_\eta(t)) = (\tilde{\sigma}_\eta)_*(t, 1) = F_1(\tilde{\sigma}_\eta(t)) \end{aligned}$$

since $\tilde{\sigma}_\eta(t) = Q(t, \eta)$.

2) Since $Q_*(V_1(t, \eta)) = F_1(\tilde{\sigma}_\eta(t))$, we get

$$(Q^*(\rho)(V_1))(t, \eta) = \rho(Q(t, \eta))(Q_*(V_1(t, \eta))) = \rho(\tilde{\sigma}_\eta(t))(F_1(\tilde{\sigma}_\eta(t))) = 0$$

and

$$(Q^*(\omega_2)(V_1))(t, \eta) = \omega_2(Q(t, \eta))(Q_*(V_1(t, \eta))) = \omega_2(\tilde{\sigma}_\eta(t))(F_1(\tilde{\sigma}_\eta(t))) = 0.$$

for all $(t, \eta) \in]0, \epsilon_{\mathbf{u}}[\times S^1$ by definition of ρ , ω_2 and F_1 . Moreover,

$$(Q^*(\omega_1)(V_1))(t, \eta) = \omega_1(Q(t, \eta))(Q_*(V_1(t, \eta))) = \omega_1(\tilde{\sigma}_\eta(t))(F_1(\tilde{\sigma}_\eta(t))) = 1$$

for all $(t, \eta) \in]0, \epsilon_{\mathbf{u}}[\times S^1$ by definition of ω_1 and F_1 .

3) Using the local representation of V_1 and V_2 provided in (9.8.1), we get

$$\begin{aligned} [V_1, V_2](t, \eta)(g) &= V_1(t, \eta)(V_2(g)) - V_2(t, \eta)(V_1(g)) \\ &= \frac{\partial}{\partial r} \left(\frac{\partial}{\partial \theta} (g \circ \phi) \right) \Big|_{(r, \theta) = \check{\phi}^{-1}(t, \eta)} - \frac{\partial}{\partial \theta} \left(\frac{\partial}{\partial r} (g \circ \phi) \right) \Big|_{(r, \theta) = \check{\phi}^{-1}(t, \eta)} = 0 \end{aligned}$$

for all $g \in C^\infty([0, \epsilon_{\mathbf{u}}[\times S^1)$. Thus $[V_1, V_2](t, \eta) = ((t, \eta), (0, 0))$

4) We proceed as in (1). Using a local chart $(\check{W}, \check{U}, \check{\phi})$ of $]0, \epsilon_{\mathbf{u}}[\times S^1$ as defined above, we get for $(t, \eta) = \check{\phi}(r, \theta)$ that

$$\begin{aligned} Q_*(V_2(t, \eta)) &= Q_*(\check{\phi}_*(\check{V}_2(r, \theta))) = (Q \circ \check{\phi})_*(\check{V}_2(r, \theta)) = (Q \circ \check{\phi})_*((t, \theta), (0, 1)) \\ &= \left(Q(\check{\phi}(t, \theta)), D_{t, \theta}(Q \circ \check{\phi})(t, \theta) \begin{pmatrix} 0 & 1 \end{pmatrix}^\top \right) = \left(Q(t, \eta), \frac{\partial}{\partial \theta}(Q \circ \check{\phi})(t, \theta) \right) \\ &= \left(Q(t, \eta), \frac{\partial}{\partial \theta} Q(t, e^{i\theta}) \right) \end{aligned}$$

Hence

$$\begin{aligned} Q_*(V_2(0, \eta)) &= \left(Q(0, \eta), \frac{\partial}{\partial \theta} Q(0, e^{i\theta}) \right) = \left(\eta(\mathbf{u}, \tilde{E}(\mathbf{u})), \frac{\partial}{\partial \theta}(\mathbf{u}, \theta. \tilde{E}(\mathbf{u})) \right) \\ &= \left(\eta(\mathbf{u}, \tilde{E}(\mathbf{u})), \left(\mathbf{0}, \frac{\partial}{\partial \theta}(\theta. \tilde{E}(\mathbf{u})) \right) \right) = H(\eta(\mathbf{u}, \tilde{E}(\mathbf{u}))) = H(\tilde{\sigma}_\eta(0)). \quad \blacksquare \end{aligned}$$

Remark 9.8.5

It is also possible to prove (3) in the previous lemma (with a lot more work) using the equivalent formulation for the Lie bracket; namely,

$$[V_1, V_2](t, \eta) = \frac{d}{ds} Y_s^*(V_2)(t, \eta) \Big|_{s=0}$$

where $Y :]-\epsilon, \epsilon[\times]0, \epsilon_{\mathbf{u}}[\times S^1 \rightarrow]0, \epsilon_{\mathbf{u}}[\times S^1$ for ϵ small enough is the flow associate to the vector field V_1 .

Using a local chart $(\check{W}, \check{U}, \check{\phi})$, we have that the flow associated to $\check{V}_1$ is the function $\check{Y} :]-\epsilon_{(r, \theta)}, \epsilon_{(r, \theta)}[\times \check{W} \rightarrow \check{W}$ defined by $\check{Y}(s, (r, \theta)) = (r + s, \theta)$ for $\epsilon_{(r, \theta)}$ small enough. Note that we effectively have that $(\check{Y}_{(r, \theta)})_*(s, 1) = ((r + s, \theta), (1, 0)) = \check{V}_1(\check{Y}_{(r, \theta)}(s))$ for $-\epsilon_{(r, \theta)} < s < \epsilon_{(r, \theta)}$.

We proceed as in Subsection 9.7.1 to get the local representation of the vector field $Q_s = Y_s^*(V_2)$. Let $\check{v}_2 : \check{W} \rightarrow \mathbb{R}^2$ be the function associated to the local representation $\check{V}_2$ of V_2 given in (9.8.1); namely, $\check{v}_2 = (0 \ 1)^\top$.

The local representation of the vector field $Q_s = Y_s^*(V_2)$ is $\tilde{Q}_s(r, \theta) = (\tilde{q}_s)_1(r, \theta) \frac{\partial}{\partial r} + (\tilde{q}_s)_2(r, \theta) \frac{\partial}{\partial \theta}$ where

$$\tilde{q}_s(r, \theta) = D\check{Y}_{-s}(\check{Y}_s(r, \theta)) \check{v}_2(\check{Y}_s(r, \theta)) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

for $(r, \theta) \in \check{W}$ and $-\epsilon_{(r, \theta)} < s < \epsilon_{(r, \theta)}$. All this work just to get that the local representation of $Y_s^*(V_2)$ is $\frac{\partial}{\partial \theta} \Big|_{(r, \theta) = \check{\phi}^{-1}(t, \eta)}$.

Therefore

$$[V_1, V_2](t, \eta)(g) = \frac{d}{ds} \left(\frac{\partial}{\partial \theta} (g \circ \check{\phi}) \Big|_{(r, \theta) = \check{\phi}^{-1}(t, \eta)} \right) \Big|_{s=0} = \frac{d}{ds} \left(\frac{\partial}{\partial \theta} g(r, e^{i\theta}) \Big|_{(r, \theta) = \check{\phi}^{-1}(t, \eta)} \right) \Big|_{s=0} = 0$$

for all $g \in C^\infty([0, \epsilon_u[\times S^1)$ since the expression between parentheses does not depend on s . ♠

Lemma 9.8.6

$P_*((r, \eta), (p, q v \eta)) = (r(\theta \cdot e_1), p(\theta \cdot e_1) + q r(\theta + \pi/2) \cdot e_1)$ for $\eta = e^{i\theta}$, $v = e^{i(\pi/2)}$ and $p, q \in \mathbb{R}$. Note that $((r, \eta), (p, q v \eta)) \in T_{(t, \eta)}[0, \epsilon_u[\times S^1$.

Proof.

We consider a local chart $(\check{W}, \check{U}, \check{\phi})$ of $[0, \epsilon_u[\times S^1$. Recall that $\phi(r, \theta) = (r, e^{i\theta})$ for $(r, \theta) \in \check{W} \subset]0, \epsilon_u[\times \mathbb{R}$. The local representation of P is given by $\tilde{P}(r, \theta) = (P \circ \phi)(r, \theta) = r(\theta \cdot e_1)$ for $(r, \theta) \in \check{W}$. We have

$$\phi_*((r, \theta), (p, q)) = ((r, e^{i\theta}), (p, q e^{i(\theta + \pi/2)})) = ((r, \eta), (p, q v \eta))$$

and

$$\tilde{P}_*((r, \theta), (p, q)) = (r(\theta \cdot e_1), p(\theta \cdot e_1) + q r((\theta + \pi/2) \cdot e_1))$$

or all $((r, \theta), (p, q)) \in T_{(r, \theta)}\check{W}$. The conclusion of the lemma comes from the fact that $P \circ \phi = \tilde{P}$ implies that $P_* \circ \phi_* = \tilde{P}_*$. ■

If we use complex representation of the elements of D_{ϵ_u} , then the conclusion of the previous lemma can be restated as $P_*((r, \eta), (p, q v \eta)) = (r\eta, p\eta + q r v \eta)$.

Before diving deeper into our study of geodesic coordinate systems, we need to restate one of our previous results. Using our knowledge of the Lie derivative of a function along a vector field learned in Remark 5.5.8, we can restate Proposition 9.7.2 in familiar terms.

Corollary 9.8.7

Suppose that ω is a smooth differential 1-form on a smooth k -dimensional manifold S , and that F_1 and F_2 are two smooth vector fields on S . Then

$$\begin{aligned} d\omega(\mathbf{u})(F_1(\mathbf{u}), F_2(\mathbf{u})) &= d(\omega(F_2))(\mathbf{u})(F_1(\mathbf{u})) - d(\omega(F_1))(\mathbf{u})(F_2(\mathbf{u})) \\ &\quad - \omega(\mathbf{u})([F_1, F_2](\mathbf{u})) . \end{aligned}$$

for all $\mathbf{u} \in S$.

Lemma 9.8.8 (Gauss)

$$\langle G_*(P_*(V_1(t, \eta))), G_*(P_*(V_2(t, \eta))) \rangle_{\pi_S(Q(t, \eta))} = 0 \text{ for all } (t, \eta) \in]0, \epsilon_u[\times S^1.$$

Proof.

We have

$$G_* \circ P_* = (G \circ P)_* = (\pi_S \circ Q)_* = (\pi_S)_* \circ Q_* .$$

Moreover, since $Q_*(V_1(t, \eta)) = F_1(\tilde{\sigma}_\eta(t)) = (\tilde{\sigma}_\eta)_*(t, 1)$ according to the first item of Lemma 9.8.4, we have

$$(\pi_S)_*(Q_*(V_1(t, \eta))) = (\pi_S)_*((\tilde{\sigma}_\eta)_*(t, 1)) = \tilde{\sigma}_\eta(t) = Q(t, \eta)$$

for all $(t, \eta) \in]0, \epsilon_u[\times S^1$. Therefore

$$\begin{aligned} &\langle G_*(P_*(V_1(t, \eta))), G_*(P_*(V_2(t, \eta))) \rangle_{\pi_S(Q(t, \eta))} \\ &= \langle (\pi_S)_*(Q_*(V_1(t, \eta))), (\pi_S)_*(Q_*(V_2(t, \eta))) \rangle_{\pi_S(Q(t, \eta))} \\ &= \langle Q(t, \eta), (\pi_S)_*(Q_*(V_2(t, \eta))) \rangle_{\pi_S(Q(t, \eta))} \\ &= \omega_1(Q(t, \eta))(Q_*(V_2(t, \eta))) = Q^*(\omega_1)(t, \eta)(V_2(t, \eta)) \end{aligned} \tag{9.8.2}$$

for all $(t, \eta) \in]0, \epsilon_u[\times S^1$. We need to prove that the last expression in the relation above is null.

Using the first of the first structural equations in Theorem 9.4.7, we have

$$d(Q^*(\omega_1)) = Q^*(d\omega_1) = Q^*(\rho \wedge \omega_2) = Q^*(\rho) \wedge Q^*(\omega_2) .$$

Hence

$$\begin{aligned} &d(Q^*(\omega_1))(t, \eta)(V_1(t, \eta), V_2(t, \eta)) \\ &= \frac{1}{2} \left(Q^*(\rho)(t, \eta)(V_1(t, \eta)) Q^*(\omega_2)(t, \eta)(V_2(t, \eta)) \right. \\ &\quad \left. - Q^*(\rho)(t, \eta)(V_2(t, \eta)) Q^*(\omega_2)(t, \eta)(V_1(t, \eta)) \right) \\ &= \frac{1}{2} \left(\rho(Q(t, \eta))(Q_*(V_1(t, \eta))) \omega_2(Q(t, \eta))(Q_*(V_2(t, \eta))) \right. \\ &\quad \left. - \rho(Q(t, \eta))(Q_*(V_2(t, \eta))) \omega_2(Q(t, \eta))(Q_*(V_1(t, \eta))) \right) \end{aligned}$$

$$- \rho(Q(t, \eta))(Q_*(V_2(t, \eta))) \omega_2(Q(t, \eta))(Q_*(V_1(t, \eta))) = 0$$

for all $(t, \eta) \in]0, \epsilon_{\mathbf{u}}[\times S^1$ because $\rho(Q(t, \eta))(Q_*(V_1(t, \eta))) = 0$ and $\omega_2(Q(t, \eta))(Q_*(V_1(t, \eta))) = 0$ according to second item of Lemma 9.8.4.

It follows from Corollary 9.8.7 that

$$\begin{aligned} 0 &= d(Q^*(\omega_1))(t, \eta)(V_1(t, \eta), V_2(t, \eta)) \\ &= d(Q^*(\omega_1)(V_2))(t, \eta)(V_1(t, \eta)) - d(Q^*(\omega_1)(V_1))(t, \eta)(V_2(t, \eta)) \\ &\quad - Q^*(\omega_1)(t, \eta)([V_1, V_2](t, \eta)) \end{aligned} \quad (9.8.3)$$

for all $(t, \eta) \in]0, \epsilon_{\mathbf{u}}[\times S^1$. However, $d(Q^*(\omega_1)(V_1)) = 0$ because $Q^*(\omega_1)(V_1) = 1$ on $]0, \epsilon_{\mathbf{u}}[\times S^1$ according to the second item of Lemma 9.8.4, and $Q^*(\omega)(t, \eta)([V_1, V_2](t, \eta)) = 0$ for all $(t, \eta) \in]0, \epsilon_{\mathbf{u}}[\times S^1$ because $[V_1, V_2](t, \eta) = ((t, \eta), (0, 0))$ according to the third item of Lemma 9.8.4. Hence, we get from (9.8.3) that

$$d(Q^*(\omega_1)(V_2))(t, \eta)(V_1(t, \eta)) = 0 \quad (9.8.4)$$

for all $(t, \eta) \in]0, \epsilon_{\mathbf{u}}[\times S^1$. It is again advantageous to use the local representation of V_1 and V_2 provided in (9.8.1), and our knowledge of the Lie derivative of a function along a vector field, to get from (9.8.4) that

$$V_1(t, \eta)(Q^*(\omega_1)(V_2)) = \frac{\partial}{\partial r} \left(Q^*(\omega_1)(V_2) \circ \phi \right) \Big|_{(r, \theta) = \phi^{-1}(t, \eta)} = 0$$

for $(t, \eta) \in]0, \epsilon_{\mathbf{u}}[\times S^1$. Thus $(Q^*(\omega_1)(V_2))(t, \eta) = Q^*(\omega_1)(t, \eta)(V_2(t, \eta))$ is constant with respect to $t = r$. Therefore, it follows from (9.8.2) that

$$\langle G_*(P_*(V_1(t, \eta))), G_*(P_*(V_2(t, \eta))) \rangle_{\pi_S(Q(t, \eta))} = \langle Q(t, \eta), G_*(P_*(V_2(t, \eta))) \rangle_{\pi_S(Q(t, \eta))}$$

is a constant function with respect to t . However, $Q(t, \eta) \rightarrow \eta(\mathbf{u}, \tilde{E}(\mathbf{u}))$ as $t \rightarrow 0$ and

$$P_*(V_2(t, \eta)) = P_*((t, \eta), (0, v\eta)) = (t\eta, t v\eta) \rightarrow (0, 0)$$

as $t \rightarrow 0$ where we have used Lemma 9.8.6 to get the last equality. Thus $\langle G_*(P_*(V_1(t, \eta))), G_*(P_*(V_2(t, \eta))) \rangle_{\pi_S(Q(t, \eta))} \rightarrow 0$ as $t \rightarrow 0$. Since this inner product is constant, it must therefore be null for all t . ■

Corollary 9.8.9

$$Q^*(\omega_1)(V_2) = 0 \text{ on }]0, \epsilon_{\mathbf{u}}[\times S^1.$$

Proof.

This follows from (9.8.2) and Gauss' lemma. ■

It follows from Gauss' lemma that $t \mapsto G_*(P_*(V_2(t, \eta)))$ is orthogonal to $t \mapsto G_*(P_*(V_1(t, \eta))) = Q(t, \eta) = \tilde{\sigma}_\eta(t)$ for all $t \in [0, \epsilon_{\mathbf{u}}[$; namely, the geodesic curve $t \mapsto \sigma_\eta(t) = \pi_S(\tilde{\sigma}_\eta(t))$ and the curve $t \mapsto G(P((t, \eta))) = \pi_S(G_*(P_*(V_2(t, \eta))))$ have an orthogonal intersect.

We now show that the length of $G_*(P_*(V_2(t, \eta)))$ along the geodesic curve σ_η satisfies a differential equation.

We have

$$\begin{aligned} G_*(P_*(V_2(t, \eta))) &= (\pi_S)_*(Q_*(V_2(t, \eta))) \\ &= \omega_1(Q(t, \eta))(Q_*(V_2(t, \eta))) Q(t, \eta) + \omega_2(Q(t, \eta))(Q_*(V_2(t, \eta))) v(Q(t, \eta)) \\ &= \omega_2(Q(t, \eta))(Q_*(V_2(t, \eta))) v(Q(t, \eta)) = Q^*(\omega_2)(t, \eta)(V_2(t, \eta)) v(Q(t, \eta)) \end{aligned}$$

because $\omega_1(Q(t, \eta))(Q_*(V_2(t, \eta))) = Q^*(\omega_1)(t, \eta)(V_2(t, \eta)) = 0$ according to Corollary 9.8.9. Thus

$$\|G_*(P_*(V_2(t, \eta)))\|_{\pi_S(Q(t, \eta))} = |Q^*(\omega_2)(t, \eta)(V_2(t, \eta))| \quad (9.8.5)$$

for all $(t, \eta) \in [0, \epsilon_{\mathbf{u}}[\times S^1$, where the norm refer to the Riemannian metric on S .

We get from the second of the first structural equations that

$$d(Q^*(\omega_2)) = Q^*(d\omega_2) = -Q^*(\rho \wedge \omega_1) = -Q^*(\rho) \wedge Q^*(\omega_1) .$$

Thus

$$\begin{aligned} d(Q^*(\omega_2))(t, \eta)(V_1(t, \eta), V_2(t, \eta)) &= \frac{1}{2} \left(Q^*(\rho)(t, \eta)(V_2(t, \eta)) Q^*(\omega_1)(t, \eta)(V_1(t, \eta)) \right. \\ &\quad \left. - Q^*(\rho)(t, \eta)(V_1(t, \eta)) Q^*(\omega_1)(t, \eta)(V_2(t, \eta)) \right) = \frac{1}{2} Q^*(\rho)(t, \eta)(V_2(t, \eta)) \end{aligned} \quad (9.8.6)$$

for all $(t, \eta) \in [0, \epsilon_{\mathbf{u}}[\times S^1$ because $Q^*(\omega_1)(t, \eta)(V_2(t, \eta)) = 0$ according to Corollary 9.8.9 and $Q^*(\omega_1)(t, \eta)(V_1(t, \eta)) = 1$ according to the second item of Lemma 9.8.4.

However, using Corollary 9.8.7, we have

$$\begin{aligned} d(Q^*(\omega_2))(t, \eta)(V_1(t, \eta), V_2(t, \eta)) &= d(Q^*(\omega_2)(V_2))(t, \eta)(V_1(t, \eta)) - d(Q^*(\omega_2)(V_1))(t, \eta)(V_2(t, \eta)) \\ &\quad - Q^*(\omega_2)(t, \eta)([V_1, V_2](t, \eta)) \\ &= d(Q^*(\omega_2)(V_2))(t, \eta)(V_1(t, \eta)) \end{aligned} \quad (9.8.7)$$

because $Q^*(\omega_2)(V_1) = 0$ and $[V_1, V_2](t, \eta) = ((t, \eta), (0, 0))$ for all $(t, \eta) \in [0, \epsilon_{\mathbf{u}}[\times S^1$ according to the second and third items of Lemma 9.8.4. We deduce from (9.8.6) and (9.8.7) that

$$d(Q^*(\omega_2)(V_2))(t, \eta)(V_1(t, \eta)) = \frac{1}{2} Q^*(\rho)(t, \eta)(V_2(t, \eta)) \quad (9.8.8)$$

for all $(t, \eta) \in [0, \epsilon_{\mathbf{u}}[\times S^1$.

We now almost repeat the previous discussion about $Q^*(\omega_2)$ but this time with $Q^*(\rho)$. From the second structural equations, we have

$$d(Q^*(\rho)) = Q^*(d\rho) = -Q^*((\kappa \circ \pi_S)\omega_1 \wedge \omega_2) = -(\kappa \circ \pi_S \circ Q)Q^*(\omega_1) \wedge Q^*(\omega_2) .$$

Thus

$$\begin{aligned} d(Q^*(\rho))(t, \eta)(V_1(t, \eta), V_2(t, \eta)) &= \frac{1}{2}\kappa(\pi_S(Q(t, \eta))) \left(Q^*(\omega_1)(t, \eta)(V_2(t, \eta)) Q^*(\omega_2)(t, \eta)(V_1(t, \eta)) \right. \\ &\quad \left. - Q^*(\omega_1)(t, \eta)(V_1(t, \eta)) Q^*(\omega_2)(t, \eta)(V_2(t, \eta)) \right) \\ &= -\frac{1}{2}\kappa(\pi_S(Q(t, \eta))) Q^*(\omega_2)(t, \eta)(V_2(t, \eta)) \end{aligned} \quad (9.8.9)$$

for all $(t, \eta) \in [0, \epsilon_u] \times S^1$ because $Q^*(\omega_1)(t, \eta)(V_2(t, \eta)) = 0$ and $Q^*(\omega_1)(t, \eta)(V_1(t, \eta)) = 1$ according to the second item of Lemma 9.8.4.

Using Corollary 9.8.7, we have

$$\begin{aligned} d(Q^*(\rho))(t, \eta)(V_1(t, \eta), V_2(t, \eta)) &= d(Q^*(\rho)(V_2))(t, \eta)(V_1(t, \eta)) - d(Q^*(\rho)(V_1))(t, \eta)(V_2(t, \eta)) \\ &\quad - Q^*(\rho)(t, \eta)([V_1, V_2](t, \eta)) \\ &= d(Q^*(\rho)(V_2))(t, \eta)(V_1(t, \eta)) \end{aligned} \quad (9.8.10)$$

because $Q^*(\rho)(V_1) = 0$ and $[V_1, V_2](t, \eta) = ((t, \eta), (0, 0))$ for all $(t, \eta) \in [0, \epsilon_u] \times S^1$ according to the second and third items of Lemma 9.8.4.

We deduce from (9.8.9) and (9.8.10) that

$$\begin{aligned} d(Q^*(\rho)(V_2))(t, \eta)(V_1(t, \eta)) &= -\frac{1}{2}\kappa(\pi_S(Q(t, \eta))) Q^*(\omega_2)(t, \eta)(V_2(t, \eta)) \\ &= -\frac{1}{2}\kappa(G(P(t, \eta))) Q^*(\omega_2)(t, \eta)(V_2(t, \eta)) \end{aligned} \quad (9.8.11)$$

for all $(t, \eta) \in [0, \epsilon_u] \times S^1$.

If we rewrite (9.8.8) and (9.8.11) using the local representation of V_1 and V_2 provided in (9.8.1), and our knowledge of the Lie derivative of a function along a vector field, then we respectively get

$$\begin{aligned} \frac{\partial}{\partial r} \left((Q^*(\omega_2)(V_2)) \circ \check{\phi} \right) \Big|_{(r, \theta) = \check{\phi}^{-1}(t, \eta)} &= \frac{1}{2} Q^*(\rho)(t, \eta)(V_2(t, \eta)) \\ &= \frac{1}{2} \left((Q^*(\rho)(V_2)) \circ \check{\phi} \right) \Big|_{(r, \theta) = \check{\phi}^{-1}(t, \eta)} \end{aligned} \quad (9.8.12)$$

and

$$\frac{\partial}{\partial r} \left((Q^*(\rho)(V_2)) \circ \check{\phi} \right) \Big|_{(r, \theta) = \check{\phi}^{-1}(t, \eta)} = -\frac{1}{2} \kappa(G(P(t, \eta))) Q^*(\omega_2)(t, \eta)(V_2(t, \eta))$$

$$= -\frac{1}{2}\kappa(G(P(t, \eta))) \left((Q^*(\omega_2)(V_2)) \circ \check{\phi} \right) \Big|_{(r, \theta) = \check{\phi}^{-1}(t, \eta)} \quad (9.8.13)$$

for all $(t, \eta) \in [0, \epsilon_{\mathbf{u}}] \times S^1$. Substituting (9.8.13) in the derivative with respect to r of (9.8.12) yields

$$\frac{\partial^2}{\partial r^2} \left((Q^*(\omega_2)(V_2)) \circ \check{\phi} \right) \Big|_{(r, \theta) = \check{\phi}^{-1}(t, \eta)} = -\frac{1}{4}\kappa(G(P(t, \eta))) \left((Q^*(\omega_2)(V_2)) \circ \check{\phi} \right) \Big|_{(r, \theta) = \check{\phi}^{-1}(t, \eta)}$$

for all $(t, \eta) \in [0, \epsilon_{\mathbf{u}}] \times S^1$. For this second order differential equation to completely determine $Q^*(\omega_2)(V_2) \circ \check{\phi}$ given η fixed (i.e. along a geodesic), we need initial condition at $t = 0$ for $Q^*(\omega_2)(V_2) \circ \check{\phi}$ and its derivative with respect to r . We have from the fourth item of Lemma 9.8.4 that

$$Q^*(\omega_2)(V_2) \circ \check{\phi} \Big|_{(r, \theta) = \check{\phi}^{-1}(0, \eta)} = \omega_2(Q(0, \eta)) \left(Q_*(V_2(t, \eta)) \Big|_{t=0} \right) = \omega_2(\tilde{\sigma}_\eta(0)) (H(\tilde{\sigma}_\eta(0))) = 0$$

by definition of ω_2 and H . We also get from (9.8.12) that

$$\begin{aligned} \frac{\partial}{\partial r} \left((Q^*(\omega_2)(V_2)) \circ \check{\phi} \right) \Big|_{(r, \theta) = \check{\phi}^{-1}(0, \eta)} &= \frac{1}{2} Q^*(\rho)(t, \eta)(V_2(t, \eta)) \Big|_{t=0} \\ &= \rho(Q(0, \eta)) (Q_*(V_2(t, \eta)) \Big|_{t=0}) = \frac{1}{2} \rho(\tilde{\sigma}_\eta(t)) (H(\tilde{\sigma}_\eta(0))) = \frac{1}{2}. \end{aligned}$$

Since $(Q^*(\omega_2)(V_2) \circ \check{\phi}) \Big|_{(r, \theta) = \check{\phi}^{-1}(0, \eta)} = 0$ and $\frac{\partial}{\partial r} \left((Q^*(\omega_2)(V_2)) \circ \check{\phi} \right) \Big|_{(r, \theta) = \check{\phi}^{-1}(0, \eta)} > 0$, we have that $(Q^*(\omega_2)(V_2) \circ \check{\phi}) \Big|_{(r, \theta) = \check{\phi}^{-1}(t, \eta)} > 0$ for $t \geq 0$ small enough. It then follows from (9.8.5) that

$$\|G_*(P_*(V_2(t, \eta)))\|_{G(P(t, \eta))} = \|G_*(P_*(V_2(t, \eta)))\|_{\pi_S(Q(t, \eta))} = Q^*(\omega_2)(t, \eta)(V_2(t, \eta))$$

for $t \geq 0$ small enough. We have proved the following theorem.

Theorem 9.8.10

Let

1. S be an oriented 2-dimensional Riemannian manifold and $\mathbf{u} \in S$,
2. $G : D_{\epsilon_{\mathbf{u}}} \rightarrow S$ be a geodesic coordinate system for S about $\mathbf{u}$,
3. $t \mapsto \sigma_\eta(t) = G(P(t, \eta))$ for $\eta = e^{i\theta} \in S^1$ fixed be a geodesic on S through $\mathbf{u}$, and
4. $h(r) = (Q^*(\omega_2)(V_2) \circ \check{\phi})(r, \theta) = \|G_*(P_*(V_2(\check{\phi}(r, \theta))))\|_{G(P(\check{\phi}(r, \theta)))}$ for $(t, \eta) = \check{\phi}(r, \theta)$.

Then

$$\frac{d^2 h}{dr^2}(r) + \frac{1}{4}\kappa(G(P(\check{\phi}(r, \theta))))h(r) = 0 \quad (9.8.14)$$

with $h(0) = 0$ and $h'(0) = 1/2$. This ordinary differential equation is called the **Jacobi's equation**.

The previous theorem states that the size of the vector field $G_*(P_*(V_2))$ along the geodesic σ_η satisfies an ordinary differential equation. Note that the Jacobi's equation (9.8.14) often stated in the literature does not include the coefficient $1/4$ and the initial condition on the first derivative of h at the origin is 1 instead of $1/2$. The difference between our formulation and the general formulation of Jacobi's equation comes from the difference between our definition of wedge product and the definition of wedge product used by some authors. They do not include the factor $1/2$ in their definition.

One of the fundamental roles that is played by geodesic curves is given in the following theorem.

Theorem 9.8.11

Let

1. S be an oriented 2-dimensional Riemannian manifold and $\mathbf{u} \in S$,
2. $G : D_{\epsilon_{\mathbf{u}}} \rightarrow S$ be a geodesic coordinate system for S about $\mathbf{u}$,
3. $t \mapsto \sigma_\eta(t) = G(P(t, \eta))$ for $\eta = e^{i\theta} \in S^1$ fixed be a geodesic on S through $\mathbf{u}$, and
4. $\tilde{\mathbf{u}} = \sigma_\eta(\epsilon)$ for some $0 < \epsilon < \epsilon_{\mathbf{u}}$.

Then the length of the geodesic curve σ_η between $\mathbf{u}$ and $\tilde{\mathbf{u}}$ is less or equal to the length of any piecewise C^∞ curves in S between $\mathbf{u}$ and $\tilde{\mathbf{u}}$.

Proof.

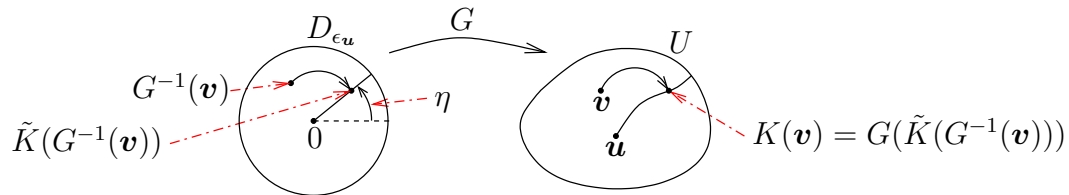
Recall that the length of a curve given by the piecewise C^∞ parametric representation $\alpha : [a, b] \rightarrow S$ is given by

$$L_\alpha = \int_a^b \|\alpha_*(t, 1)\|_{\mathbf{u}} dt .$$

Let $\tilde{K} : D_{\epsilon_{\mathbf{u}}} \rightarrow D_{\epsilon_{\mathbf{u}}}$ be the map defined by

$$\tilde{K}(z) = \begin{cases} t\eta & \text{if } z = t\mu \text{ with } t > 0 \\ 0 & \text{if } z = 0 \end{cases}$$

where $\mu \in S^1$. We have also used the complex representation of the elements of $D_{\epsilon_{\mathbf{u}}}$. Let $U = G(D_{\epsilon_{\mathbf{u}}}) \subset S$ and $K = G \circ \tilde{K} \circ G^{-1} : U \rightarrow U$.



i) We first prove that $K_* : T_{\mathbf{v}}U \rightarrow T_{K(\mathbf{v})}U$ satisfies $\|K_*(\mathbf{v}, \mathbf{x})\|_{K(\mathbf{v})} \leq \|(\mathbf{v}, \mathbf{x})\|_{\mathbf{v}}$ for all $(\mathbf{v}, \mathbf{x}) \in T_{\mathbf{v}}U$.

We have from Gauss' lemma that $\{G_*(P_*(V_1(t, \xi))), G_*(P_*(V_2(t, \xi)))\}$ is an orthogonal basis of $T_v U$ where $v = G(P(t, \xi)) = \pi_S(Q(t, \xi))$. Given $(v, x) \in T_v U$, we may write

$$(v, x) = aG_*(P_*(V_1(t, \xi))) + bG_*(P_*(V_2(t, \xi)))$$

where a and b are the orthogonal projection of (v, x) on $G_*(P_*(V_1(t, \xi)))$ and $G_*(P_*(V_2(t, \xi)))$ respectively.

Since $t \mapsto \sigma_\xi(t) = G(P(t, \xi))$ is a geodesic, we have that $(\sigma_\xi(t), \sigma'_\xi(t)) = (G \circ P)_*(V_1(t, \xi))$ is of norm 1. It follows from the orthogonality of $G_*(P_*(V_1(t, \xi)))$ and $G_*(P_*(V_2(t, \xi)))$ that

$$\begin{aligned} \|(v, x)\|_v &= |a| \|G_*(P_*(V_1(t, \xi)))\|_v + |b| \|G_*(P_*(V_2(t, \xi)))\|_v \\ &= |a| + |b| \|G_*(P_*(V_2(t, \xi)))\|_v. \end{aligned}$$

Moreover, we get from $(\tilde{K} \circ P)(t, \xi) = \tilde{K}(t\xi) = t\eta$ that $\frac{\partial}{\partial t}(\tilde{K} \circ P)(t, e^{i\nu}) = \frac{\partial}{\partial t}(te^{i\theta}) = e^{i\theta} = \eta$ and $\frac{\partial}{\partial \nu}(\tilde{K} \circ P)(t, e^{i\nu}) = \frac{\partial}{\partial \nu}(te^{i\theta}) = 0$. Thus

$$\tilde{K}_*(P_*(V_1(t, \xi))) = (\tilde{K} \circ P)_*((t, \xi), (1, 0)) = (t\eta, \eta) = P_*(V_1(t, \eta))$$

and

$$\tilde{K}_*(P_*(V_2(t, \xi))) = (\tilde{K} \circ P)_*((t, \xi), (0, 1)) = (t\eta, 0).$$

where we have use Lemma 9.8.6 for the last equality of the first relation. Since $K_*(G_*(P_*(V_i(t, \eta)))) = G_*(\tilde{K}_*(P_*(V_i(t, \eta))))$ for $i = 1, 2$, we get

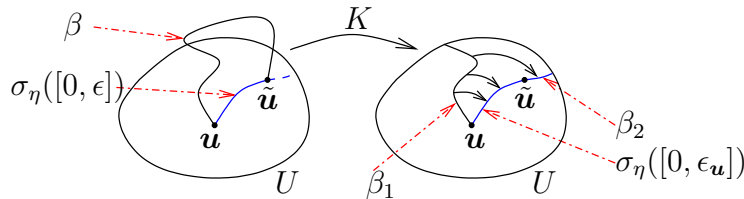
$$K_*(v, x) = aK_*(G_*(P_*(V_1(t, \xi)))) = aG_*(P_*(V_1(t, \eta))).$$

Hence

$$\begin{aligned} \|K_*(v, x)\|_{K(v)} &= |a| \|K_*(G_*(P_*(V_1(t, \xi))))\|_{K(v)} = |a| \|G_*(P_*(V_1(t, \eta)))\|_{K(v)} \\ &= |a| \leq \|(v, x)\|_v \end{aligned}$$

where again we have used the fact that $t \mapsto \sigma_\eta(t) = G(P(t, \eta))$ is a geodesic to get the third equality.

ii) Suppose that $\beta : [a, b] \rightarrow S$ is a piecewise C^∞ curve between u and $\tilde{u}$. If $\beta(t) \notin U$ for some $a \leq t < b$, then we let $c = \inf\{t \in [a, b] : \beta(t) \notin U\}$. We have that $c > a$ because $\beta(a) = u \in U$ and U is an open subset of S . Let $\beta_1 = \beta|_{[a, c]}$. We have that $\beta_1 \in U$ for $a \leq t < c$. Since $L_{\beta_1} \leq L_\beta$, we will get the conclusion of the theorem if we prove that $L_{\sigma_\eta} \leq L_{\beta_1}$.



Let $\beta_2 = K \circ \beta_1$. We have from (i) above that

$$\begin{aligned} L_{\beta_2} &= \int_a^c \|(\beta_2)_*(t, 1)\|_{\beta_2(t)} dt = \int_a^c \|K_*((\beta_1)_*(t, 1))\|_{\beta_2(t)} dt \\ &\leq \int_a^c \|(\beta_1)_*(t, 1)\|_{\beta_1(t)} dt = L_{\beta_1} . \end{aligned}$$

We have that $\beta_2([a, c]) = \sigma_\eta([0, \epsilon_u])$. Since $\sigma_\eta : [0, \epsilon_u[\rightarrow \sigma_\eta([0, \epsilon_u])$ is a bijection, we may define $g : [a, c[\rightarrow [0, \epsilon_u[$ by $g = \sigma_\eta^{-1} \circ \beta_2$. The function is piecewise C^∞ . We get

$$\begin{aligned} L_{\beta_2} &= \int_a^c \|(\beta_2)_*(t, 1)\|_{\beta_2(t)} dt = \int_a^c \|(\sigma_\eta \circ g)_*(t, 1)\|_{\beta_2(t)} dt = \int_a^c \|(\sigma_\eta)_*(g_*(t, 1))\|_{\beta_2(t)} dt \\ &= \int_a^c \|(\sigma_\eta)_*(g(t), g'(t))\|_{\beta_2(t)} dt = \int_a^c \|(\sigma_\eta(g(t)), \sigma'_\eta(g(t))g'(t))\|_{\beta_2(t)} dt \\ &= \int_a^c \|(\sigma_\eta(g(t)), \sigma'_\eta(g(t)))\|_{\beta_2(t)} |g'(t)| dt = \int_0^{\epsilon_u} \|(\sigma_\eta(s), \sigma'_\eta(s))\|_{\sigma_\eta(s)} ds \\ &\geq \int_0^{\epsilon_u} \|\sigma_\eta(s), \sigma'_\eta(s)\|_{\sigma_\eta(s)} ds = L_{\sigma_\eta} . \end{aligned}$$

Theorem 9.8.12

Let S be a connected and oriented 2-dimensional Riemannian manifold, and $\tau : S \times S \rightarrow [0, \infty[$ be the map defined by

$$\tau(\mathbf{u}_1, \mathbf{u}_2) = \inf \{ L_\beta : \beta \text{ is a piecewise } C^\infty \text{ curve in } S \text{ between } \mathbf{u}_1 \text{ and } \mathbf{u}_2 \}$$

Then τ is a metric on S and the topology generated by this metric is identical to the initial topology on the manifold S .

Proof.

We first prove that τ is a metric.

1. Given $\mathbf{u}_1, \mathbf{u}_2 \in S$, since S is connected and so path connected¹¹, there exists at list one piecewise C^∞ curve between $\mathbf{u}_1$ and $\mathbf{u}_2$. Therefore $\tau(\mathbf{u}_1, \mathbf{u}_2)$ is defined.
2. It is clear that $\tau(\mathbf{u}_1, \mathbf{u}_2) = \tau(\mathbf{u}_2, \mathbf{u}_1)$ for all $\mathbf{u}_1, \mathbf{u}_2 \in S$ by definition of τ .
3. Given $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3 \in S$ and $\epsilon > 0$. There exist two piecewise C^∞ curves β_1 and β_2 in S such that β_1 is a curve between $\mathbf{u}_1$ and $\mathbf{u}_2$ satisfying $\tau(\mathbf{u}_1, \mathbf{u}_2) \geq L_{\beta_1} - \epsilon/2$, and β_2 is a curve between $\mathbf{u}_2$ and $\mathbf{u}_3$ satisfying $\tau(\mathbf{u}_2, \mathbf{u}_3) \geq L_{\beta_2} - \epsilon/2$. We then have that $\beta_1\beta_2$ is a curve between $\mathbf{u}_1$ and $\mathbf{u}_3$ such that

$$\tau(\mathbf{u}_1, \mathbf{u}_3) \leq L_{\beta_1\beta_2} \leq L_{\beta_1} + L_{\beta_2} \leq \tau(\mathbf{u}_1, \mathbf{u}_2) + \tau(\mathbf{u}_2, \mathbf{u}_3) + \epsilon .$$

¹¹This is true because S is locally homeomorphic to $\mathbb{R}^2$. In general, connected does not imply path connected.

Since $\epsilon > 0$ is arbitrary, we get

$$\tau(\mathbf{u}_1, \mathbf{u}_3) \leq \tau(\mathbf{u}_1, \mathbf{u}_2) + \tau(\mathbf{u}_2, \mathbf{u}_3).$$

4. If $\mathbf{u}_1 = \mathbf{u}_2$, then $\tau(\mathbf{u}_1, \mathbf{u}_2) = 0$ because we can consider the path $\beta : [0, 1] \rightarrow S$ defined by $\beta(t) = \mathbf{u}_1$ for $0 \leq t \leq 1$.

Suppose that $\mathbf{u}_1 \neq \mathbf{u}_2$. Let $G : D_{\epsilon_{\mathbf{u}_1}} \rightarrow S$ be a geodesic coordinate system for S about $\mathbf{u}_1$. We may choose $\epsilon_{\mathbf{u}_1}$ small enough to have $\mathbf{u}_2 \notin U = G(D_{\epsilon_{\mathbf{u}_1}})$. If $\sigma_\eta(t) = G(P(t, \eta))$ for $0 \leq t < \epsilon_{\mathbf{u}}$ is any geodesic curve, then we have that $L_{\sigma_\eta} \leq L_\beta$ for all piecewise C^∞ curves in S between $\mathbf{u}_1$ and $\mathbf{u}_2$ because $\mathbf{u}_2 \notin U$. Note that $\epsilon_{\mathbf{u}} = L_{\sigma_\eta}$ for all geodesics σ_η because it is parameterized by arc length. We get that $0 < \epsilon_{\mathbf{u}} \leq \tau(\mathbf{u}_1, \mathbf{u}_2)$.

It remains to prove that the topology generated by τ is identical to the initial topology on the manifold S . Let $G : D_{\epsilon_{\mathbf{u}}} \rightarrow S$ be a geodesic coordinate system for S about $\mathbf{u} \in S$. Then, as we did for (4) above, we can show that

$$B_\epsilon(\mathbf{u}) \equiv \{\mathbf{v} \in S : \tau(\mathbf{u}, \mathbf{v}) < \epsilon\} = G(D_\epsilon)$$

for all $0 < \epsilon < \epsilon_{\mathbf{u}}$. Thus, the two topology have the same basis of open sets. ■

Theorem 9.8.13

For $i = 1, 2$, suppose that S_i is an oriented 2-dimensional Riemannian manifold, and that $G_i : D_\epsilon \rightarrow S_i$ is a geodesic coordinate system for S_i about $\mathbf{u}_i$. Let $U_i = G_i(D_\epsilon)$ and $\kappa_i : S_i \rightarrow \mathbb{R}$ be the curvature function on S_i for $1 \leq i \leq 2$. If $\kappa_1 = \kappa_2 \circ G_2 \circ G_1^{-1}$ on $U_1 = G_1(D_\epsilon)$, then $G_2 \circ G_1^{-1}$ is an isometry.

Proof.

We prove that $(G_2 \circ G_1^{-1})_* : TU_1 \rightarrow TU_2$ is an isometry.

We have that $G_2 \circ G_1^{-1}$ maps the geodesic $\sigma_{1,\eta}(t) = (G_1 \circ P)(t, \eta)$ onto the geodesic $\sigma_{2,\eta}(t) = (G_2 \circ P)(t, \eta)$ for $0 \leq t \leq \epsilon$ and $\eta \in S^1$. As we have seen at the beginning of the proof of Gauss' lemma, we have that $(G_i \circ P)_*(V_1(t, \eta)) = \tilde{\sigma}_{i,\eta}(t)$ for $i = 1, 2$. Hence $\|(G_i \circ P)_*(V_1(t, \eta))\|_{(G_i \circ P)(t, \eta)} = \|\tilde{\sigma}_{i,\eta}(t)\|_{\sigma_{i,\eta}(t)} = 1$ for $i = 1, 2$ because the geodesics are parameterized by arc length. Thus $G_2 \circ G_1^{-1}$ preserve the length of the tangent vectors along the geodesics.

It follow from Gauss's lemma that $(G_i \circ P)_*(V_1(t, \eta))$ is orthogonal to $(G_i \circ P)_*(V_2(t, \eta))$ for $i = 1, 2$. Thus $G_2 \circ G_1^{-1}$ preserves the orthogonality at each point along the geodesic. Therefore, to prove that $G_2 \circ G_1^{-1}$ preserves the norm of the tangent vectors at each point of U_1 , it suffices to prove that $G_2 \circ G_1^{-1}$ preserves the norm of $(G_1 \circ P)_*(V_2(t, \eta))$.

As usual, let $(\check{W}, \check{U}, \check{\phi})$ be a local chart of $[0, \epsilon] \times S^1$. Furthermore, let $h_i(r) = \|((G_i \circ P)_*(V_2) \circ \check{\phi})(r, \theta)\|_{(G_i \circ P)(\check{\phi}(r, \theta))}$ where $(t, \eta) = \check{\phi}(r, \theta)$ and $i = 1, 2$. For η fixed, we have from Theorem 9.8.14 that h_i satisfies the Jacobi's equation

$$\frac{d^2 h_i}{dr^2}(r) + \frac{1}{4} \kappa(G_i(P(\check{\phi}(r, \theta)))) h_i(r) = 0$$

with $h_i(0) = 0$ and $h'_i(0) = 1/2$. Since $\kappa_2(G_2(P(\check{\phi}(r, \theta)))) = \kappa_1(G_1(P(\check{\phi}(r, \theta))))$ for $0 \leq r \leq \epsilon$, these two differential equations are identical. By uniqueness of solutions,

$$\|((G_1 \circ P)_*(V_2) \circ \check{\phi})(r, \theta)\|_{(G_1 \circ P)(\check{\phi}(r, \theta))} = h_1(r) = h_2(r) = \|((G_2 \circ P)_*(V_2)(r, \theta) \circ \check{\phi})\|_{(G_2 \circ P)(\check{\phi}(r, \theta))}$$

for $0 \leq r < \epsilon$ and $\eta \in S^1$ arbitrary. $\blacksquare$

Corollary 9.8.14

Suppose that S is a connected and oriented 2-dimensional Riemannian manifold. Moreover, suppose that the curvature κ on S is constant. Given $(\mathbf{u}_i, \mathbf{x}_i) \in S_c(\mathbf{u}_i)$ for $i = 1, 2$, there exist open neighbourhood U_i of $\mathbf{u}_i$ for $i = 1, 2$ and an isometry $g : U_1 \rightarrow U_2$ such that $g_*(\mathbf{u}_1, \mathbf{x}_1) = (\mathbf{u}_2, \mathbf{x}_2)$.

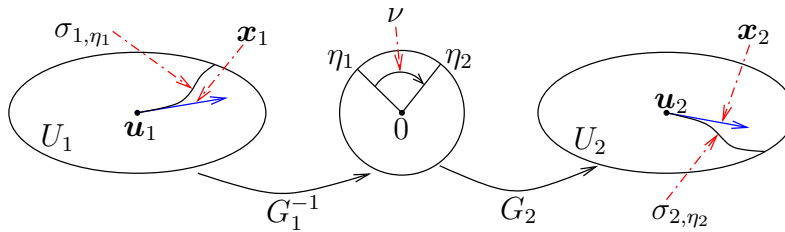
Proof.

Let $G_i : D_\epsilon \rightarrow S$ be a geodesic coordinate system for S about $\mathbf{u}_i$ and set $U_i = G_i(D_\epsilon)$ for $i = 1, 2$.

Choose $\nu \in S^1$ acting on D_ϵ such that $\nu_*((G_1^{-1})_*(\mathbf{u}_1, \mathbf{x}_1)) = (G_2^{-1})_*(\mathbf{u}_2, \mathbf{x}_2)$ and set $\tilde{G}_2 = G_2 \circ \nu$. We have that $\tilde{G}_2 : D_\epsilon \rightarrow S$ is a geodesic coordinate system for S about $\mathbf{u}_2$ because ν maps radial geodesics to radial geodesics in D_ϵ .

The map g is defined by $g = \tilde{G}_2 \circ G_1^{-1} : U_1 \rightarrow U_2$. That g is an isometry comes from the previous theorem with $S_1 = S_2 = S$. We also have

$$g_*(\mathbf{u}_1, \mathbf{x}_1) = (G_2 \circ \nu \circ G_1^{-1})_*(\mathbf{u}_1, \mathbf{x}_1) = (G_2)_*(\nu_*((G_1^{-1})_*(\mathbf{u}_1, \mathbf{x}_1))) = (\mathbf{u}_2, \mathbf{x}_2).$$



This figure summarizes the proof. For $i = 1, 2$, the angle $\eta_i \in S^1$ is such that $\tilde{\sigma}_{i, \eta_i}(0) = Q_i(0, \eta_i) = (\mathbf{u}_i, \mathbf{x}_i)$ as in the construction of G_i for $i = 1, 2$. Moreover, $\sigma_{i, \eta_i} : [0, \epsilon] \rightarrow S$ is a geodesic through $\mathbf{u}_i$. $\blacksquare$

9.9 Euclidean and Non-Euclidean Geometries

The last section of these lecture notes is an introduction to the fascinating and vast subject of non-Euclidean geometry.

9.9.1 Orientation Preserving Isometries

Suppose that S_1 and S_2 are two oriented 2-dimensional Riemannian manifolds, and that $f : S_1 \rightarrow S_2$ is an orientation preserving isometry (see Definition 9.1.2). We list some of the properties of f below.

Let $\tilde{f} = f_*|_{(S_1)_c}$. We have that $\tilde{f} : (S_1)_c \rightarrow (S_2)_c$ according to Proposition 9.2.12.

1) $\tilde{f}$ preserves the differential 1-forms ω_1 , ω_2 and ρ given in Definition 9.4.1 and Theorem 9.3.6.

More precisely, let $\omega_1^{[i]}$, $\omega_2^{[i]}$ and $\rho^{[i]}$ be the forms on S_i given in Definition 9.4.1 and Theorem 9.3.6 for $i = 1, 2$. We have

$$\begin{aligned} \tilde{f}^*(\omega_1^{[2]})(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) &= \omega_1^{[2]}(\tilde{f}(\mathbf{u}, \mathbf{x}))(\tilde{f}_*((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q}))) \\ &= \left\langle ((\pi_{S_2})_* \circ \tilde{f}_*)((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})), \tilde{f}(\mathbf{u}, \mathbf{x}) \right\rangle_{f(\mathbf{u})} \\ &= \left\langle (\tilde{f} \circ (\pi_{S_1})_*)((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})), \tilde{f}(\mathbf{u}, \mathbf{x}) \right\rangle_{f(\mathbf{u})} \\ &= \left\langle \tilde{f}((\pi_{S_1})_*((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q}))), \tilde{f}(\mathbf{u}, \mathbf{x}) \right\rangle_{f(\mathbf{u})} = \left\langle (\pi_{S_1})_*((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})), (\mathbf{u}, \mathbf{x}) \right\rangle_{\mathbf{u}} \\ &= (\omega_1^{[1]})(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \end{aligned}$$

for all $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \in T_{(\mathbf{u}, \mathbf{x})}S_c$, where we have used the second item of Proposition 9.2.2 to get the third equality.

Likewise, we have

$$\tilde{f}^*(\omega_2^{[2]})(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) = (\omega_2^{[1]})(\mathbf{u}, \mathbf{x})((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q}))$$

for all $((\mathbf{u}, \mathbf{x}), (\mathbf{p}, \mathbf{q})) \in T_{(\mathbf{u}, \mathbf{x})}S_c$.

To prove that $\tilde{f}$ preserves the differential 1-form ρ is a little bit more complex. We first prove that $\tilde{f}^*(\rho^{[2]})$ is the Riemann connection on S_1 if $\rho^{[2]}$ is the Riemann connection on S_2 .

Using Proposition 9.2.13, we find that

$$\tilde{f}^*(\rho^{[2]})(\mathbf{u}, \mathbf{x})(H(\mathbf{u}, \mathbf{x})) = \rho^{[2]}(\tilde{f}(\mathbf{u}, \mathbf{x}))(\tilde{f}_*(H(\mathbf{u}, \mathbf{x}))) = \rho^{[2]}(\tilde{f}(\mathbf{u}, \mathbf{x}))(H(\tilde{f}(\mathbf{u}, \mathbf{x}))) = 1.$$

Moreover, it follows from the third item of Proposition 9.2.2 and Proposition 9.3.8 that

$$\eta^*(\tilde{f}^*(\rho^{[2]})) = (\tilde{f} \circ \eta)^*(\rho^{[2]}) = (\eta \circ \tilde{f})^*(\rho^{[2]}) = \tilde{f}^*(\eta^*(\rho^{[2]})) = \tilde{f}^*(\rho^{[2]}).$$

Hence, we get from Proposition 9.3.9 that $\tilde{f}^*(\rho^{[2]})$ is a connection 1-form.

Since $d\omega_1^{[2]} = \rho^{[2]} \wedge \omega_2^{[2]}$ and $d\omega_2^{[2]} = -\rho^{[2]} \wedge \omega_1^{[2]}$, we get

$$d\omega_1^{[1]} = d(\tilde{f}^*(\omega_1^{[2]})) = \tilde{f}^*(d\omega_1^{[2]}) = \tilde{f}^*(\rho^{[2]} \wedge \omega_2^{[2]}) = \tilde{f}^*(\rho^{[2]}) \wedge \tilde{f}^*(\omega_2^{[2]}) = \tilde{f}^*(\rho^{[2]}) \wedge \omega_2^{[1]}$$

and similarly

$$d\omega_2^{[1]} = d(\tilde{f}^*(\omega_2^{[2]})) = \tilde{f}^*(d\omega_2^{[2]}) = -\tilde{f}^*(\rho^{[2]} \wedge \omega_1^{[2]}) = -\tilde{f}^*(\rho^{[2]}) \wedge \tilde{f}^*(\omega_1^{[2]}) = -\tilde{f}^*(\rho^{[2]}) \wedge \omega_1^{[1]}.$$

Thus, we get that $\rho^{[1]} = \tilde{f}^*(\rho^{[2]})$ by uniqueness of the Riemann connection as stated in Theorem 9.4.7.

2) f preserve the curvature. More precisely, $\kappa_1 = \kappa_2 \circ f$ where κ_i is the curvature function on S_i associated to the connection 1-form $\rho^{[i]}$ for $i = 1, 2$.

We have that $d\rho^{[1]} = -(\kappa_1 \circ \pi_{S_1})\omega_1^{[1]} \wedge \omega_2^{[1]}$ and $d\rho^{[2]} = -(\kappa_2 \circ \pi_{S_2})\omega_1^{[2]} \wedge \omega_2^{[2]}$. Hence, using the first item of Proposition 9.2.2, we get

$$\begin{aligned} -(\kappa_1 \circ \pi_{S_1})\omega_1^{[1]} \wedge \omega_2^{[1]} &= d\rho^{[1]} = d(\tilde{f}^*(\rho^{[2]})) = \tilde{f}^*(d\rho^{[2]}) = -\tilde{f}^*((\kappa_2 \circ \pi_{S_2})\omega_1^{[2]} \wedge \omega_2^{[2]}) \\ &= -(\kappa_2 \circ \pi_{S_2} \circ \tilde{f})\tilde{f}^*(\omega_1^{[2]}) \wedge \tilde{f}^*(\omega_2^{[2]}) = -(\kappa_2 \circ f \circ \pi_{S_1})\omega_1^{[1]} \wedge \omega_2^{[1]}. \end{aligned}$$

Thus $\kappa_1 \circ \pi_{S_1} = \kappa_2 \circ \tilde{f} \circ \pi_{S_1}$. Since π_{S_1} is onto, we get that $\kappa_1 = \kappa_2 \circ f$.

3) f preserves horizontal lifts. More precisely, if $\sigma : [a, b] \rightarrow S_1$ and $\tilde{\sigma} : [a, b] \rightarrow (S_1)_c$ is the horizontal lift of σ through $(\sigma(a), \mathbf{x}) \in (S_1)_c$, then $\tilde{\alpha} = \tilde{f} \circ \tilde{\sigma} : [a, b] \rightarrow (S_2)_c$ is the horizontal lift of $\alpha = f \circ \sigma : [a, b] \rightarrow S_2$ through $\tilde{f}(\sigma(a), \mathbf{x})$. This is a consequence of Proposition 9.3.13 because

1. $\tilde{\alpha}(a) = \tilde{f}(\tilde{\sigma}(a)) = \tilde{f}(\sigma(a), \mathbf{x})$,
2. $\pi_{S_2} \circ \tilde{\alpha} = \pi_{S_2} \circ \tilde{f} \circ \tilde{\sigma} = f \circ \pi_{S_1} \circ \tilde{\sigma} = f \circ \sigma = \alpha$ where the first item of Proposition 9.2.2 was used to get the second equality, and
- 3.

$$\begin{aligned} \rho^{[2]}(\tilde{\alpha}(t))(\tilde{\alpha}_*(t, 1)) &= \rho^{[2]}((\tilde{f} \circ \tilde{\sigma})(t))(\tilde{f} \circ \tilde{\sigma})_*(t, 1) = \rho^{[2]}(\tilde{f}(\tilde{\sigma}(t)))(\tilde{f}_*(\tilde{\sigma}_*(t, 1))) \\ &= \tilde{f}^*(\rho^{[2]})(\tilde{\sigma}(t))(\tilde{\sigma}_*(t, 1)) = \rho^{[1]}(\tilde{\sigma}(t))(\tilde{\sigma}_*(t, 1)) = 0 \end{aligned}$$

where the last equality comes from the fact that $\tilde{\sigma}$ is the horizontal lift of σ through $(\sigma(a), \mathbf{x})$.

4) The function f maps geodesic into geodesic. More precisely, if $\sigma : [a, b] \rightarrow S_1$ is a smooth curve and $\underline{\sigma} : [a, b] \rightarrow (S_1)_c$ is the geodesic associated to $\sigma : [a, b] \rightarrow S_1$, then by definition $\underline{\sigma}(t) = (\sigma(t), \sigma'(t))$ is the horizontal lift $\tilde{\sigma}$ of σ with $\tilde{\sigma}(a) = (\sigma(a), \sigma'(a))$. According to (3) above, $\tilde{f} \circ \underline{\sigma}$ is the horizontal lift of $f \circ \sigma$ with $(\tilde{f} \circ \underline{\sigma})(a) = \tilde{f}(\sigma(a), \sigma'(a))$. But we also have that $(\tilde{f} \circ \underline{\sigma})(t) = \tilde{f}(\sigma(t), \sigma'(t)) = f_*(\sigma(t), \sigma'(t)) = ((f \circ \sigma)(t), (f \circ \sigma)'(t))$ for $a \leq t \leq b$. Thus $\tilde{f} \circ \underline{\sigma} : [a, b] \rightarrow (S_2)_c$ is the geodesic associated to $f \circ \sigma : [a, b] \rightarrow S_2$.

9.9.2 Group of Orientation Preserving Isometries

Proposition 9.9.1

Let S be an oriented Riemannian manifold. The set $\mathcal{E}_S$ of all orientation preserving ismetries from S to itself forms a group with respect to the composition of functions. It is called the **group of orientation preserving isometries** of S .

Proof.

This proposition is a consequence of the fact that the composition of orientation preserving isometries is still an orientation preserving isometry. Note that the identity map is the unit in $\mathcal{E}_S$. ■

Definition 9.9.2

We say that a subgroup $\mathcal{E}_S^o$ of the group $\mathcal{E}_S$ is **transitive** on S if, for any $\mathbf{u}_1, \mathbf{u}_2 \in S$, there exists $f \in \mathcal{E}_S^o$ such that $f(\mathbf{u}_1) = \mathbf{u}_2$.

Proposition 9.9.3

Let S be an oriented Riemannian manifold. If the group $\mathcal{E}_S$ is transitive on S , then S has a constant curvature.

Proof.

Choose $\mathbf{u} \in S$. Given any $\mathbf{v} \in S$, there exists $f \in \mathcal{E}_S$ such that $\mathbf{v} = f(\mathbf{u})$. Since f is an orientation preserving isometry, it preserves curvature according to the item (2) of the previous subsection. Hence, we have that $\kappa(\mathbf{u}) = (\kappa \circ f)(\mathbf{u}) = \kappa(f(\mathbf{u})) = \kappa(\mathbf{v})$. Since $\mathbf{v} \in S$ is arbitrary, we get that κ is constant. ■

Proposition 9.9.4

Let S be an oriented 2-dimensional Riemannian manifold. Given $\mathbf{u} \in S$, let $\mathcal{E}_S(\mathbf{u}) = \{f \in \mathcal{E}_S : f(\mathbf{u}) = \mathbf{u}\}$. Then $\mathcal{E}_S(\mathbf{u})$ is a subgroup of $\mathcal{E}_S$. It is called the **isotropy group** of S at $\mathbf{u}$.

Proof.

If $f_1, f_2 \in \mathcal{E}_S(\mathbf{u})$, then $f_1 \circ f_2 \in \mathcal{E}_S(\mathbf{u})$ because $f_1 \circ f_2 \in \mathcal{E}_S$ since $\mathcal{E}_S$ is a group and $(f_1 \circ f_2)(\mathbf{u}) = f_1(f_2(\mathbf{u})) = f_1(\mathbf{u}) = \mathbf{u}$. We obviously have that $\text{Id} \in \mathcal{E}_S(\mathbf{u})$. ■

Given $\mathbf{u} \in S$ fixed, we have from Proposition 9.2.3 that $f_* : T_{\mathbf{u}}S \rightarrow T_{\mathbf{u}}S$ is a rotation for all $f \in \mathcal{E}_S$ such that $f(\mathbf{u}) = \mathbf{u}$. Hence, we may define a map $\Phi : \mathcal{E}_S(\mathbf{u}) \rightarrow S^1$ that assign for each $f \in \mathcal{E}_S(\mathbf{u})$ the angle of rotation of $f_* : T_{\mathbf{u}}S \rightarrow T_{\mathbf{u}}S$.

Proposition 9.9.5

Let S be a connected and oriented 2-dimensional Riemannian manifold. Given $\mathbf{u} \in S$, the map $\Phi : \mathcal{E}_S(\mathbf{u}) \rightarrow S^1$ is one-to-one.

Proof.

Suppose that $f \in \mathcal{N}(\Phi)$.

Let $G : D_{\epsilon_{\mathbf{u}}} \rightarrow S$ be a geodesic coordinate system of S about $\mathbf{u}$. We assume that $D_{\epsilon_{\mathbf{u}}} \cong B_{\epsilon_{\mathbf{u}}}(\mathbf{0}) \subset \mathbb{R}^2$. We have that f maps $G(D_{\epsilon_{\mathbf{u}}})$ into itself for all $f \in \mathcal{E}_S(\mathbf{u})$ because $f(\mathbf{u}) = \mathbf{u}$ and f is an orientation preserving isometry. To be more explicit, it follows from

item (4) in the previous subsection that the length along the geodesic from $\mathbf{u}$ to $f(\mathbf{v})$ is equal to length along the geodesic from $\mathbf{u}$ to $\mathbf{v}$ for all $\mathbf{v} \in G(D_\epsilon)$. Therefore, the length along the geodesic from $\mathbf{u}$ to $f(\mathbf{v})$ is less than ϵ_u . We then get that $f(\mathbf{v}) \in D_{\epsilon_u}$ if $\mathbf{v} \in D_{\epsilon_u}$. Recall that the geodesics are parameterized by arc length.

Since $g_f = G^{-1} \circ f \circ G : D_{\epsilon_u} \rightarrow D_{\epsilon_u}$ is an orientation preserving isometry leaving the origin fix, we get as in Proposition 9.2.3 that g_f is a rotation. Thus $(g_f)_*(\mathbf{0}, \mathbf{y}) = (\mathbf{0}, g_f(\mathbf{y}))$. Since $f_* = (G \circ g_f \circ G^{-1})_* = G_* \circ (g_f)_* \circ G_*^{-1}$, we get

$$\begin{aligned} f \in \mathcal{N}(\Phi) &\iff f_* = \text{Id}_{T_{\mathbf{u}}S} \iff (g_f)_* = \text{Id}_{T_{\mathbf{0}}D_\epsilon} \iff g_f = \text{Id}_{D_{\epsilon_u}} \\ &\iff f|_{G(D_{\epsilon_u})} = G \circ g_f \circ G^{-1} = \text{Id}_{G(D_{\epsilon_u})} . \end{aligned}$$

Therefore, we have

$$G(D_{\epsilon_u}) \subset V \equiv \{\mathbf{v} \in S : f(\mathbf{v}) = \mathbf{v} \text{ and } \text{Id}_{T_{\mathbf{v}}S} = f_* : T_{\mathbf{v}}S \rightarrow T_{\mathbf{v}}S\} .$$

Note that $f(\mathbf{v}) = \mathbf{v}$ for all $\mathbf{v} \in G(D_{\epsilon_u})$ implies that $\text{Id}_{T_{\mathbf{v}}S} = f_* : T_{\mathbf{v}}S \rightarrow T_{\mathbf{v}}S$ for all $\mathbf{v} \in G(D_{\epsilon_u})$ because f_* is defined locally. Since the previous reasoning can be repeated with $\mathbf{u}$ replaced by any $\mathbf{v} \in V$, we find that $V \subset S$ is an open set because the sets of the form $G(D_{\epsilon_v})$ are open sets according to Definition 9.8.3.

Since $f : S \rightarrow S$ is smooth, we have that V is also closed¹² This implies that $V = S$ because V is a non-empty open and closed subset of the connected set S . Therefore $f = \text{Id}_S$. ■

Definition 9.9.6

Let S be an oriented 2-dimensional Riemannian manifold. Given $\mathbf{u}, \mathbf{v} \in S$, the **orbit** of $\mathcal{E}_S(\mathbf{u})$ through $\mathbf{v}$ is the set $\mathcal{O}_{\mathcal{E}_S(\mathbf{u})}(\mathbf{v}) = \{f(\mathbf{v}) : f \in \mathcal{E}_S(\mathbf{u})\}$.

Theorem 9.9.7

Let S be an oriented 2-dimensional Riemannian manifold. Suppose that the map $\Phi : \mathcal{E}_S(\mathbf{u}) \rightarrow S^1$ is onto for some $\mathbf{u} \in S$. If $G : D_{\epsilon_u} \rightarrow S$ is a geodesic coordinate system of S about $\mathbf{u}$, then the geodesic through $\mathbf{u}$ are the orthogonal curves (parameterized by arc length) to the orbits of $\mathcal{E}_S(\mathbf{u})$ through $\mathbf{v} \in G(D_{\epsilon_u})$ (Figure 9.13).

Proof.

We assume that $D_{\epsilon_u} \cong B_{\epsilon_u}(\mathbf{0}) \subset \mathbb{R}^2$.

Given $\mathbf{v} \in G(D_\epsilon)$, let $\mathbf{w} = G(\mathbf{v})$. As we have seen in the proof of Proposition 9.9.5, $G^{-1} \circ f \circ G = g_f$ is a rotation in D_{ϵ_u} . Since $\Phi : \mathcal{E}_S(\mathbf{u}) \rightarrow S^1$ is onto, we have that $C_v = \{G^{-1}f(\mathbf{v}) = g_f(\mathbf{w}) : f \in \mathcal{E}_S(\mathbf{u})\}$ is a full circle in D_{ϵ_u} .

¹²This follows from the fact that f and f_* are continuous. If (W, U, ϕ) is a local chart about $\mathbf{u}$ and $\mathbf{u} = \phi(\mathbf{w})$, then $(\phi^{-1} \circ f \circ \phi)(\mathbf{w}) = \lim_{i \rightarrow \infty} (\phi^{-1} \circ f \circ \phi)(\mathbf{w}_i) = \lim_{i \rightarrow \infty} \mathbf{w}_i = \mathbf{w}$ for $\{\mathbf{w}_i\}_{i \in \mathbb{N}} \subset \phi^{-1}(V \cap U)$ converging to $\mathbf{w}$. Likewise, $D(\phi^{-1} \circ f \circ \phi)(\mathbf{w})\mathbf{y} = \lim_{i \rightarrow \infty} D(\phi^{-1} \circ f \circ \phi)(\mathbf{w}_i)\mathbf{y} = \lim_{i \rightarrow \infty} \text{Id}_{\mathbb{R}^k} \mathbf{y} = \mathbf{y}$ for all $\mathbf{y} \in \mathbb{R}^k$.

Since the circles in D_{ϵ_u} (i.e. C_v) are orthogonal to the radial geodesics in D_{ϵ_u} , we get that the orbits $\mathcal{O}_{\mathcal{E}_S(\mathbf{u})}(\mathbf{v}) = \{f(\mathbf{v}) : f \in \mathcal{E}_S(\mathbf{u})\} = G(C_v)$ are orthogonal to the geodesic through $\mathbf{u}$ in $G(D_{\epsilon_u})$ according to Gauss' lemma because $P_*(V_1(t, \eta))$ and $P_*(V_2(t, \eta))$ in the statement of the lemma refer to the radial and rotational directions in D_{ϵ_u} respectively. ■

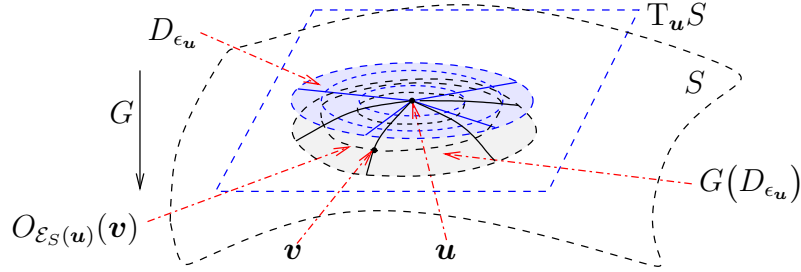


Figure 9.13: Representation of the orbits of $\mathcal{E}_S(\mathbf{u})$ through $\mathbf{v} \in S$.

The next theorem will be useful to determine the full group of orientation preserving isometry of a manifold.

Proposition 9.9.8

Let S be an oriented 2-dimensional Riemannian manifold. Suppose that $\mathcal{E}_S^o$ is a subgroup of the group $\mathcal{E}_S$ of orientation preserving isometry of S . If $\mathcal{E}_S^o$ is transitive on S and $\Phi : \mathcal{E}_S^o \cap \mathcal{E}_S(\mathbf{u}) \rightarrow S^1$ is onto for some $\mathbf{u} \in S$, then $\mathcal{E}_S^o = \mathcal{E}_S$.

Proof.

Given $f_1 \in \mathcal{E}_S$, we may use the transitivity of $\mathcal{E}_S^o$ to find $f_2 \in \mathcal{E}_S^o$ such that $f_2(f_1(\mathbf{u})) = \mathbf{u}$. Hence $f_2 \circ f_1 \in \mathcal{E}_S(\mathbf{u})$. Since Φ maps $\mathcal{E}_S^o \cap \mathcal{E}_S(\mathbf{u})$ onto S^1 , there exists $f_3 \in \mathcal{E}_S^o \cap \mathcal{E}_S(\mathbf{u})$ such that $\Phi(f_3) = \Phi(f_2 \circ f_1)$. Since $\Phi : \mathcal{E}_S(\mathbf{u}) \rightarrow S^1$ is one-to-one, we get that $f_3 = f_2 \circ f_1$. Thus $f_1 = f_2^{-1} \circ f_3 \in \mathcal{E}_S^o$. We have proved that $\mathcal{E}_S \subset \mathcal{E}_S^o$. But $\mathcal{E}_S^o$ is a subgroup of $\mathcal{E}_S$. Therefore $\mathcal{E}_S^o = \mathcal{E}_S$. ■

9.9.3 Manifolds of Constant Curvatures

Oriented Riemannian manifolds with constant curvature can be classified into three categories according to the type of geometry that they have. Let S be an oriented Riemannian manifolds with constant curvature κ .

1. If $\kappa = 0$, then the geometry on S is the Euclidean geometry.
2. If $\kappa > 0$, then the geometry on S is the elliptic geometry.
3. If $\kappa < 0$, then the geometry on S is the hyperbolic geometry.

The Euclidean geometry is well known. Its main feature is that, given a line ℓ and a point $\mathbf{p}$ not on ℓ , there exists one and only one line through $\mathbf{p}$ parallel to ℓ . This is equivalent to the property that the sum of the interior angles of a triangle is equal to π .

In elliptic geometry, the main feature is that, given a line ℓ and a point $\mathbf{p}$ not on ℓ , there does not exist any line through $\mathbf{p}$ parallel to ℓ . This is equivalent to the property that the sum of the interior angles of a triangle is greater than π .

Finally, in hyperbolic geometry, the main feature is that, given a line ℓ and a point $\mathbf{p}$ not on ℓ , there exist infinitely many lines through $\mathbf{p}$ parallel to ℓ . This is equivalent to the property that the sum of the interior angles of a triangle is less than π .

A excellent elementary introduction to the subject of Euclidean and non-Euclidean geometries is [8].

We are now going to illustrate with some example these different geometries on oriented 2-dimensional Riemannian manifolds with constant curvature.

Example 9.9.9 (Euclidean Geometry)

We consider $S = \mathbb{R}^2$ with the standard Euclidean norm and orientation. This is an oriented 2-dimensional Riemannian manifold.

We have already seen in Example 9.4.9 that S has null curvature.

The group G formed of all rotations and translation on $S = \mathbb{R}^2$. It is clearly a subgroup of the group $\mathcal{E}_S$. We may wonder if there are other orientation preserving isometries on S . Proposition 9.9.8 provided the answer to this question. It is clear that G is transitive because it contains all the translations. We also have that $\Phi : G \cap \mathcal{E}_S(\mathbf{0}) \rightarrow S^1$ is onto since G contains all the rotation about the origin. It then follows from Proposition 9.9.8 that $\mathcal{E}_S = G$ ¹³.

It follows from Theorem 9.9.7 that the geodesic through a point $\mathbf{u} \in S$ are traditional straight lines (suitably parameterized) because the orbits of $\mathcal{E}_S(\mathbf{u})$ are circles centred at $\mathbf{u}$. This should not be surprising because it is well know that in Euclidean geometry, the shortest path from one point to another is the straight line. ♣

The next example is a little surprising but not extremely surprising if we think about it.

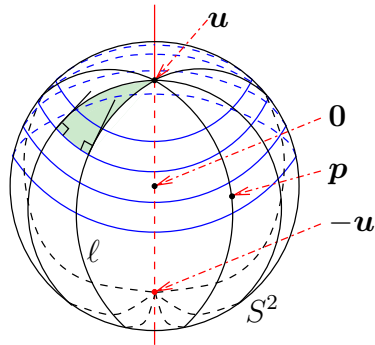
Example 9.9.10 (Elliptic Geometry)

We consider $S = S^2$ with the Riemannian metric induced as a submanifold of $\mathbb{R}^3$. The orientation on S^2 is the orientation induced by the standard outward unit normal to S^2 at each point of S^2 as defined in Definition 5.3.18.

The group G of rotations in $\mathbb{R}^3$ centred at the origin is a transitive group of orientation preserving isometries on S^2 . It therefore follows from Proposition 9.9.8 that $\mathcal{E}_S = G$. Moreover, we get from Proposition 9.9.3 that the curvature κ of S^2 is constant. It follows from Example 9.6.5 that $4\pi = \int_{S^2} \kappa \, dV = \kappa \int_{S^2} dV = 4\pi\kappa$. Thus $\kappa = 1 > 0$.

¹³It is also true that the group of all isometries (no orientation preserving required) on $\mathbb{R}^2$ is equal to the Euclidean group $E(2)$ formed of all rotations, reflections through plans containing the origin and translation on $S = \mathbb{R}^2$.

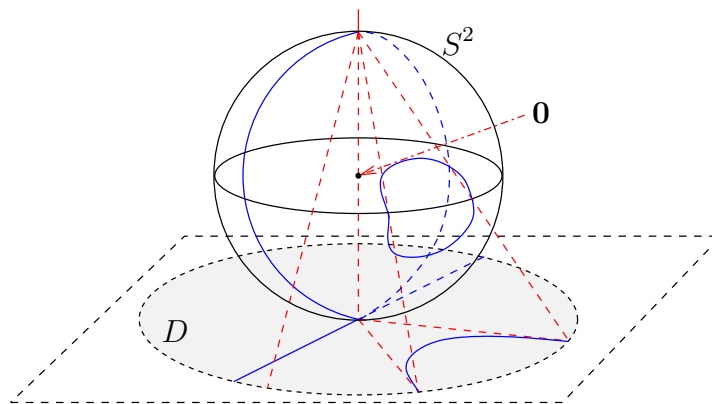
Given $\mathbf{u} \in S^2$, the isotropy group of S^2 at $\mathbf{u}$ is the group of all rotations in $\mathbb{R}^3$ whose axis of rotation is the line through $\mathbf{u}$ and $\mathbf{0}$. Thus the orbits of $\mathcal{E}_S(\mathbf{u})$ are the parallels and the geodesic through $\mathbf{u}$ are the great circle through $\mathbf{u}$. In elliptic geometry, the “straight lines” are the geodesics (i.e. the great circles) since they are associated to the path of shortest distance between two points on S^2 . Hence, give a great circle ℓ and a point $\mathbf{p}$ not on ℓ , there does not exist any great circle through $\mathbf{p}$ that does not intersect ℓ . Note also that the sum of the internal angles of a triangle is greater than π as illustrated in the figure below.



Example 9.9.11 (Hyperbolic Geometry)

It is much harder to provide a simple example of a oriented Riemannian manifold of constant negative curvature. The pseudosphere is probably the simplest example in $\mathbb{R}^3$. Nevertheless, the example that is almost always presented in the literature is the **Poincaré Model**. It is the imbedding in $\mathbb{R}^2$ of an oriented Riemannian manifold S of constant negative curvature. This imbedding is not an isometry. It has been proved that S can only be realized in $\mathbb{R}^n$ with $n > 3$.

The following figure illustrate how the Poincaré model is built.



The “straight lines” in this model are the lines through the origin and the arcs of circles that intersect perpendicularly the boundary ∂D of $D = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\| < r\}$ for some $r > 0$. They are the stereographic projections limited to the lower half of S^2 of all the circles ¹⁴

¹⁴A circle on S^2 is defined by the intersection of a sphere with S^2 .

that intersect perpendicularly the equator of S^2 .

The **Poincaré metric** on $T_{\mathbf{u}}D$ for $\mathbf{u} \in D$ is defined by $\langle (\mathbf{u}, \mathbf{x}_1), (\mathbf{u}, \mathbf{x}_2) \rangle_{\mathbf{u}} = \frac{\langle \mathbf{x}_1, \mathbf{x}_2 \rangle}{(r^2 - \|\mathbf{u}\|^2)^2}$ for all $(\mathbf{u}, \mathbf{x}_i) \in T_{\mathbf{u}}D$ for $i = 1, 2$, where $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ denote the standard inner product and Euclidean norm on $\mathbb{R}^2$.

If $\sigma : [0, 1[\rightarrow D$ is the straight line emanating from the origin given by $\sigma(t) = t(\cos \theta, \sin \theta)$ for $0 \leq t < r$ and θ fixed, then the length of σ is

$$\begin{aligned} L &= \int_0^r \|(\sigma(t), \sigma'(t))\|_{\mathbf{u}} dt = \int_0^r \|((t(\cos(\theta), \sin(\theta)), (\cos(\theta), \sin(\theta)))\|_{\mathbf{u}} dt \\ &= \int_0^r \frac{1}{r^2 - t^2} dt = \infty. \end{aligned}$$

To determine which lines are geodesics, we find the group $\mathcal{E}_D$ of orientation preserving isometry on D . The easiest way to find $\mathcal{E}_D$ is to tread D as a disk in the complex plane centred at the origin. In a first course on complex variables (see [3]), students learned about linear fractional transformations or Möbius transformations. These transformations are the composition of rotations, translations, expansions or contractions, and inversions with respect to the unit circle. A special linear fractional transformation is of the form

$$f_{w,\theta}(z) = e^{i\theta} \frac{r^2(z - w)}{r^2 - z\bar{w}}$$

for $w \in \mathbb{C}$ and $\theta \in \mathbb{R}$. Since they are the composition of transformations that preserve angles, the linear fractional transformations are isomorphism that preserve angles.

We have that $f_{w,\theta}$ maps D onto D for $w \in D$. Note that

$$|f_{w,\theta}(z)| = \left| \frac{r^2(z - w)}{r^2 - z\bar{w}} \right| = \left| \frac{z(r^2 - \bar{z}w)}{r^2 - z\bar{w}} \right| = |z| \left| \frac{r^2 - \bar{z}w}{r^2 - z\bar{w}} \right| = |z| = r \quad (9.9.1)$$

for $|z| = r$. Thus $f_{w,\theta} : \partial D \rightarrow \partial D$. Moreover, this map is onto ∂D because $f_{w,\theta}^{-1} = f_{-w,\theta}$ exists. Therefore $f_{w,\theta}(\partial D)$ splits $\mathbb{R}^2$ into two components: D and $\mathbb{R}^2 \setminus \bar{D}$. Since $f_{w,\theta}(0) = -e^{i\theta}w \in D$, we get that $f_{w,\theta}(D) \subset D$ because $f_{w,\theta}$ is a continuous function and therefore maps components to components. Again the fact that $f_{w,\theta}^{-1}$ exists implies that $f_{w,\theta}(D) = D$.

We claim that $f_{w,\theta}$ is an isometry according to the Poincaré metric. To verify that this is true, we use the isomorphism between $\mathbb{C}$ and $\mathbb{R}^2$ and first note that

$$f'_{w,\theta}(z) = e^{-i\theta} \frac{r^2(r^2 - |w|^2)}{(r^2 - z\bar{w})^2}$$

and

$$\begin{aligned} r^2 - |f_{w,\theta}(z)|^2 &= r^2 - \frac{r^4(z - w)(\bar{z} - \bar{w})}{(r^2 - z\bar{w})(r^2 - \bar{z}w)} = \frac{r^2(r^2 - z\bar{w})(r^2 - \bar{z}w) - r^4(z - w)(\bar{z} - \bar{w})}{|r^2 - z\bar{w}|^2} \\ &= \frac{r^4(r^2 - |z|^2) - r^2|w|^2(r^2 - |z|^2)}{|r^2 - z\bar{w}|^2} = \frac{r^2(r^2 - |z|^2)(r^2 - |w|^2)}{|r^2 - z\bar{w}|^2}. \end{aligned}$$

Hence

$$\frac{|f'_{w,\theta}(z)|^2}{r^2 - |f_{w,\theta}(z)|^2} = \frac{1}{r^2 - |z|^2}.$$

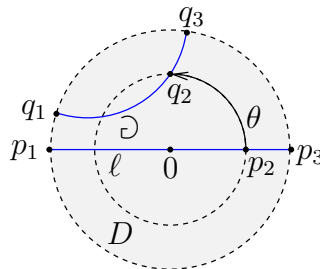
We then get

$$\begin{aligned} \langle (f_{w,\theta})_*(u, z_1), (f_{w,\theta})_*(u, z_2) \rangle_{f_{w,\theta}(u)} &= \langle (f_{w,\theta}(u), f'_{w,\theta}(u)z_1), (f_{w,\theta}(u), f'_{w,\theta}(u)z_2) \rangle_{f_{w,\theta}(u)} \\ &= \frac{|f'_{w,\theta}(u)|^2 \bar{z}_1 z_2}{(r^2 - |f_{w,\theta}(u)|^2)^2} = \frac{\bar{z}_1 z_2}{(r^2 - |u|^2)^2} = \langle (u, z_1), (u, z_2) \rangle_u \end{aligned}$$

for all $(u, z_i) \in T_u D$ and $i = 1, 2$.

Let $\mathcal{E}_D^\circ = \{f_{w,\theta} : w \in D \text{ and } \theta \in [0, 2\pi[\}$. We use Proposition 9.9.8 to prove that $\mathcal{E}_D = \mathcal{E}_D^\circ$. The group $\mathcal{E}_0^\circ$ is transitive because $f_{w,0}$ maps $w \in D$ to 0. Thus $f_{-w_2,0} \circ f_{w_1,0} = f_{w_2,0}^{-1} \circ f_{w_1,0}$ maps $w_1 \in D$ to $w_2 \in D$. Moreover $\mathcal{E}_D^\circ \cap \mathcal{E}_D(0) = \{f_{0,\theta} : \theta \in [0, 2\pi[\}$. Hence $\Phi : \mathcal{E}_D^\circ \cap \mathcal{E}_D(0) \rightarrow S^1$ is onto. Therefore, we may effectively use Proposition 9.9.8 to conclude that $\mathcal{E}_D = \mathcal{E}_D^\circ$.

We can now confirm that the “straight lines” defined at the beginning of this example are effectively the geodesics. To be precise, the orbits of $\mathcal{E}_D(0)$ are the circles centred at the origin. Therefore, according to Theorem 9.9.7, the geodesic through the origin are the tradition straight lines (suitably parameterized) through the origin. We now use the fact that isometries transform geodesics into geodesics to prove that all the arcs of circles that intersect perpendicularly ∂D are also geodesics. To be precise, we prove that all the arcs of circles that intersect perpendicularly ∂D are the images of diameters of D . We recall from basic complex analysis that linear fractional transformations preserve angles and maps circles to circles (assuming that straight lines are circles of infinite radii). Suppose that ℓ is a diameter of D . Then $\mathcal{D} = f_{w,\theta}(\ell)$ is an arc of a circle in D that may have an infinite radius. Since $f_{w,\theta} : \partial D \rightarrow \partial D$, the endpoints of $\mathcal{D}$ are on ∂D . Since ℓ intersects ∂D perpendicularly, we must also have that $\mathcal{D}$ intersect ∂D perpendicularly. Since $f_{w,\theta}(z) = 0$ implies that $z = w$, we have that $\mathcal{D}$ does not contain the origin if w is not on ℓ . Therefore, $\mathcal{D}$ cannot be a diameter and must be the arc of a circle that intersect ∂D perpendicularly. Since every arc of circle $\mathcal{D}$ that intersects perpendicularly ∂D is determined by three points, the two points of intersection with ∂D and any other point on the arc, it is a property of the special linear fractional transformations that it is possible to determine w and θ such that $f_{w,\theta}$ maps a diagonal ℓ to $\mathcal{D}$. More precisely, suppose q_1 and q_3 are the two endpoints of $\mathcal{D}$ with q_2 another point on $\mathcal{D}$. We have to determine θ , w , the endpoints p_1 and $p_3 = -p_1$ of the diagonal ℓ , and another point p_2 on ℓ such that $f_{w,\theta}(p_i) = q_i$ for $1 \leq i \leq 3$. As illustrated in the following figure, $p_2 \in \ell$ and θ should satisfy $e^{i\theta} p_2 = q_2$ to get $f_{w,\theta}(p_2) = q_2$ because $|f_{w,\theta}(p_2)| = |p_2| = |q_2|$ according to (9.9.1).



We get from $f_{w,\theta}(p_1) = q_1$ and $f_{w,\theta}(p_3) = q_3$ that

$$e^{i\theta} \frac{r^2(p_1 - w)}{r^2 - p_1 \bar{w}} = q_1 \quad \text{and} \quad e^{i\theta} \frac{r^2(-p_1 - w)}{r^2 + p_1 \bar{w}} = q_3 .$$

Hence

$$r^2 p_1 + e^{-i\theta} q_1 p_1 \bar{w} - r^2 w = r^2 e^{-i\theta} q_1 \quad \text{and} \quad r^2 p_1 + e^{-i\theta} q_3 p_1 \bar{w} + r^2 w = -r^2 e^{-i\theta} q_3 .$$

Since q_1 and q_3 are on ∂D , we have that $q_3 = e^{i\nu} q_1$ for some $\nu \in \mathbb{R}$. Thus, the previous equations become

$$r^2 p_1 + e^{-i\theta} q_1 p_1 \bar{w} - r^2 w = r^2 e^{-i\theta} q_1 \tag{9.9.2}$$

and

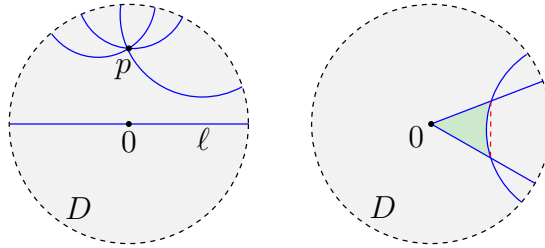
$$r^2 p_1 + e^{-i(\theta-\nu)} q_1 p_1 \bar{w} + r^2 w = -r^2 e^{-i(\theta-\nu)} q_1 . \tag{9.9.3}$$

If we subtract $e^{i\nu}$ times (9.9.2) from (9.9.3), then we get

$$r^2(1 - e^{i\nu})p_1 + r^2(1 + e^{i\nu})w = -2r^2 e^{i(\theta-\nu)} q_1 .$$

We can always solve for p_1 or w because at least one of $1 - e^{i\nu}$ or $1 + e^{i\nu}$ is non-null. We can substitute p_1 or w into (9.9.2) to get a system of linear equations (the real and imaginary parts) for w or p_1 respectively. Once we have w and p_1 in terms of θ , we can use the equation $q_2 = e^{i\theta} p_2 = \pm e^{i\theta} |q_2| p_1$ to determine θ and thus p_2 .

We cannot complete this example without illustrating two of the major differences between the hyperbolic geometry and the other geometries. The figure on the left below shows a line ℓ and a point p not on ℓ with infinitely many tangent lines through p parallel to ℓ . The figure on the right show a triangle where the sum of the internal angles is less than π .



The three examples of geometries that we have given are the three possible models of geometry on simply connected, complete and oriented 2-dimensional Riemannian manifolds with constant curvature because all such manifolds are isomorphic to one of the three models represented by our three examples (see [15]).

Bibliography

- [1] Ralph Abraham, Jerrold E. Marsden and Tudor Ratiu, **Manifolds, Tensor Analysis, and Applications**, Addison-Wesley Publishing Company, 1983.
- [2] Bourbaki, N. **Variétés différentielles et analytiques**, Springer-Verlag, 2007.
- [3] Ruel V. Churchill, James W. Brown and Roger F. Verhey, **Complex Variables and Application**, McGraw-Hill, Inc., 1974.
- [4] H. S. M. Coxeter, **Introduction to Geometry, 2nd Edition**, John Wiley & Sons, Inc., 1969.
- [5] James Dugundji, **Topology**, Allyn and Bacon, Inc., 1978.
- [6] Wendell Fleming, **Functions of Several Variables**, UTM, Springer-Verlag, 1977.
- [7] Gerald B. Folland, **Advanced Calculus**, Prentice Hall, 2002.
- [8] Marvin J. Greenberg, **Euclidean and non-Euclidean Geometries: Development and History**, W. H. Freeman and Company, 1980.
- [9] Marvin J. Greenberg and John R. Harper, **Algebraic Topology: A First Course (revised)**, Addison-Wesley Publishing Company, 1981.
- [10] Victor Guillemin and Alan Pollack, **Differential geometry**, Prentice Hall, 1974.
- [11] I. N. Herstein, **Topics in Algebra**, John Wiley & Sons, Inc. 1975.
- [12] Thomas W. Hungerford, **Algebra**, Holt, Rinehart and Winston, Inc., 1974.
- [13] Serge Lang, **Linear Algebra, 2nd Edition**, Addison-Wesley Publishing Company, 1971.
- [14] Serge Lang, **Differential Manifolds**, Addison-Wesley Publishing Company, 1972.
- [15] John M. Lee, **Riemannian Manifolds: An Introduction to Curvature**, Springer-Verlag, 1997.
- [16] W. A. J. Luxemburg, **ArselÀ's Dominated Convergence Theoem**, American Mathematics Monthly, Vol. 78, 970-979, New York, 1971.

- [17] J. E. Marsden and M. J. Hoffman, **Elementary Classical Analysis, 2nd Edition**, W. H. Freeman and Company, 1993.
- [18] J. E. Marsden and A. J. Tromba, **Vector Calculus, 2nd Edition**, W. H. Freeman and Company, 1981.
- [19] J. Milnor, **Topology from a Differential Viewpoint**, University of Virginia Press, 1965.
- [20] J. R. Munkres, **Topology, A First Course**, Printice-Hall, Inc., 1975.
- [21] J. R. Munkres, **Elements of Algebraic Topology**, Addison-Wesley Publishing Company, 1984.
- [22] Tibor Radö. **Length and Area**, American Mathematical Society, Volume XXX, New York, 1948.
- [23] Walter Rudin, **Principle of Mathematical Analysis, 3rd Edition**, McGraw-Hill, 1978.
- [24] Walter Rudin, **Real and Complex Analysis, 2nd Edition**, McGraw-Hill, 1974.
- [25] I. M. Singer and J. A. Thorpe, **Lecture Notes on Elementary Topology and Geometry**, UTM, Springer-Verlag, 1967.
- [26] Micheal Spivak, **Calculus on Manifolds**, Benjamin/Cummings Publishing Comp. 1965.
- [27] Micheal Spivak, **A Comprehensive Introduction to Differential Geometry, Volume One**, Publish or Perish, 2005.

Index

- (j, i) Face, [179](#)
- 1^{st} Homotopy Group, see Fundamental Group
- j^{th} Betti Number, [411](#)
- σ -compact, [365](#)
- j -face, [429](#)
- k -Boundary, [430](#), [437](#)
- k -Chain, [408](#)
- k -Coboundary, [488](#)
- k -Cocycle, [488](#)
- k -Cycle, [430](#), [437](#)
- m -Chain, [179](#)
- r -Skeleton, [400](#)
- (de Rham) Cohomology Module, see Cohomology Module
- (de Rham) Cohomology Module with Compact Support, see Cohomology Module with Compact Support
- Acyclic, [442](#)
- Admissible Open Cover, [82](#)
- Affine Function, [428](#)
- Affine Space, [428](#)
- Annihilator, [490](#)
- Area of a Surface, [269](#)
- Aspherical Space, [442](#)
- Atlas, [190](#)
 - Admissible Local Chart, [191](#)
 - Differential Structure, [191](#)
 - Equivalence, [191](#)
 - Maximal, [191](#)
- Augmentation, [440](#)
- Augmentation-Preserving, [441](#)
- Barycentre, [400](#)
- Barycentric Coordinates, [398](#)
 - j^{th} Barycentric Coordinates, [419](#)
- Barycentric Subdivision
 - m^{th} Barycentric Subdivision, [401](#)
 - First Barycentric Subdivision, [401](#)
- Barycentric Subdivision Operator, [466](#), [468](#)
- Boundary, [179](#), [180](#), [191](#), [408](#), [410](#), [429](#)
- Boundary Map, [409](#)
- Boundary of a Set, [3](#), [187](#)
- Boundary Operator, [429](#)
- Boundary Point, [187](#), [191](#)
- Canonical Basis, [163](#)
- Cartan Structural Formulae, [543](#)
 - First Structural Equations, [545](#)
 - Second Structural Equation, [546](#)
- Centroid, [126](#)
- Chain Complex, [436](#)
- Chain Homotopy, [439](#)
- Chain Map, [436](#)
- Circle Bundle, [520](#)
- Closed k -Simplex, [398](#)
- Closed Curve, [71](#)
- Closed Face, [398](#)
- Closure of a Set, [2](#)
- Coboundary, [415](#), [486](#)
- Coboundary Operator, [414](#), [485](#)
- Cochain Complex, [487](#)
- Cochain Map, [488](#)
- Cochaines, [414](#), [485](#)
- Cocycle, [414](#), [486](#)
- Cohomologous Forms, [352](#), [360](#)
- Cohomology Class, [352](#), [360](#)
- Cohomology Group, see Cohomology Module
- Cohomology Module, [360](#)
 - k^{th} Cohomology Module, [414](#), [486](#)
 - Relative k^{th} Cohomology Module, [417](#), [489](#)
 - Singular k^{th} Cohomology Module, [488](#)
- Cohomology Module with Compact Support, [352](#)
- Comparison Test, [52](#)
- Cone, [400](#)

- Connection, 528
- Connection 1-form, 530
- Connection Homomorphism, 460
- Consistent Orientation, 202
- Continuously Differentiable, 188
- Contractible, 385
- Contractible to a Point, 178, 343
- Convex Independent, 397
- Cotangent Bundle, 212
- Covering, 389
- Covering Transformation, 394
- Critical Point, 103
- Cross Product, 162
- Cup Product, 417, 492
- Curvature, 546
- Curve, 71, 178, 257
 - Arc Length Function, 265
 - Closed, 259
 - Curvature, 282
 - Curvature Vector, 281, 282
 - Direction, 257
 - Length, 263
 - Natural Parametric Representation, 282
 - Orientation, 257
 - Parameterized by Arc Length, 265
 - Parametric Representation, 257
 - Positive Direction, 271
 - Positive Orientation, 271
 - Radius of Curvature, 282
 - Rectifiable, 263
 - Tangent Vector, 282
 - Torsion, 284
- Cycle, 410
- Deformation Matrix, 288
- Deformation Retract, 474
- Deformation Retraction, 475
- Degree of a Function, 378
- Diameter, 2
- Diffeomorphism, 189, 201
- Differential Form
 - $\mathbf{p}$ -form, 212
 - k -Form, 167
 - Closed, 173
 - Exact, 173
- Differential Operator, 3
- Dimension, 398
- Direct Sum, 344
- Directional Derivative, 216
- Distance, 2
- Equivalent, 407
- Equivalent Curves, 207
- Euler Characteristic, 411
- Exact Sequence, 433
- Exact Triad, 479
- Excised, 464
- Excision, 464
- Extension of a Function, 188
- Finite Measure, 229
- Five Lemma, 478
- Fixed Point, 396
- Flow, 286, 564
- Functional Associated to a Differential Form, 228
- Fundamental Group, 387, 389
- Fundamental Sequence, 233
- Fundamental Theorem of Calculus, 183
- General Position, 400
- Geodesic, 548
- Geodesic Coordinate System, 568
- Geodesic Curvature, 548
- Green's Formulas, 326, 331
- Group of Orientation Preserving Isometries, 584
- Hausdorff Space, 364
- Homologous, 430, 457
- Homology Module
 - Reduced (Simplicial) k^{th} Homology Module, 413
 - Reduced (Singular) k^{th} Homology Module, 449
 - Relative (Simplicial) k^{th} Homology Module, 413
 - Relative (Singular) k^{th} Homology Module, 457
 - Simplicial k^{th} Homology Module, 410
 - Singular k^{th} Homology Module, 430, 437

- Homology Sequence, 461
- Homotopic, 385
- Homotopic Equivalence, 448
- Homotopic Equivalent, 448
- Homotopic Functions, 362, 384, 460
- Homotopy, 362, 384
- Horizontal Lift, 537

- Induced Orientation, 205
- Induced Topology, 187
- Integrable, 84
- Integral Curve, 211, 563
- Integral of a Differential Form, 182, 217, 219
- Integral of a Function, 13, 17, 86
- Integral Test, 57
- Interior, 191
- Interior of a Set, 2, 187
- Interior Point, 187, 191
- Irrotational, 289
- Isometry, 516
- Isomorphic Coverings, 394
- Isotropy Group, 585
- Iterated Integral, 75, 76

- Jacobi Identity, 560
- Jacobi's Equation, 577
- Jacobian, 88
- Join, 464
- Jordan Measurable, 17, 61

- Kronecker Delta Function, 98

- Laplacien, 326
- Length, 244
- Lie Bracket, 560, 564
- Lie Derivative, 216
- Lift, 527
- Lindelöf Space, 364
- Line Integral of a Function, 257
- Local Chart, 189, 241
- Locally Compact Space, 365
- Locally Finite, 365
- Locally Simply Connected, 393
- Long Exact Sequence, see Exact Sequence
- Loop, 386
- Lower Riemann Integral, 13, 64
- Lower Riemann Sum, 11, 63

- Manifold
 - k -Dimensional Manifold, 191
 - Manifold With Boundary, 192
 - Manifold Without Boundary, 192
 - Oriented Manifold, 203
- Manifolds with Piecewise Smooth Boundary, 242
- Matrix of Rotation, 288
- Mayer-Vietoris Sequence, 481
- Mean Value Theorem for Integrals, 110
- Measure Associated to a Differential Form, 228, 229
- Measure Zero, 16, 61
- Mesh, 402
- Module of k -Chains, 408
- Module of Relative Chains, 413
- Multilinear Mapping, 149

- Naturality, 446
- Negligible Set, 233
- Norm
 - 2-Norm, 2
 - k -Norm, 2
 - Euclidean Norm, 2
 - Supremum Norm of a Function, 3

- Open k -Simplex, 398
- Open Ball, 2
- Open Face, 398
- Orbit, 586
- Ordered k -Chain, 501
- Ordered k -Simplex, 501
- Orientable, 203
- Orientation, 161, 202, 203
- Orientation Preserving, 203, 204
- Oriented k -simplex, 407
- Outward Unit Normal, 205, 206

- Parallel Translation, 527
- Partial order, 400
- Partition, 11, 63
- Partition of Unity, 82, 218
- Path, 257, 385
- Peano's Curve, 71

- Permutation, 152
- Piecewise C^m Continuous Differentiable Function, 3
- Piecewise Continuous Function, 3
- Piecewise Smooth, 241
- Poincaré Metric, 590
- Poincaré Model, 589
- Polar Coordinate System, see Geodesic Coordinate System
- Product of Paths, 385
- Proper Function, 378
- Pull-Back, 154, 165, 168, 211, 213
- Radial Geodesic, 568
- Rectangle, 61
- Reduced Chain Complex, 440
- Refinement, 12, 64
- Regular Covering, 394
- Regular Point, 233, 241
- Regular value, 103
- Relative k -Boundary, 457
- Relative k -Cycle, 457
- Relative Coboundary, 490
- Relative Cocycle, 490
- Retract, 396
- Retraction, 396
- Riemann Integrable, 13, 64, 69
- Riemann Integral of a Function, 64, 69
- Riemann Sum, 18
- Riemannian Connection, 545
- Riemannian Manifold, 244, 515
- Riemannian Metric, 515
- Short Exact Sequence, 434
 - Split, 434
- Simplicial Approximation, 404
- Simplicial Complex, 398
- Simplicial Map, 403
- Simply Connected, 393
- Singular k -Chain, 427
- Singular k -Simplex, 427
- Singular m -Cube, 178
- Singular Point, 233, 241
- Small of Order, 471
- Smoothly Contractible to a Point, see Contractible to a Point
- Smoothly Triangulated Manifold, 418
- Standard k -simplex, 427
- Standard m -Cube, 178
- Standard Orientation, 161
- Star, 404, 420
- Star-Shaped, 176
- Subcomplex, 399
- Subdivision, 400
- Submanifold, 195
- Subrectangles of the Partition, 63
- Support of a Function, 3
- Surface, 244, 265
 - Binormal Vector, 283
 - Parametric Representation, 265
 - Positive Orientation, 269
 - Principal Normal, 283
- Surface Integral of a Function, 266
- Tangent Bundle, 201
- Tangent Bundle Projection, 201
- Tangent Space, 163, 199, 207, 208
- Tangent Unit Vector, 245
- Tangent Vector, 208
- Tensor
 - k -Tensor, 149
 - Alternating, 152
 - Inner Product, 151, 243
 - Orthonormal Basis, 151
 - Positive Defined, 151
 - Product, 149, 164
 - Symmetric, 151
 - Wedge Product, 154
- Transitive, 585
- Transposition, 152
- Universal Covering, 393
- Upper Riemann Integral, 13, 64
- Upper Riemann Sum, 11, 63
- Vanish at Infinity, 226
- Vector Bundle, 202
- Vector Field, 166, 211, 258
 - Conservative, 261
 - Curl, 255
 - Divergence, 255
 - Line Integral, 258

Surface Integral, [267](#)
Vertex, [398](#)
Vertices, [399](#)
Volume, [61](#), [244](#)
Volume Element, [162](#), [244](#), [541](#)

Zero Content, [16](#), [61](#)
Zig-Zag Lemma, [437](#), [488](#)