Monograph Series

of the Parana's Mathematical Society

Advanced Calculus in $\mathbb{R}^N$

Marcelo M. Cavalcanti and Valéria N. Domingos Cavalcanti

State University of Maringá (UEM)

Monograph Series of the Parana's Mathematical Society
©SPM – E-ISSN-2175-1188 • ISSN-2446-7146
SPM: www.spm.uem.br/bspm

Monograph 04 (2021).
doi:10.5269/bspm.81138

Advanced Calculus in $\mathbb{R}^N$

Marcelo M. Cavalcanti and Valéria N. Domingos Cavalcanti

Marcelo Moreira Cavalcanti and Valéria N. Domingos Cavalcanti),
Department of Mathematics,
State University of Maringá (UEM),
Brazil.
E-mail address: mmcavalcanti@uem.br, vndcavalcanti@uem.br

2020 *Mathematics Subject Classification:* Primary: 00-xx;
Secondary: 00-01, 00-02.

Advanced Calculus in $\mathbb{R}^n$

Marcelo M. Cavalcanti and Valéria N. Domingos Cavalcanti

Department of Mathematics
State University of Maringá (UEM)

Maringá
2025

Copyright Notice

This book originated from lecture notes used in advanced undergraduate and beginning graduate courses in Analysis and Multivariable Calculus at the Department of Mathematics of the State University of Maringá (UEM), Brazil.

Comments and suggestions from readers and instructors are very welcome.

Version: November 26, 2025

“Let us keep in mind that we are not what others think, and often not even what we think we are, but we truly are what we feel. Our feelings reveal our performance in the past, our actions in the present, and our potential for the future.” (Hamed)

Preface

Mathematics plays a central role in modern science. Beyond its intrinsic value as a discipline with its own concepts, theories, and open problems, mathematics permeates essentially all areas of human knowledge. Many mathematical theories have their roots in natural phenomena, and in turn have driven the remarkable development of physics, engineering, computer science, and many other fields.

From this perspective, mathematics should not be viewed as an isolated subject, but rather as a unifying language and framework that connects and clarifies different scientific disciplines.

This book aims to present key topics in mathematical analysis in $\mathbb{R}^n$, together with illustrative examples and figures. Our goal is to help students from a variety of programs—mathematics, physics, engineering, and related areas— develop a solid and intuitive understanding of the basic tools of real and vector analysis, while maintaining full rigor.

The material grew out of lecture notes for courses taught over several years at the State University of Maringá. We have tried to balance conceptual clarity, motivating examples, and detailed proofs, so that the text can be used both for self-study and as a companion to a classroom course.

We hope that this book will contribute to the reader's mathematical maturity and stimulate further study in analysis, differential equations, geometry, and beyond.

Maringá, 2025

Marcelo M. Cavalcanti
Valéria N. Domingos Cavalcanti

Contents

Introduction

Under the title *Advanced Calculus in* $\mathbb{R}^n$ we gather a selection of topics in real and vector analysis in several variables, complemented by figures and examples designed to support intuition and to fix the main concepts, theorems, and proofs.

Real analysis is the branch of mathematical analysis that deals with the real numbers and real-valued functions. It provides a rigorous foundation for the central notions of calculus, such as limits, continuity, differentiation, integration, and infinite series and sequences of functions.

Vector calculus extends these ideas to functions defined on $\mathbb{R}^n$ with values in $\mathbb{R}^m$. It is a fundamental tool in physics and engineering, appearing naturally in the study of fields, fluxes, and conservation laws. Topics such as gradients, divergence, curl, line integrals, surface integrals, and the classical integral theorems (Green, Gauss, Stokes) play a central role.

Differential forms provide a unified language for these ideas. They are geometric objects that can be integrated over curves, surfaces, and higher-dimensional manifolds, and they generalize many familiar notions from vector calculus. The formalism of k-forms and the exterior derivative offers a powerful and elegant framework that clarifies and extends the classical formulas.

In this book, we aim to present the main concepts involved in these areas in a coherent way, emphasizing both computational techniques and structural insights. The text is intended for students who wish to deepen their understanding of calculus and analysis, and to acquire a solid base for further studies in differential geometry, partial differential equations, and related subjects.

Chapter 1

Topology in $\mathbb{R}^n$

1.1 The Vector Space $\mathbb{R}^n$

Let $n \in \mathbb{N}$, where $\mathbb{N}$ denotes the set of natural numbers. The n-dimensional Euclidean space is the Cartesian product of n copies of $\mathbb{R}$, that is:

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ factors}}.$$

The points of $\mathbb{R}^n$ are therefore all n-tuples $x = (x_1, x_2, \ldots, x_n)$ whose coordinates $x_1, \ldots, x_n$ are real numbers. Given $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$, we have $x = y$ if and only if $x_i = y_i$ for every $i = 1, 2, \ldots, n$.

- $\mathbb{R}^1 = \mathbb{R}$ is the real line, that is, the set of real numbers;
- $\mathbb{R}^2$ is the plane, that is, the set of ordered pairs (x, y) of real numbers;
- $\mathbb{R}^3$ is three-dimensional Euclidean space, whose points are the triples (x, y, z).

Given $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$ and a scalar $\alpha \in \mathbb{R}$, we define addition and scalar multiplication by:

$$\begin{aligned} x + y &= (x_1 + y_1, \ldots, x_n + y_n), \\ \alpha x &= (\alpha x_1, \ldots, \alpha x_n). \end{aligned}$$

These operations make $\mathbb{R}^n$ into a vector space of dimension n over the field of real numbers. The elements of $\mathbb{R}^n$ are sometimes called *points* and sometimes *vectors*. From a geometric viewpoint, considering $x \in \mathbb{R}^n$ as a vector means imagining the arrow whose origin is at O and endpoint at x.

1.2 Inner Product and Norm

An inner product on a real vector space $\mathbb{V}$ is a symmetric, positive bilinear form. That is, it is a map that assigns to each pair of vectors x, y a real number $\langle x, y\rangle$ such that for all $x, y, z \in \mathbb{V}$ and all $\alpha \in \mathbb{R}$:

$$\begin{array}{ll} \text{(PI1)} & \langle x, y\rangle = \langle y, x\rangle, \\ \text{(PI2)} & \langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle, \\ \text{(PI3)} & \langle \alpha x, y\rangle = \alpha \langle x, y\rangle, \\ \text{(PI4)} & \langle x, x\rangle \geq 0 \text{ and } \langle x, x\rangle = 0 \text{ iff } x = 0. \end{array}$$

The most important example is the canonical inner product in $\mathbb{R}^n$:

$$\langle x, y\rangle = x_1y_1 + x_2y_2 + \cdots + x_ny_n,$$

where $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$.

Given $x \in \mathbb{R}^n$, we write

$$\|x\| = \sqrt{\langle x, x\rangle},$$

called the Euclidean norm of x. Indeed:

$$\|x\| = \sqrt{x_1^2 + \cdots + x_n^2},$$

and this number represents the length of the vector x.

Note that $\|x\|^2 = \langle x, x\rangle$, so that:

$$\begin{array}{ll} \text{(i)} & \|x\| = 0 \iff x = 0, \\ \text{(ii)} & \|x\| > 0 \iff x \neq 0. \end{array}$$

Two vectors $x, y \in \mathbb{R}^n$ are said to be *orthogonal* when $\langle x, y\rangle = 0$. Clearly the zero vector is orthogonal to every vector.

Given $x, y \in \mathbb{R}^n$ with $y \neq 0$ and setting $\alpha = \frac{\langle x,y\rangle}{\|y\|^2}$, the vector $z = x - \alpha y$ is orthogonal to y. Indeed:

$$\begin{aligned} \langle z, y\rangle &= \langle x - \alpha y,\, y\rangle \\ &= \langle x, y\rangle - \alpha\langle y, y\rangle \\ &= \langle x, y\rangle - \frac{\langle x, y\rangle}{\|y\|^2}\,\|y\|^2 = 0. \end{aligned}$$

Geometrically (see Figure 1.1):

$$\begin{array}{lrcl} \text{(i)} & \|\alpha y\| &=& \dfrac{|\langle x, y\rangle|}{\|y\|^2}\,\|y\| = \dfrac{|\langle x, y\rangle|}{\|y\|}, \\ \text{(ii)} & |\langle x, y\rangle| &=& \|x\|\,\|y\|\,|\cos\theta|. \end{array}$$

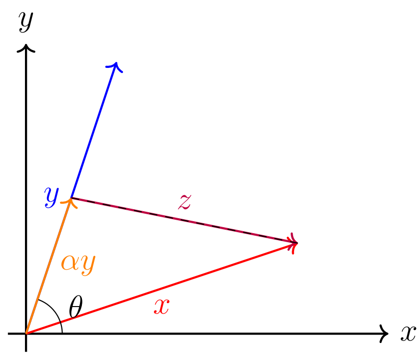

Figure 1.1: Exemplo de ortogonalidade

From (i) and (ii) we obtain

$$\|\alpha y\| = \|x\|\,|\cos\theta|.$$

Proposition 1.1 (Cauchy-Schwarz Inequality) *For all $x, y \in \mathbb{R}^n$,*

$$|\langle x, y\rangle| \le \|x\|\,\|y\|.$$

Proof: If $y = 0$ the inequality is trivial. If $y \neq 0$, with α as above, the vector $z = x - \alpha y$ is orthogonal to y. Then:

$$\begin{aligned} \|x\|^2 &= \langle x, x\rangle = \langle z + \alpha y,\, z + \alpha y\rangle \\ &= \|z\|^2 + \alpha^2\|y\|^2 \;\ge\; \alpha^2\|y\|^2 = \frac{\langle x, y\rangle^2}{\|y\|^4}\,\|y\|^2. \end{aligned}$$

Thus $\langle x, y\rangle^2 \le \|x\|^2\|y\|^2$, and the result follows. □

Remark 1. Equality holds in the Cauchy–Schwarz inequality precisely when one vector is a scalar multiple of the other. Indeed, if $y = 0$ the equality is trivial. If $y \neq 0$ and $x = \alpha y$ for some $\alpha \in \mathbb{R}$, then:

$$\begin{aligned} |\langle x, y\rangle| &= \|\alpha y\|\,\|y\| = |\alpha|\,\|y\|^2 \\ &= \frac{\|x\|}{\|y\|}\,\|y\|^2 = \|x\|\,\|y\|. \end{aligned}$$

Remark 2. The Euclidean norm $\|x\| = \sqrt{\langle x, x\rangle}$ satisfies the following properties for all $x, y \in \mathbb{R}^n$:

$$\begin{aligned} &\text{(N1)} \quad \|x\| \ge 0 \quad \text{and} \quad \|x\| = 0 \iff x = 0, \\ &\text{(N2)} \quad \|\alpha x\| = |\alpha|\,\|x\| \quad \text{for all } \alpha \in \mathbb{R}, \\ &\text{(N3)} \quad \|x + y\| \le \|x\| + \|y\| \quad \text{(triangle inequality).} \end{aligned}$$

The first two are immediate. The triangle inequality follows from the Cauchy–Schwarz inequality. Indeed:

$$\begin{aligned}\|x+y\|^2 &= \langle x+y,\, x+y\rangle \\ &= \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle \\ &\leq \|x\|^2 + \|y\|^2 + 2\|x\|\,\|y\| \\ &= (\|x\| + \|y\|)^2.\end{aligned}$$

Thus $\|x+y\| \leq \|x\| + \|y\|$, the so-called triangle inequality.

In general, a norm on a real vector space $\mathbb{V}$ is a function $\|\cdot\| : \mathbb{V} \to \mathbb{R}$ satisfying the axioms (N1), (N2), and (N3) above.

Although many norms exist on $\mathbb{R}^n$, the Euclidean norm is the most geometrically natural, since it corresponds to the usual formula for the length of a vector in Cartesian coordinates. Geometrically:

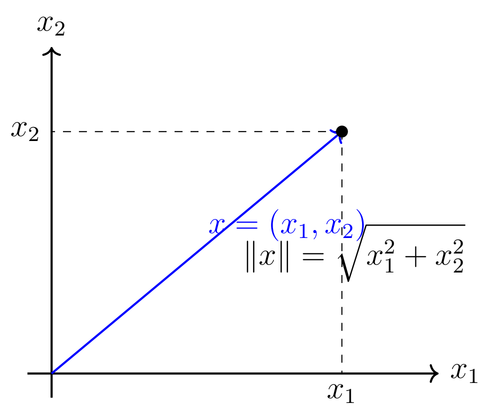

Figure 1.2: Length of a vector in the plane in Cartesian coordinates.

Unless stated otherwise, all norms in $\mathbb{R}^n$ will be assumed to be Euclidean. However, two other norms are often useful because of their simple algebraic form:

$$\begin{aligned}\|x\|_M &= \max\{|x_1|, \ldots, |x_n|\} \quad \text{(maximum norm)}, \\ \|x\|_S &= |x_1| + \cdots + |x_n| \quad \text{(sum norm)}.\end{aligned}$$

It is easy to verify that these satisfy axioms (N1)–(N3). One also shows that:

$$\|x\|_M \;\leq\; \|x\| \;\leq\; \|x\|_S \;\leq\; n\,\|x\|_M.$$

A norm on a real vector space $\mathbb{V}$ gives rise to a notion of distance:

$$\begin{aligned} d : \mathbb{V}\times\mathbb{V} &\longrightarrow \mathbb{R}, \\ (x, y) &\longmapsto d(x,y) = \|x - y\|.\end{aligned}$$

Using (N1)–(N3), one checks that the distance satisfies:

$$\begin{array}{ll} \text{(D1)} & d(x,y) \geq 0 \quad \text{and} \quad d(x,y) = 0 \iff x = y, \\ \text{(D2)} & d(x,y) = d(y,x), \\ \text{(D3)} & d(x,z) \leq d(x,y) + d(y,z) \quad \forall\, x,y,z \in \mathbb{V}. \end{array}$$

Indeed, the first two are immediate. For (D3):

$$\begin{aligned} d(x,z) &= \|x-z\| = \|(x-y)+(y-z)\| \\ &\leq \|x-y\| + \|y-z\| \\ &= d(x,y) + d(y,z). \end{aligned}$$

Remark 3. A norm $\|\cdot\|$ on a vector space $\mathbb{V}$ need not arise from an inner product. If a norm does come from an inner product, then the parallelogram identity holds:

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

Geometrically, this means that the sum of the squares of the diagonals of a parallelogram equals the sum of the squares of its four sides.

Indeed:

$$\begin{aligned} \|x+y\|^2 &= \langle x+y,\, x+y\rangle = \|x\|^2 + 2\langle x,y\rangle + \|y\|^2, \\ \|x-y\|^2 &= \langle x-y,\, x-y\rangle = \|x\|^2 - 2\langle x,y\rangle + \|y\|^2. \end{aligned}$$

Adding the two equalities yields the desired identity.

The parallelogram identity is **not** satisfied by every norm. For example, the maximum norm $\|x\|_M = \max\{|x_1|,\dots,|x_n|\}$ on $\mathbb{R}^n$ does not satisfy it. Using the canonical basis $e_1,\dots,e_n$ and setting $x = e_1$, $y = e_2$, we have:

$$\begin{aligned} \|x+y\|_M^2 + \|x-y\|_M^2 &= 2, \\ 2(\|x\|_M^2 + \|y\|_M^2) &= 4, \end{aligned}$$

which are different, showing that this norm does not come from any inner product.

A very useful inequality derived from axioms (N1)–(N3) is:

$$\big|\|x\| - \|y\|\big| \leq \|x-y\|.$$

To prove it, it suffices to show that:

$$-\|x-y\| \leq \|x\| - \|y\| \leq \|x-y\|.$$

Indeed:

$$\begin{aligned} \|y\| &\leq \|x\| + \|x-y\|, \\ \|x\| &\leq \|y\| + \|x-y\|, \end{aligned}$$

which follow directly from the triangle inequality by writing $y = x + (y-x)$ and $x = y + (x-y)$.

1.3 Balls and Bounded Sets

We now introduce some basic geometric notions in $\mathbb{R}^n$ that will be needed later.

Definition 1.2 *An open ball with center at a point $x_0 \in \mathbb{R}^n$ and radius $r > 0$, denoted by $B_r(x_0)$, is defined by*

$$B_r(x_0) = \{x \in \mathbb{R}^n \,;\, \|x - x_0\| < r\}.$$

Analogously, the closed ball $\overline{B_r(x_0)}$ and the sphere $S_r(x_0)$, both centered at x_0, are defined by

$$\begin{aligned} \overline{B_r(x_0)} &= \{x \in \mathbb{R}^n \,;\, \|x - x_0\| \le r\}, \\ S_r(x_0) &= \{x \in \mathbb{R}^n \,;\, \|x - x_0\| = r\}. \end{aligned}$$

When $n = 1$, the open ball $B_r(x_0)$ is just the open interval $(x_0 - r, x_0 + r)$; the closed ball $\overline{B_r(x_0)}$ is the closed interval $[x_0 - r, x_0 + r]$, and the sphere $S_r(x_0)$ reduces to the set consisting of the points $x_0 - r$ and $x_0 + r$. Note that in $\mathbb{R}$ the three usual norms coincide.

For $n = 2$, with the Euclidean norm, balls in the plane are called (open or closed) disks, and spheres reduce to circles (see Figure 1.3).

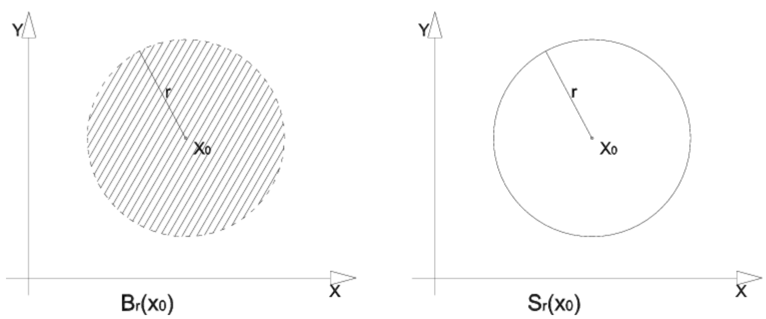

Figure 1.3: Euclidean norm in $\mathbb{R}^2$.

For $n = 3$, the Euclidean norm defines balls and spheres in space that agree with our usual geometric intuition.

Remark 1. The geometric shape of balls and spheres generally depends on the norm being used. If, instead of the Euclidean norm, we consider on $\mathbb{R}^2$ the maximum norm and the sum norm, then the ball of center $P = (x_0, y_0)$ and radius $r > 0$ is, in the first case, a square with sides parallel to the coordinate axes, each of length $2r$, and diagonals intersecting at P; in the second case, it is a square whose diagonals are parallel to the coordinate axes, both of length $2r$, and still intersecting at P.

Maximum norm.

$$\begin{aligned} B_{r,\max}(\cdot) &= \{(x,y) \in \mathbb{R}^2 \,;\, \|(x,y)\|_{\max} < r\} \\ &= \{(x,y) \in \mathbb{R}^2 \,;\, \max\{|x|, |y|\} < r\}. \end{aligned}$$

$$\begin{aligned} \text{(i) } |x| < r &\implies -r < x < r, \\ \text{(ii) } |y| < r &\implies -r < y < r. \end{aligned}$$

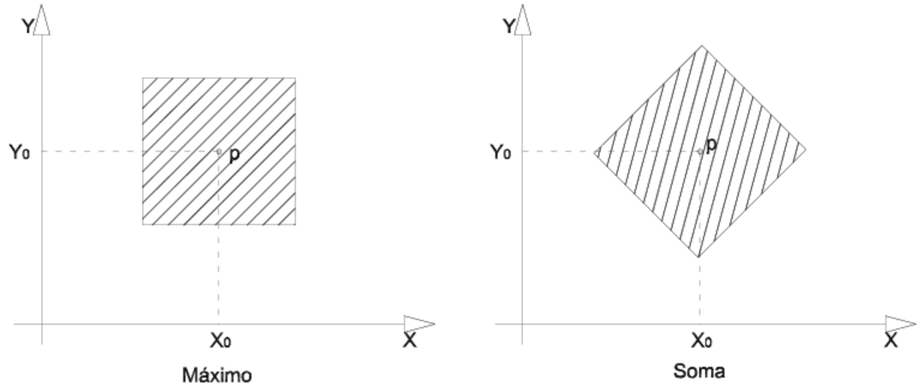

Figure 1.4: Maximum and sum norms in $\mathbb{R}^2$.

Sum norm.

$$\begin{aligned}
B_{r,\text{sum}}(\cdot) &= \{(x,y) \in \mathbb{R}^2\,;\ \|(x,y)\|_{\text{sum}} < r\} \\
&= \{(x,y) \in \mathbb{R}^2\,;\ |x| + |y| < r\}. \\
\text{(i) } x > 0,\ y > 0 &\Rightarrow x + y < r \Longrightarrow y < -x + r, \\
\text{(ii) } x > 0,\ y < 0 &\Rightarrow x - y < r \Longrightarrow y > x - r, \\
\text{(iii) } x < 0,\ y > 0 &\Rightarrow -x + y < r \Longrightarrow y < x + r, \\
\text{(iv) } x < 0,\ y < 0 &\Rightarrow -x - y < r \Longrightarrow y > -x - r.
\end{aligned}$$

Definition 1.3 *A subset $X \subset \mathbb{R}^n$ is said to be* bounded *if there exists a real number $c > 0$ such that $\|x\| \leq c$ for all $x \in X$.*

The definition above is equivalent to saying that X is contained in the closed ball of center at the origin and radius c. On the other hand, if there exists some ball $B_r(x_0)$ (with arbitrary center) containing X, then for all $x \in X$ we have $\|x - x_0\| \leq r$. Setting $c = r + \|x_0\|$, we obtain:

$$\begin{aligned}
x \in X \quad &\Longrightarrow \quad \|x\| = \|x - x_0 + x_0\| \\
&\leq \quad \|x - x_0\| + \|x_0\| \leq r + \|x_0\| = c.
\end{aligned}$$

Thus X is bounded. It follows that a set X is bounded if and only if it is contained in some ball (whose center is not necessarily the origin).

Remark 2. For the three usual norms on $\mathbb{R}^n$ we have the inequalities

$$\|x\|_M \leq \|x\| \leq \|x\|_S \leq n\,\|x\|_M,$$

which show that a set $X \subset \mathbb{R}^n$ is bounded with respect to one of these norms if and only if it is bounded with respect to each of the other two.

1.4 Sequences in $\mathbb{R}^n$

Definition 1.4 *A sequence in $\mathbb{R}^n$ is a map $x : \mathbb{N} \to \mathbb{R}^n$ defined on the set of natural numbers.*

The value of this map at $k \in \mathbb{N}$, denoted by $x(k)$ or x_k, is called the k-th term of the sequence. We use the notations (x_k), $(x_k)_{k\in\mathbb{N}}$ or $(x_1, x_2, \dots, x_k, \dots)$ to denote the sequence whose k-th term is x_k.

Definition 1.5 *A subsequence of a sequence $(x_k)_{k\in\mathbb{N}}$ is the restriction of this sequence to an infinite subset $A = \{k_1 < k_2 < \cdots < k_i < \cdots\} \subset \mathbb{N}$.*

We denote such a subsequence by $(x_k)_{k\in A}$ or $(x_{k_i})_{i\in\mathbb{N}}$.

Definition 1.6 *A sequence is said to be bounded if there exists a real number $c > 0$ such that $\|x_k\| \le c$ for all $k \in \mathbb{N}$.*

Remark 1. A sequence in $\mathbb{R}^n$ is equivalent to n sequences of real numbers. Indeed, for each $k \in \mathbb{N}$ we have $x_k = (x_{k1}, x_{k2}, \dots, x_{kn})$, where $x_{ki} = \pi_i(x_k)$ is the i-th coordinate of x_k $(i = 1, 2, \dots, n)$. The n sequences $(x_{ki})_{k\in\mathbb{N}}$ are called the coordinate sequences of (x_k). For example, in the plane $\mathbb{R}^2$, a sequence of points $z_k = (x_k, y_k)$ is equivalent to a pair of sequences (x_k) and (y_k) of real numbers.

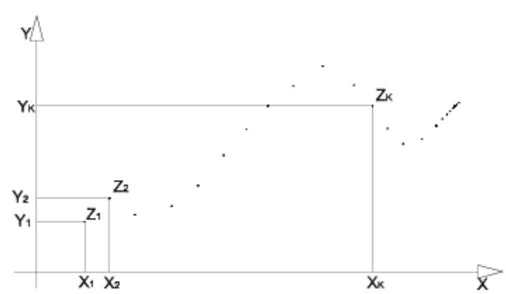

Figure 1.5: Sequence of points $z_k = (x_k, y_k)$ in $\mathbb{R}^2$.

It is easy to verify that a set $X \subset \mathbb{R}^n$ is bounded if and only if its projections $\pi_1(X), \pi_2(X), \dots, \pi_n(X)$ are bounded subsets of $\mathbb{R}$. It follows that a sequence $(x_k)_{k\in\mathbb{N}}$ is bounded if and only if each of its coordinate sequences (x_{ki}) is bounded in $\mathbb{R}$.

Definition 1.7 *A point $x_0 \in \mathbb{R}^n$ is said to be the limit of a sequence of points (x_k) in $\mathbb{R}^n$ if for every $\varepsilon > 0$ there exists $k_0 \in \mathbb{N}$ such that for all $k > k_0$ one has $\|x_k - x_0\| < \varepsilon$.*

In this case, we also say that $(x_k)_{k\in\mathbb{N}}$ converges to x_0 or that x_k tends to x_0, and we write

$$\lim_{k\to\infty} x_k = x_0, \quad \lim x_k = x_0, \quad \lim_{k\in\mathbb{N}} x_k = x_0,$$

or simply $x_k \to x_0$.

When there exists $x_0 = \lim_{k\to\infty} x_k$ we say that the sequence is *convergent*; otherwise, it is said to be *divergent*.

Note that

$$\lim_{k\to\infty} x_k = x_0 \quad \Longleftrightarrow \quad \lim_{k\to\infty} \|x_k - x_0\| = 0.$$

This reduces convergence in $\mathbb{R}^n$ to the convergence of nonnegative real numbers. Let us state some consequences of the definition of limit.

P1) In terms of balls, we have $\lim_{k\to\infty} x_k = x_0$ if and only if every open ball of center x_0 and radius $\varepsilon > 0$ contains all but finitely many terms of the sequence $(x_k)_{k\in\mathbb{N}}$ (see Figure 1.6).

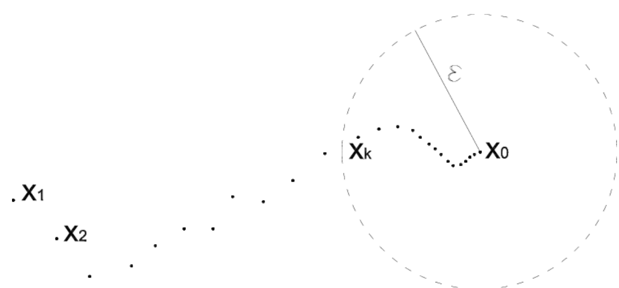

Figure 1.6: The ball $B_\varepsilon(x_0)$ containing all but finitely many terms of the sequence $(x_k)_{k\in\mathbb{N}}$.

Indeed, if $\varepsilon > 0$ is the radius of the ball and k_0 is the natural number corresponding to ε in the definition of limit, then outside the ball $B_\varepsilon(x_0)$ there can be at most the terms $x_1, \dots, x_{k_0}$ of the sequence.

P2) From the observation above, it follows that every convergent sequence is bounded. In fact, if $\lim_{k\to\infty} x_k = x_0$, then outside the open ball of center x_0 and radius 1 there can be at most the terms $x_1, \dots, x_{k_0}$. If r is the largest of the numbers $1, \|x_1 - x_0\|, \dots, \|x_{k_0} - x_0\|$, then all the terms of the sequence are contained in the ball $B_r(x_0)$.

P3) Again from the characterization via balls, if $\lim_{k\to\infty} x_k = x_0$, then every subsequence of $(x_k)_{k\in\mathbb{N}}$ has the same limit x_0.

P4) A crucial fact is the uniqueness of the limit of a sequence: if $\lim x_k = x_0$ and $\lim x_k = y_0$, then $x_0 = y_0$.

Indeed, for all $k \in \mathbb{N}$ we have

$$0 \le \|x_0 - y_0\| \le \|x_k - x_0\| + \|x_k - y_0\|.$$

Thus

$$\lim_{k\to\infty} \|x_k - x_0\| = \lim_{k\to\infty} \|x_k - y_0\| = 0 \quad \Rightarrow \quad x_0 = y_0.$$

Remark 2. The definition of the limit of a sequence in $\mathbb{R}^n$ uses a norm. From the inequalities relating the three usual norms in $\mathbb{R}^n$ we know that

$$\|x_k - x_0\|_M \leq \|x_k - x_0\| \leq \|x_k - x_0\|_S \leq n\,\|x_k - x_0\|_M.$$

It follows that

$$\lim \|x_k - x_0\|_M = 0 \iff \lim \|x_k - x_0\| = 0 \iff \lim \|x_k - x_0\|_S = 0.$$

Thus the statement $\lim_{k\to\infty} x_k = x_0$ does not depend on which of these norms we use.

Proposition 1.8 *A sequence $(x_k)_{k\in\mathbb{N}}$ in $\mathbb{R}^n$ converges to the point $x_0 = (x_{01}, \dots, x_{0n})$ if and only if, for each $i = 1, 2, \dots, n$, we have*

$$\lim_{k\to\infty} x_{ki} = x_{0i},$$

that is, each coordinate of x_k converges to the corresponding coordinate of x_0.

Proof: Assume first that $\lim_{k\to\infty} x_k = x_0$, and let $\varepsilon > 0$ be given. Then there exists $k_0 \in \mathbb{N}$ such that $\|x_k - x_0\| < \varepsilon$ for all $k \geq k_0$. For each $i = 1, \dots, n$ we have

$$|x_{ki} - x_{0i}| \leq \|x_k - x_0\|.$$

Hence $|x_{ki} - x_{0i}| < \varepsilon$ for all $k \geq k_0$, which implies $\lim_{k\to\infty} x_{ki} = x_{0i}$.

Conversely, assume that $\lim_{k\to\infty} x_{ki} = x_{0i}$ for each $i = 1, \dots, n$, and let $\varepsilon > 0$ be given. For each i there exists $k_{0i} \in \mathbb{N}$ such that $|x_{ki} - x_{0i}| < \varepsilon$ whenever $k \geq k_{0i}$. Set

$$k_0 = \max\{k_{01}, \dots, k_{0n}\}.$$

Then, if $k > k_0$,

$$\|x_k - x_0\|_M = \max\{|x_{k1} - x_{01}|, \dots, |x_{kn} - x_{0n}|\} < \varepsilon,$$

and therefore $\|x_k - x_0\| \leq \|x_k - x_0\|_S \leq n\,\|x_k - x_0\|_M < n\varepsilon$, so in particular $\|x_k - x_0\| \to 0$. Thus $\lim_{k\to\infty} x_k = x_0$. □

Similarly, given convergent sequences (x_k) and (y_k) in $\mathbb{R}^n$ and (α_k) in $\mathbb{R}$ with $\lim x_k = x_0$, $\lim y_k = y_0$ and $\lim \alpha_k = \alpha$, we have:

- $\lim\limits_{k\to\infty} (x_k + y_k) = x_0 + y_0$;
- $\lim\limits_{k\to\infty} (\alpha_k x_k) = \alpha x_0$;

- $\lim_{k\to\infty} \langle x_k, y_k\rangle = \langle x_0, y_0\rangle$;
- $\lim_{k\to\infty} \|x_k\| = \|x_0\|$.

Definition 1.9 *A sequence (x_k) in $\mathbb{R}^n$ is called a Cauchy sequence if, given $\varepsilon > 0$, there exists $k_0 \in \mathbb{N}$ such that $\|x_r - x_s\| < \varepsilon$ whenever $r, s \geq k_0$.*

Remark 3. Using the maximum norm on $\mathbb{R}^n$, we have

$$\|x_r - x_s\|_M = \max\{|x_{r1} - x_{s1}|, \ldots, |x_{rn} - x_{sn}|\}.$$

Thus (x_k) is a Cauchy sequence in $\mathbb{R}^n$ if and only if, for each $i = 1, \ldots, n$, the sequence $(x_{ki})_{k\in\mathbb{N}}$ of its i-th coordinates is a Cauchy sequence of real numbers.

This, together with the previous proposition and the fact that a sequence of real numbers is Cauchy if and only if it is convergent, yields the following immediate result.

Proposition 1.10 *A sequence (x_k) in $\mathbb{R}^n$ is a Cauchy sequence if and only if it is convergent.*

Definition 1.11 *Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on $\mathbb{R}^n$ are said to be equivalent if there exist constants $c_1, c_2 > 0$ such that*

$$\|x\|_1 \leq c_1\|x\|_2 \quad \text{and} \quad \|x\|_2 \leq c_2\|x\|_1$$

for all $x \in \mathbb{R}^n$. It is clear that if $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent, then $\lim \|x_k - x_0\|_1 = 0$ if and only if $\lim \|x_k - x_0\|_2 = 0$, that is, equivalent norms give rise to the same notion of limit in $\mathbb{R}^n$. Moreover, a set $X \subset \mathbb{R}^n$ is bounded with respect to one of them if and only if it is bounded with respect to the other.

1.5 Accumulation, Adherent, Interior, and Boundary Points

Definition 1.12 *Let $E \subset \mathbb{R}^n$. A point $x \in \mathbb{R}^n$ is called an accumulation point of E if every open ball centered at x contains some point of E different from x. In other words, x is an accumulation point of E if for every $r > 0$,*

$$(B_r(x) \setminus \{x\}) \cap E \neq \emptyset.$$

Examples.

(1) Let $E = \{\frac{1}{n}\,;\, n \in \mathbb{N}^*\}$. Then 0 is an accumulation point of E, since given $r > 0$ there exists $n(r) \in \mathbb{N}$ such that $0 < \frac{1}{n(r)} < r$.

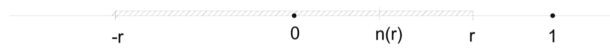

Figure 1.7:

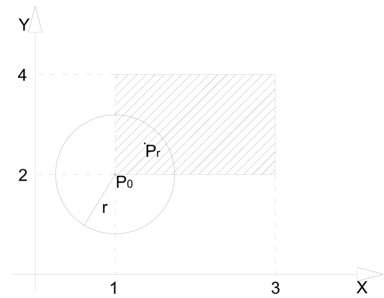

Figure 1.8:

(2) Let $E = (1, 3] \times [2, 4]$ (Figure 1.8).

The point $P_0 = (1, 2)$, for instance, is an accumulation point of E, since for any $r > 0$ there exists $p_r \in E$ such that $\|p_r - P_0\| < r$.

These examples show that an accumulation point p of a set need not belong to the set. Observe that in Example 1, no point of E is an accumulation point of E, since for each $n \in \mathbb{N}^*$ we can find $r_n > 0$ such that the open interval centered at $1/n$ and radius r_n contains no point of E other than $1/n$ itself.

On the other hand, in Example 2 we have that every point of E is an accumulation point of E.

A point $x \in E$ that is not an accumulation point is called an *isolated point* of E. Thus $x \in E$ is isolated if and only if there exists $r_0 > 0$ such that $B_{r_0}(x) \cap E = \{x\}$.

In Example 1, by the argument above, all points of E are isolated. However, in Example 2, the set E has no isolated points.

Proposition 1.13 *Let $E \subset \mathbb{R}^n$. If $x \in \mathbb{R}^n$ is an accumulation point of E, then for every $r > 0$ the ball $B_r(x)$ contains infinitely many points of E.*

Proof: Suppose, by contradiction, that there exists $r_0 > 0$ such that the ball $B_{r_0}(x)$ contains only finitely many points of E, say $x_1, \ldots, x_k$ (see Figure 1.9). Set

$$\overline{r} = \min\{\|x_1 - x\|, \ldots, \|x_k - x\|\}.$$

Then the ball $B_{\overline{r}}(x)$ contains no point of E except possibly x itself if $x \in E$. This contradicts the fact that x is an accumulation point of E, since every open ball centered at x must contain points of E different from x. $\square$

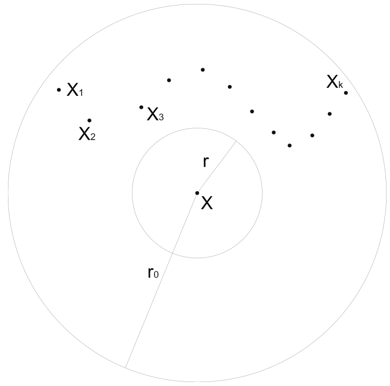

Figure 1.9:

Proposition 1.14 *Let $E \subset \mathbb{R}^n$. If x is an accumulation point of E, then there exists a sequence (x_k) of pairwise distinct elements of E converging to x.*

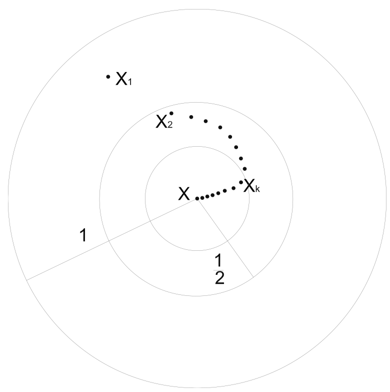

Figure 1.10:

Proof: Take $r_1 = 1$. Then there exists $x_1 \in E$ such that $0 < \|x_1 - x\| < 1$. Next, let

$$r_2 = \min\left\{\|x_1 - x\|, \frac{1}{2}\right\}.$$

Then there exists $x_2 \in E$ such that $0 < \|x_2 - x\| < r_2$. Similarly, with

$$r_3 = \min\left\{\|x_2 - x\|, \frac{1}{3}\right\},$$

there exists $x_3 \in E$ such that $0 < \|x_3 - x\| < r_3$. (See Figure 1.10)

Proceeding inductively in this way, we obtain a sequence (x_k) such that

$$0 < \|x_{k+1} - x\| < \|x_k - x\| \quad \text{and} \quad \|x_k - x\| < \frac{1}{k}.$$

Thus the x_k are pairwise distinct and, moreover, $\lim_{k\to\infty} x_k = x$. □

We denote by E' the set of all accumulation points of a set E. Returning to Examples 1 and 2 above, we obtain in the first case $E' = \{0\}$ and in the second case $E' = [1,3]\times[2,4]$.

As another example, let $E = \{1,2\}$. In this case the set E' of accumulation points of E is empty, since each point of E is isolated. It suffices to consider the balls of radius $\frac{1}{2}$ centered at each of these points. These balls contain no point of E other than their centers, as illustrated in Figure 1.11.

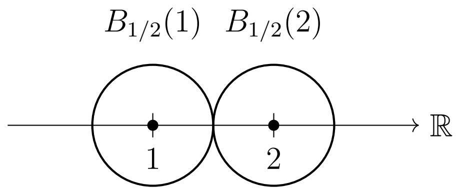

Figure 1.11:

As in this example, any set $E \subset \mathbb{R}^n$ consisting of finitely many points has $E' = \emptyset$, since all of its points are isolated. It follows that if $E' \neq \emptyset$, then E is infinite.

Definition 1.15 *A point $x \in \mathbb{R}^n$ is said to be* <u>*adherent*</u> *to a set $E \subset \mathbb{R}^n$ if every open ball centered at x contains some point of E (not necessarily different from x). In other words, $x \in \mathbb{R}^n$ is adherent to E if and only if, for every $r > 0$,*

$$B_r(x) \cap E \neq \emptyset.$$

<u>**Remark 1.**</u> If x is adherent to a set E and we consider the sequence $r_k = \frac{1}{k}$, $k \in \mathbb{N}^*$, then for each k there exists some $x_k \in B_{1/k}(x)$ with $x_k \in E$. Hence $\|x_k - x\| < \frac{1}{k}$ and therefore $x_k \to x$ as $k \to \infty$. Thus there exists a sequence $(x_k) \subset E$ such that $x_k \to x$.

Conversely, if there exists a sequence $(x_k) \subset E$ such that $x_k \to x$, then x is adherent to E. Indeed, given $r > 0$, there exists $k_0 \in \mathbb{N}$ such that

$$\|x_k - x\| < r \quad \text{for all } k \geq k_0.$$

It follows that inside the open ball of radius r there is at least one element of E, namely one of the x_k with $k \geq k_0$.

From the above discussion we obtain the following result.

Proposition 1.16 *A point $x \in \mathbb{R}^n$ is adherent to a set $E \subset \mathbb{R}^n$ if and only if there exists a sequence (x_k) of elements of E converging to x.*

We denote by $\overline{E}$ the set of adherent points of E. This set is called the <u>adherence</u> or <u>closure</u> of E. Note that every point of E is adherent to E. Indeed, if $x \in E$, then every open ball centered at x contains at least one point of E, namely x itself. Thus $E \subset \overline{E}$.

It is also easy to verify that every accumulation point of E is adherent to E. Hence $E' \subset \overline{E}$, and therefore $E \cup E' \subset \overline{E}$.

On the other hand, if $x \in \overline{E}$, then for every $r > 0$ we have $B_r(x) \cap E \neq \emptyset$. In other words, for each $r > 0$ there exists $y_r \in B_r(x)$ with $y_r \in E$.

There are two cases to consider:

(1) $y_r \neq x$.

In this case, x is an accumulation point of E.

(2) $y_r = x$.

In this case, since $y_r \in E$, we have $x \in E$.

Thus, from (1) and (2), we conclude that $\overline{E} \subset E \cup E'$, and consequently

$$\overline{E} = E \cup E'.$$

Example 4. Let

$$E = (1,3) \times [2,5] \ \cup \ \{(5,2),(5,5)\}.$$

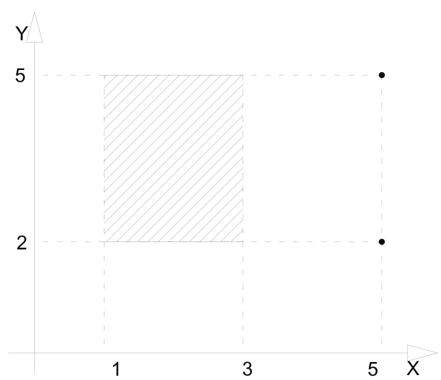

Figure 1.12:

We have $E' = [1,3] \times [2,5]$ and

$$\overline{E} = [1,3] \times [2,5] \ \cup \ \{(5,2),(5,5)\}.$$

Definition 1.17 *A point $x \in E \subset \mathbb{R}^n$ is called an* <u>*interior point*</u> *of E if there exists an open ball centered at x that is entirely contained in E. In other words, $x \in E$ is an interior point of E if there exists $r_x > 0$ such that $B_{r_x}(x) \subset E$.*

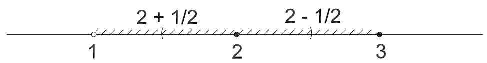

Figure 1.13:

Example 5. Let $E = (1, 3]$. The point 2 is an interior point of E, since there exists $r_0 = \frac{1}{2}$ such that $B_{1/2}(2) \subset E$ (see Figure 1.13).

Example 6. Consider

$$E = \{(x, y) \in \mathbb{R}^2\,;\ \tfrac{x^2}{4} + y^2 \leq 1\},$$

as in Figure 1.14. The point $(0, 0)$ is an interior point of E, since there exists $r_0 = \frac{1}{2}$ such that $B_{1/2}(0, 0) \subset E$.

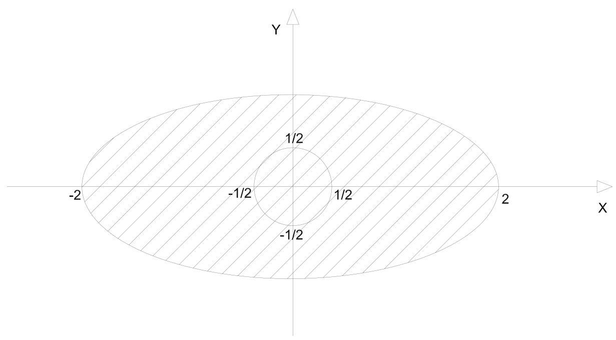

Figure 1.14:

Proposition 1.18 *Let $\varepsilon > 0$ and $x \in \mathbb{R}^n$. Then every point belonging to the ball $B_\varepsilon(x)$ is an interior point of this ball.*

Proof: Let $y \in B_\varepsilon(x)$. We must exhibit $r > 0$ such that $B_r(y) \subset B_\varepsilon(x)$.

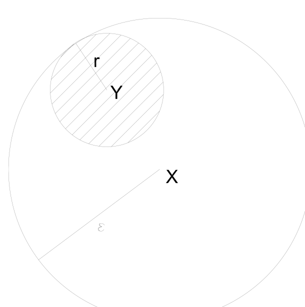

Figure 1.15:

Set $r = \varepsilon - \|y - x\|$. Note that $r > 0$, since $\|y - x\| < \varepsilon$ because $y \in B_\varepsilon(x)$. We shall prove that $B_r(y) \subset B_\varepsilon(x)$.

Indeed, let $z \in B_r(y)$. Then $\|z - y\| < r$. We want to show that $\|z - x\| < \varepsilon$. In fact,

$$\begin{aligned} \|z - x\| &\leq \|z - y\| + \|y - x\| \\ &< r + \|y - x\| \\ &= \varepsilon - \|y - x\| + \|y - x\| \\ &= \varepsilon. \end{aligned}$$

Therefore $\|z - x\| < \varepsilon$, as desired. □

We denote by E° the set of interior points of E. This set is called the interior of E. Clearly $E^\circ \subset E$.

Definition 1.19 *A point $x \in \mathbb{R}^n$ is called a boundary point of a set $E \subset \mathbb{R}^n$ if every open ball centered at x contains points of E and points of the complement of E. In other words, $x \in \mathbb{R}^n$ is a boundary point of E if and only if, for every $r > 0$,*

$$B_r(x) \cap E \neq \emptyset \quad \textit{and} \quad B_r(x) \cap E^c \neq \emptyset.$$

Example 7. Let $E = (2, 5]$. The points 2 and 5 are boundary points of E, since every open ball centered at one of these points contains points of E and points of the complement of E.

Example 8. Consider

$$E = \{(x, y) \in \mathbb{R}^2\,;\, 0 \leq x \leq 1 \text{ and } y \leq x\}.$$

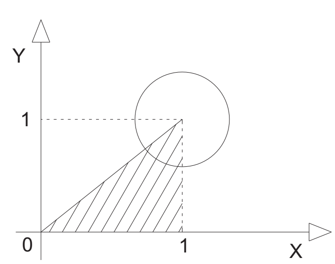

Figure 1.16:

The point $(1, 1)$, for instance, is a boundary point of E, since every ball centered at $(1, 1)$ contains points of E and points of the complement of E.

The set of all boundary points of a set E is called the boundary of E and is denoted by ∂E.

For example, if E is the closed ball $\overline{B_r(x)}$, then $\partial E = S_r(x)$, the sphere of center x and radius r. To fix ideas, consider the planar illustration in Figure 1.17.

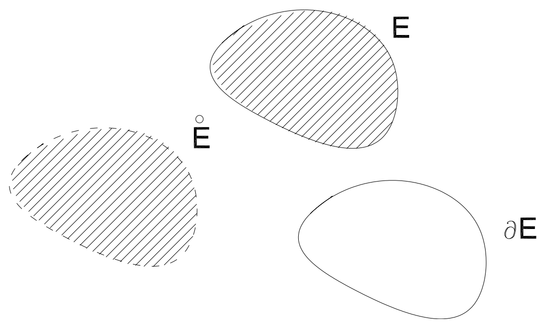

Figure 1.17:

As a more complete analysis, consider the following examples.

1. $A = \{1, 2, 3\}$.

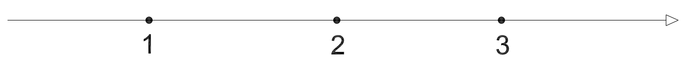

Figure 1.18:

From the figure we observe that:

- $\overline{A}$ (closure) $= A$;
- A° (interior) $= \emptyset$;
- A' (set of accumulation points) $= \emptyset$;
- A is a set of isolated points.

2. $B = [1, 2) \cup \{3\}$.

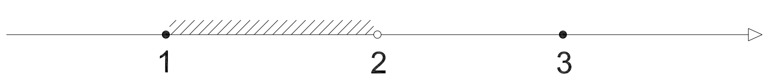

Figure 1.19:

In this case:

- $\overline{B} = [1, 2] \cup \{3\}$;
- $B^\circ = (1, 2)$;
- $B' = [1, 2]$;
- ∂B (boundary) $= \{1, 2\}$;
- $x = 3$ is an isolated point.

3. $C = \left\{\frac{1}{n}\right\}_{n\in\mathbb{N}^*}$ (Figure 1.20).

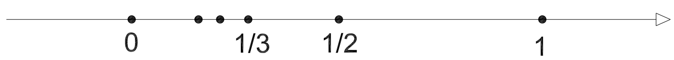

Figure 1.20:

Here:

- $\overline{C} = C \cup \{0\}$;
- $C^\circ = \emptyset$;
- $C' = \{0\}$;
- ∂C (boundary) $= \{0\}$;
- C is a set of isolated points.

4. $D = [1, 3) \times (2, 4] \cup \{(2, 5)\}$.

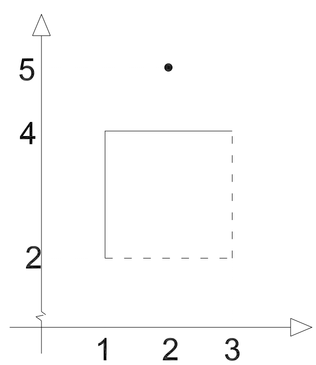

Figure 1.21:

We have:

- $\overline{D} = [1, 3] \times [2, 4] \cup \{(2, 5)\}$;

- $D^\circ = (1,3) \times (2,4)$;
- $D' = [1,3] \times [2,4]$;
- $\partial D = \{1\} \times [2,4] \cup \{3\} \times [2,4] \cup [1,3] \times \{2\} \cup [1,3] \times \{4\}$;
- $(2,5)$ is an isolated point.

5. $E = \{(x,y) \in \mathbb{R}^2\,;\, x^2+y^2 < 1 \text{ and } y \geq x\} \cup \{(2-\frac{1}{n}, 2+\frac{1}{n})\}_{n\in\mathbb{N}^*}$ (Figure 1.22).

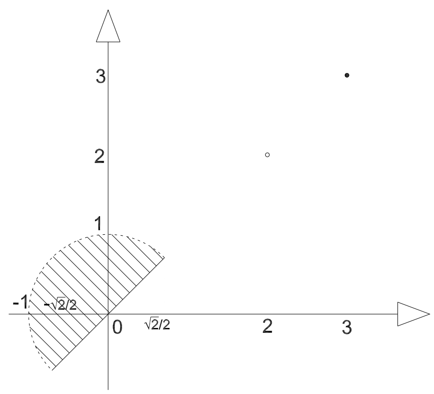

Figure 1.22:

In this case:

- $\overline{E} = \{(x,y) \in \mathbb{R}^2\,;\, x^2+y^2 \leq 1 \text{ and } y \geq x\} \cup \{(2-\frac{1}{n}, 2+\frac{1}{n})\}_{n\in\mathbb{N}^*}$;
- $E^\circ = \{(x,y) \in \mathbb{R}^2\,;\, x^2+y^2 < 1 \text{ and } y > x\}$;
- $E' = \{(x,y) \in \mathbb{R}^2\,;\, x^2+y^2 \leq 1 \text{ and } y \geq x\} \cup \{(2,2)\}$;
- $\partial E = \{(x,y)\,;\, y = x,\ -\frac{\sqrt{2}}{2} \leq x \leq \frac{\sqrt{2}}{2}\} \cup \{(x,y)\,;\, y = \sqrt{1-x^2},\ -1 \leq x \leq \frac{\sqrt{2}}{2}\}$ $\cup\ \{(x,y)\,;\, y = -\sqrt{1-x^2},\ -1 \leq x \leq -\frac{\sqrt{2}}{2}\}$.

1.6 Open and Closed Sets

Definition 1.20 *A set $E \subset \mathbb{R}^n$ is said to be* open *if every point of E is an interior point. In other words, $E \subset \mathbb{R}^n$ is open if, for each $x \in E$, there exists $r_x > 0$ such that $B_{r_x}(x) \subset E$.*

Example 1. As we have seen in the previous section, every open ball is an open set.

Example 2. The set $E = (-3,-1) \times (1,4)$ is an open set, as we can see in Figure 1.23. Note that every point of E is an interior point of E.

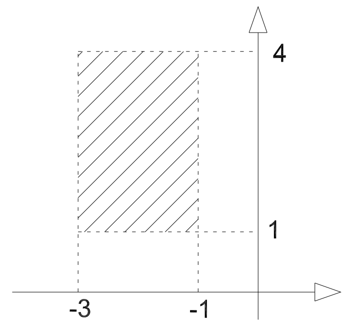

Figure 1.23:

Example 3. The empty set $\emptyset$ is open. Indeed, a set E fails to be open if there exists in E some point that is not interior. Since there is no point at all in $\emptyset$, we must accept that $\emptyset$ is open.

Example 4. The whole space $\mathbb{R}^n$ is clearly open.

Warning. The sets $A = (1, 3]$ and $B = [2, 4] \times [4, 5)$ are not open. See Figure 1.24.

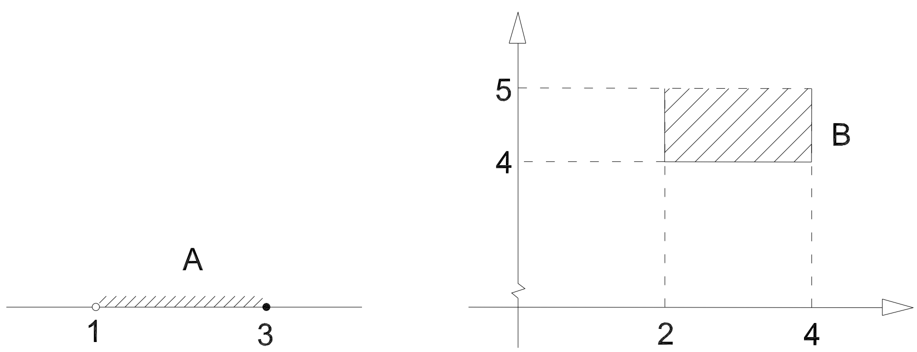

Figure 1.24:

In fact, the point $3 \in A$ is not an interior point of A, and the point $(3, 4)$, for example, is not an interior point of B.

Proposition 1.21 *Let $\{E_\alpha\}_{\alpha \in I}$ be a family of open subsets of $\mathbb{R}^n$. Then $E = \bigcup_{\alpha \in I} E_\alpha$ is an open subset of $\mathbb{R}^n$.*

Proof: Let $x \in E$. We must exhibit $r > 0$ such that $B_r(x) \subset E$. Indeed, since $x \in \bigcup_{\alpha \in I} E_\alpha$, we have $x \in E_\alpha$ for some $\alpha \in I$. As E_α is open by assumption, there exists $r > 0$ such that $B_r(x) \subset E_\alpha$, and since $E_\alpha \subset E$, we obtain $B_r(x) \subset E$. $\square$

Remark 1. The analogous statement is not true for arbitrary intersections of open sets. Consider, for example,

$$E_n = \left(-\frac{1}{n}, \frac{1}{n}\right), \quad n = 1, 2, \ldots, \quad \text{and} \quad E = \bigcap_{n=1}^{\infty} E_n.$$

Then E is not open. In fact, note first that $E = \{0\}$. Indeed, clearly $\{0\} \subset E$, since $0 \in E_n$ for all $n \in \mathbb{N}^*$. Now let $x \in E$ and suppose, by contradiction, that $x \neq 0$. Then $|x| > 0$, and there exists a natural number n_0 such that $0 < \frac{1}{n_0} < |x|$.

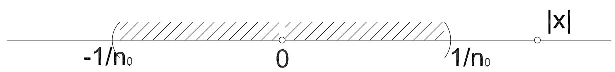

Figure 1.25:

This means that $x \notin (-\frac{1}{n_0}, \frac{1}{n_0}) = E_{n_0}$, which is a contradiction, since $x \in E_n$ for all $n \in \mathbb{N}^*$. Therefore $E = \{0\}$, and it is clear that $\{0\}$ is not open.

Definition 1.22 *A set $F \subset \mathbb{R}^n$ is said to be* closed *if F contains all of its accumulation points. In other words, F is closed if and only if $F \supset F'$.*

Example 5. Let $F = \{1, 2, 3\}$. In this case the set F' of accumulation points of F is empty, since F is a set of isolated points. Thus $F' = \emptyset$ and therefore F is a closed subset of $\mathbb{R}$.

Example 6. The closed ball $\overline{B_r(x_0)} \subset \mathbb{R}^n$ is a closed set, since $F' = F$ in this case.

Example 7. Both the empty set $\emptyset$ and the whole space $\mathbb{R}^n$ are closed.

Example 8. The set $F = [-4, -2] \times [-3, -1]$ is a closed subset of $\mathbb{R}^2$, since $F' = F$.

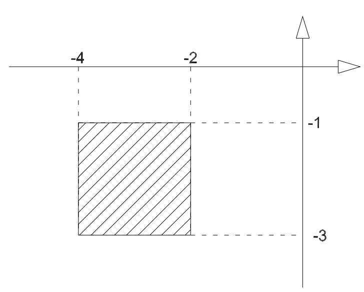

Figure 1.26:

Proposition 1.23 *A set $A \subset \mathbb{R}^n$ is open if and only if the complement $\complement A$ is closed.*

Proof: Suppose first that A is open, and assume by contradiction that $\complement A$ is not closed. Then there exists some $x_0 \in (\complement A)'$, that is, x_0 is an accumulation point of $\complement A$, but $x_0 \notin \complement A$. Since $x_0 \notin \complement A$, we have $x_0 \in A$, and because A is open there exists $r > 0$ such that $B_r(x_0) \subset A$. On the other hand, since $x_0 \in (\complement A)'$, the ball $B_r(x_0)$ must satisfy $B_r(x_0) \cap \complement A \neq \emptyset$. Thus there exists some $y \in B_r(x_0)$ with $y \notin A$, which is a contradiction, since $B_r(x_0) \subset A$.

Conversely, suppose that $\complement A$ is closed and, by contradiction, that A is not open. Then there exists some $x_0 \in A$ such that for every $\varepsilon > 0$ the ball $B_\varepsilon(x_0)$ is not contained in A. In other words, for each $\varepsilon > 0$ there exists some $y_\varepsilon \in B_\varepsilon(x_0)$ with $y_\varepsilon \notin A$ and $y_\varepsilon \neq x_0$ (for if $y_\varepsilon = x_0$, then $y_\varepsilon \in A$, which is absurd). Thus x_0 is an accumulation point of $\complement A$, and since $\complement A$ is closed, we must have $x_0 \in \complement A$, which is impossible because $x_0 \in A$. □

Corollary 1.24 *A set F is closed if and only if its complement $\complement F$ is open.*

Proof: Just take $A = \complement F$ in the previous proposition. □

Lema 1.25 *Let $\{E_\alpha\}_\alpha$ be a collection of open sets. Then*

$$\complement\Big(\bigcup_\alpha E_\alpha\Big) = \bigcap_\alpha (\complement E_\alpha).$$

Proof: If $x \in \complement(\bigcup_\alpha E_\alpha)$, then $x \notin \bigcup_\alpha E_\alpha$, that is, $x \notin E_\alpha$ for every α, or equivalently, $x \in \complement E_\alpha$ for every α. Hence $x \in \bigcap_\alpha (\complement E_\alpha)$.

Conversely, suppose that $x \in \bigcap_\alpha (\complement E_\alpha)$. Then $x \in \complement E_\alpha$ for every α, and therefore $x \notin E_\alpha$ for every α. Hence $x \notin \bigcup_\alpha E_\alpha$, and so $x \in \complement(\bigcup_\alpha E_\alpha)$. □

Remark 2. In the proof of the lemma above we used the facts that

$$x \in \bigcup_\alpha E_\alpha \iff x \in E_\alpha \text{ for some } \alpha,$$

$$x \notin \bigcup_\alpha E_\alpha \iff x \notin E_\alpha \text{ for every } \alpha.$$

Similarly, one proves that

$$\complement\Big(\bigcap_\alpha E_\alpha\Big) = \bigcup_\alpha (\complement E_\alpha).$$

Remark 3. It is worth emphasizing that there are sets in $\mathbb{R}^n$ that are neither open nor closed. For instance, the set $A = (1, 5]$ is not open, since 5 is not an interior point of A although $5 \in A$, and it is not closed, since 1 is an accumulation point of A and $1 \notin A$.

Similarly, the set

$$B = \{(x, y) \in \mathbb{R}^2\,;\, y < x \text{ and } x^2 + y^2 \leq 1\}$$

is not open, since the point $(1, 0)$, for example, is not an interior point of B although $(1, 0) \in B$, and it is not closed, since $(0, 0)$ is an accumulation point of B and does not belong to B (see Figure 1.27).

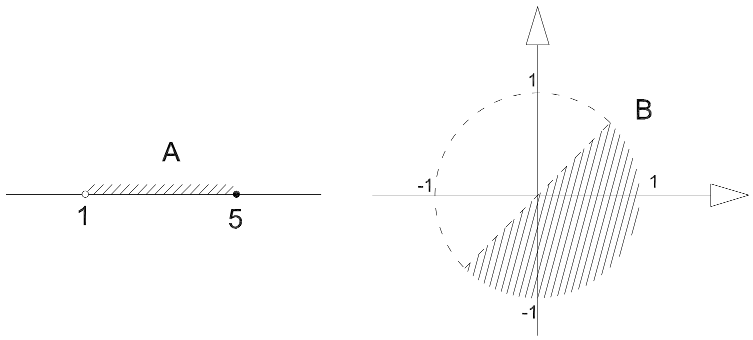

Figure 1.27:

Proposition 1.26 *Let $\{F_\alpha\}_{\alpha \in I}$ be a family of closed subsets of $\mathbb{R}^n$. Then*

$$F = \bigcap_\alpha F_\alpha$$

is a closed subset of $\mathbb{R}^n$.

Proof: It suffices to show that $\complement F$ is open.

Indeed,

$$\complement F = \complement\Big(\bigcap_\alpha F_\alpha\Big) = \bigcup_\alpha (\complement F_\alpha).$$

Since each F_α is closed, each $\complement F_\alpha$ is open. Hence $\bigcup_\alpha (\complement F_\alpha)$ is an open set, as it is an arbitrary union of open sets. □

Remark 4. Here we have an observation analogous to the one made for families of open sets: the union of an arbitrary family of closed sets need not be closed. For example, let $F \subset \mathbb{R}^n$ be a set that is not closed. Clearly

$$F = \bigcup_{x \in F} \{x\},$$

and each singleton $\{x\}$ is closed, but their union F is not closed.

Proposition 1.27 *A set $E \subset \mathbb{R}^n$ is closed if and only if $E = \overline{E}$.*

Proof: If E is closed, then $E \supset E'$, and thus $E' \cup E \subset E$, that is, $\overline{E} \subset E$. Since $E \subset \overline{E}$, it follows that $E = \overline{E}$. Conversely, if $E = \overline{E}$, then $E' \subset E$, and therefore E is closed. □

It follows from this proposition that a set E is closed if and only if for every sequence $(x_k) \subset E$ such that $x_k \to x_0$ we have $x_0 \in E$.

1.7 Compact Sets

Definition 1.28 *A* cover *of a set $E \subset \mathbb{R}^n$ is a family $\mathfrak{C} = \{C_\gamma\}_{\gamma \in I}$ of sets $C_\gamma \subset \mathbb{R}^n$ such that*

$$E \subset \bigcup_{\gamma \in I} C_\gamma.$$

Definition 1.29 *A* subcover *of $\mathfrak{C}$ is a subfamily $\mathfrak{C}' = \{C_\gamma\}_{\gamma \in I'}$ with $I' \subset I$ such that still*

$$E \subset \bigcup_{\gamma \in I'} C_\gamma.$$

Remark 1. A cover is called *open* when all sets in the family are open. Similarly, we speak of a closed cover if the sets are closed.

Definition 1.30 *A set $K \subset \mathbb{R}^n$ is called compact if every open cover of K admits a finite subcover.*

Our goal from now on is to characterize the compact sets in $\mathbb{R}^n$, since the definition above is rather abstract.

Proposition 1.31 *Let $K \subset \mathbb{R}^n$ be a compact set. Then K is closed and bounded.*

Proof: *(1) K is bounded.*

For each $n \in \mathbb{N}$, consider the open ball $G_n = B_n(0)$. Clearly,

$$K \subset \bigcup_{n=1}^{\infty} G_n,$$

since this union covers all of $\mathbb{R}^n$. Indeed, given $x \in \mathbb{R}^n$, there exists $n_0 \in \mathbb{N}$ such that $\|x\| < n_0$. Otherwise, if $\|x\| \geq n$ for every $n \in \mathbb{N}$, then the set of natural numbers would be bounded, which is absurd. Thus $x \in B_{n_0}(0) = G_{n_0} \subset \bigcup_{n=1}^{\infty} G_n$, that is, $\mathbb{R}^n \subset \bigcup_{n=1}^{\infty} G_n$.

Since K is compact, there exist $n_1, n_2, \ldots, n_k \in \mathbb{N}$ such that

$$K \subset \bigcup_{i=1}^{k} G_{n_i}.$$

We may assume, without loss of generality, that $n_1 \leq n_2 \leq \cdots \leq n_k$. Then

$$\bigcup_{i=1}^{k} G_{n_i} = B_{n_k}(0),$$

so that $K \subset B_{n_k}(0)$, which shows that K is bounded.

(2) K is closed.

It suffices to show that $\complement K$ is open. Let $x \in \complement K$. We must find $r > 0$ such that $B_r(x) \subset \complement K$.

For each $n \in \mathbb{N}$, set

$$G_n = \{y \in \mathbb{R}^n \, ; \, \|y - x\| > \tfrac{1}{n}\} = \complement \overline{B_{1/n}(x)}.$$

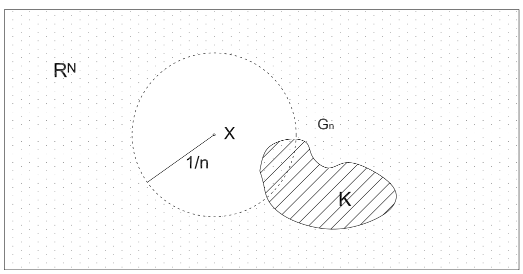

Figure 1.28:

Clearly G_n is open, since $\overline{B_{1/n}(x)}$ is closed. We claim that

$$\bigcup_{n=1}^{\infty} G_n = \mathbb{R}^n \setminus \{x\}.$$

(i) Let $y \in \bigcup_{n=1}^{\infty} G_n$. Then $y \in G_{n_0}$ for some $n_0 \in \mathbb{N}$. Thus $y \in \mathbb{R}^n$ and $\|y - x\| > \frac{1}{n_0}$, which implies $y \neq x$. Hence $y \in \mathbb{R}^n \setminus \{x\}$.

(ii) Conversely, let $y \in \mathbb{R}^n \setminus \{x\}$. Then $y \in \mathbb{R}^n$ and $y \neq x$, so $\|y - x\| > 0$. Choose $n_0 \in \mathbb{N}$ such that $\|y - x\| > \frac{1}{n_0}$. Then $y \in G_{n_0}$, and hence $y \in \bigcup_{n=1}^{\infty} G_n$, which proves the equality.

Since $x \notin K$, we have

$$K \subset \bigcup_{n=1}^{\infty} G_n.$$

Because K is compact, there exist $n_1, n_2, \ldots, n_r \in \mathbb{N}$ (suppose, without loss of generality, that $n_1 \leq n_2 \leq \cdots \leq n_r$) such that

$$K \subset \bigcup_{i=1}^{r} G_{n_i}.$$

Moreover, since $G_{n_{i+1}} \supset G_{n_i}$ for $i = 1, \ldots, r-1$, we have

$$\bigcup_{i=1}^{r} G_{n_i} = G_{n_r}.$$

Thus

$$K \subset G_{n_r} = \complement \overline{B_{1/n_r}(x)},$$

which implies

$$\complement K \supset \overline{B_{1/n_r}(x)} \supset B_{1/n_r}(x).$$

Therefore $B_{1/n_r}(x) \subset \complement K$, and $\complement K$ is open. □

Remark 2. It follows immediately from the previous proposition that if $K \subset \mathbb{R}^n$ is not closed or not bounded, then it is not compact.

1.8 Nested Intervals

Lema 1.32 *Let $(I_n)_{n\in\mathbb{N}}$ be a sequence of closed and bounded intervals in $\mathbb{R}$ such that*

$$I_n \supset I_{n+1} \qquad \forall\, n \in \mathbb{N}.$$

Then

$$\bigcap_{n=1}^{\infty} I_n \neq \emptyset.$$

Proof: For each $n \in \mathbb{N}$, write $I_n = [a_n, b_n]$.

We first show that $a_n \leq b_m$ for all $n, m \in \mathbb{N}$. Indeed, suppose the contrary: that there exist n_0, m_0 such that $a_{n_0} > b_{m_0}$. But since each interval is nonempty, we always have $a_n \leq b_n$ for all n, hence

$$a_{m_0} \leq b_{m_0} < a_{n_0} \leq b_{n_0}.$$

Thus,

$$[a_{m_0}, b_{m_0}] \cap [a_{n_0}, b_{n_0}] = \emptyset,$$

which is impossible because the intervals are nested. Hence $a_n \leq b_m$ for all n, m.

In particular,

$$a_n \leq b_1 \quad \forall n, \qquad a_1 \leq b_m \quad \forall m.$$

Thus the set $\{a_n\}$ is bounded above, and the set $\{b_m\}$ is bounded below. By the completeness axiom, the following limits exist:

$$\alpha = \sup\{a_n \ ; \ n \in \mathbb{N}\}, \qquad \beta = \inf\{b_m \ ; \ m \in \mathbb{N}\}.$$

We now prove that $\alpha \leq \beta$. Suppose by contradiction that $\alpha > \beta$ and take

$$\varepsilon = \frac{\alpha - \beta}{2} > 0.$$

Then there exist $m_0, n_0 \in \mathbb{N}$ such that

$$\beta \leq b_{m_0} < \beta + \varepsilon = \alpha - \varepsilon < a_{n_0} < \alpha.$$

Hence $b_{m_0} < a_{n_0}$, contradicting what we proved above.

Thus $\alpha \leq \beta$, and therefore

$$[\alpha, \beta] \subset \bigcap_{n=1}^{\infty} I_n.$$

Indeed, if $x \in [\alpha, \beta]$, then since $a_n \leq \alpha \leq x \leq \beta \leq b_n$ for all n, we have $x \in I_n$ for all n. Hence

$$[\alpha, \beta] \subset \bigcap_{n=1}^{\infty} I_n,$$

proving that the intersection is nonempty. □

__**Remark 3.**__ In fact, in the previous lemma we actually have

$$[\alpha, \beta] = \bigcap_{n=1}^{\infty} I_n.$$

Indeed, let $x \in I_n = [a_n, b_n]$ for all n, and suppose by contradiction that $x < \alpha$ or $x > \beta$.

(i) If $x < \alpha$, then $\alpha - x > 0$. Take $\varepsilon = \alpha - x$. Then for some n_0 we have

$$\alpha - \varepsilon < a_{n_0} \leq \alpha,$$

that is,

$$x < a_{n_0},$$

contradiction.

(ii) The case $x > \beta$ is similar.

Examples.

(a) $I_n = \left[-\frac{1}{n},\, 1 + \frac{1}{n}\right]$.

- $\bigcap_{n=1}^{\infty} I_n = [0, 1]$.

- $\alpha = \sup \left\{ -\frac{1}{n} \right\} = 0.$

- $\beta = \inf \left\{ 1 + \frac{1}{n} \right\} = 1.$

(b) $I_n = \left[-\frac{1}{n}, \frac{1}{n} \right].$

- $\bigcap_{n=1}^{\infty} I_n = \{0\}.$

- $\alpha = \sup \left\{ -\frac{1}{n} \right\} = 0.$

- $\beta = \inf \left\{ \frac{1}{n} \right\} = 0.$

Definition 1.33 *An n-*parallelepiped*, or n-*dimensional block*, or simply a *cell*, is a subset $A \subset \mathbb{R}^n$ of the form*

$$A = \prod_{i=1}^{n} [a_i, b_i] = [a_1, b_1] \times \cdots \times [a_n, b_n].$$

In one dimension an 1-parallelepiped is a closed bounded interval. In two dimensions it is a rectangle, and in three dimensions it is a usual parallelepiped.

Lema 1.34 *Let (A_k) be a sequence of n-parallelepipeds in $\mathbb{R}^n$ such that*

$$A_k \supset A_{k+1} \qquad \forall k \in \mathbb{N}.$$

Then

$$\bigcap_{k=1}^{\infty} A_k \neq \emptyset.$$

Proof: Write

$$A_k = \{(x_1, \ldots, x_n)\,;\, a_{ki} \le x_i \le b_{ki},\ 1 \le i \le n\}.$$

For each fixed i, the intervals

$$I_{k,i} = [a_{ki}, b_{ki}]$$

satisfy $I_{k,i} \supset I_{k+1,i}$. Thus by Lemma 1.32 there exists x_i^* such that

$$a_{ki} \le x_i^* \le b_{ki}, \qquad \forall k.$$

Let $x^* = (x_1^*, \ldots, x_n^*)$. Then $x^* \in A_k$ for all k, proving the lemma. □

Proposition 1.35 *Let $A \subset \mathbb{R}^n$ be an n-parallelepiped. Then A is compact.*

Proof: Suppose, by contradiction, that A is not compact. Then there exists an open cover

$$\{G_\alpha\}_{\alpha \in I}$$

of A from which no finite subcover can be extracted.

Let $A_0 = A$ and divide A_0 into 2^n equal n-parallelepipeds. At least one of them, say A_1, cannot be covered by finitely many G_α. Otherwise, A_0 itself would be covered by finitely many of the G_α, contradiction.

Now divide A_1 into 2^n equal n-parallelepipeds. Again, one of them, call it A_2, cannot be finitely covered. Proceeding inductively we obtain a nested sequence

$$A_0 \supset A_1 \supset A_2 \supset \cdots$$

of n-parallelepipeds, none of which admits a finite subcover.

Let

$$A_0 = \{(x_1, \ldots, x_n)\,;\, a_i \le x_i \le b_i\}, \qquad \delta = \Big(\sum_{i=1}^{n} (b_i - a_i)^2\Big)^{1/2}.$$

Then

$$x, y \in A_k \quad \Longrightarrow \quad \|x - y\| \le \frac{\delta}{2^k}.$$

By Lemma 1.34 there exists $x' \in \bigcap_{k=1}^{\infty} A_k$. Since $x' \in A$, and $\{G_\alpha\}$ covers A, we have $x' \in G_{\alpha_0}$ for some α_0. As G_{α_0} is open, there exists $\varepsilon_0 > 0$ such that

$$B_{\varepsilon_0}(x') \subset G_{\alpha_0}.$$

Choose $n_0 \in \mathbb{N}$ such that

$$2^{n_0} > \frac{\delta}{\varepsilon_0}.$$

If $y \in A_{n_0}$, then since $x' \in A_{n_0}$,

$$\|y - x'\| \le \frac{\delta}{2^{n_0}} < \varepsilon_0,$$

so $y \in B_{\varepsilon_0}(x') \subset G_{\alpha_0}$.

Thus

$$A_{n_0} \subset G_{\alpha_0},$$

contradicting the fact that no A_k can be finitely covered.

Hence A must be compact. □

Proposition 1.36 *Let $K \subset \mathbb{R}^n$ be a compact set and let F be a closed subset of K. Then F is compact.*

Proof: Let $\{G_\alpha\}_{\alpha\in A}$ be an open cover of F. We must exhibit a finite subcover of F.

Since F is closed, its complement $\complement F$ is open. We claim that

$$K \subset \left(\bigcup_{\alpha\in A} G_\alpha\right) \cup \complement F.$$

Indeed, let $x \in K$. Since $F \subset K$, we have two cases:

(i) If $x \in F$, then, as $\{G_\alpha\}_{\alpha\in A}$ is an open cover of F, there exists some $\alpha \in A$ such that $x \in G_\alpha$. Hence $x \in \bigcup_{\alpha\in A} G_\alpha$.

(ii) If $x \notin F$, then $x \in \complement F$.

Thus $\{G_\alpha\}_{\alpha\in A}$ together with $\complement F$ form an open cover of K. Since K is compact, there exist $\alpha_1, \dots, \alpha_n \in A$ such that

$$K \subset \left(G_{\alpha_1} \cup \cdots \cup G_{\alpha_n}\right) \cup \complement F.$$

However, $F \subset K$ and clearly $F \not\subset \complement F$, so necessarily

$$F \subset G_{\alpha_1} \cup \cdots \cup G_{\alpha_n},$$

which shows that F is compact. □

Proposition 1.37 (Heine–Borel) *A subset $K \subset \mathbb{R}^n$ is compact if and only if it is closed and bounded.*

Proof: We have already proved that every compact set is closed and bounded. It remains to show that if K is closed and bounded, then K is compact.

Since K is bounded, there exists an n-parallelepiped A such that $K \subset A$. As A is compact and K is closed, it follows from Proposition 1.36 that K is compact. □

Proposition 1.38 *Let $K \subset \mathbb{R}^n$ be compact. Then for every infinite subset $A \subset K$ there exists a point $x_A \in K$ which is an accumulation point of A.*

Proof: Suppose, by contradiction, that there exists an infinite subset $A \subset K$ such that no point of K is an accumulation point of A. Thus, for each $x \in K$ there exists $\varepsilon_x > 0$ such that

$$\left(B_{\varepsilon_x}(x) \setminus \{x\}\right) \cap A = \emptyset.$$

The collection of balls $\{B_{\varepsilon_x}(x)\}_{x\in K}$ is an open cover of K. Since K is compact, there exist $x_1, \dots, x_n \in K$ and $\varepsilon_1, \dots, \varepsilon_n > 0$ such that

$$K \subset \bigcup_{i=1}^{n} B_{\varepsilon_i}(x_i).$$

Since $A \subset K$, it follows that

$$A \subset \bigcup_{i=1}^{n} B_{\varepsilon_i}(x_i).$$

Hence

$$A = A \cap A \subset \bigcup_{i=1}^{n} \big(B_{\varepsilon_i}(x_i) \cap A\big).$$

But for each i we have either $B_{\varepsilon_i}(x_i) \cap A = \emptyset$ or $B_{\varepsilon_i}(x_i) \cap A = \{x_i\}$ (if $x_i \in A$). Thus A is contained in a finite set, which contradicts the fact that A is infinite. $\square$

Corollary 1.39 (Bolzano–Weierstrass) *Every infinite bounded subset of $\mathbb{R}^n$ has an accumulation point.*

Proof: Let A be an infinite bounded subset of $\mathbb{R}^n$. Since A is bounded, there exists an n-parallelepiped B such that $B \supset A$. As B is compact and A is an infinite subset of B, the previous proposition implies that there exists $x_A \in B$ which is an accumulation point of A. $\square$

Remark 4. If $A = (a_k)$ is a bounded sequence in $\mathbb{R}^n$, then there exists $x_A \in \mathbb{R}^n$ which is an accumulation point of A, that is, there exists a subsequence (a_{n_k}) of (a_n) such that $a_{n_k} \to x_A$.

1.9 Induced Topology

Definition 1.40 *Let $X \subset \mathbb{R}^n$. A subset $A \subset X$ is said to be open in X if for every $a \in A$ there exists $\delta > 0$ such that*

$$B_\delta(a) \cap X \subset A.$$

In other words, for each $a \in A$ there exists $\delta > 0$ such that every point $x \in X$ with $\|x - a\| < \delta$ belongs to A.

Example 1. The set $A = (0,1]$ is open in $X = [0,1]$, although it is not open in $\mathbb{R}$.

We arrive at the notion of open sets in X by ignoring the points outside X and mimicking the usual definition of an open set. When $X \subset \mathbb{R}^n$ is open, a subset $A \subset X$ is open in X if and only if it is open in $\mathbb{R}^n$ in the usual sense.

More generally, a subset $A \subset X$ is open in X if and only if there exists an open set $B \subset \mathbb{R}^n$ such that

$$A = X \cap B.$$

Indeed, if A is open in X, let B be the union of all balls $B_\delta(a)$ with center $a \in A$ and such that $B_\delta(a) \cap X \subset A$. Clearly B is open and $A = X \cap B$. Conversely, if $A = X \cap B$ with B open in $\mathbb{R}^n$, then for each $a \in A$ there exists a ball $B_\delta(a) \subset B$, and hence

$$B_\delta(a) \cap X \subset B \cap X = A,$$

so A is open in X.

Remark 1. An analogue of Proposition 1.21 for open sets in X also holds: arbitrary unions and finite intersections of sets open in X are open in X.

Definition 1.41 *Let $X \subset \mathbb{R}^n$. A subset $F \subset X$ is said to be closed in X if there exists a closed set $G \subset \mathbb{R}^n$ such that*

$$F = X \cap G.$$

Remark 2. Closed sets in X satisfy the analogue of Proposition 1.27: arbitrary intersections and finite unions of sets closed in X are closed in X.

Definition 1.42 *Let $Y \subset X \subset \mathbb{R}^n$. The* closure of Y relative to X *is the set*

$$\overline{Y}^X = \overline{Y} \cap X,$$

that is, the set of points of X which are adherent to Y.

Note that if $Y = \overline{Y}^X = \overline{Y} \cap X$, then Y is closed in X.

An important particular case occurs when the closure of Y in X is the whole of X. To describe this situation, we introduce the following notion.

Definition 1.43 *Let $Y \subset X \subset \mathbb{R}^n$. We say that Y is dense in X if*

$$\overline{Y}^X = X,$$

that is, $\overline{Y} \cap X = X$, or equivalently $X \subset \overline{Y}$.

It follows that every point of X is the limit of a sequence whose terms belong to Y. Equivalently, every open ball centered at a point of X contains points of Y.

Remark 3. When $X = \mathbb{R}^n$, a set Y is dense in X if and only if $\mathbb{R}^n \subset \overline{Y}$. For example, $\mathbb{Q}$ is dense in $\mathbb{R}$, since for every $x \in \mathbb{R}$ there exists a sequence $(x_n) \subset \mathbb{Q}$ such that $x_n \to x$, or equivalently, every open ball centered at x and of radius $r > 0$ contains rational points. Similarly, $\mathbb{Q}^n$ is dense in $\mathbb{R}^n$.

Proposition 1.44 *Every set $X \subset \mathbb{R}^n$ contains a countable subset E which is dense in X.*

Proof: Let $\mathcal{B}$ be the collection of all open balls $B_r(q)$ with center $q \in \mathbb{Q}^n$ and rational radius $r > 0$. This family is clearly countable; write

$$\mathcal{B} = \{B_1, B_2, \dots, B_k, \dots\}.$$

For each $k \in \mathbb{N}$, choose a point $x_k \in B_k \cap X$ if $B_k \cap X \neq \emptyset$. If $B_k \cap X = \emptyset$, we simply do not choose a point. The set E obtained in this way is a countable subset of X. We now show that E is dense in X, that is, $X \subset \overline{E}$.

Let $x \in X$ and $\varepsilon > 0$ be given. There exists a rational number $r > 0$ such that $2r < \varepsilon$. Since $\mathbb{Q}^n$ is dense in $\mathbb{R}^n$, we can find $q \in \mathbb{Q}^n$ such that $\|q - x\| < r$. Then $x \in B_r(q) = B_k$ for some k, so $B_k \cap X \neq \emptyset$, and thus there exists $x_k \in E$ with $x_k \in B_k$.

Because both x and x_k lie in $B_k = B_r(q)$, we have

$$\|x - x_k\| \leq \|x - q\| + \|q - x_k\| < r + r = 2r < \varepsilon.$$

Hence every open ball $B_\varepsilon(x)$ with center in X contains some point $x_k \in E$, and therefore E is dense in X. □

Proposition 1.45 (Lindelöf Property) *Let $X \subset \mathbb{R}^n$ be arbitrary. Every open cover $\{A_\lambda\}_{\lambda \in I}$ of X admits a countable subcover.*

Proof: By Proposition 1.44, let

$$E = \{x_1, x_2, \dots, x_k, \dots\} \subset X$$

be a countable dense subset of X. Let $\mathcal{B}$ be the set of all open balls $B_r(x_k)$ with center $x_k \in E$, rational radius $r > 0$, and such that each of these balls is contained in some A_λ. The family $\mathcal{B}$ is countable. We claim that the balls $B \in \mathcal{B}$ cover X.

Indeed, let $x \in X$. Since $\{A_\lambda\}_{\lambda \in I}$ is an open cover, there exists $\lambda \in I$ and $r > 0$ such that

$$B_{2r}(x) \subset A_\lambda.$$

As E is dense in X, we can choose $x_k \in E$ with $\|x - x_k\| < r$. Then $x \in B_r(x_k)$.

To show that $B_r(x_k) \in \mathcal{B}$, it remains to verify that $B_r(x_k) \subset A_\lambda$, as in Figure 1.39. If $y \in B_r(x_k)$, then

$$\|y - x_k\| < r,$$

and hence

$$\|y - x\| \leq \|y - x_k\| + \|x_k - x\| < r + r = 2r,$$

so $y \in B_{2r}(x) \subset A_\lambda$.

This proves that the balls in $\mathcal{B}$ cover X. Enumerating them as $B_1, B_2, \dots, B_k, \dots$ and choosing, for each $k \in \mathbb{N}$, an index $\lambda_k \in I$ such that

$$B_k \subset A_{\lambda_k},$$

we obtain

$$X \subset A_{\lambda_1} \cup A_{\lambda_2} \cup \cdots \cup A_{\lambda_k} \cup \cdots ,$$

which is a countable subcover. □

Remark 4. If $E \subset X \subset \mathbb{R}^n$, then E may be open relative to X without being open in $\mathbb{R}^n$. The same holds for closed sets. Compactness, however, behaves better, as we now show.

Proposition 1.46 *Let $K \subset X \subset \mathbb{R}^n$. Then K is compact in $\mathbb{R}^n$ if and only if K is compact in X.*

Proof: Assume K is compact in $\mathbb{R}^n$, and let $\{G_\alpha\}_{\alpha\in A}$ be an open cover of K in X, that is,

$$K \subset \bigcup_{\alpha\in A} G_\alpha,$$

where each G_α is open in X. For each α there exists an open set $H_\alpha \subset \mathbb{R}^n$ such that

$$G_\alpha = H_\alpha \cap X.$$

Thus

$$K \subset \bigcup_{\alpha\in A}(H_\alpha \cap X) = \Big(\bigcup_{\alpha\in A} H_\alpha\Big) \cap X,$$

and hence $K \subset \bigcup_{\alpha\in A} H_\alpha$.

Since K is compact in $\mathbb{R}^n$ and $\{H_\alpha\}_{\alpha\in A}$ is an open cover of K, there exist $\alpha_1, \dots, \alpha_k \in A$ such that

$$K \subset \bigcup_{i=1}^{k} H_{\alpha_i}.$$

As $K \subset X$, we have

$$K \subset \bigcup_{i=1}^{k}(H_{\alpha_i} \cap X) = \bigcup_{i=1}^{k} G_{\alpha_i},$$

showing that K is compact in X.

Conversely, suppose K is compact in X, and let $\{G_\alpha\}_{\alpha\in A}$ be an open cover of K in $\mathbb{R}^n$. Then

$$K \subset \bigcup_{\alpha\in A} G_\alpha,$$

and hence

$$K \cap X \subset \bigcup_{\alpha\in A}(G_\alpha \cap X).$$

Since each G_α is open in $\mathbb{R}^n$, it follows that $H_\alpha := G_\alpha \cap X$ is open in X, and

$$K \cap X \subset \bigcup_{\alpha\in A} H_\alpha.$$

But $K \subset X$, so $K \cap X = K$, and thus

$$K \subset \bigcup_{\alpha \in A} H_\alpha.$$

Since K is compact in X, there exist $\alpha_1, \ldots, \alpha_k \in A$ such that

$$K \subset \bigcup_{i=1}^{k} H_{\alpha_i} = \bigcup_{i=1}^{k} (G_{\alpha_i} \cap X),$$

which implies

$$K \subset \bigcup_{i=1}^{k} G_{\alpha_i}.$$

Therefore K is compact in $\mathbb{R}^n$. □

1.10 Continuous Functions

Definition 1.47 *Let $f : X \subset \mathbb{R}^n \to \mathbb{R}^m$ be a map. We say that f is* <u>*continuous*</u> *at a point $x_0 \in X$ if, given $\epsilon > 0$, one can find $\delta > 0$ such that every point $x \in X$ whose distance to x_0 is less than δ is mapped by f to a point $f(x)$ whose distance to $f(x_0)$ is less than ϵ. In other words,*

$$\forall\, \epsilon > 0\ \exists\, \delta > 0 \textit{ such that } \forall\, x \in X,\ \|x - x_0\| < \delta\ \Rightarrow\ \|f(x) - f(x_0)\| < \epsilon.$$

In terms of balls, the continuity of f at x_0 can be expressed as follows: for every open ball B' centred at $f(x_0)$ in $\mathbb{R}^m$ there exists an open ball B centred at x_0 in $\mathbb{R}^n$ such that

$$f(B \cap X) \subset B'.$$

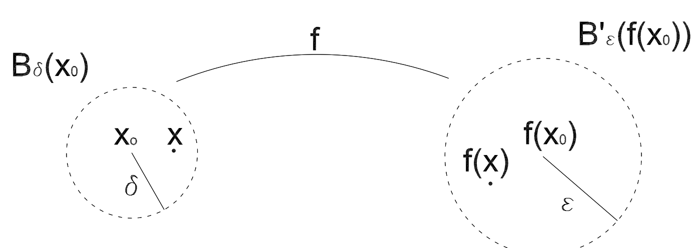

Figure 1.29:

<u>**Remark 1.**</u> Although the definition of continuity of a map $f : X \subset \mathbb{R}^n \to \mathbb{R}^m$ uses a norm in $\mathbb{R}^n$ and another in $\mathbb{R}^m$ (both denoted by the same symbol), it follows from the notion of equivalent norms that continuity of f at a point is preserved if we change one of these norms or both.

If $f : X \subset \mathbb{R}^n \to \mathbb{R}^m$ is continuous at every point of X, we simply say that f is a continuous map. It is easily seen that if f is continuous on X, then for every $Y \subset X$, the restriction $f|_Y$ is also continuous.

A trivial case of continuity is the following: if x_0 is an isolated point of X, then every map $f : X \subset \mathbb{R}^n \to \mathbb{R}^m$ is necessarily continuous at x_0. Indeed, there exists $\delta > 0$ such that $B_\delta(x_0) \cap X = \{x_0\}$. Thus, for any given $\epsilon > 0$, taking this value of δ we obtain

$$x \in X,\ \|x - x_0\| < \delta \ \Rightarrow\ x = x_0 \ \Rightarrow\ \|f(x) - f(x_0)\| = 0 < \epsilon.$$

Definition 1.48 *Given $X \subset \mathbb{R}^n$, a map $f : X \subset \mathbb{R}^n \to \mathbb{R}^m$ is said to be Lipschitz if there exists $k > 0$ (a Lipschitz constant) such that, for all $x, y \in X$,*

$$\|f(x) - f(y)\| \leq k\|x - y\|.$$

We shall prove that every Lipschitz map is continuous. Indeed, given $\epsilon > 0$, it suffices to take $\delta = \frac{\epsilon}{k}$. Then

$$\|x - x_0\| < \delta \ \Rightarrow\ \|f(x) - f(x_0)\| \leq k\|x - x_0\| < k \cdot \frac{\epsilon}{k} = \epsilon.$$

Note that the property of being Lipschitz does not depend on the particular choice of equivalent norms.

Examples of Lipschitz functions.

1) Every linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz.

Indeed, for $x \in \mathbb{R}^n$ we have

$$\|T(x)\| = \Big\|T\Big(\sum_i x_i e_i\Big)\Big\| = \Big\|\sum_i x_i\, T(e_i)\Big\| \leq \sum_i |x_i|\, \|T(e_i)\|.$$

Setting $k = \max\{\|T(e_1)\|, \ldots, \|T(e_n)\|\}$ we obtain

$$\|T(x)\| \leq k \sum_i |x_i|.$$

Taking in $\mathbb{R}^n$ the ℓ^1â€“norm (the 'sum norm'), we get $\|T(x)\| \leq k\|x\|$ for all $x \in \mathbb{R}^n$.

Thus, for arbitrary $x, y \in \mathbb{R}^n$, by linearity of T we obtain

$$\|T(x) - T(y)\| = \|T(x - y)\| \leq k\|x - y\|,$$

so T satisfies the Lipschitz condition and, in particular, is continuous.

2) The coordinate projections. For the iâ€“th projection we have

$$|\pi_i(x) - \pi_i(y)| = |x_i - y_i| \leq \|x - y\|,$$

for any of the three usual norms on $\mathbb{R}^n$.

3) The norm $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$, $x \mapsto \|x\|$.

Indeed, for any $x, y \in \mathbb{R}^n$ we have

$$\big|\|x\| - \|y\|\big| \leq \|x - y\|.$$

Proposition 1.49 *The composition of two continuous maps is continuous. More precisely, let $X \subset \mathbb{R}^n$, $Y \subset \mathbb{R}^m$, $f : X \to Y$ be continuous at $x_0 \in X$, and let $g : Y \to \mathbb{R}^p$ be continuous at $y_0 = f(x_0)$. Then $g \circ f : X \to \mathbb{R}^p$ is continuous at x_0.*

Proof: Given $\epsilon > 0$, by continuity of g there exists $\eta > 0$ such that

$$y \in Y,\ \|y - f(x_0)\| < \eta \ \Rightarrow\ \|g(y) - g(f(x_0))\| < \epsilon.$$

On the other hand, by continuity of f at x_0, corresponding to this η there exists $\delta > 0$ such that

$$x \in X,\ \|x - x_0\| < \delta \ \Rightarrow\ \|f(x) - f(x_0)\| < \eta.$$

Combining these two implications we obtain, for $x \in X$ with $\|x - x_0\| < \delta$,

$$\|g(f(x)) - g(f(x_0))\| < \epsilon.$$

Thus $g \circ f$ is continuous at x_0. □

Let $X \subset \mathbb{R}^n$. To give a map $f : X \to \mathbb{R}^m$ is the same as giving m functions $f_1, \ldots, f_m : X \to \mathbb{R}$ defined by $f_i = \pi_i \circ f$, which are called the coordinate functions of f. For every $x \in X$ we have

$$f(x) = (f_1(x), \ldots, f_m(x)).$$

Proposition 1.50 *A map $f : X \subset \mathbb{R}^n \to \mathbb{R}^m$ is continuous at a point $x_0 \in X$ if and only if each coordinate function $f_i = \pi_i \circ f : X \to \mathbb{R}$ is continuous at x_0.*

Proof: The continuity of f implies the continuity of each f_i by the previous proposition (using the projections π_i). Conversely, suppose each $f_i : X \to \mathbb{R}$ is continuous at $x_0 \in X$. Given $\epsilon > 0$, there exist $\delta_1, \ldots, \delta_m > 0$ such that

$$\|x - x_0\| < \delta_i,\ x \in X \ \Rightarrow\ |f_i(x) - f_i(x_0)| < \epsilon \quad \text{for } i = 1, \ldots, m.$$

In $\mathbb{R}^m$ take the maximum norm and set $\delta = \min\{\delta_1, \ldots, \delta_m\}$. Then

$$\|x - x_0\| < \delta,\ x \in X \ \Rightarrow\ \|f(x) - f(x_0)\| = \max_{1 \le i \le m} |f_i(x) - f_i(x_0)| < \epsilon,$$

and consequently f is continuous at x_0. □

Corollary 1.51 *Given $f : X \to \mathbb{R}^n$ and $g : X \to \mathbb{R}^m$, consider the map*

$$(f, g) : X \to \mathbb{R}^n \times \mathbb{R}^m, \qquad (f, g)(x) = (f(x), g(x)).$$

Then (f, g) is continuous if and only if both f and g are continuous.

Using the previous results it is easy to show that if $f, g : X \subset \mathbb{R}^n \to \mathbb{R}^m$ and $\alpha : X \subset \mathbb{R}^n \to \mathbb{R}$, then

$$\begin{array}{rl} \text{(i)} & f+g : X \to \mathbb{R}^m, \quad (f+g)(x) = f(x) + g(x); \\ \text{(ii)} & \alpha f : X \to \mathbb{R}^m, \quad (\alpha f)(x) = \alpha(x)\, f(x); \\ \text{(iii)} & \langle f, g\rangle : X \to \mathbb{R}, \quad \langle f, g\rangle(x) = \langle f(x), g(x)\rangle; \\ \text{(iv)} & \dfrac{1}{\alpha} : X \to \mathbb{R}, \quad \dfrac{1}{\alpha}(x) = \dfrac{1}{\alpha(x)} \quad (\alpha(x) \neq 0), \end{array}$$

are all continuous.

Proposition 1.52 *A map $f : X \subset \mathbb{R}^n \to \mathbb{R}^m$ is continuous at a point $x_0 \in X$ if and only if for every sequence $(x_k) \subset X$ with $\lim_{k\to\infty} x_k = x_0$ we have*

$$\lim_{k\to\infty} f(x_k) = f(x_0).$$

Proof: Assume f is continuous at x_0 and let $(x_k) \subset X$ with $\lim_{k\to\infty} x_k = x_0$. Given $\epsilon > 0$, there exists $\delta > 0$ such that

$$x \in X,\ \|x - x_0\| < \delta \ \Rightarrow\ \|f(x) - f(x_0)\| < \epsilon.$$

Since $\lim_{k\to\infty} x_k = x_0$, there exists $k_0 \in \mathbb{N}$ such that for all $k > k_0$,

$$\|x_k - x_0\| < \delta,$$

which implies $\|f(x_k) - f(x_0)\| < \epsilon$. Hence $\lim_{k\to\infty} f(x_k) = f(x_0)$.

For the converse, suppose, by contradiction, that f is not continuous at x_0. Then there exists $\epsilon_0 > 0$ such that for every $k \in \mathbb{N}$ we can find $x_k \in X$ with

$$\|x_k - x_0\| < \frac{1}{k} \quad \text{and} \quad \|f(x_k) - f(x_0)\| \geq \epsilon_0.$$

Then $\lim_{k\to\infty} x_k = x_0$, but $\lim_{k\to\infty} f(x_k) \neq f(x_0)$, which is a contradiction. □

Definition 1.53 *A map $f : X \subset \mathbb{R}^n \to \mathbb{R}^m$ is said to be uniformly continuous if, for every $\epsilon > 0$, one can find $\delta > 0$ such that*

$$x, y \in X,\ \|x - y\| < \delta \ \Rightarrow\ \|f(x) - f(y)\| < \epsilon.$$

For example, every Lipschitz map $f : X \subset \mathbb{R}^n \to \mathbb{R}^m$ is uniformly continuous. Indeed, if

$$\|f(x) - f(y)\| \leq k\, \|x - y\| \quad \text{for all } x, y \in X,$$

then, given $\epsilon > 0$, taking $\delta = \frac{\epsilon}{k}$ we obtain, for $\|x - y\| < \delta$,

$$\|f(x) - f(y)\| \leq k\, \|x - y\| \leq k \cdot \frac{\epsilon}{k} = \epsilon.$$

The composition $g \circ f$ of uniformly continuous functions f and g is again uniformly continuous. Hence a map $f : X \subset \mathbb{R}^n \to \mathbb{R}^m$ is uniformly continuous if and only if each of its coordinate functions $f_1, \ldots, f_m : X \to \mathbb{R}$ is uniformly continuous.

1.11 Connected Sets

Definition 1.54 *Two subsets A and B of $\mathbb{R}^n$ are said to be separated if*

$$\overline{A} \cap B = A \cap \overline{B} = \emptyset.$$

Example 1. $A = [1, 2)$ and $B = (2, 3]$ are separated.

Remark 1. Note that if A and B are separated then they are disjoint. However, two sets can be disjoint without being separated.

For example: $A = [1, 2)$ and $B = [2, 3)$ are disjoint but not separated.

There is, however, an important case in which two disjoint sets A and B are separated: when A and B are open. Indeed:

(i) $\overline{A} \cap B = \emptyset$.

Suppose, by contradiction, that there exists $x \in \overline{A} \cap B$. Since B is open, there exists $r > 0$ such that $B_r(x) \subset B$. As $x \in \overline{A}$ and $A \neq \emptyset$, there exists $y \in B_r(x)$ with $y \in A$. Because $B_r(x) \subset B$, we have $y \in B$, hence $y \in A \cap B$, which contradicts $A \cap B = \emptyset$.

(ii) $A \cap \overline{B} = \emptyset$ is proved in an analogous manner.

Definition 1.55 *A set $E \subset \mathbb{R}^n$ is said to be connected if, for every pair of separated sets $A, B \subset \mathbb{R}^n$ whose union is E, one of them is empty.*

In other words:

E is connected if and only if for every pair of separated sets $A, B \subset \mathbb{R}^n$ such that $E = A \cup B$ we have $A = \emptyset$ or $B = \emptyset$.

It follows that:

A set E is nonconnected or disconnected if and only if there exist non-empty separated sets $A, B \subset \mathbb{R}^n$ such that $E = A \cup B$.

Proposition 1.56 *A set $E \subset \mathbb{R}$ is connected if and only if, for any pair of points $x, y \in E$ and any $z \in \mathbb{R}$ with $x < z < y$, we have $z \in E$.*

Proof: First suppose that E is connected and consider $x, y \in E$ and $z \in \mathbb{R}$ with $x < z < y$. Assume, by contradiction, that $z \notin E$. Set

$$A = (-\infty, z) \cap E, \qquad B = (z, +\infty) \cap E.$$

We have:

(i) $A \neq \emptyset$, because $x \in E$ and $x \in (-\infty, z)$ since $x < z$.

$B \neq \emptyset$, because $y \in E$ and $y \in (z, +\infty)$ since $z < y$.

(ii) $E = A \cup B$.

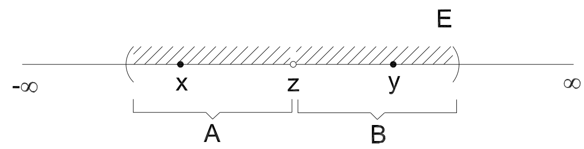

Figure 1.30:

Indeed, clearly $A \cup B \subset E$. On the other hand,

$$A \cup B = \big[(-\infty, z) \cup (z, +\infty)\big] \cap E = \big(\mathbb{R} \setminus \{z\}\big) \cap E.$$

Since $z \notin E$, we have $E \subset \mathbb{R} \setminus \{z\}$, and hence

$$E \subset \big(\mathbb{R} \setminus \{z\}\big) \cap E = A \cup B.$$

(iii) A and B are separated.

Indeed,

$$\begin{aligned} \overline{A} \cap B &= \overline{[(-\infty, z) \cap E]} \cap [(z, +\infty) \cap E] \\ &\subset \big(\overline{(-\infty, z)} \cap \overline{E}\big) \cap \big((z, +\infty) \cap E\big) \\ &= (-\infty, z] \cap (z, +\infty) \cap E = \emptyset. \end{aligned}$$

Thus $\overline{A} \cap B = \emptyset$. Similarly, $A \cap \overline{B} = \emptyset$.

By (i), (ii) and (iii), A and B form a disconnection of E, which contradicts the connectedness of E.

Conversely, suppose that for any $x, y \in E$ and $z \in \mathbb{R}$ with $x < z < y$ we have $z \in E$. We must prove that E is connected. Suppose, on the contrary, that E is disconnected. Then there exist non-empty separated sets A and B such that $E = A \cup B$.

Choose $x \in A$ and $y \in B$. Clearly $x \neq y$, otherwise both $\overline{A} \cap B$ and $A \cap \overline{B}$ would be non-empty, contradicting the fact that A and B are separated. Without loss of generality, suppose $x < y$ and set

$$z = \sup\big([x, y] \cap A\big).$$

Note that z is adherent to $[x, y] \cap A$, hence

$$z \in \overline{[x, y] \cap A} \subset [x, y] \cap \overline{A}.$$

Thus $z \notin B$, since $\overline{A} \cap B = \emptyset$. Therefore $z \neq y$, because $y \in B$. Hence

$$z \in \overline{A} \quad \text{and} \quad x \leq z < y.$$

We distinguish two cases:

(1) $z \notin A$.

In this case $z \neq x$, since $x \in A$. Hence $x < z < y$. By our hypothesis, this implies $z \in E$, which contradicts the fact that $z \notin A$, $z \notin B$ and therefore $z \notin A \cup B = E$.

(2) $z \in A$.

As $z \in A$, we have $z \notin \overline{B}$, since $A \cap \overline{B} = \emptyset$. Thus there exists $\epsilon_0 > 0$ such that the neighbourhood $(z - \epsilon_0, z + \epsilon_0)$ contains no points of B.

Let

$$\delta = \min\{\epsilon_0,\ |z - x|,\ |y - z|\}.$$

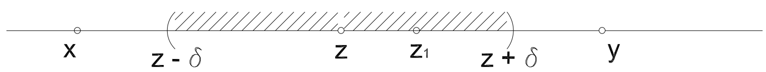

Figure 1.31:

Then $(z - \delta, z + \delta)$ contains no points of B and is contained in $[x, y] \subset E$. Choose $z_1 \in (z, z + \delta)$. Then $z < z_1 < y$. Since $z \in A$, $y \in B$ and $A \cap B = \emptyset$, by our hypothesis we must have $z_1 \in E$. However, as $z < z_1$ and z is the supremum of $[x, y] \cap A$, we have $z_1 \notin A$. Moreover, $z_1 \notin B$, because there are no points of B in $(z - \delta, z + \delta)$. Hence $z_1 \notin A \cup B = E$, again a contradiction. $\square$

Corollary 1.57 *A set $E \subset \mathbb{R}$ is connected if and only if it is one of the following:*

$$(-\infty, b),\ (-\infty, b],\ (a, +\infty),\ [a, +\infty),\ (-\infty, +\infty),\ (a, b),\ [a, b),\ (a, b],\ [a, b].$$

Thus we have characterised the connected subsets of $\mathbb{R}$: they are precisely the (possibly infinite) intervals. The next step would be to try to characterise connected subsets of $\mathbb{R}^n$. This is not possible in general. However, there is a class of sets for which this becomes possible, namely those in which the connected components are open.

Proposition 1.58 *The union of a family of connected sets having a common point is connected.*

Proof: Let $\{E_\alpha\}_{\alpha \in I}$ be a family of connected sets, all containing the same point $x \in \mathbb{R}^n$. To prove that $E = \bigcup_{\alpha \in I} E_\alpha$ is connected, let A and B be separated subsets of $\mathbb{R}^n$ such that $E = A \cup B$ with $x \in A$. For each $\alpha \in I$ we have

$$E_\alpha = E \cap E_\alpha = (A \cup B) \cap E_\alpha = (A \cap E_\alpha) \cup (B \cap E_\alpha).$$

If we show that $A \cap E_\alpha$ and $B \cap E_\alpha$ are separated, then, since E_α is connected, we must have $A \cap E_\alpha = \emptyset$ or $B \cap E_\alpha = \emptyset$. But $x \in A \cap E_\alpha$ for every α, so $A \cap E_\alpha \neq \emptyset$ and therefore $B \cap E_\alpha = \emptyset$ for all $\alpha \in I$. Hence

$$B = B \cap E = B \cap \Big(\bigcup_{\alpha \in I} E_\alpha\Big) = \bigcup_{\alpha \in I} (B \cap E_\alpha) = \emptyset,$$

which shows that E is connected.

It remains to prove that $A \cap E_\alpha$ and $B \cap E_\alpha$ are separated:

(i)

$$\overline{(A \cap E_\alpha)} \cap (B \cap E_\alpha) \subset (\overline{A} \cap B) \cap E_\alpha = \emptyset,$$

since A and B are separated.

(ii)

$$(A \cap E_\alpha) \cap \overline{(B \cap E_\alpha)} \subset (A \cap \overline{B}) \cap E_\alpha = \emptyset,$$

for the same reason. □

Corollary 1.59 *A set $E \subset \mathbb{R}^n$ is connected if and only if, for any $x, y \in E$, there exists a connected set C_{xy} such that $x, y \in C_{xy} \subset E$.*

Proof: Necessity is obvious: we can simply take $C_{xy} = E$.

For sufficiency, fix $x \in E$. By assumption, for each $y \in E$ there is a connected set C_{xy} with $x, y \in C_{xy} \subset E$. Then

$$E = \bigcup_{y \in E} C_{xy},$$

where the sets C_{xy} are connected and all contain the common point x. By the previous proposition, E is connected. □

Proposition 1.60 *Let $E \subset \mathbb{R}^n$ be a connected subset and $f : E \to \mathbb{R}^m$ a continuous map. Then $f(E)$ is connected.*

Proof: Suppose, by contradiction, that $f(E)$ is not connected. Then there exist non-empty separated sets $A, B \subset \mathbb{R}^m$ such that $f(E) = A \cup B$.

Set

$$G = E \cap f^{-1}(A) \quad \text{and} \quad H = E \cap f^{-1}(B).$$

Then:

(1°) $E = G \cup H$.

Indeed,

$$\begin{aligned} G \cup H &= (E \cap f^{-1}(A)) \cup (E \cap f^{-1}(B)) \\ &= E \cap \left(f^{-1}(A) \cup f^{-1}(B)\right) \\ &= E \cap f^{-1}(A \cup B). \end{aligned}$$

Since $f(E) = A \cup B$, we have $E \subset f^{-1}(A \cup B)$, and hence

$$G \cup H = E \cap f^{-1}(A \cup B) = E.$$

(2°) $G \neq \emptyset$ and $H \neq \emptyset$.

Indeed, since $A \neq \emptyset$, there exists $y \in A$. But $A \subset f(E)$, so $y = f(x)$ for some $x \in E$. Moreover, $x \in f^{-1}(A)$ because $f(x) = y \in A$. Thus $x \in E \cap f^{-1}(A) = G$, and therefore $G \neq \emptyset$. Similarly, using $B \neq \emptyset$, we obtain $H \neq \emptyset$.

(3°) G and H are separated.

Suppose, by contradiction, that there exists $x \in \mathbb{R}^n$ with $x \in \overline{G} \cap H$. Since $x \in \overline{G}$, there exists a sequence $(x_n) \subset G$ such that $x_n \to x$. As f is continuous, we have $f(x_n) \to f(x)$. On the other hand, $(x_n) \subset G = E \cap f^{-1}(A)$, so $(f(x_n)) \subset A$. Since $f(x_n) \to f(x)$, it follows that $f(x) \in \overline{A}$.

On the other hand, as $x \in H$, we have $x \in f^{-1}(B)$ and hence $f(x) \in B$. Thus

$$f(x) \in \overline{A} \cap B,$$

which is impossible, because A and B are separated. Similarly, one shows that $G \cap \overline{H} = \emptyset$.

From (1°), (2°) and (3°) we conclude that E is the union of two non-empty separated sets, which contradicts the connectedness of E. □

Definition 1.61 *A* path *in a set $X \subset \mathbb{R}^n$ is a continuous map*

$$\alpha : I \longrightarrow X,$$

where $I \subset \mathbb{R}$ is an interval.

For example, given $x, y \in \mathbb{R}^n$, the map

$$\lambda : [0,1] \longrightarrow \mathbb{R}^n, \quad t \longmapsto \lambda(t) = (1-t)x + ty,$$

is called the straight-line path joining x to y. Sometimes we refer to it as the path $[x, y]$.

We say that the points x and y can be joined by a path in X if there exists a path $\alpha : I \to X$ such that $x, y \in \alpha(I)$. In other words, when there is a continuous map $\alpha : I \to X$ such that $x = \alpha(t_x)$ and $y = \alpha(t_y)$ for some $t_x, t_y \in I$.

To fix ideas, consider Figure 1.32.

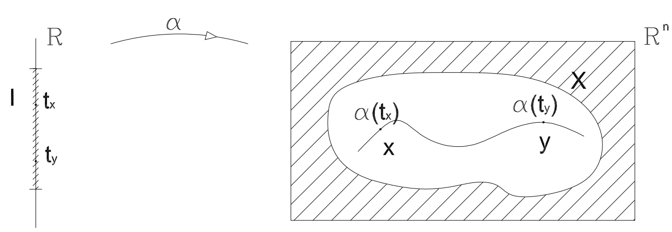

Figure 1.32:

Remark 2. If $x, y \in X$ can be joined by a path $\alpha : I \to X$, then there exists a path $\varphi : [0,1] \to X$ such that $\varphi(0) = x$ and $\varphi(1) = y$. It suffices to set

$$\varphi(s) = \alpha\big((1-s)t_x + st_y\big),$$

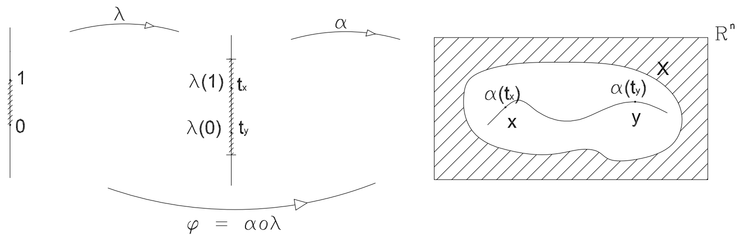

Figure 1.33:

where $x = \alpha(t_x)$ and $y = \alpha(t_y)$; see Figure 1.33.

If $\alpha, \beta : [0, 1] \to X$ are paths in X with $\alpha(1) = \beta(0)$, we define the concatenated path $\gamma = \alpha \vee \beta : [0, 1] \to X$ by

$$\gamma(t) = \begin{cases} \alpha(2t), & 0 \le t \le \frac{1}{2}, \\ \beta(2t-1), & \frac{1}{2} \le t \le 1. \end{cases}$$

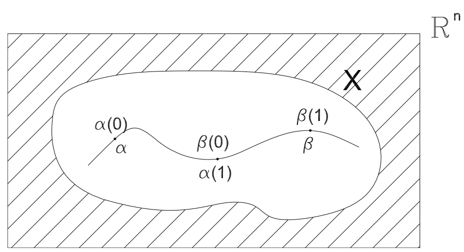

Figure 1.34:

Note that the two formulas above define the same value $\gamma\left(\frac{1}{2}\right)$. Since $\gamma|_{[0,\frac{1}{2}]}$ and $\gamma|_{[\frac{1}{2},1]}$ are continuous, it follows that γ is continuous. Intuitively, the path γ runs along the trajectory of α (with double speed) until $t = \frac{1}{2}$ and then, for $t \ge \frac{1}{2}$, it follows (with double speed) the trajectory of β, as indicated in Figure 1.34.

Let x, y, z be points of a set $X \subset \mathbb{R}^n$. If x and y can be joined by a path in X, and y and z can be joined by a path in X, then there exists a path in X joining x and z. Indeed, take paths $\alpha, \beta : [0, 1] \to X$ with $\alpha(0) = x$, $\alpha(1) = y$ and $\beta(0) = y$, $\beta(1) = z$, and set $\gamma = \alpha \vee \beta$. Then $\gamma(0) = x$ and $\gamma(1) = z$.

Definition 1.62 *A set $X \subset \mathbb{R}^n$ is said to be path-connected if any two points $x, y \in X$ can be joined by a path in X.*

Every path-connected set is connected, in view of the propositions and corollaries proved earlier, because if $\alpha : I \to X$ is a path in X joining the points x and y, then $\alpha(I) = C_{xy}$ is a connected subset of X containing x and y. Indeed, since the interval I

is a connected subset of $\mathbb{R}$ and $\alpha : I \to X$ is continuous, we have that $\alpha(I) = C_{xy}$ is a connected subset of X containing x and y.

The converse is false. The set $X \subset \mathbb{R}^2$ given by the union of the graph of the function $f(x) = \sin\left(\frac{1}{x}\right)$, $0 < x \leq 1$, with the origin $(0, 0)$ is connected but not path-connected; see Figure 1.35.

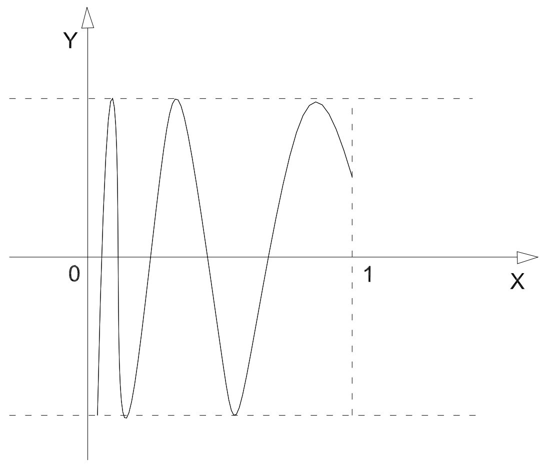

Figure 1.35:

There is, however, an important particular case in which connectedness implies path-connectedness: namely, when $X \subset \mathbb{R}^n$ is open. Before discussing this case, let us introduce an important definition.

Definition 1.63 *A set $X \subset \mathbb{R}^n$ is said to be convex if*

$$tx + (1-t)y \in X \quad \text{for all } x, y \in X \text{ and } 0 \leq t \leq 1.$$

In other words, a set $X \subset \mathbb{R}^n$ is convex when it contains every line segment whose endpoints belong to X.

To fix ideas, consider the schematic picture in Figures **??** and 1.37.

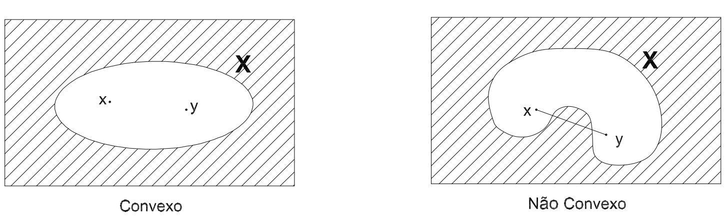

Figure 1.36:

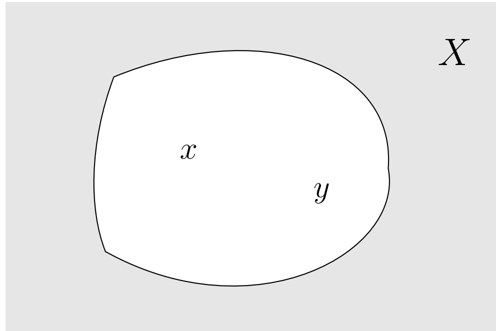

Figure 1.37: Convex set inside a region X.

The canonical example of a convex set is the ball $B_r(x_0)$. Indeed, let $x, y \in B_r(x_0)$ and consider $0 \le t \le 1$. Then

$$\begin{aligned}\|tx + (1-t)y - x_0\| &= \|tx - tx_0 + tx_0 + y - ty - x_0\| \\ &= \|t(x - x_0) + t(x_0 - y) + (y - x_0)\| \\ &= \|t(x - x_0) - t(y - x_0) + (y - x_0)\| \\ &= \|t(x - x_0) + (1-t)(y - x_0)\| \\ &\le t\|x - x_0\| + (1-t)\|y - x_0\| \\ &< tr + (1-t)r = r.\end{aligned}$$

An analogous argument applies to closed balls. If $X \subset \mathbb{R}^n$ is convex, any two points $x, y \in X$ can be joined by a path in X, namely the straight-line path $[x, y]$. Thus every convex set $X \subset \mathbb{R}^n$ is path-connected and therefore connected. In particular, every (open or closed) ball in $\mathbb{R}^n$ is path-connected.

Proposition 1.64 *An open set $E \subset \mathbb{R}^n$ is connected if and only if it is path-connected.*

Proof: The sufficiency has already been proved. We now prove the necessity.

Fix a point $x_0 \in E$ and let A be the set of all points $x \in E$ which can be joined to x_0 by a path in E. We claim that A is open. Indeed, let $x \in A$. Since E is open, there exists $r > 0$ such that $B_r(x) \subset E$. As the ball is convex, every point $y \in B_r(x)$ can be joined to x by a path in E. Hence y can be joined to x_0 by a path in E, which implies that $B_r(x) \subset A$. Consequently, A is open.

The set $B = E \setminus A$, that is, the set of all points $x \in E$ which cannot be joined to x_0 by a path in E, is also open. Indeed, take $x \in B$. Since E is open, there exists $\epsilon > 0$ such that $B_\epsilon(x) \subset E$. We claim that this ball is contained in B, i.e., that every point $z \in B_\epsilon(x)$ cannot be joined to x_0 by a path in E. Suppose, on the contrary, that there exists $z_0 \in B_\epsilon(x)$ which can be joined to x_0 by a path in E. By the convexity of the ball, the segment $[z_0, x]$ is contained in $B_\epsilon(x)$ and, therefore, in E. Concatenating a path from x_0 to z_0 with the segment $[z_0, x]$, we obtain a path in E joining x_0 to x, which would imply $x \in A$, a contradiction.

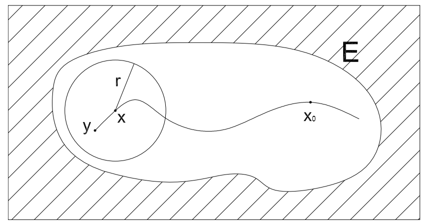

Figure 1.38:

Thus $E = A \cup B$, where A and B are disjoint open sets, hence separated. Since E is connected, one of the sets A or B must be empty. As $x_0 \in A$, it follows that $B = \emptyset$, and consequently $E = A$. This proves the proposition. □

Definition 1.65 *We say that $\alpha : [0,1] \to X$ is a polygonal path in X when α is the concatenation of a finite number of straight-line paths.*

Corollary 1.66 *If $E \subset \mathbb{R}^n$ is open and connected, then any two points of E can be joined by a polygonal path contained in E.*

1.12 Strong Relations between Continuity, Compactness and Connectedness

Proposition 1.67 *Let $K \subset \mathbb{R}^n$ be a compact set and $f : K \to \mathbb{R}^m$ a continuous function. Then $f(K)$ is compact.*

Proof: Let $\{G_\alpha\}_{\alpha \in A}$ be an open cover of $f(K)$. We must exhibit a finite subcover. Indeed, since $f(K) \subset \bigcup_{\alpha\in A} G_\alpha$, for each $y \in f(K)$ we have $y \in G_\alpha$ for some $\alpha \in A$, and moreover $y = f(x)$ for some $x \in K$. As $G_{\alpha(x)}$ is open, for each $y = f(x) \in f(K)$ there exists $\varepsilon_x > 0$ such that $B_{\varepsilon_x}(f(x)) \subset G_{\alpha(x)}$.

On the other hand, by the continuity of f, for each $\varepsilon_x > 0$ there exists $\delta_x > 0$ such that $f(B_{\delta_x}(x)) \subset B_{\varepsilon_x}(f(x))$. The family $\{B_{\delta_x}(x)\}_{x\in K}$ is an open cover of K, and since K is compact there exist $x_1, \dots, x_k \in K$ and $\delta_1, \dots, \delta_k > 0$ such that

$$K \subset \bigcup_{i=1}^{k} B_{\delta_i}(x_i),$$

and consequently

$$f(K) \subset f\Big(\bigcup_{i=1}^{k} B_{\delta_i}(x_i)\Big) \subset \bigcup_{i=1}^{k} f(B_{\delta_i}(x_i)) \subset \bigcup_{i=1}^{k} B_{\varepsilon_i}(f(x_i)) \subset \bigcup_{i=1}^{k} G_{\alpha(x_i)}.$$

Thus $f(K)$ admits a finite subcover. □

Proposition 1.68 *Let $K \subset \mathbb{R}^n$ be a compact set and $f : K \to \mathbb{R}$ a continuous function. Then f attains its absolute maximum and minimum on K.*

Proof: Since K is compact and f is continuous on K, the set $f(K)$ is a compact subset of $\mathbb{R}$ and therefore it is closed and bounded. Because it is bounded, by the completeness axiom there exist

$$M = \sup\{f(x) : x \in K\}, \qquad m = \inf\{f(x) : x \in K\}.$$

Since $f(K)$ is closed, it contains all its adherent points; hence $M, m \in f(K)$. It follows that there exist $x_1, x_2 \in K$ such that $f(x_1) = m$ and $f(x_2) = M$, as desired. □

Proposition 1.69 *Let $K \subset \mathbb{R}^n$ be a compact set and $f : K \to \mathbb{R}^m$ a continuous function. Then f is uniformly continuous.*

Proof: Let $\varepsilon > 0$ be given. Since f is continuous, for this $\varepsilon > 0$ and for each $x \in K$ there exists $\delta_x > 0$ such that, if $y \in K$ and $\|y - x\| < \delta_x$, then

$$\|f(x) - f(y)\| < \frac{\varepsilon}{2}.$$

Note that the family $\{B_{\delta_x/2}(x)\}_{x \in K}$ is an open cover of K. As K is compact, there exist $x_1, \ldots, x_k \in K$ and $\delta_1, \ldots, \delta_k > 0$ such that

$$K \subset \bigcup_{i=1}^{k} B_{\delta_i/2}(x_i).$$

Set

$$\delta = \min\left\{\frac{\delta_1}{2}, \ldots, \frac{\delta_k}{2}\right\},$$

and let $x, y \in K$ with $\|x - y\| < \delta$. We must show that $\|f(x) - f(y)\| < \varepsilon$.

Indeed, since $x \in K$, there exists $i_0 \in \{1, \ldots, k\}$ such that $x \in B_{\delta_{i_0}/2}(x_{i_0})$. Then

(i) $x \in B_{\delta_{i_0}}(x_{i_0})$,

(ii)

$$\|y - x_{i_0}\| \leq \|y - x\| + \|x - x_{i_0}\| < \delta + \frac{\delta_{i_0}}{2} \leq \frac{\delta_{i_0}}{2} + \frac{\delta_{i_0}}{2} = \delta_{i_0},$$

so $y \in B_{\delta_{i_0}}(x_{i_0})$ as well.

By continuity of f at x_{i_0}, we have

$$\|f(x) - f(x_{i_0})\| < \frac{\varepsilon}{2} \quad \text{and} \quad \|f(y) - f(x_{i_0})\| < \frac{\varepsilon}{2}.$$

Hence

$$\|f(x) - f(y)\| \leq \|f(x) - f(x_{i_0})\| + \|f(x_{i_0}) - f(y)\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

which proves the proposition. □

Proposition 1.70 (Intermediate Value Theorem) *Let $f : E \subset \mathbb{R}^n \to \mathbb{R}$ be a continuous function defined on a connected set $E \subset \mathbb{R}^n$. If there exist $x, y \in E$ and $c \in \mathbb{R}$ such that $f(x) < c < f(y)$, then there exists $z \in E$ such that $f(z) = c$.*

Proof: The set E is connected and f is continuous, so $f(E)$ is a connected subset of $\mathbb{R}$. By the first proposition about connected subsets of $\mathbb{R}$, since $f(x), f(y) \in f(E)$ and $f(x) < c < f(y)$, we must have $c \in f(E)$. Thus $c = f(z)$ for some $z \in E$, as claimed. $\square$

Chapter 2

Differentiation in $\mathbb{R}^n$

2.1 The Norm of a Linear Transformation

Definition 2.1 *A map T from a vector space X into a vector space Y (both over the same field) is called a* <u>*linear transformation*</u> *if*

$$T(x_1 + x_2) = T(x_1) + T(x_2),$$
$$T(\alpha x_1) = \alpha\, T(x_1),$$

for all $x_1, x_2 \in X$ and all scalars $\alpha \in K$.

Since $T : X \to Y$ is linear, it is customary to write Tx instead of $T(x)$.

Definition 2.2 *Let X and Y be vector spaces over the same field K. The set of all linear transformations $T : X \to Y$, denoted by $\mathcal{L}(X, Y)$, is a vector space with the operations*

(i) $+ : \mathcal{L}(X,Y) \times \mathcal{L}(X,Y) \to \mathcal{L}(X,Y)$, $(T,S) \mapsto T + S$, where

$$(T + S)(x) = T(x) + S(x), \quad \forall x \in X;$$

(ii) $\cdot : K \times \mathcal{L}(X,Y) \to \mathcal{L}(X,Y)$, $(\alpha, T) \mapsto \alpha T$, where

$$(\alpha T)(x) = \alpha\, T(x), \quad \forall x \in X.$$

When $X = Y$, instead of $\mathcal{L}(X,X)$ we simply write $\mathcal{L}(X)$. If X, Y, Z are vector spaces (all over the same field K) and $T \in \mathcal{L}(X,Y)$ and $S \in \mathcal{L}(Y,Z)$, we define the product ST as the composition of T and S, that is,

$$(ST)(x) = (S \circ T)(x) = S(T(x)), \quad \forall x \in X.$$

Then $ST \in \mathcal{L}(X,Z)$.

Definition 2.3 *Given $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, we define a norm $\|A\|$ of A as the map*

$$\|\cdot\| : \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \to \mathbb{R}, \qquad A \longmapsto \|A\|,$$

where

$$\|A\| = \sup\{\|Ax\| : x \in \mathbb{R}^n,\ \|x\| \leq 1\}.$$

The following proposition shows that this map indeed defines a norm on $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$.

Proposition 2.4 *Let $A, B \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ and $\alpha \in \mathbb{R}$. Then:*

(a) *$\|A\| < +\infty$, and therefore $\|A\|$ is well-defined.*

(b)

(i) $\|A\| \geq 0$ *and* $\|A\| = 0 \iff A \equiv 0$;

(ii) $\|\alpha A\| = |\alpha|\,\|A\|$;

(iii) $\|A + B\| \leq \|A\| + \|B\|$.

Proof: **(a)** Let $\beta = \{e_1, \ldots, e_n\}$ be a basis of $\mathbb{R}^n$, and let $x \in \mathbb{R}^n$. We can write $x = \sum_{i=1}^n x_i e_i$. For $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ we have

$$A(x) = A\Big(\sum_{i=1}^n x_i e_i\Big) = \sum_{i=1}^n x_i A(e_i).$$

Hence

$$\|A(x)\| = \Big\|\sum_{i=1}^n x_i A(e_i)\Big\| \leq \sum_{i=1}^n \|x_i A(e_i)\| = \sum_{i=1}^n |x_i|\,\|A(e_i)\| \leq \|x\| \sum_{i=1}^n \|A(e_i)\|.$$

If $\|x\| \leq 1$, then $\|A(x)\| \leq c$, where $c > 0$ is a constant. Thus the set $\{\|Ax\| : x \in \mathbb{R}^n,\ \|x\| \leq 1\}$ is bounded above and, by the completeness axiom, its supremum exists:

$$\|A\| = \sup\{\|Ax\| : x \in \mathbb{R}^n,\ \|x\| \leq 1\} < +\infty.$$

(b)

(i) Clearly $\|A\| \geq 0$, since $\|Ax\| \geq 0$ for all $x \in \mathbb{R}^n$ with $\|x\| \leq 1$. Moreover, if $A = 0$, then $\|A\| = 0$. Conversely, if $\|A\| = 0$, then

$$0 \leq \|Ax\| \leq 0, \quad \forall x \in \mathbb{R}^n,\ \|x\| \leq 1,$$

which implies $Ax = 0$ for all such x. If $x \neq 0$, then

$$A\Big(\frac{x}{\|x\|}\Big) = 0 \quad \Longrightarrow \quad A(x) = 0.$$

Since $A(0) = 0$, it follows that $A(x) = 0$ for all $x \in \mathbb{R}^n$, that is, $A \equiv 0$.

(ii)

$$\begin{aligned}\|\alpha A\| &= \sup\{\|(\alpha A)x\| : x \in \mathbb{R}^n,\ \|x\| \le 1\}\\ &= \sup\{|\alpha|\,\|Ax\| : x \in \mathbb{R}^n,\ \|x\| \le 1\}\\ &= |\alpha| \sup\{\|Ax\| : x \in \mathbb{R}^n,\ \|x\| \le 1\}\\ &= |\alpha|\,\|A\|.\end{aligned}$$

(iii)

$$\|A + B\| = \sup\{\|Ax + Bx\| : x \in \mathbb{R}^n,\ \|x\| \le 1\}.$$

We shall show that

$$\sup\{\|Ax + Bx\| : \|x\| \le 1\} \le \sup\{\|Ax\| : \|x\| \le 1\} + \sup\{\|Bx\| : \|x\| \le 1\},$$

that is, that the right-hand side is an upper bound for $\{\|Ax + Bx\| : x \in \mathbb{R}^n,\ \|x\| \le 1\}$.

For any $x \in \mathbb{R}^n$ with $\|x\| \le 1$ we have

$$\|Ax\| \le \sup\{\|Ax\|_{\mathbb{R}^m} : x \in \mathbb{R}^n,\ \|x\|_{\mathbb{R}^n} \le 1\}$$

and

$$\|Bx\| \le \sup\{\|Bx\|_{\mathbb{R}^m} : x \in \mathbb{R}^n,\ \|x\|_{\mathbb{R}^n} \le 1\}.$$

Therefore

$$\|Ax + Bx\| \le \|Ax\| + \|Bx\| \le \|A\| + \|B\|,$$

and hence $\|A + B\| \le \|A\| + \|B\|$. □

Proposition 2.5 *(a) If $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, then*

$$\|Ax\| \le \|A\|\,\|x\|, \quad \forall x \in \mathbb{R}^n.$$

(b) If $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ and $B \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^k)$, then

$$\|BA\| \le \|B\|\,\|A\|.$$

Proof: (a) Let $x \in \mathbb{R}^n$ and $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$.

If $x = 0$, then $A(0) = 0$ and the inequality is trivial. If $x \ne 0$, then $y = \dfrac{x}{\|x\|}$ is a unit vector in $\mathbb{R}^n$ and hence $\|Ay\| \le \|A\|$. But

$$\|Ay\| = \left\|A\Big(\frac{x}{\|x\|}\Big)\right\| = \left\|A\Big(x \cdot \frac{1}{\|x\|}\Big)\right\| = \left|\frac{1}{\|x\|}\right|\,\|Ax\| = \frac{1}{\|x\|}\,\|Ax\|.$$

Therefore $\|Ax\| \le \|A\|\,\|x\|$.

(b) Let $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ and $B \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^k)$. Then, for all $x \in \mathbb{R}^n$, by part (a),

$$\|(BA)x\| = \|B(A(x))\| \leq \|B\| \, \|Ax\| \leq \|B\| \, \|A\| \, \|x\|.$$

We will show that $\|BA\| \leq \|B\| \, \|A\|$.

Since

$$\|BA\| = \sup\{\|(BA)x\| : x \in \mathbb{R}^n, \ \|x\| \leq 1\},$$

it suffices to show that $\|B\| \, \|A\|$ is an upper bound for the set $\{\|(BA)x\| : x \in \mathbb{R}^n, \ \|x\| \leq 1\}$. But from the inequality above, for any $x \in \mathbb{R}^n$ with $\|x\| \leq 1$ we have

$$\|B(A(x))\| \leq \|B\| \, \|A\|,$$

and the result follows. □

Note 1. The norm $\|A\|$ of a linear transformation $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ induces a distance or metric on $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, that is, an application

$$d : \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \times \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \to \mathbb{R}, \qquad (A, B) \mapsto d(A, B) = \|A - B\|,$$

which satisfies:

(i) $d(A, B) \geq 0$ and $d(A, B) = 0 \iff A = B$;

(ii) $d(A, B) = d(B, A)$;

(iii) $d(A, C) \leq d(A, B) + d(B, C)$ for all $A, B, C \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$.

Proposition 2.6 *Let Ω be the set of all invertible linear maps from $\mathbb{R}^n$ into $\mathbb{R}^n$, that is,*

$$\Omega = \{A \in \mathcal{L}(\mathbb{R}^n) : \exists \, A^{-1}\}.$$

(a) Ω is an open subset of $\mathcal{L}(\mathbb{R}^n)$. In other words: if $A \in \Omega$, then there exists $r = \dfrac{1}{\|A^{-1}\|} > 0$ such that, whenever $B \in \mathcal{L}(\mathbb{R}^n)$ and $\|B - A\| < r$, we have $B \in \Omega$.

(b) The map

$$\psi : \Omega \to \Omega, \qquad A \mapsto A^{-1},$$

is continuous on Ω.

Proof: (a) Let $A \in \Omega$ and let $B \in \mathcal{L}(\mathbb{R}^n)$ be such that

$$\|B - A\| < \frac{1}{\|A^{-1}\|}.$$

We shall prove that $B \in \Omega$. Since B is a linear map between spaces of the same (finite) dimension, it suffices to show that B is injective, for this will imply that B is surjective as well.

Set $\|B - A\| = \beta$ and $\|A^{-1}\| = \dfrac{1}{\alpha}$. Then, by hypothesis, $\beta < \alpha$, that is $(\alpha - \beta) > 0$.

For a generic $x \in \mathbb{R}^n$ consider the expression

$$(\alpha - \beta)\|x\| = \alpha\|x\| - \beta\|x\|.$$

Observe that

(i) $\|x\| = \|A^{-1}(A(x))\| \leq \|A^{-1}\|\,\|A(x)\| = \dfrac{1}{\alpha}\|A(x)\|$, hence

$$\alpha\|x\| \leq \|A(x)\|;$$

(ii) $\|B(x) - A(x)\| = \|(B - A)(x)\| \leq \|B - A\|\,\|x\| = \beta\|x\|$, so

$$-\beta\|x\| \leq -\|B(x) - A(x)\|.$$

Thus, from (i) and (ii),

$$(\alpha - \beta)\|x\| = \alpha\|x\| - \beta\|x\| \leq \|A(x)\| - \|B(x) - A(x)\|. \tag{2.2}$$

Moreover,

$$\|B(x) - A(x)\| = \|A(x) - B(x)\| \geq \|A(x)\| - \|B(x)\|,$$

which implies

$$-\|B(x) - A(x)\| = -\|A(x) - B(x)\| \leq -\|A(x)\| + \|B(x)\|. \tag{2.3}$$

Combining (2.2) and (2.3) we obtain

$$(\alpha - \beta)\|x\| \leq \|A(x)\| - \|A(x)\| + \|B(x)\| = \|B(x)\|.$$

Since $(\alpha - \beta) > 0$, we conclude that

$$0 \leq (\alpha - \beta)\|x\| \leq \|B(x)\|. \tag{2.4}$$

If $B(x) = 0$ (and hence $\|B(x)\| = 0$), then by (2.4) we must have $\|x\| = 0$ (because $\alpha - \beta > 0$), so $x = 0$. By linearity, $\ker(B) = \{0\}$, and therefore B is injective, as required.

(b) We now prove that

$$\psi : \Omega \to \Omega, \qquad A \mapsto A^{-1},$$

is continuous on Ω.

Let $\varepsilon > 0$ and $A \in \Omega$ be given. We must find $\delta > 0$ such that, if $B \in \Omega$ and $\|B - A\| < \delta$, then $\|B^{-1} - A^{-1}\| < \varepsilon$. Indeed, for $x \in \mathbb{R}^n$ and $A, B \in \Omega$ we have

$$\begin{aligned}(B^{-1}(A - B)A^{-1})(x) &= B^{-1}\big((A - B)(A^{-1}(x))\big)\\ &= B^{-1}(A(A^{-1}(x))) - B^{-1}(B(A^{-1}(x)))\\ &= B^{-1}(x) - A^{-1}(x).\end{aligned}$$

Thus

$$B^{-1}(x) - A^{-1}(x) = B^{-1}(A - B)\big(A^{-1}(x)\big),$$

and consequently

$$\|B^{-1} - A^{-1}\| \le \|B^{-1}\| \, \|A - B\| \, \|A^{-1}\|. \tag{2.5}$$

As before, set $\alpha = \dfrac{1}{\|A^{-1}\|}$ and $\beta = \|B - A\|$, and choose

$$\delta < \min\left\{\alpha, \ \frac{\alpha^2\varepsilon}{1+\alpha\varepsilon}\right\}.$$

Suppose now that $\|B - A\| < \delta$. Arguing as in (2.4) we obtain

$$(\alpha - \beta)\|x\| \le \|B(x)\|, \quad \forall x \in \mathbb{R}^n. \tag{2.6}$$

Since $B \in \Omega$, for each $x \in \mathbb{R}^n$ there exists a unique $y \in \mathbb{R}^n$, and conversely, such that $y = B(x)$ or $x = B^{-1}(y)$. From (2.6) we get

$$(\alpha - \beta)\|B^{-1}(y)\| \le \|B(B^{-1}(y))\| = \|y\|, \quad \forall y \in \mathbb{R}^n,$$

which implies

$$\|B^{-1}(y)\| \le \frac{\|y\|}{\alpha - \beta}, \quad \forall y \in \mathbb{R}^n.$$

Hence

$$\|B^{-1}\| = \sup\{\|B^{-1}(y)\| : y \in \mathbb{R}^n, \ \|y\| \le 1\} \le \frac{1}{\alpha - \beta}. \tag{2.7}$$

Therefore, from (2.5) and (2.7),

$$\|B^{-1} - A^{-1}\| \le \frac{1}{\alpha - \beta} \, \|A - B\| \, \frac{1}{\alpha}. \tag{2.8}$$

Since $\|A - B\| < \delta$, we have

$$-\|A - B\| > -\delta \quad \Longrightarrow \quad \alpha - \|A - B\| > \alpha - \delta > 0$$

(because $\delta < \alpha$), and hence

$$\frac{1}{\alpha - \|A - B\|} < \frac{1}{\alpha - \delta} \quad \Longrightarrow \quad \frac{1}{\alpha - \beta} < \frac{1}{\alpha - \delta}. \tag{2.9}$$

Moreover,

$$\delta < \frac{\alpha^2\varepsilon}{1+\alpha\varepsilon} \Longleftrightarrow \delta + \alpha\delta\varepsilon < \alpha^2\varepsilon \Longleftrightarrow \delta < (\alpha^2 - \alpha\delta)\varepsilon \Longleftrightarrow \delta < \alpha(\alpha - \delta)\varepsilon. \tag{2.10}$$

Combining (2.8), (2.9) and (2.10), we obtain

$$\|B^{-1} - A^{-1}\| < \frac{1}{\alpha - \delta} \, \delta \, \frac{1}{\alpha} < \frac{1}{\alpha - \delta} \, \alpha(\alpha - \delta)\varepsilon \, \frac{1}{\alpha} = \varepsilon.$$

Thus ψ is continuous at A, and since $A \in \Omega$ was arbitrary, ψ is continuous on Ω. □

2.2 Differentiability of a Map

Before defining what it means for a map to be differentiable as a function from $\mathbb{R}^n$ to $\mathbb{R}^m$, let us first consider the particular case where $n = m = 1$.

So let $f : (a, b) \to \mathbb{R}$ be a function differentiable at a point $x_0 \in (a, b)$. Then the limit

$$\lim_{h\to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

exists and is denoted, as usual, by $f'(x_0)$. Equivalently, we also have

$$\lim_{h\to 0} \frac{f(x_0 + h) - f(x_0) - f'(x_0)h}{h} = 0.$$

Now, setting

$$r(h) = f(x_0 + h) - f(x_0) - f'(x_0)h, \tag{2.11}$$

and viewing h as the variable near x_0, we obtain

$$\lim_{h\to 0} \frac{r(h)}{h} = 0.$$

Because of the relation $\lim_{h\to 0} \frac{r(h)}{h} = 0$, we say that the remainder $r(h)$ tends to zero faster than h. We also say that $r(h)$ is an infinitesimal (a function whose limit is zero) of order higher than 1, relative to h.

Conversely, given f, suppose that there exists a constant L such that we can write

$$f(x_0 + h) = f(x_0) + Lh + r(h) \quad \text{with} \quad \lim_{h\to 0} \frac{r(h)}{h} = 0. \tag{2.12}$$

In this case,

$$\frac{f(x_0 + h) - f(x_0)}{h} = L + \frac{r(h)}{h},$$

and therefore

$$\lim_{h\to 0} \frac{f(x_0 + h) - f(x_0)}{h} = L,$$

that is, the derivative of f exists at $x_0 \in (a, b)$ and is equal to the number L. Condition (2.12) is thus necessary and sufficient for the existence of the derivative $f'(x_0)$. Under these conditions, (2.11) and (2.12) are equivalent.

We can now interpret the existence of the derivative $f'(x_0)$, in a neighbourhood of x_0, as meaning that the function f can be expressed as an affine map T plus a remainder which is 'very small' in a precise sense. Indeed, to fix ideas, consider Figure 2.1.

The equation of the tangent line to the graph of f at the point $(x_0, f(x_0))$ is

$$T(x) = f'(x_0)(x - x_0) + f(x_0) = f'(x_0)x + \big(f(x_0) - f'(x_0)x_0\big).$$

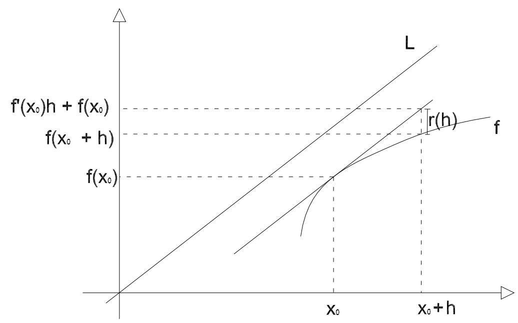

Figure 2.1:

In particular,

$$T(x_0 + h) = f'(x_0)(x_0 + h - x_0) + f(x_0) = f'(x_0)h + f(x_0). \tag{2.13}$$

On the other hand, from (2.11) we obtain

$$f(x_0 + h) = f(x_0) + f'(x_0)h + r(h). \tag{2.14}$$

Thus, from (2.13) and (2.14) we see that

$$f(x_0 + h) = T(x_0 + h) + r(h),$$

that is, near x_0, $f(x_0+h)$ is equal to an affine map $T(x_0+h)$ plus a 'very small' remainder, which becomes smaller as $|h|$ becomes smaller. Indeed,

$$|T(x_0 + h) - f(x_0 + h)| = |r(h)|,$$

and since $r(h) \to 0$, the points $T(x_0+h)$ and $f(x_0+h)$ become arbitrarily close as $h \to 0$. If we now consider the linear map $L(x) = f'(x_0)x$, then from (2.11) we can write

$$f(x_0 + h) - f(x_0) = L(h) + r(h)$$

with $\dfrac{r(h)}{h} \to 0$ as $h \to 0$.

In this way, we may regard the derivative of f at x_0 not as a real number, but rather as a linear map L which sends h to $f'(x_0)h$. Let us now generalise this new point of view on the derivative to maps $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$ with $m, n > 1$.

Definition 2.7 *Let E be a non-empty open subset of $\mathbb{R}^n$ and let $x_0 \in E$. Consider a map $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$. We say that f is differentiable at x_0 if there exists a linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ such that*

$$\lim_{h \to 0} \frac{\|f(x_0 + h) - f(x_0) - Lh\|}{\|h\|} = 0. \tag{2.15}$$

In the definition above, we allow $h \in \mathbb{R}^n$. Since E is open, we may take $\|h\|$ sufficiently small so that $(x_0 + h) \in E$. Thus $f(x_0 + h)$ is defined and the definition is meaningful. Note that

$$f(x_0 + h) - f(x_0) - Lh \in \mathbb{R}^m,$$

so the norm in the numerator of the expression above is taken in $\mathbb{R}^m$, while the norm in the denominator is taken in $\mathbb{R}^n$. However, from a topological point of view, it makes no difference which of the three usual norms (or, more generally, which equivalent norm) we use in each of these spaces, according to what we saw in the previous chapter.

It follows from the definition that a map $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $x_0 \in E$ if and only if there exists a linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$f(x_0 + h) - f(x_0) = Lh + r(h) \quad \text{where} \quad \lim_{h\to 0} \frac{\|r(h)\|}{\|h\|} = 0. \tag{2.16}$$

Indeed, if f is differentiable at $x_0 \in E$, there exists a linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ such that (2.15) holds. Setting

$$r(h) = f(x_0 + h) - f(x_0) - Lh,$$

we obtain

$$\frac{\|f(x_0 + h) - f(x_0) - Lh\|}{\|h\|} = \frac{\|r(h)\|}{\|h\|} \quad (h \neq 0).$$

Taking the limit as $h \to 0$ gives

$$\lim_{h\to 0} \frac{\|f(x_0 + h) - f(x_0) - Lh\|}{\|h\|} = 0.$$

Consequently,

$$\lim_{h\to 0} \frac{\|r(h)\|}{\|h\|} = \lim_{h\to 0} \frac{\|f(x_0 + h) - f(x_0) - Lh\|}{\|h\|} = 0.$$

Conversely, suppose that there exists a linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ such that (2.16) holds. Then

$$\lim_{h\to 0} \frac{\|f(x_0 + h) - f(x_0) - Lh\|}{\|h\|} = \lim_{h\to 0} \frac{\|r(h)\|}{\|h\|} = 0,$$

and hence

$$\lim_{h\to 0} \frac{\|f(x_0 + h) - f(x_0) - Lh\|}{\|h\|} = 0.$$

We can interpret (2.16) in the same way as in the real case, saying that, for small h, the left-hand side of the equality

$$f(x_0 + h) - f(x_0) = Lh + r(h)$$

is approximately equal to Lh, that is, to the value of a linear transformation applied to h.

Definition 2.8 *We say that a map $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$, where E is an open subset of $\mathbb{R}^n$, is differentiable on E if f is differentiable at every point of E.*

Proposition 2.9 *Let $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$ be a map differentiable at $x_0 \in E$, where E is an open subset of $\mathbb{R}^n$. Then the linear transformation $L : \mathbb{R}^n \to \mathbb{R}^m$ which provides a good approximation of the increment $f(x_0 + h) - f(x_0)$ in a neighbourhood of x_0 is unique.*

Proof: Suppose that there exist $L_1, L_2 \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ satisfying

$$\lim_{h\to 0} \frac{\|f(x_0 + h) - f(x_0) - L_1 h\|}{\|h\|} = \lim_{h\to 0} \frac{\|f(x_0 + h) - f(x_0) - L_2 h\|}{\|h\|} = 0.$$

Then

$$\begin{aligned} \|L_1 h - L_2 h\| &= \|L_1 h - f(x_0 + h) + f(x_0 + h) - f(x_0) + f(x_0) - L_2 h\| \\ &\leq \|L_1 h - f(x_0 + h) + f(x_0)\| + \|f(x_0 + h) - f(x_0) - L_2 h\| \\ &= \|f(x_0 + h) - f(x_0) - L_1 h\| + \|f(x_0 + h) - f(x_0) - L_2 h\|. \end{aligned}$$

Hence

$$0 \leq \frac{\|L_1 h - L_2 h\|}{\|h\|} \leq \frac{\|f(x_0 + h) - f(x_0) - L_1 h\|}{\|h\|} + \frac{\|f(x_0 + h) - f(x_0) - L_2 h\|}{\|h\|}.$$

Passing to the limit as $h \to 0$, we obtain

$$\lim_{h\to 0} \frac{\|L_1 h - L_2 h\|}{\|h\|} = 0. \tag{2.17}$$

Let $v \in \mathbb{R}^n$ be arbitrary but fixed, and consider $h = tv$, $t \in \mathbb{R}$. Then $h \to 0$ if and only if $t \to 0$, and from (2.17) we obtain

$$\lim_{t\to 0} \frac{\|(L_1 - L_2)(tv)\|}{\|tv\|} = 0.$$

Thus

$$\lim_{t\to 0} \left(\frac{|t|\, \|(L_1 - L_2)(v)\|}{\|tv\|} \right) = 0,$$

and consequently $(L_1 - L_2)(v) = 0$. Since v was arbitrary, $L_1 = L_2$. □

Note 1. If $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$, defined on the open set $E \subset \mathbb{R}^n$, is differentiable at $x_0 \in E$, then there exists a unique linear transformation $L : \mathbb{R}^n \to \mathbb{R}^m$ which provides a good approximation to the increment $f(x_0 + h) - f(x_0)$ in a neighbourhood of x_0. This linear transformation is called the derivative of f at x_0 and is denoted by $f'(x_0)$.

Therefore, if $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$, defined on an open set $E \subset \mathbb{R}^n$, is differentiable at $x_0 \in E$, its derivative is the linear map $f'(x_0) : \mathbb{R}^n \to \mathbb{R}^m$ characterised by

$$\lim_{h\to 0} \frac{\|f(x_0 + h) - f(x_0) - f'(x_0)h\|}{\|h\|} = 0, \tag{2.18}$$

or equivalently

$$f(x_0 + h) - f(x_0) = f'(x_0)h + r(h), \quad \text{with} \quad \lim_{h\to 0} \frac{\|r(h)\|}{\|h\|} = 0. \tag{2.19}$$

When $n = m = 1$, the linear transformation $f'(x_0) : \mathbb{R} \to \mathbb{R}$ coincides with the real number $f'(x_0)$, and for every $h \in \mathbb{R}$, $f'(x_0)h$ is simply the product of the number $f'(x_0)$ by the number h.

Example. If $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ and $x \in \mathbb{R}^n$, then $L'(x) = L$, that is, L is differentiable and its derivative is itself. Indeed, taking $L'(x) : \mathbb{R}^n \to \mathbb{R}^m$ to be L itself, then for every $x \in \mathbb{R}^n$, by linearity of L,

$$L(x + h) - L(x) = L(x + h - x) = L(h),$$

and consequently, from (2.18),

$$\lim_{h\to 0} \frac{\|L(x + h) - L(x) - L'(x)h\|}{\|h\|} = \lim_{h\to 0} \frac{\|Lh - Lh\|}{\|h\|} = 0.$$

When a map $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$ is differentiable on the open set $E \subset \mathbb{R}^n$, we can define the *derivative map*

$$f' : E \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \tag{2.20}$$

which associates to each point $x \in E$ the linear transformation $f'(x) : \mathbb{R}^n \to \mathbb{R}^m$, the derivative of f at that point.

Note 2. If $E \subset \mathbb{R}^n$ is open and $f : E \to \mathbb{R}^m$ is differentiable at $x_0 \in E$, then it follows from (2.19) that f is continuous at x_0. Indeed,

$$f(x_0 + h) - f(x_0) = f'(x_0)h + r(h),$$

where $\lim_{h\to 0} \frac{\|r(h)\|}{\|h\|} = 0$.

But:

(i)

$$\lim_{h\to 0} r(h) = \lim_{h\to 0} \left(\frac{r(h)}{\|h\|} \|h\| \right) = 0;$$

(ii)

$$\lim_{h\to 0} f'(x_0)h = 0,$$

since the linear map $f'(x_0) : \mathbb{R}^n \to \mathbb{R}^m$ is continuous at 0, and therefore $\lim_{h\to 0} f'(x_0)h = 0$.

Thus, from (i) and (ii),

$$\lim_{h\to 0} \big(f(x_0 + h) - f(x_0)\big) = 0 \quad \Longleftrightarrow \quad \lim_{h\to 0} f(x_0 + h) = f(x_0).$$

Setting $x = x_0 + h$, when $h \to 0$ we have $x \to x_0$, and therefore $\lim_{x\to x_0} f(x) = f(x_0)$, which implies that f is continuous at x_0.

Proposition 2.10 (Chain Rule) *Let $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$ and $g : F \subset \mathbb{R}^m \to \mathbb{R}^k$ be maps, where E and F are open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, and $f(E) \subset F$. If f is differentiable at x_0 and g is differentiable at $y_0 = f(x_0)$, then $H = g \circ f : E \to \mathbb{R}^k$ is differentiable at x_0 and, moreover,*

$$H'(x_0) = g'(y_0)\, f'(x_0).$$

Proof: We shall prove that there exists a linear map $C : \mathbb{R}^n \to \mathbb{R}^k$ such that

$$\lim_{h\to 0} \frac{\|H(x_0+h) - H(x_0) - Ch\|}{\|h\|} = 0. \tag{2.21}$$

Indeed, by hypothesis there exist linear maps $A : \mathbb{R}^n \to \mathbb{R}^m$ and $B : \mathbb{R}^m \to \mathbb{R}^k$ such that

$$\lim_{h\to 0} \frac{\|f(x_0+h) - f(x_0) - Ah\|}{\|h\|} = 0, \tag{2.22}$$

$$\lim_{k\to 0} \frac{\|g(y_0+k) - g(y_0) - Bk\|}{\|k\|} = 0. \tag{2.23}$$

Setting $C = BA$ (that is, $C = B \circ A$), C is clearly a linear map from $\mathbb{R}^n$ to $\mathbb{R}^k$, and we shall prove that C satisfies (2.21). In fact:

In (2.22) we are considering $h \in \mathbb{R}^n$ with $\|h\|$ sufficiently small so that $(x_0 + h) \in E$. This is always possible because E is open. Thus $f(x_0 + h) \in f(E) \subset F$. Setting

$$k = f(x_0 + h) - f(x_0),$$

and using the linearity of B, we obtain

$$\begin{aligned} H(x_0+h) - H(x_0) - (BA)(h) &= g(f(x_0+h)) - g(f(x_0)) - (BA)h \\ &= g(f(x_0)+k) - g(f(x_0)) - (BA)h - Bk + Bk \\ &= \big(g(y_0+k) - g(y_0) - Bk\big) + B\big(k - Ah\big) \\ &= \big(g(y_0+k) - g(y_0) - Bk\big) + B\big(f(x_0+h) - f(x_0) - Ah\big). \end{aligned}$$

Therefore

$$\|H(x_0+h)-H(x_0)-(BA)(h)\| \le \|g(y_0+k)-g(y_0)-Bk\|+\|B(f(x_0+h)-f(x_0)-Ah)\|.$$

Consequently,

$$\frac{\|H(x_0+h) - H(x_0) - (BA)(h)\|}{\|h\|} \le \frac{\|g(y_0+k) - g(y_0) - Bk\|}{\|h\|} + \frac{\|B(f(x_0+h) - f(x_0) - Ah)\|}{\|h\|}. \tag{2.24}$$

We claim that

(i) $\lim_{h\to 0} \dfrac{\|g(y_0+k)-g(y_0)-Bk\|}{\|h\|} = 0$;

(ii) $\lim_{h\to 0} \dfrac{\|B(f(x_0+h)-f(x_0)-Ah)\|}{\|h\|} = 0$.

From (i) and (ii) it will follow that (2.21) holds, as desired.

Proof of (i). From (2.22), taking $\varepsilon = 1$, there exists $\delta_1 > 0$ such that if $\|h\| < \delta_1$, then

$$\|f(x_0+h)-f(x_0)-Ah\| < \|h\|.$$

Thus

$$\|h\| \geq \|f(x_0+h)-f(x_0)\| - \|Ah\|,$$

and therefore

$$\|f(x_0+h)-f(x_0)\| \leq \|h\| + \|Ah\| \leq \|h\| + \|A\|\,\|h\| = c\|h\|,$$

where $c = 1 + \|A\|$. Since $k = f(x_0+h) - f(x_0)$, this last inequality implies that

$$\|k\| \leq c\|h\| \quad \text{whenever} \quad \|h\| < \delta_1. \tag{2.25}$$

On the other hand, from (2.23), given $\eta > 0$ there exists $\delta_2 > 0$ such that for every $k \in \mathbb{R}^m$ with $\|k\| < \delta_2$ we have

$$\|g(y_0+k)-g(y_0)-Bk\| < \frac{\eta}{c}\,\|k\|.$$

Set $\delta = \min\{\delta_1, \delta_2/c\}$. Then, if $\|h\| < \delta$, from (2.25) we obtain

$$\|k\| \leq c\|h\| < c\delta \leq c\frac{\delta_2}{c} = \delta_2.$$

Consequently,

$$\|g(y_0+k)-g(y_0)-Bk\| < \frac{\eta}{c}\,\|k\| < \frac{\eta}{c}\,c\|h\| = \eta\|h\|.$$

This proves (i).

Proof of (ii). We have

$$\frac{\|B(f(x_0+h)-f(x_0)-Ah)\|}{\|h\|} \leq \|B\|\,\frac{\|f(x_0+h)-f(x_0)-Ah\|}{\|h\|}.$$

From (2.22),

$$\lim_{h\to 0} \frac{\|f(x_0+h)-f(x_0)-Ah\|}{\|h\|} = 0,$$

and therefore

$$\lim_{h\to 0} \frac{\|B(f(x_0+h)-f(x_0)-Ah)\|}{\|h\|} = 0.$$

This proves (ii), and hence the proposition. □

2.3 Partial Derivatives

Definition 2.11 *Let $E \subset \mathbb{R}^n$ be open and $f : E \to \mathbb{R}^m$ a map. Let $\{e_1, \ldots, e_n\}$ and $\{u_1, \ldots, u_m\}$ be the canonical bases of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. The components of f, as we know, are the functions $f_1, \ldots, f_m : E \to \mathbb{R}$ defined by $f_i(x) = \pi_i(f(x))$, where π_i is the i-th coordinate projection, so that*

$$f(x) = \sum_{i=1}^{m} f_i(x)\, u_i. \tag{2.25}$$

For each $x \in E$, $1 \le i \le m$ and $1 \le j \le n$, we define

$$\frac{\partial f_i}{\partial x_j}(x) = \lim_{t \to 0} \frac{f_i(x + te_j) - f_i(x)}{t},$$

whenever this limit exists. This quantity is called the partial derivative of the function f_i in the direction e_j.

Given $x \in E$, the image of the path $\lambda : \mathbb{R} \to \mathbb{R}^n$ defined by $\lambda(t) = x + te_j$ is what we call â€œthe line that passes through the point x and is parallel to the j-th axis. Since E is open, there exists $\varepsilon > 0$ such that, if $-\varepsilon < t < \varepsilon$, then $\lambda(t) = x + te_j \in E$. We can then say that the partial derivative of f_i in the direction e_j is the derivative, at $t = 0$, of the map $f_i \circ \lambda : (-\varepsilon, \varepsilon) \to \mathbb{R}$.

Indeed, to fix ideas, consider the figure:

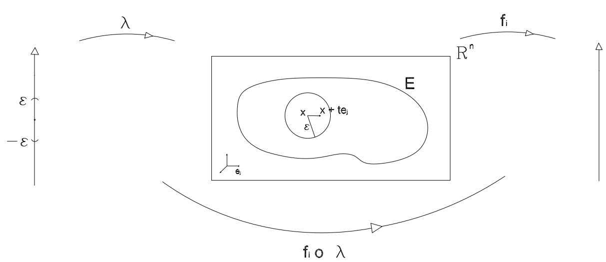

Figure 2.2:

We have

$$(f_i \circ \lambda)'(0) = \lim_{t \to 0} \frac{(f_i \circ \lambda)(t) - (f_i \circ \lambda)(0)}{t} = \lim_{t \to 0} \frac{f_i(x + te_j) - f_i(x)}{t} = \frac{\partial f_i}{\partial x_j}(x).$$

Thus we may say that f_i, when restricted to the open line segment $(x - \varepsilon e_j, x + \varepsilon e_j)$, becomes a real-valued function, namely $f_i(x + te_j)$ of the real variable t, and $\dfrac{\partial f_i}{\partial x_j}(x)$ is the derivative of this function at $t = 0$.

Example. When $n = 2$ and $m = 1$, that is, when $f : E \subset \mathbb{R}^2 \to \mathbb{R}$, the graph of f is a surface in $\mathbb{R}^3$. The restriction of f to the line segment that passes through $(x_0, y_0) \in E$ and is parallel to the x-axis has as its graph the plane curve obtained on this surface by keeping y constant and equal to y_0. Thus $\dfrac{\partial f}{\partial x}(x_0, y_0)$ is the slope of the tangent line to this curve at the point $(x_0, y_0, f(x_0, y_0))$, as illustrated in Figure 2.3.

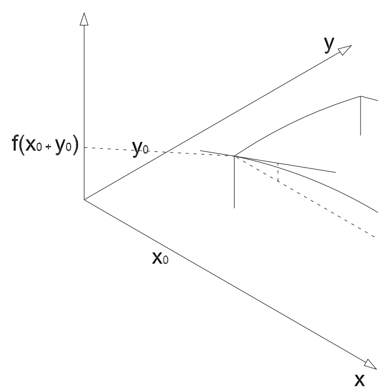

Figure 2.3:

Writing $f_i(x_1, \ldots, x_n)$ instead of $f_i(x)$, we see that the practical computation of the j-th partial derivative of f_i is carried out by treating all variables as constants except the j-th, and then applying the usual rules of differentiation.

The existence of the partial derivatives $\dfrac{\partial f_i}{\partial x_j}(x)$ does not imply the differentiability of f at x, as can be seen in an exercise at the end of the chapter. However, it is known that differentiability at a point x implies the existence of the partial derivatives at x, and that these determine the linear transformation $f'(x)$ completely, as we shall now see in the following result.

Proposition 2.12 *Let $E \subset \mathbb{R}^n$ be open and $f : E \to \mathbb{R}^m$ be differentiable at $x_0 \in E$. Then the partial derivatives $\dfrac{\partial f_i}{\partial x_j}(x_0)$ exist and, moreover,*

$$f'(x_0)e_j = \sum_{i=1}^{m} \frac{\partial f_i}{\partial x_j}(x_0)\, u_i, \quad 1 \le j \le n, \tag{2.26}$$

where $\{e_1, \ldots, e_n\}$ and $\{u_1, \ldots, u_m\}$ are the canonical bases of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively.

Proof: Since f is differentiable at $x_0 \in E$, there exists a linear transformation $f'(x_0) : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$f(x_0 + h) - f(x_0) = f'(x_0)h + r(h), \quad \text{where} \quad \lim_{h \to 0} \frac{\|r(h)\|}{\|h\|} = 0.$$

For t sufficiently small and for each e_j we can write

$$f(x_0 + te_j) - f(x_0) = f'(x_0)(te_j) + r(te_j),$$

where $\lim_{t\to 0} \frac{\|r(te_j)\|}{\|te_j\|} = 0$. By linearity of $f'(x_0)$,

$$\frac{f(x_0 + te_j) - f(x_0)}{t} = f'(x_0)e_j + \frac{r(te_j)}{t}. \tag{2.27}$$

However,

$$\lim_{t\to 0} \frac{r(te_j)}{t} = \lim_{t\to 0} \left(\frac{r(te_j)}{\|te_j\|} \frac{\|te_j\|}{t} \right) = 0,$$

and therefore, from (2.27),

$$\lim_{t\to 0} \frac{f(x_0 + te_j) - f(x_0)}{t} = f'(x_0)e_j. \tag{2.28}$$

If we write f in terms of its components, as in (2.25), we obtain

$$f(x_0 + te_j) - f(x_0) = \sum_{i=1}^{m} \big(f_i(x_0 + te_j) - f_i(x_0)\big) u_i. \tag{2.29}$$

Thus, from (2.28) and (2.29),

$$\lim_{t\to 0} \sum_{i=1}^{m} \frac{f_i(x_0 + te_j) - f_i(x_0)}{t}\, u_i = f'(x_0)e_j. \tag{2.30}$$

On the other hand, by the Chain Rule, each $f_i = \pi_i \circ f$ is differentiable as a composition of differentiable maps. Thus we can apply to each f_i the same reasoning used for f, and obtain, as in (2.28),

$$\lim_{t\to 0} \frac{f_i(x_0 + te_j) - f_i(x_0)}{t} = f_i'(x_0)e_j,$$

where $f_i'(x_0) : \mathbb{R}^n \to \mathbb{R}$ is the linear map associated to f_i. Hence

$$\frac{\partial f_i}{\partial x_j}(x_0) = f_i'(x_0)e_j.$$

It then follows from (2.30) that

$$\sum_{i=1}^{m} \frac{\partial f_i}{\partial x_j}(x_0)\, u_i = f'(x_0)e_j, \quad \forall j,\ 1 \le j \le n,$$

which proves the claim. □

As consequences, we have

$$\begin{aligned}
f'(x_0)(e_1) &= \frac{\partial f_1}{\partial x_1}(x_0)u_1 + \frac{\partial f_2}{\partial x_1}(x_0)u_2 + \cdots + \frac{\partial f_m}{\partial x_1}(x_0)u_m,\\
f'(x_0)(e_2) &= \frac{\partial f_1}{\partial x_2}(x_0)u_1 + \frac{\partial f_2}{\partial x_2}(x_0)u_2 + \cdots + \frac{\partial f_m}{\partial x_2}(x_0)u_m,\\
&\vdots\\
f'(x_0)(e_n) &= \frac{\partial f_1}{\partial x_n}(x_0)u_1 + \frac{\partial f_2}{\partial x_n}(x_0)u_2 + \cdots + \frac{\partial f_m}{\partial x_n}(x_0)u_m.
\end{aligned}$$

The matrix associated to the linear transformation $f'(x_0) : \mathbb{R}^n \to \mathbb{R}^m$, obtained by transposing the matrix of coefficients of this system, is called the *Jacobian matrix* and is denoted by $Jf(x_0)$ or simply $f'(x_0)$. It is given by

$$Jf(x_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \frac{\partial f_1}{\partial x_2}(x_0) & \cdots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \frac{\partial f_2}{\partial x_1}(x_0) & \frac{\partial f_2}{\partial x_2}(x_0) & \cdots & \frac{\partial f_2}{\partial x_n}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x_0) & \frac{\partial f_m}{\partial x_2}(x_0) & \cdots & \frac{\partial f_m}{\partial x_n}(x_0) \end{pmatrix}.$$

Note. It is worth observing that if

$$Jf(a) = \left(\frac{\partial f_i}{\partial x_j}(a)\right), \quad Jg(f(a)) = \left(\frac{\partial g_i}{\partial x_j}(f(a))\right), \quad J(g\circ f)(a) = \left(\frac{\partial (g_i \circ f)}{\partial x_j}(a)\right)$$

are the Jacobian matrices of the maps f, g and $g\circ f$ at the indicated points, then, assuming that f is differentiable at a and g is differentiable at $f(a)$, it follows from the Chain Rule that

$$J(g\circ f)(a) = Jg(f(a))\, Jf(a).$$

2.4 Directional Derivatives

Definition 2.13 *Let $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$, where E is an open subset of $\mathbb{R}^n$. The directional derivative of f at a point $x_0 \in E$ in the direction of a vector $v \in \mathbb{R}^n$ is, by definition,*

$$\frac{\partial f}{\partial v}(x_0) = \lim_{t\to 0} \frac{f(x_0 + tv) - f(x_0)}{t},$$

whenever this limit exists.

We can interpret $\dfrac{\partial f}{\partial v}(x_0)$ as follows: since E is open, there exists $\varepsilon > 0$ such that the line segment $(x_0 - \varepsilon v, x_0 + \varepsilon v)$ is contained in E. The straightâ€“line path $\lambda : (-\varepsilon, \varepsilon) \to E$ defined by $\lambda(t) = x_0 + tv$ is mapped by f onto the path $f \circ \lambda : (-\varepsilon, \varepsilon) \to \mathbb{R}^m$ that associates to each t the point $f(x_0 + tv)$ in $\mathbb{R}^m$. The directional derivative $\dfrac{\partial f}{\partial v}(x_0)$ is the velocity vector $(f \circ \lambda)'(0)$. Indeed,

$$(f \circ \lambda)'(0) = \lim_{t \to 0} \frac{(f \circ \lambda)(t) - (f \circ \lambda)(0)}{t} = \lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0)}{t} = \frac{\partial f}{\partial v}(x_0).$$

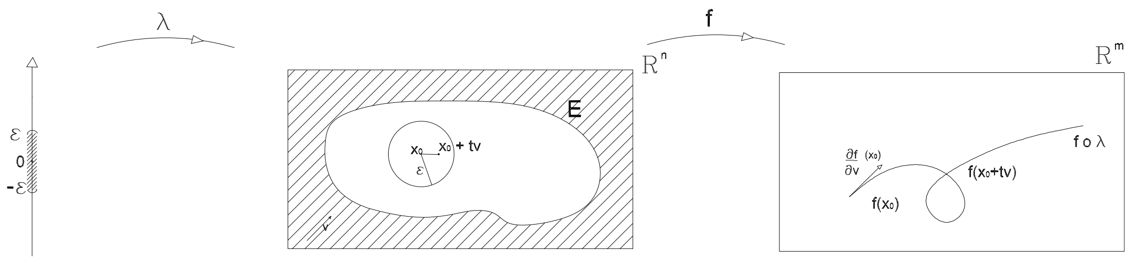

Figure 2.4:

If $f = (f_1, \ldots, f_m)$ then

$$\frac{\partial f}{\partial v}(x_0) = \Big(\frac{\partial f_1}{\partial v}(x_0), \ldots, \frac{\partial f_m}{\partial v}(x_0) \Big).$$

Indeed, suppose $\dfrac{\partial f}{\partial v}(x_0) = y \in \mathbb{R}^m$ and write $y = (y_1, \ldots, y_m)$. Let $\{u_1, \ldots, u_m\}$ be the canonical basis of $\mathbb{R}^m$. Then

$$\lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0)}{t} = \lim_{t \to 0} \frac{\sum_{i=1}^m \big(f_i(x_0 + tv) - f_i(x_0)\big) u_i}{t} = \sum_{i=1}^m y_i u_i,$$

and therefore

$$\frac{\partial f_i}{\partial v}(x_0) = \lim_{t \to 0} \frac{f_i(x_0 + tv) - f_i(x_0)}{t} = y_i.$$

Note 1. If $v = e_j$ for some $j = 1, \ldots, n$, where $\{e_j\}_{1 \le j \le n}$ is the canonical basis of $\mathbb{R}^n$, then $\dfrac{\partial f}{\partial v}(x_0) = \dfrac{\partial f}{\partial x_j}(x_0)$. Thus partial derivatives are special cases of directional derivatives, when the vector v is one of the canonical basis vectors.

Now suppose $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at the point $x_0 \in E$. Then, for every $v \in \mathbb{R}^n$ and every $t \in \mathbb{R}$ sufficiently small, we obtain from (2.16)

$$f(x_0 + tv) - f(x_0) = L(tv) + r(tv), \quad \text{with} \quad \lim_{t \to 0} \frac{r(tv)}{\|tv\|} = 0.$$

Since $L(tv) = tL(v)$ and

$$\lim_{t\to 0} \frac{r(tv)}{t} = \lim_{t\to 0} \left[\frac{r(tv)}{\|tv\|}\frac{\|tv\|}{t}\right] = 0,$$

we obtain

$$\lim_{t\to 0} \frac{f(x_0+tv)-f(x_0)}{t} = L(v).$$

Consequently,

$$Lv = \frac{\partial f}{\partial v}(x_0) \quad \text{or} \quad f'(x_0)v = \frac{\partial f}{\partial v}(x_0). \tag{2.32}$$

The relation in (2.32) allows us to conclude that if a map $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at a point $x_0 \in E$, then all directional derivatives exist at x_0 and, moreover, they can be computed simply by evaluating $f'(x_0)v$. On the other hand, the existence of all directional derivatives does not imply the differentiability of f.

Since $v \in \mathbb{R}^n$, we can write $v = \sum_{j=1}^n \alpha_j e_j$. Thus, from (2.26),

$$\begin{aligned} f'(x_0)v &= f'(x_0)\Big(\sum_{j=1}^n \alpha_j e_j\Big) = \sum_{j=1}^n f'(x_0)(\alpha_j e_j) \\ &= \sum_{j=1}^n \alpha_j f'(x_0)e_j = \sum_{j=1}^n \alpha_j \Big(\sum_{i=1}^m \frac{\partial f_i}{\partial x_j}(x_0)u_i\Big) \\ &= \sum_{j=1}^n\sum_{i=1}^m \frac{\partial f_i}{\partial x_j}(x_0)\alpha_j u_i = \sum_{i=1}^m\Big(\sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(x_0)\alpha_j\Big)u_i. \end{aligned}$$

Thus, if f is differentiable at $x_0 \in E$, the derivative of f applied to a vector $v \in \mathbb{R}^n$ is given by

$$f'(x_0)v = \sum_{i=1}^m\Big(\sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(x_0)\alpha_j\Big)u_i, \quad \text{where } v = \sum_{j=1}^n \alpha_j e_j. \tag{2.33}$$

As a consequence, from (2.32) and (2.33) we obtain

$$\frac{\partial f}{\partial v}(x_0) = \sum_{i=1}^m\Big(\sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(x_0)\alpha_j\Big)u_i. \tag{2.34}$$

In fact, the relation in (2.33) or (2.34) is simply the matrix product of the Jacobian matrix in (2.31) with the vector v, that is,

$$f'(x_0)v = \frac{\partial f}{\partial v}(x_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \cdots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x_0) & \cdots & \frac{\partial f_m}{\partial x_n}(x_0) \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \displaystyle\sum_{j=1}^n \frac{\partial f_1}{\partial x_j}(x_0)\alpha_j \\ \vdots \\ \displaystyle\sum_{j=1}^n \frac{\partial f_m}{\partial x_j}(x_0)\alpha_j \end{pmatrix} = \begin{pmatrix} f_1'(x_0)v \\ \vdots \\ f_m'(x_0)v \end{pmatrix}. \tag{2.35}$$

The vector equality

$$f(x_0 + v) - f(x_0) = f'(x_0)v + r(v)$$

is equivalent to the m scalar equalities

$$f_i(x_0 + v) - f_i(x_0) = f_i'(x_0)v + r_i(v),$$

where $r(v) = (r_1(v), \dots, r_m(v))$, while the vector limit

$$\lim_{v \to 0} \frac{r(v)}{\|v\|} = 0$$

corresponds to the m scalar limits

$$\lim_{v \to 0} \frac{r_i(v)}{\|v\|} = 0.$$

This yields the following result.

Proposition 2.14 *A map $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at a point $x_0 \in E$ if and only if each coordinate function $f_1, \dots, f_m$ is differentiable at this point.*

Note 2. It follows from the Chain Rule that if $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $x_0 \in E$, then, in order to calculate the directional derivative

$$\frac{\partial f}{\partial v}(x_0) = (f \circ \lambda)'(0),$$

it is not necessary to take $\lambda(t) = x_0 + tv$. Instead of restricting ourselves to a straightâ€œline path, we may consider any path $\lambda : (-\varepsilon, \varepsilon) \to E$ differentiable at 0, with $\lambda(0) = x_0$ and $\lambda'(0) = v = (v_1, \dots, v_n)$, and we still have

$$\frac{\partial f}{\partial v}(x_0) = (f \circ \lambda)'(0) = f'(\lambda(0))\lambda'(0) = f'(x_0)v.$$

Thus we may regard the derivative $f'(x_0) : \mathbb{R}^n \to \mathbb{R}^m$ as the linear map which assigns to each tangent vector $v = \lambda'(0)$ (to any differentiable curve $\lambda : (-\varepsilon, \varepsilon) \to E$ such that $\lambda(0) = x_0$) the tangent vector

$$(f \circ \lambda)'(0) = \frac{\partial f}{\partial v}(x_0)$$

to the curve $f \circ \lambda$.

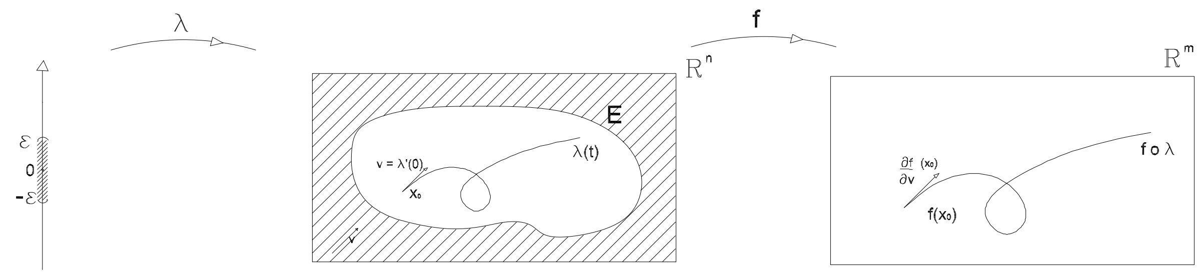

Figure 2.5:

2.5 Mean Value Inequalities

Proposition 2.15 (Mean Value Inequality for Vectorâ€“Valued Functions) *Let $f : [a, b] \to \mathbb{R}^m$ be a map continuous on $[a, b]$ and differentiable on (a, b). Then there exists $x \in (a, b)$ such that*

$$\|f(b) - f(a)\| \leq (b - a)\, \|f'(x)\|.$$

Proof: Set $z = f(b) - f(a) \in \mathbb{R}^m$ and define the map

$$\varphi : [a, b] \to \mathbb{R}, \qquad t \mapsto \varphi(t) = \langle z, f(t) \rangle.$$

Then φ is a realâ€“valued function, continuous on $[a, b]$ and differentiable on (a, b). Indeed, writing $f(t) = (f_1(t), \ldots, f_m(t))$ and $z = (z_1, \ldots, z_m)$, we have

$$\varphi(t) = \langle z, f(t) \rangle = \sum_{i=1}^{m} z_i f_i(t).$$

Since f is differentiable on (a, b), the functions $f_1, \ldots, f_m$ are differentiable on (a, b) and continuous on $[a, b]$. Thus φ satisfies the hypotheses of the Mean Value Theorem for real functions of one variable. Hence, there exists $x \in (a, b)$ such that

$$\varphi(b) - \varphi(a) = (b - a)\, \varphi'(x) = (b - a)\, \langle z, f'(x) \rangle. \tag{2.36}$$

On the other hand,

$$\varphi(b) - \varphi(a) = \langle z, f(b) \rangle - \langle z, f(a) \rangle = \langle z, z \rangle = \|z\|^2. \tag{2.37}$$

Thus, from (2.36) and (2.37) we obtain

$$\|z\|^2 = (b - a)\, \langle z, f'(x) \rangle.$$

By the Cauchy-Schwarz inequality,

$$\|z\|^2 \leq (b - a)\, \|z\|\, \|f'(x)\|,$$

and therefore

$$\|z\| \le (b-a)\,\|f'(x)\|.$$

(If $z = 0$, the inequality is trivial.) In other words,

$$\|f(b) - f(a)\| \le (b-a)\,\|f'(x)\|.$$

□

Proposition 2.16 (Mean Value Inequality) *Let $E \subset \mathbb{R}^n$ be an open convex set and let $f : E \to \mathbb{R}^m$ be a differentiable map on E. If there exists a constant $M \ge 0$ such that $\|f'(x)\| \le M$ for every $x \in E$, then*

$$\|f(b) - f(a)\| \le M\,\|b-a\| \quad \textit{for all } a, b \in E.$$

Proof: Let $a, b \in E$ be arbitrary. Consider the straightâ€“line path $\gamma : [0,1] \to E$ defined by $\gamma(t) = (1-t)a + tb$, which joins the points a and b. This path is well defined because E is convex.

Define the map

$$g : [0,1] \to \mathbb{R}^m, \qquad t \mapsto g(t) = (f \circ \gamma)(t),$$

that is, g is the restriction of f to the straightâ€“line path γ. Then, by the Chain Rule,

$$g'(t) = f'(\gamma(t))\,\gamma'(t) = f'(\gamma(t))(b-a), \quad \forall\, t \in [0,1].$$

Hence

$$\|g'(t)\| \le \|f'(\gamma(t))\|\,\|b-a\| \le M\,\|b-a\|, \quad \forall\, t \in [0,1]. \tag{2.38}$$

On the other hand, by the mean value inequality for vectorâ€“valued functions, there exists $t_0 \in (0,1)$ such that

$$\|g(1) - g(0)\| \le \|g'(t_0)\|. \tag{2.39}$$

Thus, from (2.38) and (2.39),

$$\|g(1) - g(0)\| \le M\,\|b-a\|.$$

Since $g(1) = f(\gamma(1)) = f(b)$ and $g(0) = f(\gamma(0)) = f(a)$, it follows that

$$\|f(b) - f(a)\| \le M\,\|b-a\|.$$

□

Corollary 2.17 *If $f : E \to \mathbb{R}^n$ is differentiable on the open convex set E and $f'(x) = 0$ for every $x \in E$, then f is constant.*

Proof: To prove this, simply note that the hypotheses of the previous proposition are now satisfied with $M = 0$. □

2.6 Continuously Differentiable Functions

Definition 2.18 *A differentiable function $f : E \to \mathbb{R}^m$, defined on an open set $E \subset \mathbb{R}^n$, is said to be continuously differentiable on E if its derivative*

$$f' : E \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m), \qquad x \mapsto f'(x),$$

is a continuous map. More precisely, for each $x \in E$ and for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\|f'(y) - f'(x)\| < \varepsilon$ whenever $y \in E$ and $\|x - y\| < \delta$.

Note. Recall that on $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ we are using the operator norm (supremum norm), while on E we use the norm induced from $\mathbb{R}^n$. When f satisfies the definition above, we also say that f is of class $C^1(E)$.

Proposition 2.19 *Let $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$, where E is an open subset of $\mathbb{R}^n$. Then*

$$f \text{ is of class } C^1(E) \iff f_i \text{ is of class } C^1(E) \text{ for } i = 1, \dots, m.$$

Proof: First suppose that f is of class $C^1(E)$. Then each f_i is differentiable. We shall prove that $f_i' : E \to \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ is continuous for every $i = 1, \dots, m$. Indeed, let $\varepsilon > 0$ be given and let $x_0 \in E$. By hypothesis, there exists $\delta > 0$ such that if $x \in E$ and $\|x - x_0\| < \delta$, then $\|f'(x) - f'(x_0)\| < \varepsilon$.

However, for each $i = 1, \dots, m$ we have

$$|f_i'(x)h - f_i'(x_0)h| \le \|f'(x)h - f'(x_0)h\|,$$

since $f'(x)h = (f_1'(x)h, \dots, f_m'(x)h)$. Consequently, for each $i = 1, \dots, m$,

$$\begin{aligned} \|f_i'(x) - f_i'(x_0)\| &= \sup\big\{|f_i'(x)h - f_i'(x_0)h|;\ h \in \mathbb{R}^n,\ \|h\| \le 1\big\} \\ &\le \sup\big\{\|f'(x)h - f'(x_0)h\|;\ h \in \mathbb{R}^n,\ \|h\| \le 1\big\} \\ &= \|f'(x) - f'(x_0)\|. \end{aligned}$$

Therefore $\|f_i'(x) - f_i'(x_0)\| < \varepsilon$ whenever $\|x - x_0\| < \delta$.

Conversely, suppose that each f_i, $i = 1, \dots, m$, is of class $C^1(E)$. Then f is differentiable. We shall show that the map $f' : E \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ is continuous.

Let $\varepsilon > 0$ and $x_0 \in E$. For each $i = 1, \dots, m$ there exists $\delta_i > 0$ such that if $x \in E$ and $\|x - x_0\| < \delta_i$, then

$$\|f_i'(x) - f_i'(x_0)\| < \varepsilon.$$

Set $\delta = \min\{\delta_1, \dots, \delta_m\}$. We shall show that $\|f'(x) - f'(x_0)\| < \varepsilon$ whenever $\|x - x_0\| < \delta$. In fact, since

$$\|f'(x) - f'(x_0)\| = \sup\big\{\|f'(x)h - f'(x_0)h\|;\ h \in \mathbb{R}^n,\ \|h\| \le 1\big\},$$

it suffices to prove that ε is an upper bound for the set

$$A = \big\{\|f'(x)h - f'(x_0)h\|;\ h \in \mathbb{R}^n,\ \|h\| \le 1\big\},$$

whenever $\|x - x_0\| < \delta$. Indeed,

$$\begin{aligned}\|f'(x)h - f'(x_0)h\| &= \|(f_1'(x)h - f_1'(x_0)h, \ldots, f_m'(x)h - f_m'(x_0)h)\| \\ &= \max\Big\{|f_1'(x)h - f_1'(x_0)h|, \ldots, |f_m'(x)h - f_m'(x_0)h|\Big\} \\ &< \varepsilon,\end{aligned}$$

whenever $\|x - x_0\| < \delta$, which proves the continuity of f'. □

Teorema 2.20 *Let $f : E \to \mathbb{R}^m$ be defined on an open set $E \subset \mathbb{R}^n$. Then f is of class $C^1(E)$ if and only if the partial derivatives $\dfrac{\partial f_i}{\partial x_j}$ exist and are continuous on E for $1 \leq i \leq m$ and $1 \leq j \leq n$.*

Proof: First suppose that $f \in C^1(E)$. Then f is differentiable, and hence $\dfrac{\partial f_i}{\partial x_j}(x)$ exist for all $x \in E$. Moreover,

$$f'(x)e_j = \sum_{i=1}^{m} \frac{\partial f_i}{\partial x_j}(x)\, u_i,$$

where $\{e_1, \ldots, e_n\}$ and $\{u_1, \ldots, u_m\}$ are the canonical bases of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Thus

$$\big\langle f'(x)e_j, u_k\big\rangle = \left\langle \sum_{i=1}^{m} \frac{\partial f_i}{\partial x_j}(x)\, u_i, u_k \right\rangle = \frac{\partial f_k}{\partial x_j}(x)\, \langle u_k, u_k\rangle = \frac{\partial f_k}{\partial x_j}(x), \quad \forall\, x \in E.$$

We now show that the partial derivatives are continuous on E. In fact, if $x, y \in E$, then

$$\begin{aligned}\frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(y) &= \big\langle f'(x)e_j, u_i\big\rangle - \big\langle f'(y)e_j, u_i\big\rangle \\ &= \big\langle f'(x)e_j - f'(y)e_j, u_i\big\rangle \\ &= \big\langle (f'(x) - f'(y))e_j, u_i\big\rangle. \end{aligned} \tag{2.40}$$

From (2.40), using the Cauchy-Schwarz inequality,

$$\begin{aligned}\big|\tfrac{\partial f_i}{\partial x_j}(x) - \tfrac{\partial f_i}{\partial x_j}(y)\big| &= \big|\big\langle (f'(x) - f'(y))e_j, u_i\big\rangle\big| \\ &\leq \|(f'(x) - f'(y))e_j\|\, \|u_i\| \\ &= \|(f'(x) - f'(y))e_j\| \\ &\leq \|f'(x) - f'(y)\|\, \|e_j\| \\ &= \|f'(x) - f'(y)\|.\end{aligned}$$

Therefore,

$$\big|\tfrac{\partial f_i}{\partial x_j}(x) - \tfrac{\partial f_i}{\partial x_j}(y)\big| \leq \|f'(x) - f'(y)\|, \quad \forall\, x, y \in E. \tag{2.41}$$

Since f is of class $C^1(E)$, given $\varepsilon > 0$ and $x \in E$ there exists $\delta > 0$ such that if $y \in E$ and $\|x - y\| < \delta$, then $\|f'(x) - f'(y)\| < \varepsilon$. It follows from (2.41) that

$$\left|\tfrac{\partial f_i}{\partial x_j}(x) - \tfrac{\partial f_i}{\partial x_j}(y)\right| < \varepsilon$$

whenever $\|x - y\| < \delta$, which proves the necessity.

To prove the sufficiency, suppose that the partial derivatives $\dfrac{\partial f_i}{\partial x_j}$ exist and are continuous on E. We shall prove that f is of class $C^1(E)$. Without loss of generality, by the previous proposition, it suffices to consider the case $m = 1$, that is, the case of a map $f : E \subset \mathbb{R}^n \to \mathbb{R}$.

Let $\varepsilon > 0$ and $x \in E$ be arbitrary. Since $\dfrac{\partial f}{\partial x_j}$ is continuous on E for each $j = 1, \ldots, n$, then, for the given $\varepsilon > 0$ and $x \in E$, there exists, for each $j = 1, \ldots, n$, a $\delta_j > 0$ such that if $y \in E$ and $\|y - x\| < \delta_j$, then

$$\left|\frac{\partial f}{\partial x_j}(y) - \frac{\partial f}{\partial x_j}(x)\right| < \frac{\varepsilon}{n}.$$

Set $\delta = \min\{\delta_1, \ldots, \delta_n\}$. Thus

$$\left|\frac{\partial f}{\partial x_j}(y) - \frac{\partial f}{\partial x_j}(x)\right| < \frac{\varepsilon}{n}, \quad \forall\, j = 1, \ldots, n, \tag{2.42}$$

whenever $\|y - x\| < \delta$.

We shall first prove that f is differentiable at x, by exhibiting a candidate $L_x : \mathbb{R}^n \to \mathbb{R}$ for the derivative of f at x. For this, we need some preliminary results.

Let $h \in \mathbb{R}^n$ with $\|h\| < \delta$. Suppose $h = (h_1, \ldots, h_n)$ and set $v_0 = 0$ and $v_k = h_1 e_1 + \cdots + h_k e_k$ for $1 \leq k \leq n$, that is,

$$v_0 = (0, \ldots, 0), \quad v_1 = (h_1, 0, \ldots, 0), \quad v_2 = (h_1, h_2, 0, \ldots, 0), \quad \ldots, \quad v_n = (h_1, \ldots, h_n) = h.$$

Note that $\|v_k\| < \delta$ for all $k = 0, 1, \ldots, n$. We claim that

$$[x + v_{k-1}, x + v_k] \subset B_\delta(x), \quad \forall\, k \geq 1. \tag{2.43}$$

Indeed,

$$[x + v_{k-1}, x + v_k] = \big\{(x + v_{k-1})(1 - t) + t(x + v_k);\ t \in [0, 1]\big\}.$$

Let $z \in [x + v_{k-1}, x + v_k]$. Then $z = (x + v_{k-1})(1 - t) + t(x + v_k)$ for some $t \in [0, 1]$, so

$$\begin{aligned}
\|z - x\| &= \|(x + v_{k-1})(1 - t) + t(x + v_k) - x\| \\
&= \|x - tx + v_{k-1}(1 - t) + tx + tv_k - x\| \\
&= \|(1 - t)v_{k-1} + tv_k\| \\
&\leq (1 - t)\|v_{k-1}\| + t\|v_k\| \\
&< (1 - t)\delta + t\delta = \delta,
\end{aligned}$$

which proves (2.43).

We also claim that

$$f(x+h) - f(x) = \sum_{k=1}^{n} \big[f(x+v_k) - f(x+v_{k-1}) \big]. \tag{2.44}$$

In fact,

$$\begin{aligned}\sum_{k=1}^{n} \big[f(x+v_k) - f(x+v_{k-1}) \big] &= f(x+v_1) - f(x+v_0) + f(x+v_2) - f(x+v_1) + \cdots \\ &\quad + f(x+v_{n-1}) - f(x+v_{n-2}) + f(x+v_n) - f(x+v_{n-1}) \\ &= f(x+v_n) - f(x+v_0) \\ &= f(x+h) - f(x),\end{aligned}$$

which proves (2.44).

Now define the auxiliary function

$$\delta_k : [0,1] \to \mathbb{R}, \qquad t \mapsto \delta_k(t) = f\big(x + v_{k-1} + t(v_k - v_{k-1})\big).$$

Notice that δ_k is the restriction of f to the straightâ€“line path passing through the point $x + v_{k-1}$ and parallel to the vector $v_k - v_{k-1} = h_k e_k$. We may also write

$$\delta_k(t) = f(x + v_{k-1} + t h_k e_k).$$

Therefore

$$\begin{aligned}\delta_k'(t) &= \lim_{s\to 0} \frac{\delta_k(t+s) - \delta_k(t)}{s} \\ &= \lim_{s\to 0} \frac{f\big(x + v_{k-1} + (t+s)h_k e_k\big) - f\big(x + v_{k-1} + t h_k e_k\big)}{s} \\ &= \lim_{s\to 0} \frac{f\big((x + v_{k-1} + t h_k e_k) + s h_k e_k\big) - f\big(x + v_{k-1} + t h_k e_k\big)}{s} \\ &= \frac{\partial f}{\partial (h_k e_k)}\big(x + v_{k-1} + t h_k e_k\big) \\ &= h_k \frac{\partial f}{\partial x_k}\big(x + v_{k-1} + t h_k e_k\big).\end{aligned}$$

By the Mean Value Theorem for real functions of one variable, there exists $t_k \in (0,1)$ such that

$$\delta_k(1) - \delta_k(0) = \delta_k'(t_k).$$

Equivalently,

$$f(x+v_k) - f(x+v_{k-1}) = \frac{\partial f}{\partial x_k}\big(x + v_{k-1} + t_k h_k e_k\big)\, h_k,$$

which implies

$$\sum_{k=1}^{n}\big[f(x+v_k)-f(x+v_{k-1})\big] = \sum_{k=1}^{n}\frac{\partial f}{\partial x_k}\big(x+v_{k-1}+t_kh_ke_k\big)\,h_k.$$

From (2.44) we obtain

$$f(x+h)-f(x) = \sum_{k=1}^{n}\frac{\partial f}{\partial x_k}\big(x+v_{k-1}+t_kh_ke_k\big)\,h_k. \tag{2.45}$$

Now consider the linear map $L_x:\mathbb{R}^n\to\mathbb{R}$ defined by

$$L_x(h)=\sum_{k=1}^{n}\frac{\partial f}{\partial x_k}(x)\,h_k.$$

From (2.45) we obtain

$$\begin{aligned}
|f(x+h)-f(x)-L_xh| &= \left|\sum_{k=1}^{n}\Big(\frac{\partial f}{\partial x_k}(x+v_{k-1}+t_kh_ke_k)-\frac{\partial f}{\partial x_k}(x)\Big)h_k\right| \\
&\le \sum_{k=1}^{n}\left|\frac{\partial f}{\partial x_k}(x+v_{k-1}+t_kh_ke_k)-\frac{\partial f}{\partial x_k}(x)\right||h_k| \\
&\le \|h\|\sum_{k=1}^{n}\left|\frac{\partial f}{\partial x_k}(x+v_{k-1}+t_kh_ke_k)-\frac{\partial f}{\partial x_k}(x)\right| \\
&= \|h\|\sum_{k=1}^{n}\left|\frac{\partial f}{\partial x_k}(x+v_{k-1}+t_k(v_k-v_{k-1}))-\frac{\partial f}{\partial x_k}(x)\right| \\
&= \|h\|\sum_{k=1}^{n}\left|\frac{\partial f}{\partial x_k}(x+v_{k-1}+t_k(x+v_k-(x+v_{k-1})))-\frac{\partial f}{\partial x_k}(x)\right|.
\end{aligned}$$

However,

$$y_k=(x+v_{k-1})+t_k(x+v_k-(x+v_{k-1}))\in[x+v_{k-1},x+v_k],$$

since $t_k\in(0,1)$, and $[x+v_{k-1},x+v_k]\subset B_\delta(x)$. Hence $\|y_k-x\|<\delta$. Therefore, from (2.42),

$$\begin{aligned}
|f(x+h)-f(x)-L_xh| &\le \|h\|\sum_{k=1}^{n}\big|\tfrac{\partial f}{\partial x_k}(y_k)-\tfrac{\partial f}{\partial x_k}(x)\big| \\
&< \|h\|\big(\tfrac{\varepsilon}{n}+\cdots+\tfrac{\varepsilon}{n}\big)=\|h\|\,\varepsilon,
\end{aligned}$$

which proves that f is differentiable at x.

It remains to prove that f' is continuous. Let $y\in E$ be such that $\|y-x\|<\delta$. We know that

$$\|f'(x)-f'(y)\|=\sup\big\{|f'(x)h-f'(y)h|;\ h\in\mathbb{R}^n,\ \|h\|\le 1\big\}.$$

To obtain the desired result, it suffices to show that the given $\varepsilon > 0$ is an upper bound for the set

$$\{|f'(x)h - f'(y)h|;\ \|h\| \le 1,\ h \in \mathbb{R}^n\}.$$

Indeed, let $h \in \mathbb{R}^n$ with $\|h\| \le 1$ and write $h = (h_1, \dots, h_n)$. Then, using (2.42),

$$\begin{aligned} |f'(x)h - f'(y)h| &= \left|\sum_{k=1}^{n}\Big(\frac{\partial f}{\partial x_k}(x)\,h_k - \frac{\partial f}{\partial x_k}(y)\,h_k\Big)\right| \\ &\le \sum_{k=1}^{n}\left|\frac{\partial f}{\partial x_k}(x) - \frac{\partial f}{\partial x_k}(y)\right|\,|h_k| \\ &\le \|h\|\sum_{k=1}^{n}\left|\frac{\partial f}{\partial x_k}(x) - \frac{\partial f}{\partial x_k}(y)\right| \\ &\le \frac{\varepsilon}{n} + \cdots + \frac{\varepsilon}{n} = \varepsilon, \end{aligned}$$

as required. □

2.7 Differentiable Maps from $\mathbb{R}^n$ into $\mathbb{R}$

In this section we shall consider specifically differentiable maps $f : E \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, where E is an open subset of $\mathbb{R}^n$.

If f is differentiable at the point $x_0 \in E$, then, in accordance with (2.35), given $v \in \mathbb{R}^n$, $v = (v_1, \dots, v_n)$, we have:

$$f'(x_0)v = \frac{\partial f}{\partial v}(x_0) = \left(\frac{\partial f}{\partial x_1}(x_0), \dots, \frac{\partial f}{\partial x_n}(x_0)\right)\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(x_0)\,v_i. \tag{2.1}$$

Definition 2.21 *Let $f : E \to \mathbb{R}$ be defined on the open set $E \subset \mathbb{R}^n$. We define the gradient of f at the point $x_0 \in E$ as the vector in $\mathbb{R}^n$ given by*

$$\nabla f(x_0) = \left(\frac{\partial f}{\partial x_1}(x_0), \dots, \frac{\partial f}{\partial x_n}(x_0)\right).$$

Thus, from (2.46) we may write

$$\begin{aligned} f'(x_0)v = \frac{\partial f}{\partial v}(x_0) = \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(x_0)\,v_i = \langle\nabla f(x_0), v\rangle = \\ = \lim_{t\to 0}\frac{f(x_0 + tv) - f(x_0)}{t} = (f\circ\lambda)'(0), \end{aligned} \tag{2.2}$$

where $\lambda : (-\varepsilon, \varepsilon) \longrightarrow E$ is any differentiable path such that $\lambda(0) = x_0$ and $\lambda'(0) = v$.

The derivative of a path is a vector. In the dual situation, the role of the derivative of a function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is played by a linear functional, namely $f'(x_0) : \mathbb{R}^n \longrightarrow \mathbb{R}$, which assigns to each vector $v \in \mathbb{R}^n$ the value $f'(x_0)v$ given as in (2.47).

Definition 2.22 *If $f : E \longrightarrow \mathbb{R}$, defined on the open set $E \subset \mathbb{R}^n$, is differentiable at the point $x_0 \in E$, the differential of f at x_0 is the linear functional*

$$df(x_0) : \mathbb{R}^n \longrightarrow \mathbb{R}$$

whose value on a vector $v = (v_1, \ldots, v_n)$ is given by

$$df(x_0)v = \frac{\partial f}{\partial v}(x_0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0)\, v_i = \langle \nabla f(x_0), v \rangle.$$

We call the *dual space* of $\mathbb{R}^n$ the space $(\mathbb{R}^n)^*$ consisting of all linear maps, that is, $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$. A basis for this space is given by the linear functionals $\pi_i : \mathbb{R}^n \longrightarrow \mathbb{R}$, where $\pi_i(x) = x_i$, the i-th projection of x. Indeed, note that if $\{e_1, \ldots, e_n\}$ is the canonical basis of $\mathbb{R}^n$, then

$$\pi_i(e_j) = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

Hence:

(i) If $\sum_{i=1}^n \lambda_i \pi_i = 0$ (where 0 denotes the identically zero functional), then for every $j = 1, \ldots, n$ we have

$$\Big(\sum_{i=1}^{n} \lambda_i \pi_i\Big)(e_j) = 0 \implies \sum_{i=1}^{n} \lambda_i \pi_i(e_j) = 0 \implies \lambda_j = 0.$$

Thus $\pi_1, \ldots, \pi_n$ are linearly independent.

(ii) Given $w \in (\mathbb{R}^n)^*$, for every $x = \sum_{i=1}^n x_i e_i \in \mathbb{R}^n$ we have

$$w(x) = w\Big(\sum_{i=1}^{n} x_i e_i\Big) = \sum_{i=1}^{n} x_i w(e_i) = \sum_{i=1}^{n} \lambda_i\, \pi_i(x) = \Big(\sum_{i=1}^{n} \lambda_i \pi_i\Big)(x),$$

where $\lambda_i = w(e_i)$. Therefore, there exist $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ such that $w = \sum_{i=1}^n \lambda_i \pi_i$. This proves that $\pi_1, \ldots, \pi_n$ span $(\mathbb{R}^n)^*$.

It is common to denote the canonical basis of $(\mathbb{R}^n)^*$ by $\{dx_1, \ldots, dx_n\}$ instead of $\{\pi_1, \ldots, \pi_n\}$. Thus

$$dx_i \cdot v = v_i, \quad \text{if } v = (v_1, \ldots, v_n).$$

The reason for this notation is the following: since, for each point $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, the i-th projection $\pi_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ takes the value $\pi_i(x) = x_i$, computing the differential of the i-th projection we obtain

$$d\pi_i(x_0)v = \sum_{j=1}^{n} \frac{\partial \pi_i}{\partial x_j}(x_0)\, v_j = v_i.$$

We write x_i instead of π_i. Hence $dx_i(x_0)v = v_i$.

Writing $dx_j v$ instead of v_j in the definition of the differential, we obtain

$$df(x_0)v = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x_0)\, dx_j v.$$

Since this equality holds for every $v \in \mathbb{R}^n$, we have

$$df(x_0) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x_0)\, dx_j.$$

This means that the linear functional $df(x_0)$ can be written as a linear combination of the functionals dx_i, with coefficients $\dfrac{\partial f}{\partial x_i}(x_0)$.

If we use only orthonormal bases in $\mathbb{R}^n$, the coordinates of the gradient vector $\nabla f(x_0)$ with respect to the basis $\{e_1, \ldots, e_n\}$ coincide with the coordinates of the functional $df(x_0)$ with respect to the dual basis $\{dx_1, \ldots, dx_n\}$. Under these conditions, the gradient becomes practically indistinguishable from the differential. This vector exhibits very convenient geometric features, providing information about the behaviour of the function, as we shall see below.

From now on we shall assume $\nabla f(x_0) \neq 0$. We highlight the three most important properties of the gradient, namely:

(1) The gradient points in a direction along which the function f is increasing.

(2) Among all directions along which the function f increases, the direction of the gradient is the one of fastest increase.

(3) The gradient of f at the point x_0 is perpendicular to the level surface of f passing through this point.

Indeed:

First, if $w = \nabla f(x_0)$, then, from (3.47) we have

$$\frac{\partial f}{\partial w}(x_0) = \langle \nabla f(x_0), w \rangle = \|\nabla f(x_0)\|^2 > 0.$$

This means that if $\lambda : (-\varepsilon, \varepsilon) \longrightarrow E$ is a differentiable path taking values in the domain E of f, such that $\lambda(0) = x_0$ and $\lambda'(0) = \nabla f(x_0)$, then the real function $t \longmapsto f(\lambda(t))$ has positive derivative at the point $t = 0$. If we assume that f and λ are of class C^1, then the derivative of $f \circ \lambda$ will remain positive at every point of some open interval centred at 0. That is, if we choose $\varepsilon > 0$ sufficiently small, then $f \circ \lambda : (-\varepsilon, \varepsilon) \longrightarrow \mathbb{R}$ is an increasing function. This means that f *increases in the direction of the gradient* (Figure 2.6).

Of course, we do not have $\dfrac{\partial f}{\partial v}(x_0) > 0$ only when $v = \nabla f(x_0)$. Since

$$\frac{\partial f}{\partial v}(x_0) = \langle \nabla f(x_0), v \rangle,$$

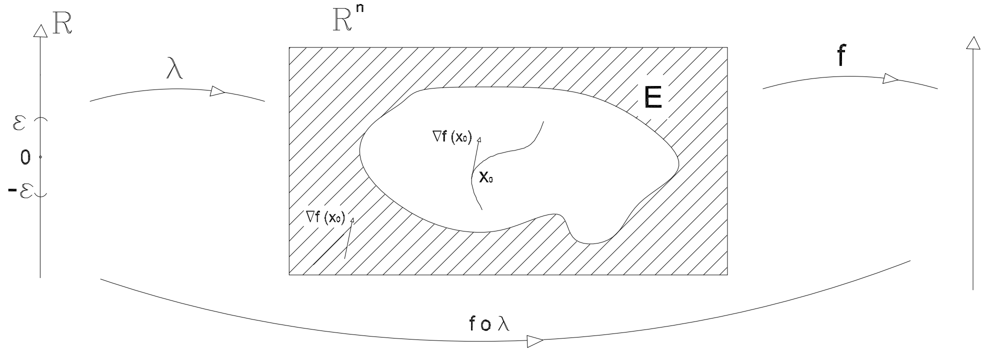

Figure 2.6:

the vectors v that point in directions along which the function f increases are precisely those that form an acute angle with $\nabla f(x_0)$, that is, those for which the inner product $\langle \nabla f(x_0), v\rangle$ is positive. What distinguishes the gradient is the fact that in its direction the growth of f is faster than in any other direction. Indeed, if $v \in \mathbb{R}^n$ is such that $\|v\| = \|\nabla f(x_0)\|$, then, by the Cauchy-Schwarz inequality, we have

$$\frac{\partial f}{\partial v}(x_0) = \langle \nabla f(x_0), v\rangle \leq \|\nabla f(x_0)\|\,\|v\| = \|\nabla f(x_0)\|^2 = \frac{\partial f}{\partial(\nabla f(x_0))}(x_0),$$

that is,

$$\frac{\partial f}{\partial v}(x_0) \leq \frac{\partial f}{\partial(\nabla f(x_0))}(x_0).$$

Finally, we clarify the third property. Given a real number c, we say that <u>$x \in E$ is at level c</u> with respect to f when $f(x) = c$. Once c is fixed, the set

$$f^{-1}(\{c\}) = \{x \in E;\ f(x) = c\}$$

is called the level surface c of the function f. In particular, when $n = 2$, $f^{-1}(\{c\})$ is called the <u>level curve c of f</u>. It is worth noting that the inverse image $f^{-1}(\{c\})$ does not always look like a curve or a surface (for example, if $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is a constant function equal to c). It would be more accurate to call $f^{-1}(\{c\})$ a 'level set'. However, the terminology is standard and is justified by the fact that $f^{-1}(\{c\})$ is indeed a surface (or a curve) whenever $\nabla f(x) \neq 0$ for all $x \in E$ with $f(x) = c$, as can be proved, as we shall see later, with the aid of the Implicit Function Theorem.

To say that a vector w is perpendicular to the level surface (or curve) $f^{-1}(\{c\})$ at the point x_0 means that w is perpendicular to the velocity vector, at x_0, of any differentiable path $\lambda : (-\varepsilon, \varepsilon) \longrightarrow \mathbb{R}^n$ at $t = 0$, with $\lambda(0) = x_0$ and $\lambda(t) \in f^{-1}(\{c\})$, that is, $f(\lambda(t)) = c$ for all $t \in (-\varepsilon, \varepsilon)$. From this last equality we obtain

$$0 = (f \circ \lambda)'(0) = f'(\lambda(0))\,\lambda'(0) = f'(x_0)\,\lambda'(0) = \langle \nabla f(x_0), \lambda'(0)\rangle.$$

Hence $\nabla f(x_0)$ is perpendicular to $\lambda'(0)$ (the velocity vector at the point $\lambda(0) = x_0$ of any differentiable path λ contained in the level surface of f through x_0).

Proposition 2.23 (Mean Value Theorem) *Let $f : E \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ be a map defined on an open set. If*

$$[x, y] = \{(1-t)x + ty;\ t \in [0,1]\} \subset E$$

and f is differentiable on $[x, y]$, then there exists $z \in [x, y]$ such that

$$f(y) - f(x) = f'(z)(y - x).$$

Proof: Consider the straightâ€“line path joining the points x and y:

$$\gamma : [0,1] \longrightarrow \mathbb{R}^n, \qquad t \longmapsto \gamma(t) = (1-t)x + ty.$$

Define

$$H : [0,1] \longrightarrow \mathbb{R}, \qquad t \longmapsto H(t) = (f \circ \gamma)(t).$$

By the Mean Value Theorem for real functions of one variable, there exists $t_0 \in (0,1)$ such that $H'(t_0) = H(1) - H(0)$. More precisely,

$$H'(t_0) = f(y) - f(x). \tag{2.3}$$

On the other hand, by the Chain Rule we have

$$H'(t_0) = f'(\gamma(t_0))\,\gamma'(t_0) = f\big((1-t_0)x + t_0 y\big)'(y - x) = f'((1-t_0)x + t_0 y)(y - x). \tag{2.4}$$

Setting $z = (1-t_0)x + t_0 y$, it is clear that $z \in [x, y]$, and from (2.48) and (2.49) we obtain

$$f(y) - f(x) = f'(z)(y - x).$$

Corollary 2.24 *Let $f : E \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ be a differentiable map defined on a convex open set E. Given $x, y \in E$, there exists $z \in [x, y]$ such that*

$$f(y) - f(x) = f'(z)(y - x).$$

Definition 2.25 *Let $f : E \longrightarrow \mathbb{R}$ be differentiable on the open set $E \subset \mathbb{R}^n$. A natural question arises as to whether the functions*

$$\frac{\partial f}{\partial x_j} : E \longrightarrow \mathbb{R}$$

are differentiable at a point x_0. If all of them are, we say that f is twice differentiable at x_0. In this case, for all integers $i, j = 1, 2, \ldots, n$ there exist the secondâ€“order partial derivatives

$$\frac{\partial}{\partial x_j}\Big(\frac{\partial f}{\partial x_i}(x_0)\Big) = \frac{\partial^2 f(x_0)}{\partial x_j\,\partial x_i}.$$

When f is twice differentiable at every point of E, there are n^2 functions of the form

$$\frac{\partial^2 f}{\partial x_j\,\partial x_i} : E \longrightarrow \mathbb{R}, \qquad 1 \leq i, j \leq n.$$

If all these functions are differentiable at a point $x_0 \in E$, we say that f is three times differentiable at x_0, and so on.

□

Teorema 2.26 *(Schwarz) Let $f : E \longrightarrow \mathbb{R}$ be a map defined on an open set $E \subset \mathbb{R}^n$. If f is twice differentiable at a point $p \in E$, then for any $0 \leq i, j \leq n$ one has*

$$\frac{\partial^2 f(p)}{\partial x_i\,\partial x_j} = \frac{\partial^2 f(p)}{\partial x_j\,\partial x_i}.$$

Proof: Without loss of generality, assume that $E \subset \mathbb{R}^2$ and set $p = (x_0, y_0)$. Since E is an open set, there exists $\delta > 0$ such that the ball $B_\delta(x_0, y_0) \subset E$. Consequently, there exists $\varepsilon > 0$ such that the square

$$(x_0 - \varepsilon, x_0 + \varepsilon) \times (y_0 - \varepsilon, y_0 + \varepsilon)$$

is contained in $B_\delta(x_0, y_0)$ (for instance, take $\varepsilon = \frac{\delta\sqrt{2}}{2}$).

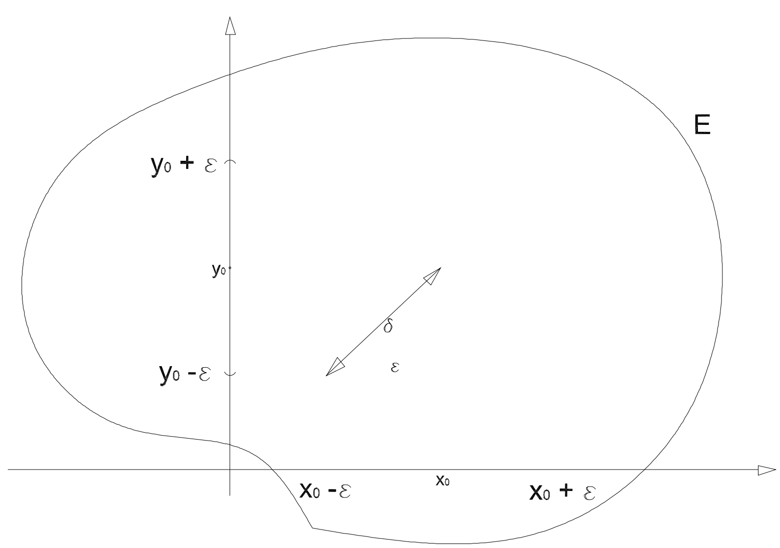

Figure 2.7:

For every $t \in (-\varepsilon, \varepsilon)$ define

$$\varphi(t) = f(x_0 + t, y_0 + t) - f(x_0 + t, y_0) - f(x_0, y_0 + t) + f(x_0, y_0).$$

For each fixed $t \in (-\varepsilon, \varepsilon)$ and every $x \in [x_0 - t, x_0 + t]$, set

$$h(x) = f(x, y_0 + t) - f(x, y_0).$$

In this way we see that

$$\varphi(t) = h(x_0 + t) - h(x_0).$$

Since, by hypothesis, f is differentiable on E, in particular h is differentiable on the interval $(x_0, x_0 + t)$ and continuous on $[x_0, x_0 + t]$, for every $0 < t < \varepsilon$. Hence, by the Mean Value Theorem for real functions of one variable, there exists $\xi \in (x_0, x_0 + t)$ such that

$$h'(\xi) = \frac{h(x_0 + t) - h(x_0)}{t} = \frac{\varphi(t)}{t}.$$

Moreover, since $\xi \in (x_0, x_0 + t) = \{x_0 + st;\ s \in (0, 1)\}$, we have $\xi = x_0 + \theta t$ for some $\theta \in (0, 1)$. Thus, from (2.50) we obtain

$$\varphi(t) \quad = \quad h'(\xi)\, t = \left[\frac{\partial f}{\partial x}(x_0 + \theta t, y_0 + t) - \frac{\partial f}{\partial x}(x_0 + \theta t, y_0)\right] t. \tag{2.5}$$

On the other hand, since the map

$$\frac{\partial f}{\partial x} : E \longrightarrow \mathbb{R}$$

is differentiable at the point $\varphi = (x_0, y_0)$, we have

$$\frac{\partial f}{\partial x}\big((x_0, y_0) + (\theta t, t)\big) = \frac{\partial f}{\partial x}(x_0, y_0) + \frac{\partial^2 f(x_0, y_0)}{\partial x^2}(\theta t) + \frac{\partial^2 f(x_0, y_0)}{\partial y\, \partial x}\, t + r_1, \tag{2.6}$$

where $\lim\limits_{t \to 0} \dfrac{r_1}{t} = 0$.

$$\frac{\partial f}{\partial x}(x_0 + \theta t, y_0) = \frac{\partial f(x_0, y_0)}{\partial x} + \frac{\partial^2 f(x_0, y_0)}{\partial x^2}(\theta t) + r_2, \tag{2.7}$$

where $\lim\limits_{t \to 0} \dfrac{r_2}{t} = 0$.

Therefore, from (2.51), (2.52) and (2.53) we obtain

$$\varphi(t) = \frac{\partial^2 f(x_0, y_0)}{\partial y\, \partial x}\, t^2 + (r_1 - r_2)\, t. \tag{2.8}$$

Hence, from (2.54) we have

$$\lim_{t\to 0}\frac{\varphi(t)}{t^2}=\frac{\partial^2 f(x_0,y_0)}{\partial y\,\partial x}+\lim_{t\to 0}\frac{r}{t},$$

where $r=r_1-r_2$, and therefore $\lim_{t\to 0}\frac{r}{t}=0$.

Thus,

$$\lim_{t\to 0}\frac{\varphi(t)}{t^2}=\frac{\partial^2 f(x_0,y_0)}{\partial y\,\partial x}. \tag{2.9}$$

Similarly, if we consider the function

$$g(y)=f(x_0+t,y)-f(x_0,y)$$

and argue as before, we obtain

$$\lim_{t\to 0}\frac{\varphi(t)}{t^2}=\frac{\partial^2 f(x_0,y_0)}{\partial x\,\partial y}. \tag{2.10}$$

Thus, from (2.55) and (2.56) we obtain the desired result. □

Note: We saw earlier that a map $f:E\subset\mathbb{R}^n\longrightarrow\mathbb{R}$ is said to be of class $C^1(E)$ when the differential

$$df:E\longrightarrow(\mathbb{R}^n)^*$$

is a continuous map on the open set $E\subset\mathbb{R}^n$. Equivalently (see Theorem 2.20), we proved that f is of class $C^1(E)$ if and only if the partial derivatives $\frac{\partial f}{\partial x_i}$ exist on E and are continuous. In the same way, we say that a function f is of class $C^2(E)$ when the mixed partial derivatives $\frac{\partial^2 f}{\partial x_i\,\partial x_j}$ exist and are continuous on E, or, equivalently, when the first-order partial derivatives $\frac{\partial f}{\partial x_i}$ exist on E and are of class $C^1(E)$.

Corollary 2.27 *Let $f:E\subset\mathbb{R}^n\longrightarrow\mathbb{R}$ be a map of class $C^2(E)$. Then, for any $0\le i,j\le n$ one has*

$$\frac{\partial^2 f(x)}{\partial x_i\,\partial x_j}=\frac{\partial^2 f(x)}{\partial x_j\,\partial x_i},\qquad \forall x\in E.$$

Proof: Immediate. □

2.8 The Inverse Function Theorem

In the proof of the Inverse Function Theorem we shall use the *Method of Successive Approximations*, a principle of great usefulness in proving existence and uniqueness of solutions for differential equations, integral equations, and so on.

Definition 2.28 *Let $X \subset \mathbb{R}^n$. A map $f : X \longrightarrow \mathbb{R}^m$ is called a contraction if there exist $\lambda \in \mathbb{R}$, $0 \leq \lambda < 1$, and norms on $\mathbb{R}^n$ and $\mathbb{R}^m$ such that*

$$\|f(x) - f(y)\| \leq \lambda \|x - y\|, \qquad \forall x, y \in X.$$

Every contraction is uniformly continuous. For example, let $E \subset \mathbb{R}^n$ be an open and convex set. If $f : E \longrightarrow \mathbb{R}^m$ is a differentiable map such that $\|f'(x)\| \leq \lambda < 1$ for some constant λ and every $x \in E$, then the Mean Value Inequality guarantees that $\|f(x) - f(y)\| \leq \lambda \|x - y\|$, and therefore f is a contraction.

Definition 2.29 *Let $X \subset \mathbb{R}^n$. A fixed point of a map $f : X \longrightarrow \mathbb{R}^n$ is a point $x \in X$ such that $f(x) = x$.*

Teorema 2.30 (Fixed Point Theorem for Contractions) *Let $F \subset \mathbb{R}^n$ be a closed subset and $\varphi : F \longrightarrow F$ a contraction. Given any point $x_0 \in F$, the sequence*

$$x_1 = \varphi(x_0), \quad x_2 = \varphi(x_1), \quad \ldots, \quad x_{k+1} = \varphi(x_k), \ldots$$

converges to a point $x \in F$, which is the unique fixed point of φ.

Proof: *(1) Existence.*

Let $x_0 \in F$ and consider the sequence (x_k) defined by the recurrence relation

$$x_{k+1} = \varphi(x_k), \qquad k = 0, 1, 2, \ldots. \tag{2.11}$$

Since φ is a contraction, there exist $0 \leq C < 1$ and a norm on $\mathbb{R}^n$ such that

$$\|\varphi(x) - \varphi(y)\| \leq C\|x - y\|, \qquad \forall x, y \in F. \tag{2.12}$$

Note that

$$\|x_{k+1} - x_k\| = \|\varphi(x_k) - \varphi(x_{k-1})\| \leq C\,\|x_k - x_{k-1}\|, \qquad k = 1, 2, \ldots.$$

Hence

$$\|x_{k+1} - x_k\| \leq C\|x_k - x_{k-1}\|, \qquad k = 1, 2, \ldots. \tag{2.13}$$

From (2.59) we may write

$$\|x_2 - x_1\| \leq C\|x_1 - x_0\|,$$

$$\|x_3 - x_2\| \leq C\|x_2 - x_1\| \leq C^2\|x_1 - x_0\|,$$

$$\vdots$$

We claim that

$$\|x_{k+1} - x_k\| \leq C^k\|x_1 - x_0\| \qquad \forall k \geq 1. \tag{2.14}$$

We prove this by induction on k.

(i) $k = 1$ (already proved).

(ii) Suppose the statement holds for k, that is,

$$\|x_{k+1} - x_k\| \leq C^k\|x_1 - x_0\|.$$

(iii) We prove it for $k+1$. From (2.59) and (ii) we have

$$\|x_{k+2} - x_{k+1}\| \leq C\|x_{k+1} - x_k\| \leq C^{k+1}\|x_1 - x_0\|,$$

which proves (3.60).

We now show that (x_k) is a Cauchy sequence in $\mathbb{R}^n$. Indeed, let $r, s \in \mathbb{N}$ with $s < r$. Then $r = s + p$ for some $p \in \mathbb{N}$. Hence

$$\|x_s - x_r\| = \|x_s - x_{s+p}\| \leq \|x_s - x_{s+1}\| + \|x_{s+1} - x_{s+2}\| + \cdots + \|x_{s+p-1} - x_{s+p}\|.$$

Therefore, by (2.60) we obtain

$$\|x_s - x_{s+p}\| \leq C^s\|x_1 - x_0\| + C^{s+1}\|x_1 - x_0\| + \cdots + C^{s+p-1}\|x_1 - x_0\|$$

$$= \big(C^s + C^{s+1} + \cdots + C^{s+p-1}\big)\|x_1 - x_0\| = \frac{C^s(1 - C^p)}{1 - C}\,\|x_1 - x_0\|.$$

However, since $0 \leq C < 1$, we have $1 - C^p \leq 1$ and hence

$$\|x_s - x_{s+p}\| \leq K\,C^s,$$

where $K = \dfrac{\|x_1 - x_0\|}{1 - C}$ is a constant. As

$$\lim_{s \to +\infty} (KC^s) = 0,$$

it follows that, for any $\varepsilon > 0$, there exists $s_0 \in \mathbb{N}$ such that, for all $s \geq s_0$, $KC^s < \varepsilon$. Hence, given $\varepsilon > 0$ we have

$$\|x_s - x_{s+p}\| < \varepsilon, \qquad \forall s \geq s_0,$$

that is,

$$\lim_{s\to+\infty} \|x_s - x_{s+p}\| = 0$$

for every $p \in \mathbb{N}$. Thus (x_k) is a Cauchy sequence in $\mathbb{R}^n$, and since $\mathbb{R}^n$ is complete, we have

$$\lim_{k\to+\infty} x_k = x^*$$

and clearly $x^* \in F$, since F is closed. As φ is a contraction, it is continuous on F. Hence

$$\varphi(x^*) = \varphi\big(\lim x_k\big) = \lim \varphi(x_k) = \lim x_{k+1} = x^*.$$

(2) Uniqueness.

Suppose there is also $y^* \in F$ such that $\varphi(y^*) = y^*$. Then, from (2.58) we have

$$\|x^* - y^*\| = \|\varphi(x^*) - \varphi(y^*)\| \le C\|x^* - y^*\|.$$

Hence

$$\|x^* - y^*\|(1 - C) \le 0.$$

Since $(1-C) > 0$, from the last inequality we obtain $x^* = y^*$, which proves uniqueness and completes the proof of the theorem. □

Definition 2.31 *Given sets $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$, a homeomorphism between X and Y is a continuous bijection $f : X \longrightarrow Y$ whose inverse $f^{-1} : Y \longrightarrow X$ is also continuous. In this case, X and Y are said to be homeomorphic.*

A bijection may be continuous without its inverse being continuous. A canonical example is given by the map

$$\begin{aligned} f : [0, 2\pi) &\longrightarrow S^1 = \{(x, y) \in \mathbb{R}^2;\ x^2 + y^2 = 1\} \\ t &\longmapsto f(t) = (\cos t, \sin t). \end{aligned}$$

The map f is clearly continuous and is evidently bijective. However, its inverse $f^{-1} : S^1 \longrightarrow [0, 2\pi)$ is discontinuous at the point $p = (1, 0)$. Indeed, for each $k \in \mathbb{N}$ let $t_k = 2\pi - \dfrac{1}{k}$ and $z_k = f(t_k)$. Then $\lim\limits_{k\to\infty} z_k = p$, but it is not true that

$$\lim_{k\to\infty} f^{-1}(z_k) = \lim_{k\to\infty} t_k$$

is equal to $f^{-1}(p) = 0$.

As examples of homeomorphisms of $\mathbb{R}^n$ onto itself we have:

(i) The translations:

$$T_\alpha : \mathbb{R}^n \longrightarrow \mathbb{R}^n, \qquad T_\alpha(x) = \alpha + x.$$

Indeed, T_α is actually an isometry, that is, a continuous surjective map that preserves distance:

$$\|T_\alpha(x) - T_\alpha(y)\| = \|x - y\|.$$

It follows that it is also injective, since if $T_\alpha(x) = T_\alpha(y)$ then $\|x - y\| = 0$, hence $x = y$. As $(T_\alpha)^{-1} = T_{-\alpha}$, we see that its inverse is also continuous.

(ii) The homotheties:

$$H_\lambda : \mathbb{R}^n \longrightarrow \mathbb{R}^n, \qquad H_\lambda(x) = \lambda x, \quad \lambda \neq 0.$$

Each homothety H_λ is an invertible linear transformation with $(H_\lambda)^{-1} = H_{\lambda^{-1}}$.

As examples of homeomorphic subsets of $\mathbb{R}^n$, we may consider two open (or closed) balls. Indeed, given $B_r(a)$ and $B_s(b)$, it suffices to consider the map

$$\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

defined by

$$\varphi = T_b \circ H_{s/r} \circ T_{-a}.$$

A more sophisticated example of a homeomorphism is given by the map

$$\begin{aligned} \psi : \Omega &\longrightarrow \Omega \\ A &\longmapsto A^{-1}, \end{aligned}$$

where

$$\Omega = \{A \in \mathcal{L}(\mathbb{R}^n);\ \exists A^{-1}\},$$

as mentioned in the first section of this chapter.

Definition 2.32 *Let U and V be open subsets of $\mathbb{R}^n$. A bijection $f : U \to V$ is called a diffeomorphism from U onto V if it is differentiable and its inverse $f^{-1} : V \to U$ is also differentiable.*

We say that a differentiable map $f : E \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a 'local diffeomorphism' if for each $x \in E$ there exists an open set V_x with $x \in V_x \subset E$ such that the restriction of f to V_x is a diffeomorphism onto an open set $W_x \subset \mathbb{R}^n$, that is, the map $f : V_x \longrightarrow W_x$ is a diffeomorphism.

Care must be taken not to confuse a diffeomorphism with a differentiable homeomorphism. An example of a homeomorphism whose inverse is not differentiable (at 0) is the function $f : \mathbb{R} \longrightarrow \mathbb{R}$ given by $f(x) = x^3$.

Note: It follows from the Chain Rule that if a map $f : U \longrightarrow \mathbb{R}^n$, where U is an open subset of $\mathbb{R}^n$, is differentiable at a point $a \in U$ and admits an inverse $g = f^{-1} : V \longrightarrow \mathbb{R}^n$ defined on the open set $V \subset \mathbb{R}^n$, differentiable at $b = f(a)$, then $f'(a) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is an isomorphism whose inverse is $g'(b) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$. Indeed:

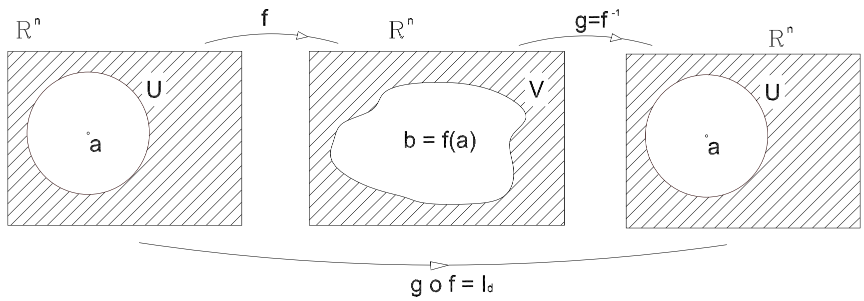

Figure 2.8:

From $g \circ f = I_U$ and $f \circ g = I_V$ it follows that

$$\begin{aligned} g'(b) \circ f'(a) &= I_d : \mathbb{R}^n \longrightarrow \mathbb{R}^n, \\ f'(a) \circ g'(b) &= I_d : \mathbb{R}^n \longrightarrow \mathbb{R}^n. \end{aligned}$$

Hence

$$g'(b) = (f'(a))^{-1}.$$

As a consequence of the observation above, if $f : U \longrightarrow V$ is a diffeomorphism between open subsets of $\mathbb{R}^n$, then for every $x \in U$ the derivative

$$f'(x) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

is an isomorphism. In terms of the Jacobian determinant, this means that

$$\det(Jf(x)) \neq 0, \qquad \forall x \in U,$$

where $Jf(x) = \left(\frac{\partial f_i}{\partial x_j}(x)\right)$ is the Jacobian matrix of f at x. It is natural to ask whether the converse holds.

The Inverse Function Theorem will provide the converse in the case where $f \in C^k$ $(k \geq 1)$, in the sense of 'local diffeomorphism', as we now state.

Teorema 2.33 (Inverse Function Theorem) *Let $f : E \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a continuously differentiable function on the open set E, and suppose that $f'(a)$ is invertible for some $a \in E$. (In other words, we are assuming that $f'(a) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is an isomorphism, or equivalently, that $Jf(a) \neq 0$.) Set $b = f(a)$. Then:*

(i) There exist open sets U and V in $\mathbb{R}^n$ such that $a \in U$, $b \in V$, f is one-to-one on U, and $f(U) = V$, that is, f is a bijection from U onto V.

(ii) If g denotes the inverse of f (which exists by (i)), defined by $g(f(x)) = x$ for $x \in U$, then $g \in C^1(V)$ and

$$g'(y) = f'(g(y))^{-1}, \qquad \forall y \in V.$$

In other words, f is a local diffeomorphism from U onto V, with f^{-1} of class C^1.

Proof: (i) Set $A = f'(a)$. Since $f \in C^1(E)$, given $\varepsilon > 0$ there exists $\delta > 0$ such that for every $x \in E$ with $\|x - a\| < \delta$ one has

$$\|f'(x) - f'(a)\| < \varepsilon. \tag{2.15}$$

For each fixed $y \in \mathbb{R}^n$ define the auxiliary map

$$\begin{aligned} \varphi : E &\longrightarrow \mathbb{R}^n \\ x &\longmapsto \varphi(x) = x + A^{-1}(y - f(x)). \end{aligned} \tag{2.16}$$

Note that $y = f(x)$ if and only if x is a fixed point of φ. Indeed, if $y = f(x)$ then $\varphi(x) = x + A^{-1}(0)$. Since A^{-1} is linear, $A^{-1}(0) = 0$, thus $\varphi(x) = x$. Conversely, if $\varphi(x) = x$, then $\varphi(x) = x + A^{-1}(y - f(x))$, and therefore $A^{-1}(y - f(x)) = 0$. As A^{-1} is linear and injective, it follows that $y = f(x)$.

On the other hand, from the expression for φ we may write

$$\varphi(x) = I(x) + A^{-1}(y - f(x)),$$

whose derivative is, by the Chain Rule,

$$\varphi'(x) = I'(x) + (A^{-1})'(y - f(x)) \circ (y - f(x))'. \tag{2.17}$$

Recall that if $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is linear, then $T'(x) = T$ for every $x \in \mathbb{R}^n$. Thus

$$I'(x) = I \quad \text{and} \quad (A^{-1})'(y - f(x)) = A^{-1}.$$

Therefore, from (2.63) we obtain

$$\begin{aligned} \varphi'(x) &= I + A^{-1} \circ (-f'(x)) = A^{-1} \circ A - A^{-1} \circ f'(x) \\ &= A^{-1}(A - f'(x)). \end{aligned} \tag{2.18}$$

From (2.64) we deduce that

$$\begin{aligned} \|\varphi'(x)\| &= \|A^{-1}(A - f'(x))\| \leq \|A^{-1}\|\|A - f'(x)\| \\ &= \|A^{-1}\|\|f'(a) - f'(x)\|. \end{aligned} \tag{2.19}$$

If we take $\varepsilon = \dfrac{1}{2\|A^{-1}\|}$ in (2.61), then from (2.65) we obtain

$$\|\varphi'(x)\| \leq \|A^{-1}\|\|f'(a) - f'(x)\| < \frac{1}{2}, \tag{2.20}$$

for all $x \in E$ with $\|x - a\| < \delta$.

Now let $U = B_\delta(a)$, which is clearly a convex open set. In view of (2.66) and by the Mean Value Inequality, we have

$$\|\varphi(x_1) - \varphi(x_2)\| \leq \frac{1}{2}\|x_1 - x_2\| \tag{2.21}$$

for all $x_1, x_2 \in U$, that is,

$$\|\varphi(x_1) - \varphi(x_2)\| = \|x_1 + A^{-1}(y - f(x_1)) - x_2 - A^{-1}(y - f(x_2))\| \leq \frac{1}{2}\|x_1 - x_2\|$$

for all $x_1, x_2 \in U$.

Thus

$$\|(x_1 - x_2) + A^{-1}(f(x_2) - f(x_1))\| \leq \frac{1}{2}\|x_1 - x_2\|, \qquad \forall x_1, x_2 \in U.$$

Note that if $f(x_1) = f(x_2)$ then

$$0 \leq \|x_1 - x_2\| \leq \frac{1}{2}\|x_1 - x_2\|,$$

which implies $x_1 = x_2$ and consequently $f|_U$ is injective.

Let $V = f(U)$. Then the map $f : U \longrightarrow V$ is a bijection. It remains to prove that V is an open subset of $\mathbb{R}^n$. Indeed, take $y_0 \in V$. We shall show that there exists $R > 0$ such that $B_R(y_0) \subset V$. In fact:

Since $y_0 \in V$, we have $y_0 = f(x_0)$ for some $x_0 \in U$. Let $r > 0$ be sufficiently small so that the closure of the ball $B = B_r(x_0)$ is contained in U, that is, $\overline{B} \subset U$.

Set

$$\lambda = \frac{1}{2\|A^{-1}\|} > 0 \quad \text{and} \quad R = \lambda r.$$

We claim that $B_R(y_0) \subset V$. Indeed, let $y \in B_R(y_0)$. Then

$$\|y - y_0\| < \frac{r}{2\|A^{-1}\|}.$$

For this $y \in B_R(y_0)$, consider the function φ defined by (2.62). In particular, for $x = x_0$ we have

$$\begin{aligned}\|\varphi(x_0) - x_0\| &= \|x_0 + A^{-1}(y - f(x_0)) - x_0\| \leq \|A^{-1}\|\|y - f(x_0)\| \\ &= \|A^{-1}\|\|y - y_0\| < \|A^{-1}\|\frac{r}{2\|A^{-1}\|} = \frac{r}{2}.\end{aligned}$$

On the other hand, for any $x \in \overline{B} \subset U$ we have $\|x - x_0\| \leq r$, and from (2.67) we obtain

$$\|\varphi(x) - x_0\| \leq \|\varphi(x) - \varphi(x_0)\| + \|\varphi(x_0) - x_0\| < \frac{1}{2}\|x - x_0\| + \frac{r}{2} < \frac{r}{2} + \frac{r}{2} = r.$$

Hence $\varphi(x) \in B = B_r(x_0)$ for every $x \in \overline{B}$. This allows us to define the map

$$\begin{aligned}\varphi : \overline{B} &\longrightarrow \overline{B} \\ x &\longmapsto \varphi(x) = x + A^{-1}(y - f(x)),\end{aligned}$$

which is a contraction, as we saw in (2.67).

Since $\overline{B}$ is a closed subset of $\mathbb{R}^n$, it follows from the Fixed Point Theorem that φ has a unique fixed point $x^* \in \overline{B}$, that is, there exists $x^* \in \overline{B}$ such that $\varphi(x^*) = x^*$.

However, as we proved earlier, $y = f(x)$ if and only if x is a fixed point of φ. Thus, since x^* is a fixed point of φ, we have $f(x^*) = y$, and since $x^* \in \overline{B}$, it follows that $y = f(x^*) \in f(\overline{B}) \subset f(U) = V$, which completes the proof of item (i). □

Proof: (ii) Let us now prove the differentiability of the inverse. Set $g = f^{-1}$.

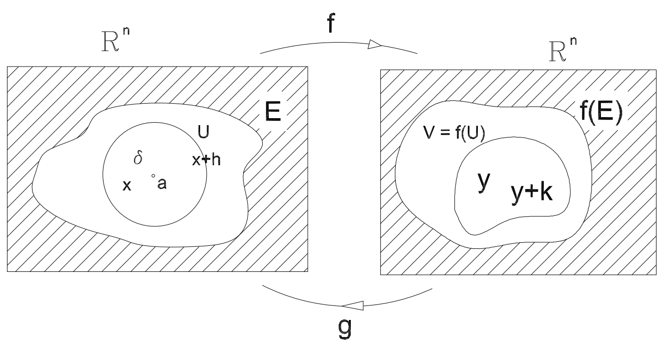

Figure 2.9:

Let $y \in V$ and $k \in \mathbb{R}^n$ be such that $\|k\|$ is sufficiently small to ensure that $y + k \in V$. Since f is a bijection, there exist $x, x + h \in U$ (with $h \neq 0$) such that $y = f(x)$ and $y + k = f(x + h)$. For the chosen $y \in V$, define, as before, the map

$$\varphi(x) = x + A^{-1}(y - f(x)), \qquad x \in U.$$

Then

$$\begin{aligned} \varphi(x+h) - \varphi(x) &= (x+h) + A^{-1}(y - f(x+h)) - x - A^{-1}(y - f(x)) \\ &= h + A^{-1}(y - f(x+h) - y + f(x)) \\ &= h - A^{-1}(f(x+h) - f(x)) \\ &= h - A^{-1}(k). \end{aligned} \tag{2.22}$$

On the other hand, from (2.67) we have

$$\|\varphi(x+h) - \varphi(x)\| \leq \frac{1}{2}\|x + h - x\| = \frac{\|h\|}{2}. \tag{2.23}$$

Thus, from (2.68) and (2.69) we obtain

$$\|h - A^{-1}(k)\| \leq \frac{\|h\|}{2},$$

which implies

$$\|h\| \leq \|A^{-1}(k)\| + \frac{\|h\|}{2}.$$

Equivalently,

$$\|h\| \leq 2\|A^{-1}\|\|k\|. \tag{2.24}$$

Hence

$$\frac{1}{\|k\|} \leq \frac{2\|A^{-1}\|}{\|h\|}. \tag{2.25}$$

On the other hand, as in item (i), we have

$$\|f'(x) - A\| < \frac{1}{2\|A^{-1}\|}, \qquad \forall x \in B_\delta(a) = U,$$

which implies

$$\|f'(x) - A\|\|A^{-1}\| < \frac{1}{2} < 1,$$

and by Proposition 2.6 we obtain $f'(x) \in \Omega$, that is, $f'(x)$ is invertible for every $x \in U$.

Consider $x \in U = B_\delta(a)$ and $f'(x)$, and denote

$$T = (f'(x))^{-1}.$$

Let $y = f(x)$ and consider $k \in \mathbb{R}^n$ such that $f(x+h) = y + k$. Then

$$\begin{aligned} g(y+k) - g(y) - Tk &= f^{-1}(y+k) - f^{-1}(y) - Tk \\ &= f^{-1}(f(x+h)) - f^{-1}(f(x)) - Tk \\ &= x + h - x - Tk \\ &= h - Tk \\ &= h - T(f(x+h) - f(x)). \end{aligned}$$

Therefore

$$g(y+k) - g(y) - Tk = h - T(f(x+h) - f(x)),$$

and since $h = T(T^{-1}(h)) = T(f'(x)h)$, we obtain

$$\begin{aligned} g(y+k) - g(y) - Tk &= T(f'(x)h) - T(f(x+h) - f(x)) \\ &= T(f'(x)h - f(x+h) + f(x)) \\ &= -T\big(f(x+h) - f(x) - f'(x)h\big). \end{aligned} \tag{2.26}$$

Thus, from (2.71) and (2.72) we get

$$\begin{aligned} 0 &\leq \frac{\|g(y+k) - g(y) - Tk\|}{\|k\|} \leq \frac{\|T\|\|f(x+h) - f(x) - f'(x)h\|}{\|k\|} \\ &\leq 2\|A^{-1}\|\|T\|\frac{\|f(x+h) - f(x) - f'(x)h\|}{\|h\|}. \end{aligned} \tag{2.27}$$

Since (by (2.70)) $k \to 0$ implies $h \to 0$, it follows from (2.73) that $g'(y)$ exists and, moreover, $g'(y) = T$, that is, $g'(y) = (f'(x))^{-1}$. This implies

$$g'(y) = (f'(f^{-1}(y)))^{-1},$$

or equivalently

$$g'(y) = (f'(g(y)))^{-1}, \qquad \forall y \in V.$$

Finally, since $g : V \to U$ is continuous (as it is differentiable), $f' : U \to \Omega \subset \mathcal{L}(\mathbb{R}^n)$ is continuous (because f is of class $C^1(U)$), and $\psi : \Omega \to \Omega$ is continuous (Proposition 2.6), it follows that g' is continuous on V, that is, $g \in C^1(V)$. □

To fix ideas, consider the diagram below:

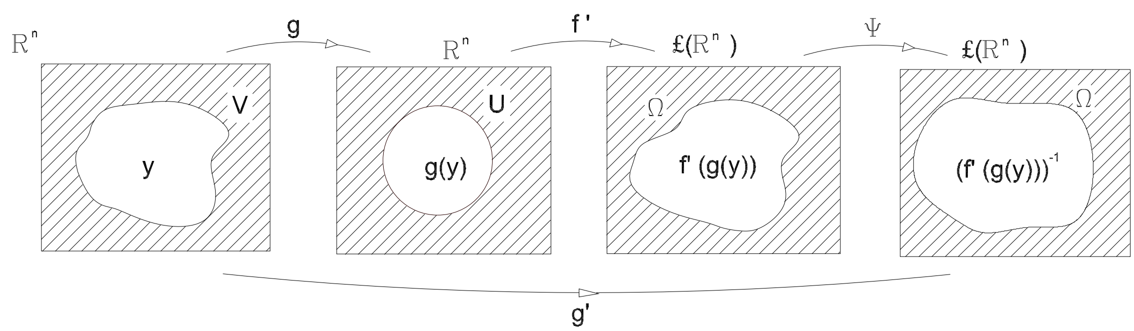

Figure 2.10:

Corollary 2.34 *A map $f : E \to \mathbb{R}^n$ of class C^k on the open set $E \subset \mathbb{R}^n$ (where $1 \leq k \leq +\infty$) is a local diffeomorphism if and only if, for every $x \in E$, the derivative $f'(x) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is an isomorphism (that is, $\det Jf(x) \neq 0$).*

We now introduce some notation that will be useful in the proof of the Implicit Function Theorem.

If $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$, we write (x, y) for the point

$$(x_1, \ldots, x_n, y_1, \ldots, y_m) \in \mathbb{R}^{n+m}.$$

Every $A \in \pounds(\mathbb{R}^{n+m}, \mathbb{R}^n)$ can be written in terms of two linear maps A_x and A_y defined by

$$A_x h = A(h, 0), \qquad A_y k = A(0, k),$$

for all $h \in \mathbb{R}^n$, $k \in \mathbb{R}^m$. Then $A_x \in \pounds(\mathbb{R}^n, \mathbb{R}^n)$, $A_y \in \pounds(\mathbb{R}^m, \mathbb{R}^n)$ and

$$A(h, k) = A_x h + A_y k. \tag{2.28}$$

Proposition 2.35 *Let $A \in \mathcal{L}(\mathbb{R}^{n+m}, \mathbb{R}^n)$ and assume that A_x is invertible. Then, for each $k \in \mathbb{R}^m$, there exists a unique $h \in \mathbb{R}^n$ such that $A(h,k) = 0$. This h is given by*

$$h = -A_x^{-1}(A_y k). \tag{2.29}$$

Proof: Let $A : \mathbb{R}^{n+m} \longrightarrow \mathbb{R}^n$ be linear. Then, by (2.74), if $A(h,k) = 0$ we have $A_x h + A_y k = 0$. Since A_x is invertible, there exists A_x^{-1}, and therefore

$$A_x^{-1}\big(A_x h + A_y k\big) = A_x^{-1}(0),$$

so

$$h + A_x^{-1}(A_y k) = 0 \quad \Rightarrow \quad h = -A_x^{-1}(A_y k).$$

Conversely, it is clear that if $h = -A_x^{-1}(A_y k)$, then $A(h,k) = 0$.

Moreover, for each $k \in \mathbb{R}^m$ there is a unique $A_y k \in \mathbb{R}^n$, since A_y is a function; and to each such $A_y k$ there corresponds a unique $A_x^{-1}(A_y k)$, because A_x^{-1} is a bijection. It follows that to each $k \in \mathbb{R}^m$ there corresponds a unique $h \in \mathbb{R}^n$, given by (2.75), such that $A(h,k) = 0$. □

Teorema 2.36 (Implicit Function Theorem) *Let f be a C^1 function defined on an open set $E \subset \mathbb{R}^{n+m}$ with values in $\mathbb{R}^n$, such that $f(a,b) = 0$ for some point $(a,b) \in E$. Let $A = f'(a,b)$ and assume that A_x is invertible. Then:*

(a) *There exist open sets $U \subset \mathbb{R}^{n+m}$ and $W \subset \mathbb{R}^m$, with $(a,b) \in U$ and $b \in W$, having the following property: for each $y \in W$ there exists a unique $x \in \mathbb{R}^n$ such that $(x,y) \in U$ and $f(x,y) = 0$.*

(b) *There exists (by (a)) a map $g : W \to \mathbb{R}^n$, with $g \in C^1(W)$, $g(b) = a$ and $f(g(y), y) = 0$ for all $y \in W$. Moreover,*

$$g'(y) = -A_x^{-1} A_y.$$

Proof: **(a)** Define the auxiliary map

$$\begin{aligned} F : E \subset \mathbb{R}^{n+m} &\longrightarrow \mathbb{R}^{n+m} \\ (x,y) &\longmapsto F(x,y) = (f(x,y), y), \end{aligned}$$

which is of class $C^1(E)$, since each coordinate function is C^1.

Now consider the linear map

$$\begin{aligned} L : \mathbb{R}^{n+m} &\longrightarrow \mathbb{R}^{n+m} \\ (h,k) &\longmapsto L(h,k) = (A(h,k), k), \end{aligned}$$

which we shall show is the derivative of F at (a,b). Indeed, let $h \in \mathbb{R}^n$ and $k \in \mathbb{R}^m$ have sufficiently small norms so that $(a+h, b+k) \in E$. Then

$$\begin{aligned} \|F(a+h,b+k) - F(a,b) - L(h,k)\| &= \|(f(a+h,b+k), b+k) - (f(a,b),b) - (A(h,k),k)\| \\ &= \|(f(a+h,b+k) - f(a,b) - A(h,k), 0)\| \\ &= \|f(a+h,b+k) - f(a,b) - A(h,k)\|. \end{aligned}$$

Since f is differentiable at (a, b), by hypothesis,

$$\lim_{(h,k)\to(0,0)} \frac{f(a+h, b+k) - f(a,b) - A(h,k)}{\|(h,k)\|} = 0.$$

It follows that

$$\lim_{(h,k)\to(0,0)} \frac{F(a+h, b+k) - F(a,b) - L(h,k)}{\|(h,k)\|} = 0,$$

and hence $F'(a,b) = L$.

We now show that L is an isomorphism. In fact, if $L(h,k) = (0,0)$ then $(A(h,k), k) = (0,0)$, so $A(h,k) = 0$ and $k = 0$. Since A_x is invertible, and for each $k \in \mathbb{R}^m$ there is a unique $h \in \mathbb{R}^n$ such that $A(h,k) = 0$, given by

$$h = -A_x^{-1}(A_y k),$$

we obtain, for $k = 0$, that

$$h = -A_x^{-1}(A_y 0) = 0,$$

that is, $(h,k) = (0,0)$. Thus L is injective, and by the Rankâ€œNullity Theorem L is an isomorphism.

Therefore, by the Inverse Function Theorem, there exist open sets $U, V \subset \mathbb{R}^{n+m}$ such that $F : U \to V$ is a bijection, with $F'(x,y)$ invertible for all $(x,y) \in U$. Moreover, $G = F^{-1} : V \to U$ is a C^1 map, with $(a,b) \in U$ and $F(a,b) = (0,b) \in V$.

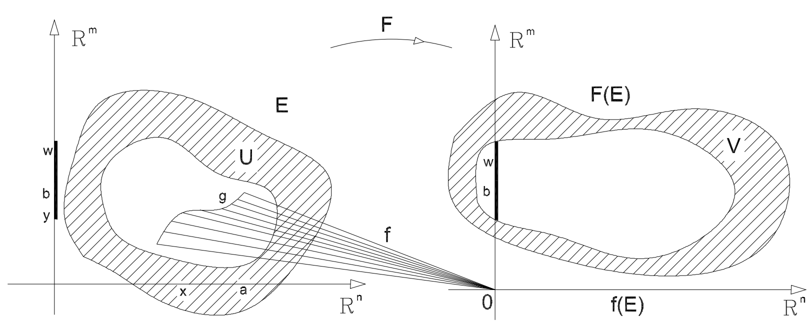

Figure 2.11:

Define

$$W = \{y \in \mathbb{R}^m;\ (0,y) \in V\}.$$

Note that W is an open subset of $\mathbb{R}^m$, since W is the image of $V \cap (\{0\} \times \mathbb{R}^m)$ under the natural identification $\{0\} \times \mathbb{R}^m \simeq \mathbb{R}^m$, and V is an open subset of $\mathbb{R}^{n+m}$. Now, if $y \in W$, then $(0,y) \in V$ and consequently

$$(0,y) = F(x,y)$$

for some $(x, y) \in U$. Since $F(x, y) = (f(x, y), y)$, it follows that $f(x, y) = 0$ for this x.

Suppose now that for the same y we had $x' \in \mathbb{R}^n$ with $(x', y) \in U$ and $f(x', y) = 0$. Then

$$F(x', y) = (f(x', y), y) = (f(x, y), y) = F(x, y),$$

and by the injectivity of F we would have $x' = x$. Thus, to each $y \in W$ there corresponds a unique $x \in \mathbb{R}^n$ such that $(x, y) \in U$ and $f(x, y) = 0$. Moreover, $b \in W$, since $(0, b) \in V$. This completes the proof of part (a).

(b) By part (a) we obtain a well-defined map

$$\begin{aligned} g : W &\longrightarrow \mathbb{R}^n \\ y &\longmapsto g(y) = x, \end{aligned}$$

such that $(x, y) \in U$ and $f(g(y), y) = f(x, y) = 0$ for all $y \in W$. Clearly $g(b) = a$, since $(a, b) \in U$, $b \in W$ and $f(a, b) = 0$.

On the other hand, note that for every $y \in W$ we have

$$F(g(y), y) = F(x, y) = (f(x, y), y) = (0, y).$$

Hence

$$G(0, y) = F^{-1}(0, y) = (g(y), y).$$

Since G is C^1 on V and $(0, y) \in V$ for all $y \in W$, the restriction g is C^1 on W.

Finally, set

$$\Phi(y) = (g(y), y) = (g(y), I_d(y)).$$

Since

$$(f \circ \Phi)(y) = f(g(y), y) \equiv 0,$$

it follows from the Chain Rule that

$$f'(\Phi(y)) \circ \Phi'(y) = 0, \qquad \forall y \in W.$$

In particular, for $y = b$ we have

$$f'(\Phi(b)) \circ \Phi'(b) = 0.$$

Since $\Phi(b) = (g(b), b) = (a, b)$, this can be written as

$$f'(a, b) \circ \Phi'(b) = 0,$$

that is,

$$A \circ \Phi'(b) = 0.$$

Now, for $k \in \mathbb{R}^m$,

$$\Phi'(b)k = \big(g'(b), I_d\big)k = (g'(b)k, k),$$

and therefore

$$A(g'(b)k, k) = 0, \qquad \forall k \in \mathbb{R}^m.$$

Using (2.74), we obtain

$$A_x(g'(b)k) + A_y k = 0, \qquad \forall k \in \mathbb{R}^m,$$

which implies

$$g'(b)k = -A_x^{-1}(A_y k), \qquad \forall k \in \mathbb{R}^m.$$

Thus

$$g'(b) = -A_x^{-1} A_y,$$

as claimed. □

Chapter 3

Multiple Integrals

3.1 The Definition of the Integral

According to Definition 1.33, an n-parallelepiped, or n-dimensional block, or cell, is the subset A of $\mathbb{R}^n$ given by the Cartesian product

$$A = \prod_{i=1}^{n} [a_i, b_i] = [a_1, b_1] \times \cdots \times [a_n, b_n]$$

of n compact intervals $[a_i, b_i]$, each of which is called an edge of the block A. When all the edges have the same length $b_i - a_i = a$, the block is called an nâ€“dimensional cube. When $n = 1$, A is an interval; for $n = 2$, the block reduces to a rectangle and the cube to a square. The point whose coordinates are $c_i = \dfrac{a_i + b_i}{2}$ is called the centre of the block A. The vertices of the block A are the points $p = (c_1, \ldots, c_n)$, where for each $i = 1, \ldots, n$ we have $c_i = a_i$ or $c_i = b_i$. The faces of the block A are the Cartesian products

$$F = L_1 \times \cdots \times L_n$$

such that, for each $i = 1, \ldots, n$, we have $L_i = \{a_i\}$, or $L_i = \{b_i\}$, or $L_i = [a_i, b_i]$. We say that the face F has dimension k when there are precisely k indices i for which $L_i = [a_i, b_i]$. In particular, each vertex of the block A is a face of dimension zero, while the block A itself is a face of dimension n.

Definition 3.1 *The nâ€“dimensional volume of the block $A = \prod_{i=1}^{n}[a_i, b_i]$ is*

$$vol\, A = \prod_{i=1}^{n} (b_i - a_i).$$

If A is an nâ€“dimensional cube whose edges have length a, then $vol\, A = a^n$.

The n-dimensional block $A = \prod_{i=1}^{n}[a_i, b_i]$ is a compact subset of $\mathbb{R}^n$ whose interior is the Cartesian product

$$\operatorname{int} A = \prod_{i=1}^{n}(a_i, b_i)$$

of open intervals (a_i, b_i), which we call the open nâ€“dimensional block. By definition, the volume of an open block is the same as that of the corresponding closed block.

Definition 3.2 *A partition of the block $A = \prod_{i=1}^{n}[a_i, b_i]$ is a finite set of the form*

$$P = P_1 \times \cdots \times P_n,$$

where each P_i is a partition of the interval $[a_i, b_i]$.

The elements of P are called the vertices of the partition P. A partition $P = P_1 \times \cdots \times P_n$ of the block A determines a decomposition of A into sub-blocks of the form $B = I_1 \times \cdots \times I_n$, where each I_j is an interval of the partition P_j. Each of these sub-blocks B will be called a block of the partition P. We write $B \in P$. Let us look at an example for $n = 2$.

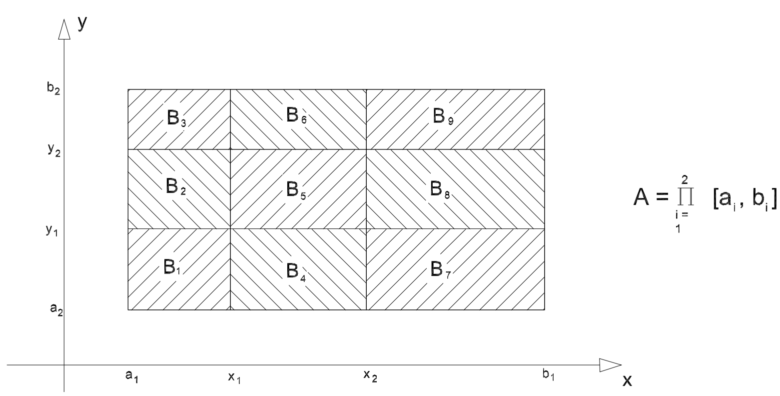

Figure 3.1:

Let $P_1 = \{a_1, x_1, x_2, b_1\}$ and $P_2 = \{a_2, y_1, y_2, b_2\}$ be partitions of the intervals $[a_1, b_1]$ and $[a_2, b_2]$, respectively. Then $P = P_1 \times P_2$ is a partition of A. Each B_i is a sub-block, called a block of the partition P; we denote this by $B_i \in P$.

If, for each $j = 1, \ldots, n$, the partition P_j decomposes the interval $[a_j, b_j]$ into k_j subintervals, then the partition P decomposes the block $A = \prod_{j=1}^{n}[a_j, b_j]$ into $k_1, k_2, \ldots, k_n$ sub-blocks.

In the example above, P_1 decomposes the interval $[a_1, b_1]$ into 3 subintervals, while P_2 decomposes $[a_2, b_2]$ also into 3 subintervals. Hence the partition P decomposes the block $A = \prod_{i=1}^{2}[a_i, b_i]$ into 3×3, that is, 9 sub-blocks.

If B_1 and B_2 are blocks of the same partition, then either their intersection is empty or it is a common kâ€“dimensional face of B_1 and B_2 ($k = 0, 1, \ldots, n-1$).

Given a partition $P = P_1 \times \cdots \times P_n$ of the block $A = \prod_{j=1}^n [a_j, b_j]$, the length of each edge is the sum of the lengths of the intervals in the partition P_i. It follows, by the distributive law of multiplication, that the volume of A is the sum of the volumes of all the blocks into which P decomposes A.

Returning to the example, we have

$$\operatorname{vol} A = \prod_{i=1}^{n}(b_i - a_i) = (b_1 - a_1)(b_2 - a_2).$$

However,

$$\begin{aligned} b_1 - a_1 &= (x_1 - a_1) + (x_2 - x_1) + (b_1 - x_2), \\ b_2 - a_2 &= (y_1 - a_2) + (y_2 - y_1) + (b_2 - y_2). \end{aligned}$$

Therefore,

$$\begin{aligned} \operatorname{vol} A = \; & (x_1 - a_1)(y_1 - a_2) + (x_1 - a_1)(y_2 - y_1) + (x_1 - a_1)(b_2 - y_2) + \\ & + (x_2 - x_1)(y_1 - a_2) + (x_2 - x_1)(y_2 - y_1) + (x_2 - x_1)(b_2 - y_2) + \\ & + (b_1 - x_2)(y_1 - a_2) + (b_1 - x_2)(y_2 - y_1) + (b_1 - x_2)(b_2 - y_2), \end{aligned}$$

that is, $\operatorname{vol} A = \sum_{i=1}^{9} \operatorname{vol} B_i$.

In general, we may write

$$\operatorname{vol} A = \sum_{B \in P} \operatorname{vol} B.$$

Definition 3.3 *Let P and Q be partitions of the block A. We say that Q is finer than P if $P \subset Q$. If $P = P_1 \times \cdots \times P_n$ and $Q = Q_1 \times \cdots \times Q_n$, then $P \subset Q$ if and only if $P_1 \subset Q_1, \ldots, P_n \subset Q_n$. In this case, each block of the partition Q is contained in a unique block of P, and each block of P is the union of those blocks of Q that it contains. More precisely, if $P \subset Q$, then Q induces a partition of each block of P, and therefore the volume of a block $B \in P$ is the sum of the volumes of the blocks of Q that are contained in B.*

Let us look at an example:

Let $P_1 = \{a_1, x_1, x_2, b_1\}$ and $P_2 = \{a_2, y_2, b_2\}$ be partitions of $[a_1, b_1]$ and $[a_2, b_2]$, respectively, and let $P = P_1 \times P_2$ be a partition of $A = \prod_{i=1}^2 [a_i, b_i]$. Consider $Q_1 = \{a_1, x_1, x_2, x_3, b_1\}$ and $Q_2 = \{a_2, y_1, y_2, b_2\}$, partitions of $[a_1, b_1]$ and $[a_2, b_2]$, respectively, with $P_1 \subset Q_1$ and $P_2 \subset Q_2$. Clearly $P \subset Q$. Notice that:

Each block B_i' of Q is contained in a unique block B_i of P.

Each block B_i of P is the union of those blocks B_i' of Q that are contained in it.

In this case:

$$B_1 = B_1' \cup B_2', \quad B_2 = B_3', \quad B_3 = B_4' \cup B_5', \quad B_4 = B_6', \quad B_5 = B_7' \cup B_8' \cup B_{10}' \cup B_{11}', \quad B_6 = B_9' \cup B_{12}'.$$

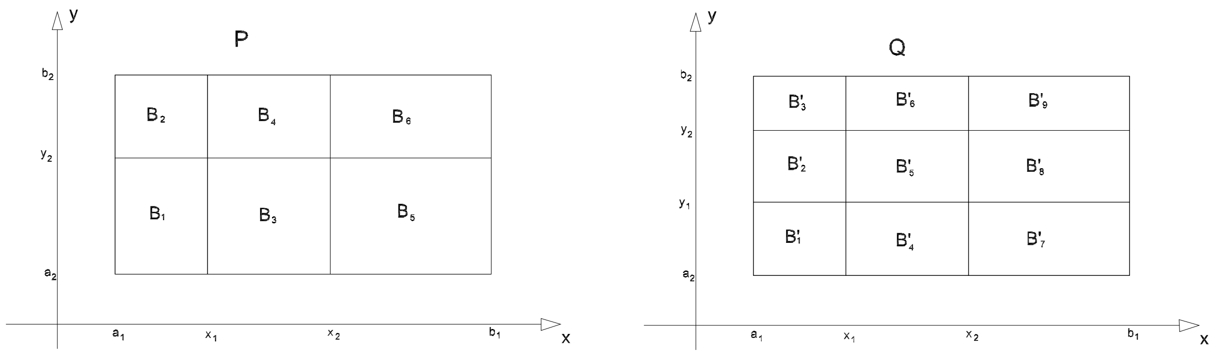

Figure 3.2:

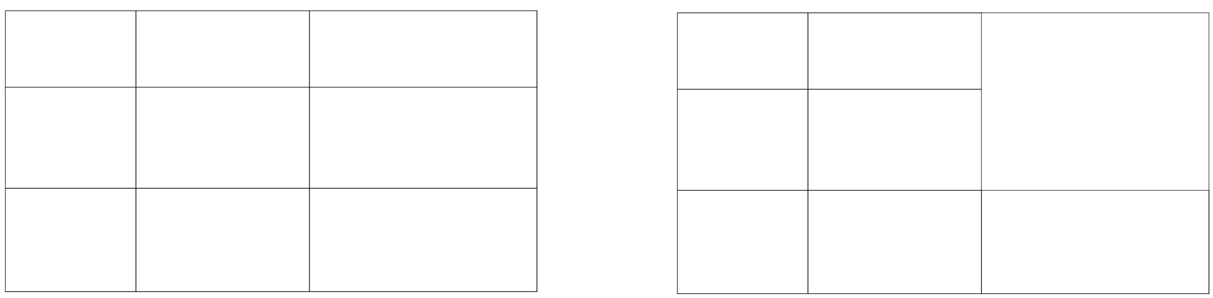

Figure 3.3:

The volume of each block B_i of P is the sum of the volumes of the blocks B_i' of Q contained in it.

In general, if $P \subset Q$ and $B \in P$ is a block of the partition P, and $B' \in Q$ is a block of the partition Q contained in B, then

$$\operatorname{vol} B = \sum_{B' \subset B} \operatorname{vol} B'.$$

It follows that

$$\operatorname{vol} A = \sum_{B \in P} \operatorname{vol} B = \sum_{B \in P} \sum_{B' \subset B} \operatorname{vol} B'. \tag{3.1}$$

<u>Note 1:</u> In Figure 3.3 we see two blocks (rectangles) decomposed as unions of sub-blocks:

The decomposition on the left arises from a partition, but the one on the right does not.

If $P = \prod_{i=1}^n P_i$ and $Q = \prod_{i=1}^n Q_i$ are partitions of a block A, the union $P \cup Q$ is not, in general, a partition of A. However, there exists a partition

$$P + Q = \prod_{i=1}^{n} (P_i \cup Q_i)$$

that refines both P and Q.

Definition 3.4 *The norm $|P|$ of a partition $P = \prod_{i=1}^n P_i$ is the largest length of a subinterval of any of the partitions P_i, that is, the largest length of the edges of the blocks $B \in P$.*

Note 2: If we equip $\mathbb{R}^n$ with the maximum norm, then the diameter of a block will be the length of its longest edge. To fix ideas, consider the diagram below (Figure 3.4):

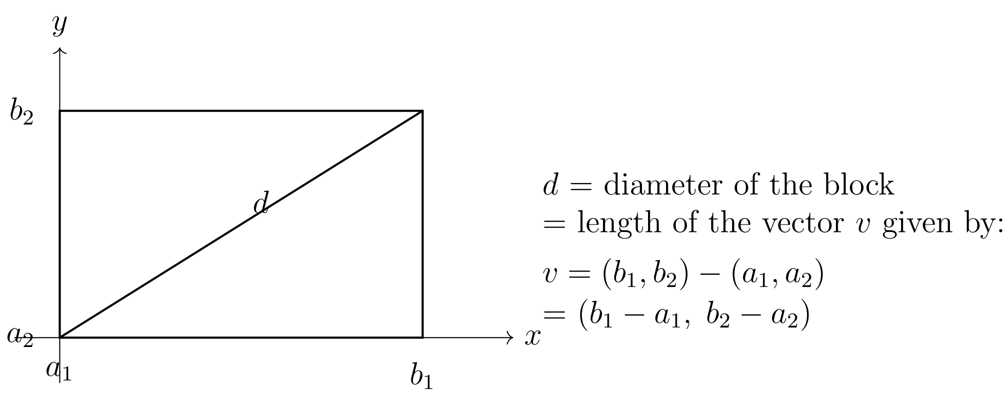

Figure 3.4: Diameter of a rectangular block in $\mathbb{R}^2$.

Thus

$$\|v\| = \max\{\, |b_1 - a_1|,\ |b_2 - a_2| \,\} = \text{length of the longest edge of the block.}$$

In general,

$$\operatorname{diam} A = \max\{\, |b_1 - a_1|, \ldots, |b_n - a_n| \,\}.$$

In this case, the norm $|P|$ of the partition P will be the largest diameter of the blocks $B \in P$.

Definition 3.5 *Let $f : A \to \mathbb{R}$ be a bounded real function defined on a block $A \subset \mathbb{R}^n$. Given a partition P of A, to each block $B \in P$ we associate the numbers*

$$m_B = \inf\{f(x);\, x \in B\} \quad \text{and} \quad M_B = \sup\{f(x);\, x \in B\},$$

with which we define, respectively, the lower sum and upper sum of f with respect to the partition P by setting

$$s(f;P) = \sum_{B \in P} m_B \operatorname{vol} B \quad \text{and} \quad S(f;P) = \sum_{B \in P} M_B \operatorname{vol} B. \tag{3.2}$$

The sums above extend over all blocks B of the partition P. Since $m_B \le M_B$ for each B, we have $s(f;P) \le S(f;P)$.

As in the case of functions of one variable, we shall now show that, when a partition is refined, the lower sum does not decrease and the upper sum does not increase. Indeed, consider the following:

Teorema 3.6 *Let P and Q be partitions of the block $A \subset \mathbb{R}^n$, with $P \subset Q$, and let $f : A \to \mathbb{R}$ be a bounded function. Then*

$$s(f;P) \leq s(f;Q) \leq S(f;Q) \leq S(f;P). \tag{3.3}$$

Proof: It is enough to prove the first and third inequalities, since the second has already been established. We denote by B and B' the blocks of the partitions P and Q, respectively. From (3.1) and (3.2), we have

$$\begin{aligned} s(f;P) &= \sum_{B\in P} m_B \operatorname{vol} B = \sum_{B\in P} m_B \Big[\sum_{B'\subset B} \operatorname{vol} B'\Big] \\ &= \sum_{B\in P}\Big[\sum_{B'\subset B} m_B \operatorname{vol} B'\Big] \leq \sum_{B\in P}\Big[\sum_{B'\subset B} m_{B'} \operatorname{vol} B'\Big] \\ &= \sum_{B'\in Q} m_{B'} \operatorname{vol} B'. \end{aligned}$$

The inequality above follows from the fact that $B' \subset B$, and therefore $m_{B'} \geq m_B$. The third inequality is analogous, using the fact that $M_{B'} \leq M_B$. □

Corollary 3.7 *Let $f : A \to \mathbb{R}$ be bounded. For any partitions P and Q of the block A, one has*

$$s(f;P) \leq S(f;Q).$$

Proof: Since $P \subset P+Q$ and $Q \subset P+Q$, it follows from the theorem above that

$$s(f;P) \leq s(f;P+Q) \leq S(f;P+Q) \leq S(f;Q).$$

□

Note 3: Any partition Q of the block A refines the trivial partition P, whose only sub-block is A. It follows from Theorem 3.6 that if $m \leq f(x) \leq M$ for all $x \in A$, then

$$s(f;P) \leq s(f;Q) \leq S(f;Q) \leq S(f;P).$$

However,

$$\begin{aligned} s(f;P) = \sum_{B\in P} m_B \operatorname{vol} B &= m_B \operatorname{vol} A \quad \text{(since } P \text{ is trivial)}, \\ S(f;P) = \sum_{B\in P} M_B \operatorname{vol} B &= M_B \operatorname{vol} A. \end{aligned}$$

Thus

$$m_B \operatorname{vol} A \le s(f;Q) \le S(f;Q) \le M_B \operatorname{vol} A,$$

and therefore

$$m \operatorname{vol} A \le s(f;Q) \le S(f;Q) \le M \operatorname{vol} A. \tag{3.4}$$

Let $\wp$ be the set of all partitions of A. Then the set

$$\sigma = \{s(f;Q);\ Q \in \wp\}$$

is bounded above, and the set

$$\Sigma = \{S(f;Q);\ Q \in \wp\}$$

is bounded below.

Definition 3.8 *Let $f : A \to \mathbb{R}$ be a bounded function defined on a block $A \subset \mathbb{R}^n$. We define the lower integral, denoted by $\underline{\int_A} f(x)\,dx$, and the upper integral, denoted by $\overline{\int_A} f(x)\,dx$, of the function f over the block A by setting*

$$\underline{\int_A} f(x)\,dx = \sup \sigma, \qquad \overline{\int_A} f(x)\,dx = \inf \Sigma. \tag{3.5}$$

Note 4: It follows from Corollary 3.7 and Note 3 that, if $m \le f(x) \le M$, then

$$m \operatorname{vol} A \le \underline{\int_A} f(x)\,dx \le \overline{\int_A} f(x)\,dx \le M \operatorname{vol} A. \tag{3.6}$$

Indeed, the first and third inequalities follow from the fact that $m \operatorname{vol} A$ and $M \operatorname{vol} A$ are, respectively, lower and upper bounds for the sets σ and Σ. We now prove the middle inequality.

Proof: Suppose, by contradiction, that

$$\underline{\int_A} f(x)\,dx > \overline{\int_A} f(x)\,dx,$$

that is,

$$\sup \sigma > \inf \Sigma.$$

Take

$$\epsilon = \frac{\sup \sigma - \inf \Sigma}{2} > 0.$$

Then, for this $\epsilon > 0$, there exist partitions P_0 and P_1 in $\wp$ such that

$$\inf \Sigma \le S(f;P_0) < \inf \Sigma + \epsilon = \sup \sigma - \epsilon < s(f;P_1) \le \sup \sigma,$$

that is,

$$S(f;P_0) < s(f;P_1),$$

which contradicts Corollary 3.7. □

Note 5: Let P_0 be an arbitrary partition of the block A. In order to compute the lower and upper integrals of a bounded function $f : A \to \mathbb{R}$, it suffices to consider the partitions that refine P_0, that is,

$$\underline{\int_A} f(x)\,dx = \sup\{s(f;P);\ P \supset P_0\} \quad \text{and} \quad \overline{\int_A} f(x)\,dx = \inf\{S(f;P);\ P \supset P_0\}.$$

Indeed, for every partition Q of the block A there exists a partition P that refines both Q and P_0, namely $P = Q + P_0$. Then $s(f;Q) \leq s(f;P)$ and $S(f;P) \leq S(f;Q)$, with $P \supset P_0$, and the claim follows.

Definition 3.9 *Let $f : A \to \mathbb{R}$ be bounded on the block $A \subset \mathbb{R}^n$. We say that f is integrable if its lower and upper integrals coincide. In that case we define the integral of f by*

$$\int_A f(x)\,dx = \underline{\int_A} f(x)\,dx = \overline{\int_A} f(x)\,dx. \tag{3.7}$$

Teorema 3.10 *A bounded function $f : A \to \mathbb{R}$ is integrable on the block $A \subset \mathbb{R}^n$ if and only if, for every $\epsilon > 0$, there exists a partition P of A such that*

$$S(f;P) - s(f;P) < \epsilon.$$

Proof: Suppose that f is integrable and let $\epsilon > 0$ be given. Then, by hypothesis,

$$\inf \Sigma = \sup \sigma,$$

where, as we have seen,

$$\sigma = \{s(f;Q);\ Q \in \wp\}, \qquad \Sigma = \{S(f;Q);\ Q \in \wp\}.$$

Thus, for the given $\epsilon > 0$, there exist partitions $P_0, P_1 \in \wp$ such that

$$\sup \sigma - \frac{\epsilon}{2} < s(f;P_0) \leq \sup \sigma = \inf \Sigma \leq S(f;P_1) < \inf \Sigma + \frac{\epsilon}{2}.$$

To fix ideas, see Figure 3.5:

From the last relation we obtain

$$S(f;P_1) - s(f;P_0) < \epsilon.$$

Set $P = P_0 + P_1$. Then

$$S(f;P) - s(f;P) \leq S(f;P_1) - s(f;P_0) < \epsilon.$$

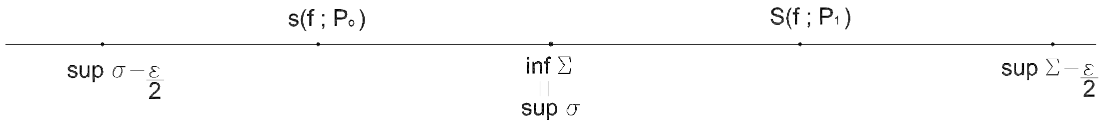

Figure 3.5:

Conversely, suppose that, for each $\epsilon > 0$, there exists $P_\epsilon \in \wp$ such that

$$S(f;P_\epsilon) - s(f;P_\epsilon) < \epsilon.$$

We must prove that $\inf\Sigma = \sup\sigma$. Indeed, we have

$$\begin{aligned} S(f;P_\epsilon) < \epsilon + s(f;P_\epsilon) &\Rightarrow \inf\Sigma \le S(f;P_\epsilon) < \epsilon + s(f;P_\epsilon) \Rightarrow \\ &\Rightarrow \inf\Sigma < \epsilon + s(f;P_\epsilon) \le \epsilon + \sup\sigma \Rightarrow \inf\Sigma - \sup\sigma < \epsilon. \end{aligned} \tag{3.8}$$

On the other hand, since $\sup\sigma \le \inf\Sigma$, that is, $\inf\Sigma - \sup\sigma \ge 0$, it follows from (3.8) that, for every $\epsilon > 0$,

$$0 \le \inf\Sigma - \sup\sigma < \epsilon,$$

and, by the arbitrariness of ϵ, we obtain the equality. □

Definition 3.11 *Let $f : X \to \mathbb{R}$ be bounded on a set $X \subset A$. We call the oscillation of f on X the number*

$$\omega_X = \omega(f;X) = \sup\{|f(x) - f(y)|;\ x, y \in X\}. \tag{3.9}$$

Lema 3.12 *Let $f : X \to \mathbb{R}$ be a bounded function. Set*

$$m_X = \inf\{f(x);\ x \in X\}, \qquad M_X = \sup\{f(x);\ x \in X\}.$$

Then $\omega_X = M_X - m_X$.

Proof: For each $x \in X$ we have $f(x) \le M_X$ and $m_X \le f(x)$. It follows that

$$f(x) - f(y) \le M_X - m_X, \qquad \forall x, y \in X. \tag{3.10}$$

In particular,

$$\begin{aligned} f(y) - f(x) &\le M_X - m_X, \quad \text{and therefore} \\ f(x) - f(y) &\ge -(M_X - m_X). \end{aligned} \tag{3.11}$$

Since $M_X - m_X \ge 0$, from (3.10) and (3.11) we obtain

$$|f(x) - f(y)| \le M_X - m_X, \qquad \forall x, y \in X.$$

Thus $M_X - m_X$ is an upper bound for the set

$$\{|f(x) - f(y)|;\ x, y \in X\},$$

and hence $\omega_X \leq M_X - m_X$.

We now show that in fact $M_X - m_X$ is the least upper bound. Indeed, given $\epsilon > 0$, there exist $x, y \in X$ such that

$$m_X \leq f(x) < m_X + \frac{\epsilon}{2} \quad \text{and} \quad M_X - \frac{\epsilon}{2} < f(y) \leq M_X,$$

and we may assume $f(x) < f(y)$. Then

$$\begin{aligned} |f(y) - f(x)| &= f(y) - f(x) \\ &> \left(M_X - \frac{\epsilon}{2}\right) - \left(m_X + \frac{\epsilon}{2}\right) \\ &= (M_X - m_X) - \epsilon. \end{aligned}$$

Consequently,

$$M_X - m_X < |f(y) - f(x)| + \epsilon,$$

and therefore

$$M_X - m_X < \omega_X + \epsilon.$$

By the arbitrariness of ϵ we obtain $M_X - m_X \leq \omega_X$, as required. □

Note 6: In view of the previous lemma, we can write

$$\begin{aligned} S(f;P) - s(f;P) &= \sum_{B \in P} M_B \operatorname{vol} B - \sum_{B \in P} m_B \operatorname{vol} B \\ &= \sum_{B \in P} (M_B - m_B) \operatorname{vol} B \\ &= \sum_{B \in P} \omega_B \operatorname{vol} B. \end{aligned} \tag{3.12}$$

In view of this last observation and of Theorem 3.10, for a function $f : A \to \mathbb{R}$ defined on a block $A \subset \mathbb{R}^n$ to be integrable it is necessary and sufficient that, for every $\epsilon > 0$, one can find a partition P of the block A such that

$$\sum_{B \in P} \omega_B \operatorname{vol} B < \epsilon. \tag{3.13}$$

Proposition 3.13 *Every continuous function $f : A \to \mathbb{R}$ defined on a block A of $\mathbb{R}^n$ is integrable.*

Proof: Let $\epsilon > 0$ be given. Since the block A is compact, the function f is uniformly continuous. Hence, for the given $\epsilon > 0$ there exists $\delta > 0$ such that, if $x, y \in A$ and $\|x - y\| < \delta$, then

$$|f(x) - f(y)| < \frac{\epsilon}{\operatorname{vol} A}.$$

Using on $\mathbb{R}^n$ the maximum norm, and letting $P \in \wp$ be a partition of A with $|P| < \delta$, we have that the largest diameter of the blocks $B \in P$ does not exceed δ. Thus, given $x, y \in B$ we have $\|x - y\| \leq |P| < \delta$ and hence $|f(x) - f(y)| < \dfrac{\epsilon}{\operatorname{vol} A}$, for every $B \in P$. Therefore, $\dfrac{\epsilon}{\operatorname{vol} A}$ is an upper bound of the set $\{|f(x) - f(y)|;\, x, y \in B\}$, for every $B \in P$. It follows that $\omega_B < \dfrac{\epsilon}{\operatorname{vol} A}$ for all $B \in P$. Consequently,

$$\sum_{B \in P} \omega_B \operatorname{vol} B < \frac{\epsilon}{\operatorname{vol} A} \sum_{B \in P} \operatorname{vol} B = \epsilon,$$

which proves the proposition (by (3.13)). □

Example 1.

Let $A \subset \mathbb{R}^n$ be a block and let X be the subset of A whose coordinates are rational numbers. Define $\chi_X : A \to \mathbb{R}$ by

$$\chi_X(x) = \begin{cases} 1, & \text{if } x \in X, \\ 0, & \text{if } x \notin X. \end{cases}$$

The function χ_X is called the characteristic function of X. For every partition P of the block A and every $B \in P$, we have $m_B = 0$ and $M_B = 1$. It follows that $s(\chi_X; P) = 0$ and $S(\chi_X; P) = \operatorname{vol} A$, for every partition P. Thus the characteristic function χ_X is not integrable.

Proposition 3.14 *Let $f, g : A \to \mathbb{R}$ be integrable functions. Then:*

(a) $f + g$ is integrable and, moreover,

$$\int_A (f(x) + g(x))\, dx = \int_A f(x)\, dx + \int_A g(x)\, dx.$$

(b) For every $c \in \mathbb{R}$, the function $c\, f$ is integrable and

$$\int_A c\, f(x)\, dx = c \int_A f(x)\, dx.$$

(c) If $f(x) \geq 0$ for all $x \in A$, then $\int_A f(x)\, dx \geq 0$. Equivalently, if $f(x) \leq g(x)$ for all $x \in A$, then

$$\int_A f(x)\, dx \leq \int_A g(x)\, dx.$$

(d) The function $|f(x)|$ is integrable and

$$\left| \int_A f(x)\, dx \right| \leq \int_A |f(x)|\, dx.$$

In particular, if $|f(x)| \leq C$ for all $x \in A$, then

$$\left| \int_A f(x)\, dx \right| \leq C \operatorname{vol} A.$$

(e) If f is continuous, there exists $\xi \in A$ such that

$$\int_A f(x)\,dx = f(\xi)\,vol\,A.$$

Proof:

(a) For every $B \subset A$ we have $m_B(f) + m_B(g) \leq m_B(f+g)$ and $M_B(f+g) \leq M_B(f) + M_B(g)$. Indeed, note that

$$\begin{array}{lll} m_B(f)+m_B(g) & \text{is a lower bound of the set} & \{(f+g)(x);\ x \in B\}, \\ M_B(f)+M_B(g) & \text{is an upper bound of the set} & \{(f+g)(x);\ x \in B\}, \end{array}$$

since

$$\begin{array}{lll} m_B(f) \leq f(x) \text{ and } m_B(g) \leq g(x) & \Rightarrow & m_B(f)+m_B(g) \leq f(x)+g(x), \\ M_B(f) \geq f(x) \text{ and } M_B(g) \geq g(x) & \Rightarrow & M_B(f)+M_B(g) \geq f(x)+g(x), \end{array}$$

for all $x \in B$.

It follows that, for any partitions P, Q of the block A, we have

$$s(f;P) + s(g;P) \leq s(f+g;P) \leq S(f+g;Q) \leq S(f;Q) + S(g;Q).$$

Consequently,

$$\begin{gathered} \sup\{s(f;P);\ P \in \wp\} + \sup\{s(g;P);\ P \in \wp\} \leq \\ \leq \sup\{s(f+g;P);\ P \in \wp\} \leq \inf\{S(f+g;Q);\ Q \in \wp\} \leq \\ \leq \inf\{S(f;Q);\ Q \in \wp\} + \inf\{S(g;Q);\ Q \in \wp\}, \end{gathered}$$

or, equivalently,

$$\begin{gathered} \int_A f(x)\,dx + \int_A g(x)\,dx \leq \\ \leq \underline{\int_A}[f(x)+g(x)]\,dx \leq \overline{\int_A}[f(x)+g(x)]\,dx \leq \\ \leq \int_A f(x)\,dx + \int_A g(x)\,dx, \end{gathered}$$

which gives the desired equality and shows that $f+g$ is integrable.

(b) We have

$$m_B(c\,f) = \inf\{(c\,f)(x);\ x \in B\} = \begin{cases} c\,\inf\{f(x);\ x \in B\} = c\,m_B(f), & \text{if } c \geq 0, \\ c\,\sup\{f(x);\ x \in B\} = c\,M_B(f), & \text{if } c < 0, \end{cases}$$

and similarly

$$M_B(c\,f) = \begin{cases} c\,M_B(f), & \text{if } c \geq 0, \\ c\,m_B(f), & \text{if } c < 0. \end{cases}$$

Thus

$$s(c\,f;P) = c\,s(f;P) \quad \text{and} \quad S(c\,f;P) = c\,S(f;P), \quad \text{if } c \geq 0,$$
$$s(c\,f;P) = c\,S(f;P) \quad \text{and} \quad S(c\,f;P) = c\,s(f;P), \quad \text{if } c < 0.$$

Hence we have two cases:

(i) $c < 0$:

$$\begin{aligned}
\overline{\int_A}(c\,f)(x)\,dx &= \inf\{S(c\,f;P);\ P \in \wp\} = \inf\{c\,s(f;P);\ P \in \wp\} \\
&= c\,\sup\{s(f;P);\ P \in \wp\} = c\int_A f(x)\,dx, \\
\underline{\int_A}(c\,f)(x)\,dx &= \sup\{s(c\,f;P);\ P \in \wp\} = \sup\{c\,S(f;P);\ P \in \wp\} \\
&= c\,\inf\{S(f;P);\ P \in \wp\} = c\int_A f(x)\,dx.
\end{aligned}$$

(ii) $c > 0$:

$$\begin{aligned}
\overline{\int_A}(c\,f)(x)\,dx &= \inf\{S(c\,f;P);\ P \in \wp\} = \inf\{c\,S(f;P);\ P \in \wp\} \\
&= c\,\inf\{S(f;P);\ P \in \wp\} = c\int_A f(x)\,dx, \\
\underline{\int_A}(c\,f)(x)\,dx &= \sup\{s(c\,f;P);\ P \in \wp\} = \sup\{c\,s(f;P);\ P \in \wp\} \\
&= c\,\sup\{s(f;P);\ P \in \wp\} = c\int_A f(x)\,dx.
\end{aligned}$$

In either case, $c\,f$ is integrable and

$$\int_A (c\,f)(x)\,dx = c\int_A f(x)\,dx.$$

(c) If $f(x) \geq 0$ for all $x \in A$, then $m_B \geq 0$ for every block $B \subset A$, hence $s(f;P) = \sum_{B\in P} m_B \operatorname{vol} B \geq 0$ for all $P \in \wp$. Consequently, $\sup\{s(f;P);\ P \in \wp\} \geq 0$, that is, $\int_A f(x)\,dx \geq 0$.

If $f(x) \leq g(x)$ for all $x \in A$, then $(g - f)(x) \geq 0$ for all x, and therefore

$$\int_A (g - f)(x)\,dx = \int_A g(x)\,dx - \int_A f(x)\,dx \geq 0,$$

that is,

$$\int_A f(x)\,dx \leq \int_A g(x)\,dx.$$

(d) Let $P \in \wp$ be a generic partition of A, and let $B \in P$ be any block of P. Then

$$\begin{aligned}\omega(f;B) &= \sup\{|f(x)-f(y)|;\ x,y\in B\},\\ \omega(|f|;B) &= \sup\big\{\,\big||f(x)|-|f(y)|\big|;\ x,y\in B\big\}.\end{aligned}$$

However, since $||f(x)|-|f(y)|| \le |f(x)-f(y)|$, it follows that $\omega(|f|;B) \le \omega(f;B)$ for every $B \in P$.

Thus

$$\sum_{B\in P}\omega(|f|;B)\operatorname{vol}B \le \sum_{B\in P}\omega(f;B)\operatorname{vol}B,$$

and since f is integrable, (3.13) implies that $|f|$ is integrable. Moreover, from $-|f(x)| \le f(x) \le |f(x)|$ we obtain, by item (c),

$$-\int_A |f(x)|\,dx \le \int_A f(x)\,dx \le \int_A |f(x)|\,dx,$$

which means that

$$\Big|\int_A f(x)\,dx\Big| \le \int_A |f(x)|\,dx.$$

If $|f(x)| \le C$ for all $x \in A$, then

$$\Big|\int_A f(x)\,dx\Big| \le \int_A |f(x)|\,dx \le \int_A C\,dx = C\int_A dx = C\operatorname{vol}A.$$

(e) Let $m = \inf\{f(x);\ x\in A\}$ and $M = \sup\{f(x);\ x\in A\}$, which exist since f is continuous on A. As f is integrable, we have

$$m\operatorname{vol}A \le \int_A f(x)\,dx \le M\operatorname{vol}A.$$

Consequently,

$$m \le \frac{\int_A f(x)\,dx}{\operatorname{vol}A} \le M.$$

Hence, by the Intermediate Value Theorem (and the fact that A is convex), there exists $\xi \in A$ such that

$$\frac{\int_A f(x)\,dx}{\operatorname{vol}A} = f(\xi).$$

□

The first two items of the proposition above show that the set of integrable functions on a block $A \subset \mathbb{R}^n$ is a real vector space, and the map $f \mapsto \int_A f(x)\,dx$ is a linear functional on that space. The third item says that this functional is positive, and the fourth implies that the functional is continuous when we consider uniform convergence in the space of integrable functions. In other words:

Let

$$\mathcal{R}(A) = \{\, f : A \to \mathbb{R};\ f \text{ Riemann integrable}\,\}$$

and consider the functional

$$\begin{aligned} T:\mathcal{R}(A) &\longrightarrow \mathbb{R} \\ f &\longmapsto T(f)=\int_A f(x)\,dx. \end{aligned}$$

If we equip $\mathcal{R}(A)$ with the metric $d(f,g)=\sup\{|f(x)-g(x)|;\ x\in A\}$, then

$$|T(f)|\leq \int_A |f(x)|\,dx \leq \sup_{x\in A}|f(x)|\operatorname{vol}A = C\|f\|,$$

where $\|f\|=\sup_{x\in A}|f(x)|$ and $C=\operatorname{vol}A$. Therefore T is continuous.

Item (c) is further complemented by the remark that, if $f\geq 0$, then $\int_A f(x)\,dx=0$ can only occur when $f(x)=0$ at every point $x\in A$ at which f is continuous.

Every integrable function $f:A\to\mathbb{R}$ can be written as the difference $f=f_+-f_-$ of two nonâ€“negative integrable functions. The function $f_+:A\to\mathbb{R}$ is called the positive part of f, whilst the function $f_-:A\to\mathbb{R}$ is its negative part. For every $x\in A$ we set

$$f_+(x)=\max\{f(x),0\} \qquad \text{and} \qquad f_-(x)=-\min\{f(x),0\}. \tag{3.14}$$

Thus, when $f(x)\geq 0$ we have $f_+(x)=f(x)$ and $f_-(x)=0$. On the other hand, if $f(x)\leq 0$, then $f_+(x)=0$ and $f_-(x)=-f(x)$. The equality $f=f_+-f_-$ is evident. Hence, if f_+ and f_- are integrable, then f is also integrable. Conversely, since

$$f_+(x)=\frac{f(x)+|f(x)|}{2} \qquad \text{and} \qquad f_-(x)=\frac{|f(x)|-f(x)}{2}, \tag{3.15}$$

for every $x\in A$, the integrability of f implies that of f_+ and f_-, by item (d) of the previous proposition.

<u>Note 7:</u> Let A,B be blocks in $\mathbb{R}^n$ with $B\subset A$, and let $\chi_B:A\to\mathbb{R}$ be the characteristic function of B, that is,

$$\chi_B(x)=\begin{cases} 1, & \text{if } x\in B,\\ 0, & \text{if } x\in A \text{ and } x\notin B.\end{cases}$$

We claim that χ_B is integrable and that

$$\int_A \chi_B(x)\,dx=\operatorname{vol}B.$$

Indeed, let P_0 be a partition of A that has B as one of its blocks. By Note 5, in order to obtain the lower and upper integrals of a bounded function it suffices to consider the partitions that refine P_0, that is,

$$\underline{\int_A}\chi_B(x)\,dx=\sup_{P\supset P_0}s(\chi_B;P), \qquad \overline{\int_A}\chi_B(x)\,dx=\inf_{P\supset P_0}S(\chi_B;P).$$

If $P \supset P_0$ and C is a block of the partition P, then

$$s(\chi_B; P) = \sum_{C \in P} m_C \operatorname{vol} C = \operatorname{vol} B + \sum_{C \in P \setminus \{B\}} m_C \operatorname{vol} C = \operatorname{vol} B,$$

since $m_C = \inf\{\chi_B(x);\ x \in C\} = 0$ for all $C \in P \setminus \{B\}$.

Similarly,

$$S(\chi_B; P) = \sum_{C \in P} M_C \operatorname{vol} C = \operatorname{vol} B + \sum_{C \in P \setminus \{B\}} M_C \operatorname{vol} C = \operatorname{vol} B,$$

because $M_C = \sup\{\chi_B(x);\ x \in C\} = 0$ for all $C \in P \setminus \{B\}$.

It follows that

$$\underline{\int_A} \chi_B(x)\, dx = \overline{\int_A} \chi_B(x)\, dx = \operatorname{vol} B.$$

3.2 Sets of Measure Zero

Definition 3.15 *We say that a set $X \subset \mathbb{R}^n$ has measure zero, and we write* $\operatorname{med} X = 0$*, if, for every $\epsilon > 0$, it is possible to find a sequence of open nâ€“dimensional cubes $C_1, C_2, \ldots$ such that*

$$X \subset \bigcup_{i=1}^{+\infty} C_i \quad \textit{and} \quad \sum_{i=1}^{+\infty} vol C_i < \epsilon.$$

When necessary, to be more precise, we shall say in this situation that X has nâ€“dimensional measure zero or that X has measure zero in $\mathbb{R}^n$.

Example 1. Let $X = \{r_1, r_2, \ldots, r_n, \ldots\}$ be a countable subset of the real line $\mathbb{R}$. For each $\epsilon > 0$ consider the intervals

$$I_n = \left\{x \in \mathbb{R};\ r_n - \frac{\epsilon}{2^{n+2}} < x < r_n + \frac{\epsilon}{2^{n+2}}\right\}, \quad n = 1, 2, \ldots.$$

The family $\{I_n\}_{n \in \mathbb{N}}$ is a countable covering of X and, in addition,

$$\sum_{n=1}^{+\infty} \operatorname{vol} I_n = \sum_{n=1}^{+\infty} \frac{\epsilon}{2^{n+2}} < \epsilon.$$

We conclude that any countable subset of the real line has measure zero. As a consequence, any finite set has measure zero.

We now present a list of nine propositions on sets of measure zero.

Proposition 3.16 *Every subset of a set of measure zero has measure zero.*

Proof: Let $Z \subset \mathbb{R}^n$ be a set of measure zero and let $X \subset Z$. By the definition above, for each $\epsilon > 0$ there exists a sequence of open nâ€“dimensional cubes $C_1, C_2, \ldots$ such that $Z \subset \bigcup_{i=1}^{+\infty} C_i$ and $\sum_{i=1}^{+\infty} \operatorname{vol} C_i < \epsilon$. Since $X \subset Z$, the same sequence of cubes shows that X also has measure zero. □

Proposition 3.17 *Every countable union of sets of measure zero is again a set of measure zero.*

Proof: Let $Z = \bigcup_{\alpha \in I} Z_\alpha$, with $I \subset \mathbb{N}$, and suppose $\operatorname{med}(Z_\alpha) = 0$ for all $\alpha \in I$. Given $\epsilon > 0$, we can obtain, for each $\alpha \in I$, a sequence of open cubes $C_{\alpha 1}, C_{\alpha 2}, \ldots$ such that $Z_\alpha \subset \bigcup_{j=1}^{+\infty} C_{\alpha j}$ and

$$\sum_{j=1}^{+\infty} \operatorname{vol}(C_{\alpha j}) < \frac{\epsilon}{2^\alpha}.$$

It follows that Z is contained in the countable union of all the $C_{\alpha j}$. Moreover, given any finite subset $F \subset I \times \mathbb{N}$, there exist $m \in I$ and $n \in \mathbb{N}$ such that $(\alpha, j) \in F \Rightarrow \alpha \leq m$ and $j \leq n$, and hence

$$\sum_{(\alpha,j) \in F} \operatorname{vol}(C_{\alpha j}) \leq \sum_{\alpha=1}^{m} \Big[\sum_{j=1}^{n} \operatorname{vol}(C_{\alpha j}) \Big] < \sum_{\alpha=1}^{m} \frac{\epsilon}{2^\alpha} < \epsilon.$$

Therefore, regardless of the way in which the $C_{\alpha j}$ are enumerated in a sequence, we have

$$\sum_{\{\alpha,j\}} \operatorname{vol}(C_{\alpha j}) \leq \epsilon.$$

Hence $\operatorname{med}(Z) = 0$. □

In particular, since every point has measure zero, it follows that any countable subset of $\mathbb{R}^n$ has measure zero in $\mathbb{R}^n$.

Proposition 3.18 *Let $A \subset \mathbb{R}^n$ be a block. Given any countable cover $A \subset \bigcup_{i=1}^{+\infty} B_i$ by open blocks, we have*

$$\sum_{i=1}^{+\infty} vol\, B_i \geq vol\, A.$$

Proof: Assume first that A is closed. Being bounded, A is compact. Since $\{B_i\}_{i \in \mathbb{N}}$ is an open cover, there exist $B_1, \ldots, B_k$ in this cover such that $A \subset \bigcup_{i=1}^{k} B_i$. Take a closed block B such that $B_1 \subset B, \ldots, B_k \subset B$.

We have

$$\chi_A \leq \chi_{B_1 \cup \cdots \cup B_k} \leq \sum_{i=1}^{k} \chi_{B_i}.$$

It follows from Note 7 of the previous section and from Proposition 3.14 (items (a) and (c)) that

$$\begin{aligned}\operatorname{vol} A &= \int_B \chi_A(x)\,dx \leq \int_B \chi_{B_1}(x)\,dx + \cdots + \int_B \chi_{B_k}(x)\,dx \\ &= \operatorname{vol} B_1 + \cdots + \operatorname{vol} B_k \leq \sum_{i=1}^{k} \operatorname{vol} B_i.\end{aligned}$$

If the block A is open, then for every closed block D contained in A we have

$$\operatorname{vol} D \leq \sum_{i=1}^{+\infty} \operatorname{vol} B_i.$$

However,

$$\operatorname{vol} A = \sup\{\operatorname{vol} D;\ D \text{ is a closed block contained in } A\}. \tag{3.16}$$

Indeed, let

$$S = \sup\{\operatorname{vol} D;\ D \text{ is a closed block contained in } A\}.$$

Clearly $S \leq \operatorname{vol} A$, since $\operatorname{vol} D \leq \operatorname{vol} A$ for all closed $D \subset A$. It remains to show that $S \geq \operatorname{vol} A$. In fact, let $A = \prod_{j=1}^{n}(a_j, b_j)$. Note that

$$\lim_{k\to+\infty} \prod_{j=1}^{n}\Big(b_j - a_j - \frac{1}{k}\Big) = \prod_{j=1}^{n}(b_j - a_j) = \operatorname{vol} A.$$

Thus, given $\epsilon > 0$, there exists $k_0 \in \mathbb{N}$ such that, for all $k \geq k_0$,

$$\left|\operatorname{vol} A - \prod_{j=1}^{n}\Big(b_j - a_j - \frac{1}{k}\Big)\right| < \epsilon,$$

that is,

$$\operatorname{vol} A < \epsilon + \prod_{j=1}^{n}\Big(b_j - a_j - \frac{1}{k}\Big).$$

Setting

$$D_{k(\epsilon)} = \prod_{j=1}^{n}\Big[a_j + \frac{1}{2k},\ b_j - \frac{1}{2k}\Big],$$

we have $D_{k(\epsilon)} \subset A$ and

$$\operatorname{vol} A < \operatorname{vol} D_{k(\epsilon)} + \epsilon \leq S + \epsilon.$$

By the arbitrariness of ϵ we deduce that $\operatorname{vol} A \leq S$, hence $\operatorname{vol} A = S$.

Since $\sum_{i=1}^{+\infty} \operatorname{vol} B_i$ is an upper bound for the set

$$\{\operatorname{vol} D;\ D \text{ is a closed block contained in } A\},$$

it follows from (3.16) that $\operatorname{vol} A \leq \sum_{i=1}^{+\infty} \operatorname{vol} B_i$. □

Note 1: It follows from the previous result that A does not have measure zero, since there exists $\epsilon_0 = \operatorname{vol} A$ such that, for every countable cover $A \subset \bigcup_{i=1}^{+\infty} B_i$ by open cubes, we have

$$\sum_{i=1}^{+\infty} \operatorname{vol} B_i \geq \operatorname{vol} A.$$

From Propositions 3.16 and 3.18 it follows that every set X of measure zero has empty interior; otherwise, if $\operatorname{int} X \neq \emptyset$, then X would contain some block, and hence X would not have measure zero.

Proposition 3.19 *In the definition of a set of measure zero, we may use closed cubes instead of open cubes.*

Proof: Suppose $X \subset \bigcup_i C_i$, where each C_i is a closed cube, and $\sum_i \operatorname{vol} C_i < \epsilon$.

On the other hand, let $C = \prod_{j=1}^{n} [a_j, a_{j+1}]$ be a closed cube of edge length 1. For each $\delta > 0$, the set

$$D(\delta) = \prod_{j=1}^{n} \left(a_j - \frac{\delta}{2},\ a_j + 1 + \frac{\delta}{2} \right)$$

is an open cube of edge length $(1 + \delta)$ that contains C. Clearly,

$$\lim_{\delta \to 0} \operatorname{vol} D(\delta) = \lim_{\delta \to 0} \prod_{j=1}^{n} \left[(a_j + 1 + \tfrac{\delta}{2}) - (a_j - \tfrac{\delta}{2}) \right] = 1^n = \operatorname{vol} C.$$

It follows that, given $\epsilon' = \operatorname{vol} C > 0$, there exists $\delta_0 > 0$ such that, if $0 < \delta < \delta_0$, then $|\operatorname{vol} D(\delta) - \operatorname{vol} C| < \operatorname{vol} C$. Since $D(\delta) \supset C$, we have $\operatorname{vol} D(\delta) - \operatorname{vol} C \geq 0$. Thus there exists $\delta > 0$ such that $\operatorname{vol} D(\delta) < 2 \operatorname{vol} C$.

Therefore, for each i we can choose an open cube D_i containing C_i with $\operatorname{vol} D_i < 2 \operatorname{vol} C_i$, so that $X \subset \bigcup_i D_i$ and, in addition,

$$\sum_i \operatorname{vol} D_i < \sum_i 2 \operatorname{vol} C_i < 2\epsilon.$$

□

Proposition 3.20 *Let $X \subset \mathbb{R}^n$ be such that, for every $\epsilon > 0$, there exists a sequence of closed blocks $A_1, \ldots, A_i, \ldots$ with $X \subset \bigcup_i A_i$ and $\sum_i \operatorname{vol} A_i < \epsilon$. Then $\operatorname{med} X = 0$. In other words, in the definition of a set of measure zero we may use closed blocks instead of cubes.*

Proof: We first show the following: given a closed block $A \subset \mathbb{R}^n$ and $\epsilon > 0$, there exist closed cubes $C_1, \ldots, C_k$ such that $A \subset \bigcup_{i=1}^k C_i$ and $\sum_{i=1}^k \operatorname{vol} C_i < \operatorname{vol} A + \epsilon$.

Indeed, let $A = \prod_{j=1}^n [a_j, b_j]$. Take $q \in \mathbb{N}$. For every $j = 1, \ldots, n$ there exists an integer $p_j \geq 0$ such that

$$p_j < q(b_j - a_j) \leq p_j + 1,$$

that is,

$$\frac{p_j}{q} < b_j - a_j \leq \frac{p_j + 1}{q}.$$

The block

$$A' = \prod_{j=1}^n \Big[a_j,\, a_j + \frac{p_j + 1}{q}\Big]$$

contains A, since $b_j \leq a_j + \dfrac{p_j + 1}{q}$. Moreover, A' admits a natural partition $P = P_1 \times \cdots \times P_n$, where

$$P_j = \Big\{a_j,\; a_j + \frac{1}{q},\; a_j + \frac{2}{q}, \ldots, a_j + \frac{p_j + 1}{q}\Big\}$$

is a partition of the interval $\big[a_j,\, a_j + \dfrac{p_j + 1}{q}\big]$.

Note that the blocks associated with the partition P are cubes C_i, all with edges of length $\dfrac{1}{q}$. We have

$$\begin{aligned} A \subset A' &= \bigcup_i C_i, \quad \text{with} \\ \sum_i \operatorname{vol} C_i = \operatorname{vol} A' &= \prod_{j=1}^n \Big(\frac{p_j}{q} + \frac{1}{q}\Big) < \prod_{j=1}^n \Big(b_j - a_j + \frac{1}{q}\Big). \end{aligned}$$

Since

$$\lim_{q \to +\infty} \prod_{j=1}^n \Big(b_j - a_j + \frac{1}{q}\Big) = \prod_{j=1}^n (b_j - a_j) = \operatorname{vol} A,$$

for the given $\epsilon > 0$ there exists $q_0 \in \mathbb{N}$ such that, for all $q \geq q_0$,

$$\left| \prod_{j=1}^n \Big(b_j - a_j + \frac{1}{q}\Big) - \operatorname{vol} A \right| < \epsilon.$$

Since $A' \supset A$, we have $\operatorname{vol} A' \geq \operatorname{vol} A$ and therefore

$$\prod_{j=1}^n \Big(b_j - a_j + \frac{1}{q}\Big) < \operatorname{vol} A + \epsilon$$

for sufficiently large q. Consequently,

$$\sum_i \operatorname{vol} C_i < \operatorname{vol} A + \epsilon.$$

Returning to the statement of the proposition, for each $i \in \mathbb{N}$ we choose closed cubes C_{ij} such that $A_i \subset \bigcup_j C_{ij}$ and

$$\sum_j \operatorname{vol} C_{ij} < \operatorname{vol} A_i + \frac{\epsilon}{2^i}.$$

Then $X \subset \bigcup_{i,j} C_{ij}$ and, moreover,

$$\sum_{i,j} \operatorname{vol} C_{ij} < \sum_i \operatorname{vol} A_i + \sum_i \frac{\epsilon}{2^i} < \epsilon + \epsilon = 2\epsilon.$$

By the previous proposition, we conclude that $\operatorname{med}(X) = 0$. □

Note 2: A fortiori, we may use open blocks to define measure zero.

Indeed, let $X \subset \bigcup_i C_i$, where the C_i are open blocks and $\sum_i \operatorname{vol} C_i < \epsilon$. Clearly $X \subset \bigcup_i \overline{C_i}$ and, in addition,

$$\sum_i \operatorname{vol} \overline{C_i} = \sum_i \operatorname{vol} C_i < \epsilon.$$

On some occasions, in order to prove that a set has measure zero, instead of covering it with a sequence of blocks whose sum of volumes can be made arbitrarily small, it may be convenient to leave uncovered a subset that we already know to have measure zero. This is the content of the proposition below. It is worth noting that the set Y may depend on ϵ.

Proposition 3.21 *Let $X \subset \mathbb{R}^n$. Suppose that, for every $\epsilon > 0$, there exists a sequence of blocks A_i (open or closed) such that $\sum_i \operatorname{vol} A_i < \epsilon$ and*

$$X \subset \Big(\bigcup_i A_i\Big) \cup Y,$$

where $\operatorname{med}(Y) = 0$. *Then X has measure zero.*

Proof: Given $\epsilon > 0$, by hypothesis we obtain blocks A_i and a set Y of measure zero such that

$$\sum_i \operatorname{vol} A_i < \frac{\epsilon}{2} \quad \text{and} \quad X \subset \Big(\bigcup_i A_i\Big) \cup Y.$$

Since $\operatorname{med}(Y) = 0$, we can find blocks B_j such that $Y \subset \bigcup_j B_j$ and, moreover,

$$\sum_j \operatorname{vol} B_j < \frac{\epsilon}{2}.$$

It follows that

$$X \subset \Big(\bigcup_i A_i\Big) \cup \Big(\bigcup_j B_j\Big)$$

and

$$\sum_i \operatorname{vol} A_i + \sum_j \operatorname{vol} B_j < \epsilon.$$

Hence X has measure zero. □

Definition 3.22 *A map $f : X \subset \mathbb{R}^n \to \mathbb{R}^m$ is said to be* locally Lipschitz *if, for every $x \in X$, there exist an open set $V_x \subset \mathbb{R}^n$ containing x and a constant $k_x > 0$ such that, whenever $y, z \in V_x$, one has*

$$\|f(y) - f(z)\| \leq k_x \|y - z\|.$$

In other words, there exists an open covering $X \subset \bigcup V_x$ such that each restriction $f|_{V_x \cap X}$ is Lipschitz.

The next proposition shows that the notion of a set of measure zero is invariant under locally Lipschitz maps. It is important to note that X and $f(X)$ are required to lie in the same Euclidean space $\mathbb{R}^n$.

Proposition 3.23 *If $X \subset \mathbb{R}^n$ has measure zero and $f : X \to \mathbb{R}^n$ is locally Lipschitz, then $f(X)$ has measure zero in $\mathbb{R}^n$.*

Proof: First consider the case in which f is globally Lipschitz. Then there exists $K > 0$ such that

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

for all x, y. We shall work with the maximum norm on $\mathbb{R}^n$. Thus, given $\epsilon > 0$, there exists a countable covering

$$X \subset \bigcup_i C_i,$$

where each C_i is an open cube of edge length a_i, and

$$\sum_{i=1}^{+\infty} a_i^n < \frac{\epsilon}{K^n}.$$

For each $i \in \mathbb{N}$, take $x, y \in C_i \cap X$. Write $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$. Then

$$\|x - y\| = \max\{|x_1 - y_1|, \ldots, |x_n - y_n|\} < a_i,$$

which implies

$$\|f(x) - f(y)\| \leq K\|x - y\| < K\, a_i.$$

It follows that each of the n coordinate projections of $f(X \cap C_i)$ is contained in an interval of length $K\, a_i$, since if $f(x) = (z_1, \ldots, z_n)$ and $f(y) = (w_1, \ldots, w_n)$, then

$$\|f(x) - f(y)\| = \max\{|z_1 - w_1|, \ldots, |z_n - w_n|\} < K\, a_i.$$

Hence $f(X \cap C_i)$ is contained in the Cartesian product of these intervals, which is a cube D_i of volume $K^n a_i^n$.

To fix ideas, consider the picture below in the particular case $n = 2$.

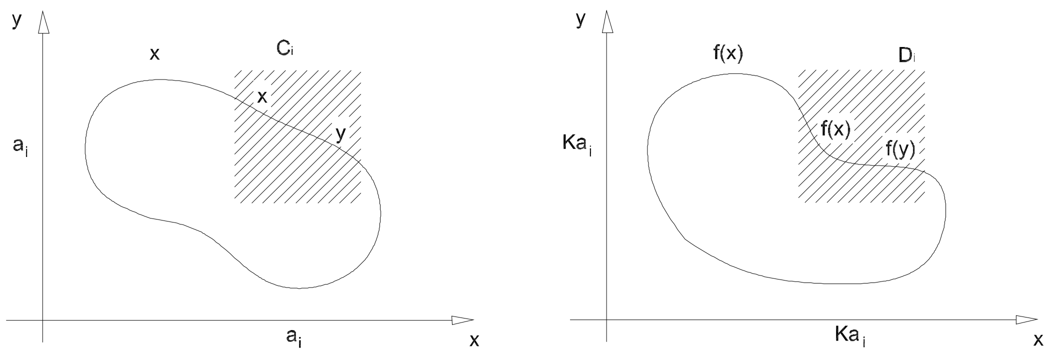

Figure 3.6:

Since $X = \bigcup_i (X \cap C_i)$, we have

$$f(X) = \bigcup_i f(X \cap C_i) \subset \bigcup_i D_i,$$

and

$$\sum_i \operatorname{vol} D_i = \sum_i K^n a_i^n = K^n \sum_i a_i^n < K^n \frac{\epsilon}{K^n} = \epsilon.$$

Therefore $\operatorname{med}(f(X)) = 0$.

In the general case, we have $X \subset \bigcup V_x$, where each V_x is open and each restriction $f\ |_{V_x \cap X}$ is Lipschitz. By Lindelöf's theorem (Induced Topology, Chapter 1, Proposition 1.45), X admits a countable subcover, that is,

$$X \subset \bigcup_{j \in I} V_j,$$

with $I \subset \mathbb{N}$. By the first part of the proof, $f(V_j \cap X)$ has measure zero for each $j \in \mathbb{N}$. Hence

$$f(X) = \bigcup_{j=1}^{+\infty} f(V_j \cap X)$$

is a countable union of sets of measure zero, and therefore $\operatorname{med}(f(X)) = 0$. □

Proposition 3.24 *Let $f : U \subset \mathbb{R}^n \to \mathbb{R}^n$ be of class C^1 on the open set U. If $X \subset U$ has measure zero in $\mathbb{R}^n$, then $f(X) \subset \mathbb{R}^n$ also has measure zero.*

Proof: For each $x \in X$, let V_x be a ball centred at x with $V_x \subset U$, and set

$$k_x = \sup\{\|f'(y)\|;\ y \in \overline{V_x}\}.$$

Recall that the derivative map f' is given by

$$\begin{aligned} f' : U &\longrightarrow \pounds(\mathbb{R}^n) \\ y &\longmapsto f'(y), \end{aligned}$$

and that on $\pounds(\mathbb{R}^n)$ we are using the operator norm (supremum norm). Since f is of class C^1, the map f' is continuous; and as $\overline{V_x}$ is compact and contained in U, the set $f'(\overline{V_x})$ is compact in $\pounds(\mathbb{R}^n)$. Hence there exists $C > 0$ such that $\|f'(y)\| \leq C$ for all $y \in \overline{V_x}$, so k_x is well defined.

On the other hand, since V_x is convex and $\|f'(y)\| \leq C$ for all $y \in V_x$, the mean value inequality yields

$$\|f(x) - f(y)\| \leq k_x \|x - y\| \quad \forall\, x, y \in V_x.$$

This implies that f is locally Lipschitz. Therefore, by Proposition 3.23, we obtain the desired conclusion. □

Note 3: Let $n < m$ and $a \in \mathbb{R}^{m-n}$. Then every subset

$$X \subset \mathbb{R}^n \times \{a\} \subset \mathbb{R}^m$$

has measure zero in $\mathbb{R}^m$. To fix ideas, consider Figure 3.7.

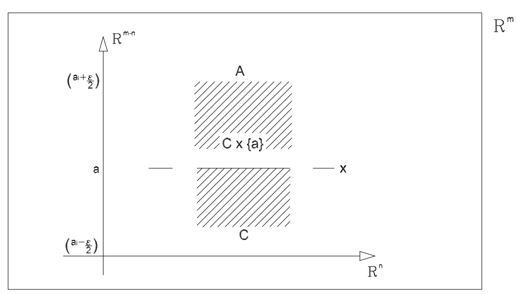

Figure 3.7:

Indeed, since $\mathbb{R}^n \times \{a\}$ is a union of nâ€“dimensional cubes, it suffices to prove this for one such cube, say $C \times \{a\}$. Now, for every $\epsilon > 0$, the set $C \times \{a\}$ is contained in the mâ€“dimensional block

$$A = C \times \left(a_i - \frac{\epsilon}{2}, a_i + \frac{\epsilon}{2}\right)^{m-n} \quad \text{with} \quad \operatorname{vol} A = \epsilon^{m-n} \operatorname{vol} C.$$

In view of Proposition 3.20 we have $\mathrm{med}(C \times \{a\}) = 0$ and, consequently, $\mathrm{med}(X) = 0$.

Proposition 3.25 *If $n < m$ and $f : U \to \mathbb{R}^m$ is of class C^1 on the open set $U \subset \mathbb{R}^n$, then $f(U)$ has measure zero in $\mathbb{R}^m$.*

Proof: Consider $0 \in \mathbb{R}^{m-n}$. By the previous note, $U \times \{0\}$ has measure zero in $\mathbb{R}^m$. On the open set

$$W = U \times \mathbb{R}^{m-n} \subset \mathbb{R}^m$$

define the map

$$\begin{array}{rcl} g : W \subset \mathbb{R}^m & \longrightarrow & \mathbb{R}^m \\ (x,y) & \longmapsto & g(x,y) = f(x). \end{array}$$

Clearly, g is of class C^1 and, moreover,

$$g\,|_{U\times\{0\}} \equiv f, \quad \text{that is,} \quad g(U \times \{0\}) = f(U).$$

By Proposition 3.24 it follows that $f(U)$ has measure zero in $\mathbb{R}^m$. □

Proposition 3.25 shows, in particular, that there are no Peano curves of class C^1, that is, the unpleasant phenomenon of a map f, defined on a subset $X \subset \mathbb{R}^n$, whose image contains a cube of $\mathbb{R}^m$, with $n < m$, may occur in class C^0 but not in class C^1.

Definition 3.26 *We say that a set $X \subset \mathbb{R}^n$ is* locally of measure zero *if, for each $x \in X$, there exists an open set $V_x \subset \mathbb{R}^n$ containing x such that* $\operatorname{med}(V_x \cap X) = 0$.

Note 4: Let $X \subset \mathbb{R}^n$ be such that X is locally of measure zero. From the open covering $X \subset \bigcup_{x\in X} V_x$, Lindelöf's theorem yields a countable subcover

$$X \subset \bigcup_i V_i.$$

Thus

$$X = \bigcup_i (V_i \cap X)$$

is a countable union of sets of measure zero; hence $\operatorname{med}(X) = 0$.

Therefore, a set $X \subset \mathbb{R}^n$ is locally of measure zero if and only if it has measure zero.

Definition 3.27 *An n-dimensional C^k surface in $\mathbb{R}^m$ is a set $S \subset \mathbb{R}^m$ that can be covered by a family of open sets $U \subset \mathbb{R}^m$ such that each $V = U \cap S$ admits a parametrisation $\varphi : V_0 \to V$, of class C^k, defined on an open set $V_0 \subset \mathbb{R}^n$. Each such set $V = U \cap S$ is open in S. For each $p \in S$, V is called a parametrised neighbourhood of p.*

Thus, an n-dimensional C^k surface in $\mathbb{R}^m$ is a subset such that each of its points has a parametrised neighbourhood via an n-dimensional C^k parametrisation.

Note 5: Let $S \subset \mathbb{R}^m$ be a C^1 surface of dimension $n < m$ in $\mathbb{R}^m$. Given a parametrisation $\varphi : V_0 \to V$ of S, it follows from Proposition 3.25 that the parametrised neighbourhood $V \subset S$ has measure zero in $\mathbb{R}^m$. Since $V = A \cap S$, where A is open in $\mathbb{R}^m$, we see that S is locally of measure zero and, consequently, $\operatorname{med}(S) = 0$ in $\mathbb{R}^m$. More generally, if $X \subset \mathbb{R}^m$ is a countable union of C^1 surfaces of dimension $< m$, then $\operatorname{med}(X) = 0$ in $\mathbb{R}^m$.

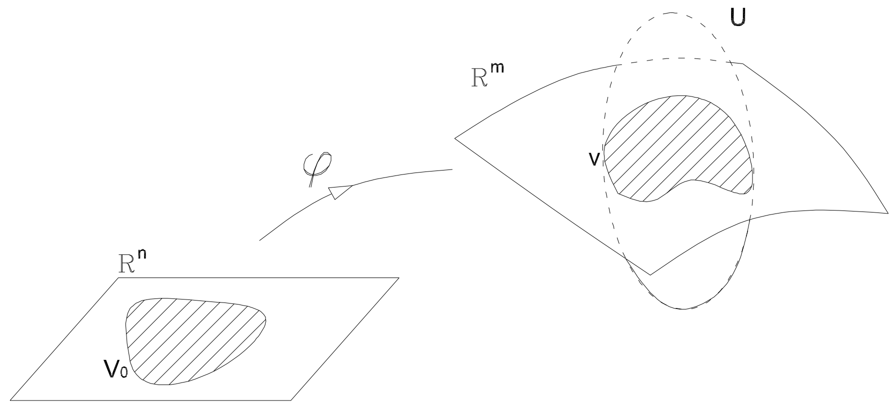

Figure 3.8:

3.3 Characterisation of Integrable Functions

Let $X \subset \mathbb{R}^n$ and $f : X \to \mathbb{R}$ be a bounded function. Fix $x \in X$ and, for each $\delta > 0$, set

$$\Omega(\delta) = \omega[f; X \cap B_\delta(x)] = \sup\{|f(y) - f(z)|;\ y, z \in X \cap B_\delta(x)\}, \tag{3.17}$$

which is the oscillation of f on the set of points of X whose distance from x is less than δ. This defines a nonâ€“negative function

$$\begin{aligned} \Omega : (0, +\infty) &\longrightarrow \mathbb{R} \\ \delta &\longmapsto \Omega(\delta). \end{aligned} \tag{3.18}$$

Since f is bounded, so is Ω. Moreover, if $\delta \leq \delta'$ then $\Omega(\delta) \leq \Omega(\delta')$. Hence the limit

$$\omega(f; x) = \lim_{\delta\to 0} \omega[f; X \cap B_\delta(x)] = \lim_{\delta\to 0} \Omega(\delta) = \inf_{\delta>0} \Omega(\delta)$$

exists, by the Monotone Sequence Theorem. We call this limit the oscillation of the function f at the point x.

The oscillation enjoys the following properties:

I. $\omega(f; x) \geq 0$ for every $x \in X$.

This is evident, since the infimum can only be nonâ€“negative.

II. $\omega(f; x) = 0$ if and only if f is continuous at x.

Note that $\omega(f; x) = 0$ means $\inf_{\delta>0} \Omega(\delta) = 0$. Thus, given $\epsilon > 0$, there exists $\delta > 0$ such that

$$0 \leq \sup\{|f(y) - f(z)|;\ y, z \in X \cap B_\delta(x)\} < \epsilon,$$

or, equivalently,

$$\text{if } y, z \in X \cap B_\delta(x) \text{ then } |f(y) - f(z)| < \epsilon.$$

This is clearly equivalent to the continuity of f at x.

III. If $x \in \dot{Y}$ and $Y \subset X$, then $\omega(f;x) \le \omega(f;Y)$.

Indeed, since $x \in \dot{Y}$, there exists $\delta > 0$ such that $B_\delta(x) \subset Y$. Hence

$$\omega[f; X \cap B_\delta(x)] = \omega[f; B_\delta(x)],$$

since $B_\delta(x) \subset X$, and consequently

$$\begin{aligned} \omega(f;x) &= \inf_{\delta>0} \Omega(\delta) = \inf_{\delta>0} \omega[f; X \cap B_\delta(x)] \\ &\le \omega[f; X \cap B_\delta(x)] = \omega[f; B_\delta(x)] \le \omega(f;Y). \end{aligned}$$

The last inequality follows from the fact that $B_\delta(x) \subset Y$ and therefore

$$\sup\{|f(y) - f(z)|;\ y, z \in B_\delta(x)\} \le \sup\{|f(y) - f(z)|;\ y, z \in Y\}.$$

IV. If $\omega(f;x) < C$, then there exists $\delta > 0$ such that $\omega(f;y) < C$ for every $y \in X$ with $\|y - x\| \le \delta$.

Since $\lim_{\delta \to 0} \Omega(\delta) < C$, by the definition of limit there exists $\delta > 0$ such that

$$\Omega(\delta) = \omega[f; X \cap B_\delta(x)] < C.$$

Now, given $y \in X$ with $\|y - x\| \le \delta$, that is, $y \in B_\delta(x)$, take $\eta > 0$ such that $B_\eta(y) \subset B_\delta(x)$; then

$$\begin{aligned} \omega(f;y) &= \inf_{\eta>0} \omega[f; X \cap B_\eta(y)] \le \omega[f; X \cap B_\eta(y)] \\ &\le \omega[f; X \cap B_\delta(x)] < C. \end{aligned}$$

V. If $X \subset \mathbb{R}^n$ is closed (respectively compact), then for every $C \ge 0$ the set

$$\{x \in X;\ \omega(f;x) \ge C\}$$

is closed (respectively compact).

In fact, let (y_k) be a sequence in $\{x \in X;\ \omega(f;x) \ge C\}$ such that $y_k \to y$ in $\mathbb{R}^n$. We shall prove that $y \in \{x \in X;\ \omega(f;x) \ge C\}$, and hence that this set is closed.

Indeed, since $(y_k) \subset X$ and X is closed, we have $y \in X$. On the other hand, it cannot happen that $\omega(f;y) < C$, because otherwise, by item IV, there would exist $\delta > 0$ such that $\omega(f,z) < C$ for all $z \in X$ with $z \in B_\delta(y)$, and in particular, since $y_k \to y$, there would exist $k_0 \in \mathbb{N}$ such that $\omega(f, y_k) < C$ for all $k \ge k_0$, which contradicts the fact that $\omega(f, y_k) \ge C$ for all k, because

$$(y_k) \subset \{x \in X;\ \omega(f;x) \ge C\}.$$

Hence $\omega(f;y) \ge C$ and therefore y belongs to this set.

With these preliminaries in place, we now prove the characterisation theorem for integrable functions.

Teorema 3.28 *A function $f : A \to \mathbb{R}$, bounded on the block $A \subset \mathbb{R}^n$, is integrable if and only if the set D_f of its points of discontinuity has measure zero.*

Proof: Assume first that $\operatorname{med}(D_f) = 0$ and let $\epsilon > 0$. We must exhibit a partition P of the block A such that

$$\sum_{B \in P} \omega_B \operatorname{vol} B < \epsilon,$$

as in (3.13). Indeed:

For the given $\epsilon > 0$, and using the fact that the set D_f has measure zero, there exists a countable covering $\{C_i'\}$ of D_f by open cubes such that

$$\sum_i \operatorname{vol} C_i' < \frac{\epsilon}{2K},$$

where $K = M - m$ (the difference between the supremum and infimum of f on A) is the oscillation of f on the block A. For each $x \in A \setminus D_f$, take an open cube C_x'' containing x such that the oscillation of f on the closure $\overline{C_x''}$ is less than $\dfrac{\epsilon}{2 \operatorname{vol} A}$. Since A is compact, from the open covering

$$A \subset \Big(\bigcup_i C_i'\Big) \cup \Big(\bigcup_{x \in A \setminus D_f} C_x''\Big)$$

we can extract a finite subcover

$$A \subset C_1' \cup \cdots \cup C_r' \cup C_1'' \cup \cdots \cup C_s''.$$

Let P be a partition of A such that each (open) block $B \in P$ is contained either in one of the cubes C_i' or in one of the cubes C_j''. More precisely, if

$$A = \prod_{k=1}^{n} [a_k, b_k],$$

then $P = P_1 \times \cdots \times P_n$, where, for each $k = 1, \ldots, n$, P_k is the set consisting of a_k, b_k, together with the k-th coordinates of the vertices of the cubes C_i' and C_j''. The picture below illustrates how to 'prolong the faces' of five cubes in order to obtain a partition of the block A.

We denote generically by B' those blocks of P that are contained in some cube C_i'. The remaining blocks (necessarily contained in cubes C_j'') will be denoted by B''. The sum of the volumes of the blocks B' is less than $\dfrac{\epsilon}{2K}$, and on each block B'' the oscillation of f does not exceed $\dfrac{\epsilon}{2 \operatorname{vol} A}$. Therefore, the partition P gives

$$\begin{aligned}
\sum_{B \in P} \omega_B \operatorname{vol} B &= \sum_{B'} \omega_{B'} \operatorname{vol} B' + \sum_{B''} \omega_{B''} \operatorname{vol} B'' \\
&\leq K \sum_{B'} \operatorname{vol} B' + \frac{\epsilon}{2 \operatorname{vol} A} \sum_{B''} \operatorname{vol} B'' \\
&< K \frac{\epsilon}{2K} + \frac{\epsilon}{2 \operatorname{vol} A} \operatorname{vol} A = \epsilon.
\end{aligned}$$

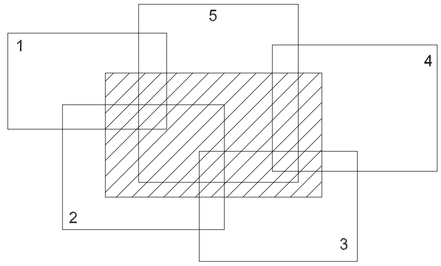

Figure 3.9:

Hence f is integrable.

Conversely, suppose that f is integrable. We shall prove that D_f has measure zero. Let $\epsilon > 0$ be given. For each $i \in \mathbb{N}^*$ set

$$D_i = \left\{ x \in A;\ \omega(f;x) \geq \frac{1}{i} \right\}.$$

We shall prove that

$$D_f = \bigcup_{i=1}^{+\infty} D_i. \tag{3.19}$$

Indeed, as observed above, if $\omega(f;x) > 0$ then f is discontinuous at x. Since D_i collects the points of A such that $\omega(f;x) \geq \dfrac{1}{i} > 0$, the function f is discontinuous at every point of D_i, for any i. Therefore $D_i \subset D_f$ for all i, and hence $\bigcup_{i=1}^{+\infty} D_i \subset D_f$.

On the other hand, if $x \in D_f$, then $\omega(f;x) > 0$. Choose $i_0 \in \mathbb{N}^*$ such that $0 < \dfrac{1}{i_0} \leq \omega(f;x)$. Then

$$x \in D_{i_0} = \left\{ x \in A;\ \omega(f;x) \geq \frac{1}{i_0} \right\},$$

and thus $D_f \subset \bigcup_{i=1}^{+\infty} D_i$, which proves (3.19).

To show that D_f has measure zero, it suffices to prove that, for each $i \in \mathbb{N}^*$, $\mathrm{med}(D_i) = 0$. Indeed, since f is integrable, for the given $\epsilon > 0$ there exists a partition P of the block A such that

$$\sum_{B \in P} \omega_B \operatorname{vol} B < \frac{\epsilon}{i}.$$

Let B' denote those blocks of the partition P that contain some point of D_i in their interior. By item III, since $x \in B'^{\circ}$ and $B'^{\circ} \subset B'$, we have $\omega(f;x) \leq \omega(f;B')$ and therefore $\omega(f;B') \geq \dfrac{1}{i}$. Hence

$$\frac{1}{i} \sum_{B' \in P} \operatorname{vol} B' \leq \sum_{B' \in P} \omega_{B'} \operatorname{vol} B' \leq \sum_{B \in P} \omega_B \operatorname{vol} B < \frac{\epsilon}{i}.$$

Multiplying by i we obtain

$$\sum_{B' \in P} \operatorname{vol} B' < \epsilon.$$

Now, clearly

$$D_i \subset \Big(\bigcup_{B' \in P} B' \Big) \cup Y,$$

where Y is the union of the proper faces of the blocks $B \in P$ that contain some point of D_i. We know that Y has measure zero. By Proposition 3.21 we conclude that $\operatorname{med}(D_i) = 0$, which completes the proof. □

Definition 3.29 *A bounded set $X \subset \mathbb{R}^n$ is said to be <u>J-measurable</u> (Jordan-measurable) if, taking a block $A \subset \mathbb{R}^n$ that contains X, the characteristic function $\chi_X : A \to \mathbb{R}$ is integrable.*

When X is Jâ€“measurable, its volume is, by definition, the integral of its characteristic function:

$$\operatorname{vol} X = \int_A \chi_X(x)\, dx. \tag{3.20}$$

As a consequence of Lebesgue's theorem, we shall now prove an important characterisation of J-measurable sets. Before doing so, recall that the boundary δX of a set $X \subset \mathbb{R}^n$ is the set of points $x \in \mathbb{R}^n$ such that every neighbourhood of x contains points of X and points of $\mathbb{R}^n \setminus X$. One has the disjoint union

$$\mathbb{R}^n = X^\circ \cup \delta X \cup (\mathbb{R}^n \setminus X)^\circ.$$

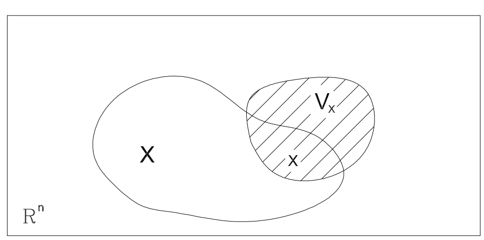

Figure 3.10:

Teorema 3.30 *A bounded set $X \subset \mathbb{R}^n$ is J-measurable if and only if its boundary has measure zero.*

Proof: Given a block $A \supset X$ in $\mathbb{R}^n$, let D be the set of points of discontinuity of the characteristic function $\chi_X : A \to \mathbb{R}$. Note that the possible discontinuities of the characteristic function occur on the boundary of X, that is, $D \subset \delta X$, because

$$A = X^\circ \cup \delta X \cup (A \setminus X)^\circ$$

and, moreover, the characteristic function is continuous on X° and on $(A \setminus X)^\circ$.

On the other hand, a point of δX which is not a point of discontinuity of χ_X must belong to δA. Indeed, suppose otherwise that $x \notin D$, $x \in \delta X$ and yet $x \notin \delta A$. Since x is a point of discontinuity of χ_X, it is clear that $x \in A$; and as $x \notin \delta A$, it follows that $x \in A^\circ$. Hence there exists $\epsilon_0 > 0$ such that $B_{\epsilon_0}(x) \subset A$.

Choose $n_0 \in \mathbb{N}$ sufficiently large so that $\frac{1}{n_0} < \epsilon_0$. Then, for every $n \geq n_0$, we have

$$B_{1/n}(x) \subset B_{\epsilon_0}(x) \subset A.$$

Since we are assuming that $x \in \delta X$, for each $n \geq n_0$ there exist points $y_n, z_n \in B_{1/n}(x)$ with $y_n \in X$ and $z_n \in A \setminus X$. The sequences (y_n) and (z_n) then converge to x, with $\chi_X(y_n) = 1$ and $\chi_X(z_n) = 0$. It follows that x is a point of discontinuity of χ_X, which is a contradiction. Thus, a point of δX is either a point of discontinuity of χ_X or a point that belongs to the boundary of A. In other words,

$$\delta X = D \cup (\delta X \cap \delta A).$$

The figure below (Figure 3.11) illustrates the case where $x \in \delta X$ and $x \notin D$.

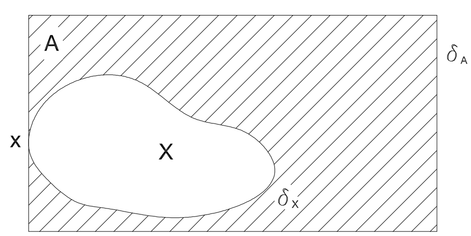

Figure 3.11:

Note that if $X \subset A^\circ$ then $\delta X = D$. Since δA has measure zero in $\mathbb{R}^n$, it follows that

$$\text{med}(\delta X) = 0 \iff \text{med}(D) = 0.$$

Indeed, if $\text{med}(\delta X) = 0$, then $D \subset \delta X$ implies $\text{med}(D) = 0$, since every subset of a set of measure zero has measure zero. Conversely, assume that $\text{med}(D) = 0$. To prove that $\text{med}(\delta X) = 0$, it suffices to show that $\text{med}(\delta X \cap \delta A) = 0$, since a countable union of sets of measure zero has measure zero. However, $\delta X \cap \delta A \subset \delta A$ and $\text{med}(\delta A) = 0$, so $\text{med}(\delta X \cap \delta A) = 0$.

Thus $\text{med}(\delta X) = 0 \iff \text{med}(D) = 0$, and, in view of Lebesgueâ€™s theorem, X is J-measurable $\iff \text{med}(D) = 0$, which proves the theorem. □

Note 1: We have just proved that the characterisation of J-measurable sets does not depend on the choice of the block containing them. Likewise, the value of $\operatorname{vol} X$ is also independent of the block A taken in the definition.

Note 2: A block is a J-measurable set, and its volume has already been determined in Note 7 of Section 3.1. A ball (open or closed) is Jâ€“measurable because its boundary is a sphere, which has measure zero, according to Note 3 of Section 3.2. More generally, if $X \subset \mathbb{R}^n$ is a bounded set whose boundary is a countable union of C^1-surfaces of dimension $< n$, then X is J-measurable.

Definition 3.31 *Given a bounded set $X \subset \mathbb{R}^n$, we may consider a block $A \supset X$ in $\mathbb{R}^n$ and define the* inner volume *and* outer volume *of X, respectively, by*

$$\operatorname{vol\,int} X = \underline{\int_A} \chi_X(x)\,dx \quad \textit{and} \quad \operatorname{vol\,ext} X = \overline{\int_A} \chi_X(x)\,dx.$$

Recalling the definitions of the lower and upper integrals, we see that

$$\begin{aligned} \operatorname{vol\,int} X &= \sup\{s(\chi_X; P);\ P \in \wp\}, \\ \operatorname{vol\,ext} X &= \inf\{S(\chi_X; P);\ P \in \wp\}, \end{aligned}$$

where $\wp$ is the set of all possible partitions of the block A.

We denote by P a partition of the block A and by B the blocks of P. Then

$$s(\chi_X; P) = \sum_{B \in P} m_B \operatorname{vol} B \quad \text{and} \quad S(\chi_X; P) = \sum_{B \in P} M_B \operatorname{vol} B,$$

where $m_B = \inf\{\chi_X(x);\ x \in B\}$ and $M_B = \sup\{\chi_X(x);\ x \in B\}$.

Observe that:

$s(\chi_X; P)$ is the sum of the volumes of those blocks of P which are contained in X, since if $B \subset X$ then $m_B = 1$ and if B is not contained in X then $m_B = 0$. Consequently,

$$s(\chi_X; P) = \sum_{\substack{B \in P \\ B \subset X}} \operatorname{vol} B.$$

Similarly:

$S(\chi_X; P)$ is the sum of the volumes of the blocks of P that have a non-empty intersection with X, because if $B \in P$ is such that $B \cap X \neq \emptyset$ then $M_B = 1$, whereas if $B \cap X = \emptyset$ then $M_B = 0$. Thus

$$S(\chi_X; P) = \sum_{\substack{B \in P \\ B \cap X \neq \emptyset}} \operatorname{vol} B.$$

To fix ideas, see the figure below (Figure 3.12):

Note 3: Let $X \subset \mathbb{R}^n$ be a J-measurable set. Then X has volume zero if and only if it has outer volume zero. Indeed, if $\operatorname{vol} X = 0$, then

$$\int_A \chi_X(x)\,dx = 0,$$

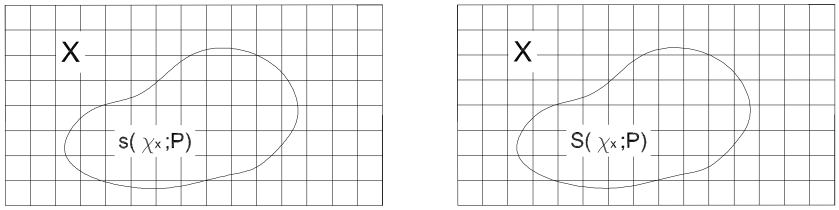

Figure 3.12:

and consequently $\overline{\int_A} \chi_X(x)\,dx = 0$, that is, vol ext $X = 0$.

Conversely, if vol ext $X = 0$, then, since

$$0 \leq \underline{\int_A} \chi_X(x)\,dx \leq \overline{\int_A} \chi_X(x)\,dx,$$

it follows that vol int $X = 0$. Therefore vol $X = 0$.

From what we have seen above,

$$\text{vol ext}\, X = \inf\{S(\chi_X; P);\ P \in \wp^*\},$$

where $\wp^*$ is the set of partitions P of the block A such that the blocks of P have some point in common with X. In this case

$$S(\chi_X; P) = \sum_{\substack{B \in P \\ B \cap X \neq \emptyset}} \text{vol}\, B.$$

Since vol ext $X = 0$, for every $\epsilon > 0$ one can find a block $A \supset X$ and a partition P of A such that the sum of the volumes of the blocks of P that intersect X is less than ϵ. This is equivalent to saying that, given $\epsilon > 0$, there exist blocks $B_1, \ldots, B_k$ in $\mathbb{R}^n$ with

$$X \subset B_1 \cup \cdots \cup B_k \quad \text{and} \quad \sum_{i=1}^{k} \text{vol}\, B_i < \epsilon.$$

Now consider a compact J-measurable set $K \subset \mathbb{R}^n$. Then vol $K = 0$ if and only if med$(K) = 0$. Indeed:

If vol $K = 0$, then vol ext $K = 0$ and, as seen above, given $\epsilon > 0$ there exist blocks $B_1, \ldots, B_k$ in $\mathbb{R}^n$ such that

$$K \subset \bigcup_{i=1}^{k} B_i \quad \text{and} \quad \sum_{i=1}^{k} \text{vol}\, B_i < \epsilon,$$

that is, med$(K) = 0$.

Conversely, if $\operatorname{med}(K) = 0$, then given $\epsilon > 0$ there exists a countable open covering

$$K \subset \bigcup_{i=1}^{\infty} B_i \quad \text{with} \quad \sum_{i=1}^{\infty} \operatorname{vol} B_i < \epsilon.$$

Since K is compact, there exist blocks $B_1, \ldots, B_k$ with

$$K \subset \bigcup_{i=1}^{k} B_i \quad \text{and} \quad \sum_{i=1}^{k} \operatorname{vol} B_i < \epsilon,$$

which implies that vol ext $K = 0$ and consequently $\operatorname{vol} K = 0$.

For example, the boundary ∂X of a bounded set is always compact (because $\partial X \subset \overline{X}$, ∂X is closed, and $\overline{X}$ is compact). Hence the following statements are equivalent:

(a) X is J-measurable;

(b) ∂X has measure zero;

(c) ∂X has volume zero.

Note 4: A set which is not compact and has measure zero is not necessarily J-measurable. For instance, the set of points of a block A whose coordinates are rational is countable, hence has measure zero, but is not J-measurable, as seen in Note 7 of Section 3.1, since its characteristic function is not integrable in any block containing A. However, if $X \subset \mathbb{R}^n$ is J-measurable and has measure zero (or more generally has empty interior), then its volume must be zero, because, as no block is contained in X, we have $\operatorname{vol} X^\circ = 0$, and therefore vol ext $X = 0$; and since X is J-measurable, vol int $X =$ vol ext X. Thus, if $X \subset \mathbb{R}^n$ is J-measurable, we have

$$\operatorname{vol} X = 0 \quad \Longleftrightarrow \quad X^\circ = \emptyset.$$

Indeed:

If $\operatorname{vol} X = 0$, then vol ext $X = 0$, which implies $\operatorname{med}(X) = 0$ and consequently $X^\circ = \emptyset$. Conversely, if $X^\circ = \emptyset$, then, for the reason given above (no block is contained in X and hence $\operatorname{vol} X^\circ = 0$), we have $\operatorname{vol} X = 0$.

Teorema 3.32 *Let X, Y be J-measurable subsets of a block $A \subset \mathbb{R}^n$. Then:*

(a) $X \cup Y$, $X \cap Y$ and $A \setminus X$ are J-measurable;

(b) $\operatorname{vol}(X \cup Y) + \operatorname{vol}(X \cap Y) = \operatorname{vol} X + \operatorname{vol} Y$.

Proof: Assertion (a) follows immediately from the three inclusions below, which in turn follow from the definition of boundary:

$$\partial(X \cup Y) \subset \partial X \cup \partial Y, \quad \partial(X \cap Y) \subset \partial X \cup \partial Y, \quad \partial(A \setminus X) \subset \partial A \cup \partial X.$$

Assertion (b) follows from the equality

$$\chi_{X\cup Y} + \chi_{X\cap Y} = \chi_X + \chi_Y,$$

which is straightforward to verify. □

Corollary 3.33 *If X, Y are J-measurable and $X^\circ \cap Y^\circ = \emptyset$, then*

$$\operatorname{vol}(X \cup Y) = \operatorname{vol} X + \operatorname{vol} Y.$$

Since $(X \cap Y)^\circ = X^\circ \cap Y^\circ$, the hypothesis of the corollary means that the J-measurable sets X and Y have at most points of their boundaries in common.

Definition 3.34 *We now define the integral of a bounded function $f : X \to \mathbb{R}$ whose domain is a J-measurable set $X \subset \mathbb{R}^n$. To this end, consider a block $A \subset \mathbb{R}^n$ that contains X and extend f to a function $\tilde{f} : A \to \mathbb{R}$ by setting*

$$\tilde{f}(x) = \begin{cases} f(x), & \text{if } x \in X, \\ 0, & \text{if } x \in A \setminus X. \end{cases}$$

We say that f is integrable on X if the function $\tilde{f}$ is integrable on A and we define

$$\int_X f(x)\,dx = \int_A \tilde{f}(x)\,dx.$$

Evidently, the above definition does not depend on the choice of the block $A \supset X$.

Note 5: One sometimes writes $\tilde{f} = f\chi_X$, meaning that $\tilde{f}(x) = f(x)\chi_X(x)$ for all $x \in A$. This is an abuse of notation corresponding to treating a product as zero when one factor is zero and the other is not defined. What is actually true is that $\tilde{f} = \tilde{f}\chi_X$.

Teorema 3.35 *Let $X \subset \mathbb{R}^n$ be a J-measurable set. A bounded function $f : X \to \mathbb{R}$ is integrable if and only if the set D_f of its points of discontinuity has measure zero.*

Proof: First note that every point of discontinuity of f is also a point of discontinuity of $\tilde{f}$. In fact, if f is discontinuous at a point $x \in X$, there exists a sequence (x_k) of points in X such that $f(x_k)$ does not converge to $f(x)$. It is then clear that x is a point of discontinuity of $\tilde{f}$, and thus $D_f \subset D_{\tilde{f}}$.

Hence the points of discontinuity of $\tilde{f}$ are either points of discontinuity of f or lie on the boundary of X, since

$$A = X^\circ \cup \partial X \cup (A \setminus X)^\circ,$$

and $\tilde{f}$ is continuous on $(A \setminus X)^\circ$. Thus

$$D_f \subset D_{\tilde{f}} \subset D_f \cup \partial X.$$

Since ∂X has measure zero, we see that $\operatorname{med}(D_f) = 0$ if and only if $\operatorname{med}(D_{\tilde{f}}) = 0$. By definition, f is integrable on X if and only if $\tilde{f}$ is integrable on A, that is, if and only if $\operatorname{med}(D_{\tilde{f}}) = 0$. This proves the theorem, as summarised in the scheme below:

$$\begin{array}{ccc} f \text{ integrable on } X & \Longleftrightarrow & \tilde{f} \text{ integrable on } A \\ \updownarrow & & \updownarrow \\ \operatorname{med}(D_f) = 0 & \Longleftrightarrow & \operatorname{med}(D_{\tilde{f}}) = 0. \end{array}$$

□

Definition 3.36 *Let $f : X \subset \mathbb{R}^n \to \mathbb{R}$ be bounded on the J-measurable set $X \subset \mathbb{R}^n$. We define the lower integral and the upper integral of f on X by*

$$\underline{\int_X} f(x)\, dx = \underline{\int_A} \tilde{f}(x)\, dx \quad \text{and} \quad \overline{\int_X} f(x)\, dx = \overline{\int_A} \tilde{f}(x)\, dx,$$

where A is a block in $\mathbb{R}^n$ containing X and $\tilde{f} : A \to \mathbb{R}$ is the extension of f which vanishes on $A \setminus X$.

Clearly, f is integrable on X if and only if its lower and upper integrals coincide on X.

Proposition 3.37 (Properties) *Let $f, g : X \to \mathbb{R}$ be integrable functions on the J-measurable set $X \subset \mathbb{R}^n$ and let $C \in \mathbb{R}$. Then:*

(a) The functions Cf and $f + g$ are integrable on X, and

$$\int_X (Cf)(x)\, dx = C \int_X f(x)\, dx.$$

(b) If $f(x) \leq g(x)$ for all $x \in X$, then

$$\int_X f(x)\, dx \leq \int_X g(x)\, dx.$$

In particular, if $m \leq f(x) \leq M$ for all $x \in X$, then

$$m \operatorname{vol} X \leq \int_X f(x)\, dx \leq M \operatorname{vol} X.$$

(c) The function $x \mapsto |f(x)|$ is integrable and

$$\left| \int_X f(x)\, dx \right| \leq \int_X |f(x)|\, dx.$$

In particular, if $|f(x)| \leq K$ for all $x \in X$, then

$$\left| \int_X f(x)\, dx \right| \leq K \operatorname{vol} X.$$

(d) *If f is continuous and X is connected, then there exists $x_0 \in X$ such that*

$$\int_X f(x)\,dx = f(x_0)\ \mathrm{vol}\,X.$$

Proof:

(a) Let $A \subset \mathbb{R}^n$ be a block containing X and let $\tilde{f}, \tilde{g} : A \to \mathbb{R}$ be the extensions of f and g which are equal to zero on $A \backslash X$. The extensions of Cf and $f+g$ are, respectively, $C\tilde{f}$ and $\tilde{f}+\tilde{g}$. By hypothesis, $\tilde{f}$ and $\tilde{g}$ are integrable on A, and therefore so are $C\tilde{f}$ and $\tilde{f}+\tilde{g}$. Hence Cf and $f+g$ are integrable on X and, moreover,

$$\begin{aligned}
\int_X (Cf)(x)\,dx &= \int_A C\tilde{f}(x)\,dx = C\int_A \tilde{f}(x)\,dx = C\int_X f(x)\,dx,\\
\int_X [f(x)+g(x)]\,dx &= \int_A [\tilde{f}(x)+\tilde{g}(x)]\,dx = \int_A \tilde{f}(x)\,dx + \int_A \tilde{g}(x)\,dx\\
&= \int_X f(x)\,dx + \int_X g(x)\,dx.
\end{aligned}$$

(b) If $f(x) \le g(x)$ for all $x \in X$, then $\tilde{f}(x) \le \tilde{g}(x)$ for all $x \in A$. Thus

$$\int_X f(x)\,dx = \int_A \tilde{f}(x)\,dx \le \int_A \tilde{g}(x)\,dx = \int_X g(x)\,dx.$$

If $m \le f(x) \le M$ for all $x \in X$, then

$$m\,\chi_X(x) \le \tilde{f}(x) \le M\,\chi_X(x) \quad \text{for all } x \in A.$$

Hence

$$m\ \mathrm{vol}\,X = m\int_A \chi_X(x)\,dx = \int_A m\,\chi_X(x)\,dx \le \int_A \tilde{f}(x)\,dx = \int_X f(x)\,dx,$$

and similarly

$$\int_X f(x)\,dx \le M\ \mathrm{vol}\,X.$$

(c) Let $g : X \to \mathbb{R}$ be defined by $g(x) = |f(x)|$. Clearly $D_g \subset D_f$, because if $x \in D_g$ then there exists a sequence $(x_n) \subset X$ such that $x_n \to x$ and $|f(x_n)| \not\to |f(x)|$. It follows that $f(x_n) \not\to f(x)$, that is, $x \in D_f$. Hence g is integrable. The extension $\tilde{g} : A \to \mathbb{R}$ is given by $\tilde{g}(x) = |\tilde{f}(x)|$, where $\tilde{f} : A \to \mathbb{R}$ is the extension of f. Therefore

$$\begin{aligned}
\left|\int_X f(x)\,dx\right| &= \left|\int_A \tilde{f}(x)\,dx\right| \le \int_A |\tilde{f}(x)|\,dx\\
&= \int_A \tilde{g}(x)\,dx = \int_X |f(x)|\,dx.
\end{aligned}$$

If $|f(x)| \le K$ for all $x \in X$, then

$$\left|\int_X f(x)\,dx\right| \le \int_X |f(x)|\,dx \le K\ \mathrm{vol}\,X.$$

(d) Since X is connected, $f(X)$ is an interval whose endpoints are m and M, where $M = \max_{x\in X} f(x)$ and $m = \min_{x\in X} f(x)$. As

$$\frac{1}{\operatorname{vol} X}\int_X f(x)\,dx$$

belongs to this interval, it equals $f(x_0)$ for some $x_0 \in X$.

$\square$

Proposition 3.38 *Let $X, Y \subset \mathbb{R}^n$ be J-measurable sets. A function $f : X \cup Y \to \mathbb{R}$ is integrable if and only if its restrictions $f|_X$ and $f|_Y$ are integrable. In this case,*

$$\int_{X\cup Y} f(x)\,dx + \int_{X\cap Y} f(x)\,dx = \int_X f(x)\,dx + \int_Y f(x)\,dx.$$

In particular, if X and Y have no interior points in common, then

$$\int_{X\cup Y} f(x)\,dx = \int_X f(x)\,dx + \int_Y f(x)\,dx.$$

Proof: Let D, D_X and D_Y denote the sets of points of discontinuity of f, $f|_X$ and $f|_Y$, respectively. Then

$$D_X \cup D_Y \subset D \subset D_X \cup D_Y \cup \partial X \cup \partial Y.$$

We know that ∂X and ∂Y have measure zero. Therefore

$$\operatorname{med}(D) = 0 \quad\Longleftrightarrow\quad \operatorname{med}(D_X) = \operatorname{med}(D_Y) = 0,$$

that is, f is integrable if and only if $f|_X$ and $f|_Y$ are integrable.

In this case, let A be a block in $\mathbb{R}^n$ containing $X \cup Y$ and let $\tilde{f} : A \to \mathbb{R}$ be the extension of f which vanishes on $A \setminus (X \cup Y)$. Then $\tilde{f} = \tilde{f}\chi_{X\cup Y}$. From the equality

$$\chi_{X\cup Y} + \chi_{X\cap Y} = \chi_X + \chi_Y$$

we obtain

$$\tilde{f} + \tilde{f}\chi_{X\cap Y} = \tilde{f}\chi_X + \tilde{f}\chi_Y,$$

and hence

$$\begin{aligned}\int_{X\cup Y} f(x)\,dx + \int_{X\cap Y} f(x)\,dx &= \int_A \tilde{f}(x)\,dx + \int_A \tilde{f}(x)\chi_{X\cap Y}(x)\,dx \\ &= \int_A \tilde{f}(x)\chi_X(x)\,dx + \int_A \tilde{f}(x)\chi_Y(x)\,dx \\ &= \int_X f(x)\,dx + \int_Y f(x)\,dx.\end{aligned}$$

If X and Y have no interior points in common, then $\text{vol}(X \cap Y) = 0$, in view of Note 4 of this section. On the other hand, there exists $K > 0$ such that $|f(x)| \leq K$ for all $x \in X \cup Y$. Thus

$$\left| \int_{X \cap Y} f(x)\, dx \right| \leq K \text{ vol}(X \cap Y) = 0.$$

Consequently, $\int_{X \cap Y} f(x)\, dx = 0$, which completes the proof. □

Corollary 3.39 *Let $f : X \to \mathbb{R}$ be integrable on the J-measurable set $X \subset \mathbb{R}^n$. If $Y \subset X$ is Jâ€“measurable and $X \setminus Y$ has empty interior, then*

$$\int_X f(x)\, dx = \int_Y f(x)\, dx.$$

In particular, if $U = X^\circ$, then

$$\int_X f(x)\, dx = \int_U f(x)\, dx.$$

Proof: The corollary follows from the last part of Proposition 3.38, applied to the equality $X = (X \setminus Y) \cup Y$, together with the following observations:

1. The set $X \setminus Y$ is J-measurable, since if we take a block $A \supset X$, then

   $$X \setminus Y = X \cap (A \setminus Y),$$

   and we may apply Proposition 3.32 to X and $A \setminus Y$.

2. As the J-measurable set $X \setminus Y$ has empty interior, its volume is zero (by Note 4). Hence

   $$\int_{X \setminus Y} f(x)\, dx = 0.$$

   Therefore

   $$\int_X f(x)\, dx = \int_{X \setminus Y} f(x)\, dx + \int_Y f(x)\, dx = \int_Y f(x)\, dx.$$

3. If $U = X^\circ$, then $\partial U \subset \partial X$, so U is J-measurable. Moreover,

   $$X \setminus U = \partial X$$

   has empty interior, and thus

   $$\int_X f(x)\, dx = \int_U f(x)\, dx.$$

□

Note 6: The corollary above shows that, when considering integrals over J-measurable sets in $\mathbb{R}^n$, there is no loss of generality in assuming that such sets are open. It should be noted, however, that not every bounded open set is J-measurable. There is a particular case in which we can define the integral of a function $f : U \to \mathbb{R}$ defined on an open set $U \subset \mathbb{R}^n$, even if U is not J-measurable: this happens when f has compact support, where

$$\text{supp}(f) = \overline{\{x \in U;\ f(x) \neq 0\}}^{U}.$$

If supp(f) is compact, we can define

$$\int_U f(x)\,dx = \int_K f(x)\,dx,$$

where K is any J-measurable set such that

$$\text{supp}(f) \subset K \subset U.$$

The integral exists provided the set of points of discontinuity of f has measure zero. Such a compact J-measurable set K also exists. Observe that

$$d(\text{supp}(f), \mathbb{R}^n \setminus U) = \delta > 0.$$

Using the maximum norm, we can cover the compact set supp(f) by finitely many closed cubes of edge length $< \delta$ and take K as the union of these cubes.

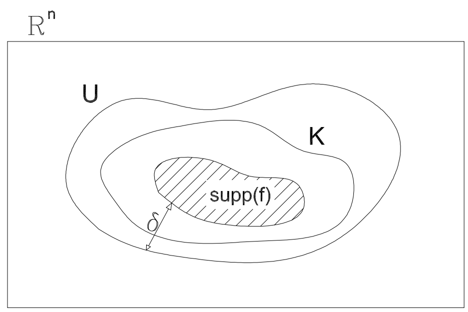

Figure 3.13:

The integral $\int_K f(x)\,dx$ clearly does not depend on the set K chosen under these conditions, since f vanishes outside supp(f).

3.4 Repeated Integration

The reduction of an integral over an n-dimensional block (multiple integral) to a sequence of n integrals of functions of one variable (repeated or iterated integral) is an effective computational tool.

Clearly, to reduce an integral over a block to successive integrals over intervals, it suffices to consider $A = A_1 \times A_2 \subset \mathbb{R}^{m+n}$, where $A_1 \subset \mathbb{R}^m$ and $A_2 \subset \mathbb{R}^n$ are blocks, and to show that any integral over A can be obtained by integrating first over A_2 and then over A_1 (or vice versa). This is what we shall do. Points of $A_1 \times A_2$ will be written as (x, y), where $x \in A_1$, $y \in A_2$. If $f : A_1 \times A_2 \to \mathbb{R}$ is integrable, its integral will be denoted by

$$\int_{A_1 \times A_2} f(x, y)\, d(x, y).$$

Given $f : A_1 \times A_2 \to \mathbb{R}$, for each $x \in A_1$ we write $f_x : A_2 \to \mathbb{R}$, where $f_x(y) = f(x, y)$, $y \in A_2$. Thus f_x is essentially the restriction of f to the nâ€“dimensional block $\{x\} \times A_2$. Even assuming f to be integrable, the function f_x may, for some values of $x \in A_1$, fail to be integrable. Indeed, the set of points of discontinuity of f has measure zero in $\mathbb{R}^{m+n}$, but its intersection with some block $\{x\} \times A_2$ may fail to have nâ€“dimensional measure zero.

Example. Let $f : [0, 1] \times [0, 1] \to \mathbb{R}$ be defined by

$$f(x, y) = \begin{cases} 0, & \text{if } x \neq \frac{1}{2}, \\ 1, & \text{if } x = \frac{1}{2} \text{ and } y \text{ is rational}, \\ 0, & \text{if } x = \frac{1}{2} \text{ and } y \text{ is irrational}. \end{cases}$$

The set of points of discontinuity of f is the vertical segment $\{1/2\} \times [0, 1]$, which has measure zero in $\mathbb{R}^2$. Thus f is integrable. For every $x \neq \frac{1}{2}$, $f_x : [0, 1] \to \mathbb{R}$ is identically zero and hence integrable. But $f_{1/2}$ is discontinuous at every point of the interval $[0, 1]$, and therefore is not integrable.

Teorema 3.40 (Repeated Integration) *Let $f : A_1 \times A_2 \to \mathbb{R}$ be integrable on the product of blocks $A_1 \subset \mathbb{R}^m$, $A_2 \subset \mathbb{R}^n$. For each $x \in A_1$, let $f_x : A_2 \to \mathbb{R}$ be defined by $f_x(y) = f(x, y)$ and set*

$$\varphi(x) = \underline{\int}_{A_2} f_x(y)\, dy, \qquad \psi(x) = \overline{\int}_{A_2} f_x(y)\, dy.$$

The functions $\varphi, \psi : A_1 \to \mathbb{R}$ thus defined are integrable and satisfy

$$\int_{A_1} \varphi(x)\, dx = \int_{A_1} \psi(x)\, dx = \int_{A_1 \times A_2} f(x, y)\, dx\, dy,$$

that is,

$$\int_{A_1 \times A_2} f(x, y)\, dx\, dy = \int_{A_1} dx \left(\underline{\int}_{A_2} f(x, y)\, dy \right) = \int_{A_1} dx \left(\overline{\int}_{A_2} f(x, y)\, dy \right).$$

Proof: Let $P = P_1 \times P_2$ be an arbitrary partition of $A_1 \times A_2$. The blocks of P are the products $B_1 \times B_2$, where $B_1 \in P_1$ and $B_2 \in P_2$. The lower sum of f with respect to the partition P is

$$\begin{aligned} s(f;P) &= \sum m_{B_1\times B_2} \operatorname{vol} B_1 \operatorname{vol} B_2 \\ &= \sum_{B_1\in P_1}\left(\sum_{B_2\in P_2} m_{B_1\times B_2} \operatorname{vol} B_2\right)\operatorname{vol} B_1. \end{aligned}$$

For every $x \in B_1$, one has $m_{B_1\times B_2} = m_{B_1\times B_2}(f) \leq m_{B_2}(f_x)$. Hence

$$\sum_{B_2\in P_2} m_{B_1\times B_2} \operatorname{vol} B_2 \leq \sum_{B_2\in P_2} m_{B_2}(f_x) \operatorname{vol} B_2 \leq \varphi(x).$$

Since this inequality holds for every $x \in B_1$, we conclude that

$$\sum_{B_2\in P_2} m_{B_1\times B_2} \operatorname{vol} B_2 \leq m_{B_1}(\varphi).$$

Therefore

$$s(f;P) \leq \sum_{B_1\in P_1} m_{B_1}(\varphi) \operatorname{vol} B_1 = s(\varphi;P_1).$$

Similarly, one proves the inequality $S(\varphi;P_1) \leq S(f;P)$. Thus

$$s(f;P) \leq s(\varphi;P_1) \leq S(\varphi;P_1) \leq S(f;P)$$

for any partition $P = P_1 \times P_2$. Since f is integrable, it follows immediately that φ is integrable and

$$\int_{A_1} \varphi(x)\,dx = \int_{A_1\times A_2} f(x,y)\,dx\,dy.$$

The assertion concerning ψ is proved in the same way. □

Corollary 3.41 *If $f : A_1 \times A_2 \to \mathbb{R}$ is integrable, then*

$$\int_{A_1\times A_2} f(x,y)\,dx\,dy = \int_{A_1} dx\left(\overline{\int_{A_2}} f(x,y)\,dy\right) = \int_{A_2} dy\left(\overline{\int_{A_1}} f(x,y)\,dx\right),$$

and there are three further analogous equalities, obtained by taking lower and upper integrals inside the parentheses. In particular, if f_x and f_y are continuous for all $x \in A_1$ and $y \in A_2$ (for example, if f is continuous), then

$$\int_{A_1\times A_2} f(x,y)\,dx\,dy = \int_{A_1} dx\left(\int_{A_2} f(x,y)\,dy\right) = \int_{A_2} dy\left(\int_{A_1} f(x,y)\,dx\right).$$

Proof: Indeed, everything we did with x in the previous theorem can equally well be done with y. □

Chapter 4

Differential Forms

4.1 k-Forms

In what follows, $\mathbb{R}^n$ will denote a vector space of dimension n (not necessarily the usual Euclidean space we work with). We fix an arbitrary basis $\{e_1, \dots, e_n\}$ of $\mathbb{R}^n$. Thus, if $v \in \mathbb{R}^n$, then

$$v = a_1 e_1 + \cdots + a_n e_n,$$

that is, we can write any vector of $\mathbb{R}^n$ as a linear combination of the basis elements, and this combination is unique.

Definition 4.1.1 *A **1-form** on $\mathbb{R}^n$ is a map*

$$\omega : \mathbb{R}^n \longrightarrow \mathbb{R}$$

which is linear, or, in other words, a linear functional. Hence

$$\omega(\lambda_1\xi_1 + \lambda_2\xi_2) = \lambda_1\omega(\xi_1) + \lambda_2\omega(\xi_2), \quad \forall\, \xi_1, \xi_2 \in \mathbb{R}^n,\ \forall\, \lambda_1, \lambda_2 \in \mathbb{R}.$$

We denote by $(\mathbb{R}^n)^*$ the **set of all 1-forms**. This set is called the **dual** of $\mathbb{R}^n$ and is a vector space over $\mathbb{R}$, endowed with the operations

$$\text{i)} \quad (\omega_1 + \omega_2)(\xi) = \omega_1(\xi) + \omega_2(\xi),$$

$$\text{ii)} \quad (\lambda\omega)(\xi) = \lambda\,\omega(\xi),$$

for all $\omega_1, \omega_2 \in (\mathbb{R}^n)^*$ and $\lambda \in \mathbb{R}$.

We shall now show that $(\mathbb{R}^n)^*$ is a vector space of **dimension** n by exhibiting a basis $\{X_1, \ldots, X_n\}$ of linear functionals, defined for each $i = 1, \ldots, n$ by

$$\begin{array}{rcl} X_i : \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ \xi = a_1 e_1 + \cdots + a_n e_n & \longmapsto & X_i(\xi) = a_i. \end{array}$$

In fact, $X_i(\xi) = a_i$ is the i-th coordinate of the point $x \in \mathbb{R}^n$. Note that, since

$$e_j = 0e_1 + \cdots + 1e_j + \cdots + 0e_n,$$

we have

$$X_i(e_j) = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

In view of this, we shall prove that $\{X_1, \ldots, X_n\}$ is a basis of $(\mathbb{R}^n)^*$. Indeed:

1) $X_1, \ldots, X_n$ are linearly independent, because if

$$\lambda_1 X_1 + \cdots + \lambda_n X_n = \mathbf{0},$$

(where $\mathbf{0}$ denotes the zero 1-form, the neutral element of $(\mathbb{R}^n)^*$), then

$$(\lambda_1 X_1 + \cdots + \lambda_n X_n)(e_i) = 0,$$

that is,

$$\lambda_1 X_1(e_i) + \cdots + \lambda_i X_i(e_i) + \cdots + \lambda_n X_n(e_i) = 0,$$

whence $\lambda_i = 0$ for all $i = 1, \ldots, n$.

2) $X_1, \ldots, X_n$ span $(\mathbb{R}^n)^*$, since, given $\omega \in (\mathbb{R}^n)^*$, we have, for every

$$\xi = a_1 e_1 + \cdots + a_n e_n \in \mathbb{R}^n,$$

$$\begin{aligned} \omega(\xi) &= \omega(a_1 e_1 + \cdots + a_n e_n) \\ &= a_1 \omega(e_1) + \cdots + a_n \omega(e_n) \\ &= X_1(\xi)\, \omega(e_1) + \cdots + X_n(\xi)\, \omega(e_n) \\ &= (\lambda_1 X_1 + \cdots + \lambda_n X_n)(\xi), \end{aligned}$$

where $\lambda_i = \omega(e_i)$ are real numbers.

Therefore, there exist $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ such that

$$\omega = (\lambda_1 X_1 + \cdots + \lambda_n X_n),$$

which proves that the subspace generated by $X_1, \ldots, X_n$ is precisely $(\mathbb{R}^n)^*$, that is,

$$[X_1, \ldots, X_n] = (\mathbb{R}^n)^*.$$

Example 4.1.2 *Consider a uniform force field F in $\mathbb{R}^3$ and define*

$$\omega : \mathbb{R}^3 \longrightarrow \mathbb{R}, \qquad \xi \longmapsto \omega(\xi) = \langle F, \xi \rangle.$$

The value $\omega(\xi)$ represents the work done by the field F in displacing a particle through

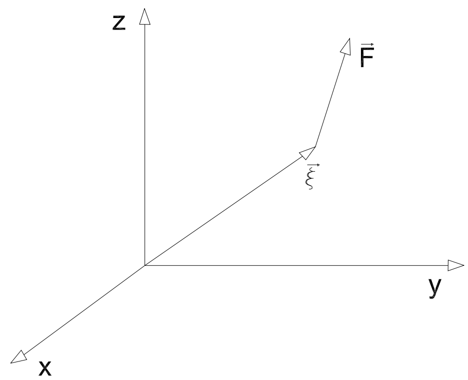

Figure 4.1: Force field

a distance equal to the modulus of ξ. The map ω is clearly a linear functional, by the linearity of the inner product.

Definition 4.1.3 *A* ***2-form*** *on $\mathbb{R}^n$ is a map*

$$\omega^2 : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$$

which is bilinear and antisymmetric.

Thus, for all $\xi_1, \xi_1', \xi_2 \in \mathbb{R}^n$ and all $\lambda_1, \lambda_1' \in \mathbb{R}$, we have

$$\omega^2(\lambda_1\xi_1 + \lambda_1'\xi_1', \xi_2) = \lambda_1\omega^2(\xi_1, \xi_2) + \lambda_1'\omega^2(\xi_1', \xi_2),$$

and

$$\omega^2(\xi_1, \xi_2) = -\omega^2(\xi_2, \xi_1).$$

Example 4.1.4 *Let $\mathbb{R}^2$ be the plane. Given $\xi = \xi_1 e_1 + \xi_2 e_2$ and $\eta = \eta_1 e_1 + \eta_2 e_2$ in $\mathbb{R}^2$, set*

$$\omega^2 : \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}, \qquad (\xi, \eta) \longmapsto \omega^2(\xi, \eta) = \begin{vmatrix} \xi_1 & \xi_2 \\ \eta_1 & \eta_2 \end{vmatrix} = \xi_1\eta_2 - \eta_1\xi_2.$$

Then ω^2 *is a 2-form on* $\mathbb{R}^2$*. In fact, writing* $\xi = \xi_1 e_1 + \xi_2 e_2$, $\eta = \eta_1 e_1 + \eta_2 e_2$, $\nu = \nu_1 e_1 + \nu_2 e_2$, *we have*

$$\begin{aligned}
\omega^2(\xi+\eta,\nu) &= \begin{vmatrix} \xi_1+\eta_1 & \xi_2+\eta_2 \\ \nu_1 & \nu_2 \end{vmatrix} \\
&= (\xi_1+\eta_1)\nu_2 - (\xi_2+\eta_2)\nu_1 \\
&= (\xi_1\nu_2 - \xi_2\nu_1) + (\eta_1\nu_2 - \eta_2\nu_1) \\
&= \begin{vmatrix} \xi_1 & \xi_2 \\ \nu_1 & \nu_2 \end{vmatrix} + \begin{vmatrix} \eta_1 & \eta_2 \\ \nu_1 & \nu_2 \end{vmatrix} \\
&= \omega^2(\xi,\nu) + \omega^2(\eta,\nu).
\end{aligned}$$

Similarly, one proves that

$$\omega^2(\lambda\xi,\eta) = \lambda\,\omega^2(\xi,\eta), \quad \forall\,\lambda \in \mathbb{R}.$$

It remains to prove antisymmetry. Indeed,

$$\omega^2(\xi,\eta) = \xi_1\eta_2 - \eta_1\xi_2 = -(\eta_1\xi_2 - \xi_1\eta_2) = -\omega^2(\eta,\xi).$$

Remark 4.1.5 *Geometrically, the modulus of* $\omega^2(\xi,\eta)$ *is the area of the parallelogram generated by the vectors* ξ *and* η*.*

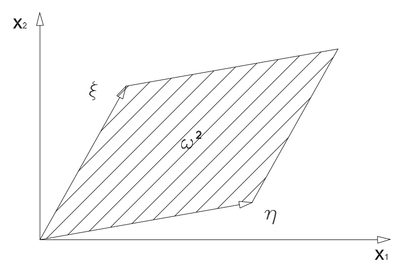

Figure 4.2:

Recall this fact:

The area of the parallelogram above is given by

$$|\eta|\cdot h = |\eta|\cdot|\xi|\sin\theta = \|\xi\wedge\eta\|.$$

On the other hand, writing $\xi = \xi_1 e_1 + \xi_2 e_2 + 0e_3$, $\eta = \eta_1 e_1 + \eta_2 e_2 + 0e_3$, *we have*

$$\xi\wedge\eta = \begin{vmatrix} i & j & k \\ \xi_1 & \xi_2 & 0 \\ \eta_1 & \eta_2 & 0 \end{vmatrix} = (0,0,\xi_1\eta_2 - \xi_2\eta_1).$$

Hence $\|\xi\wedge\eta\| = |\xi_1\eta_2 - \xi_2\eta_1|$.

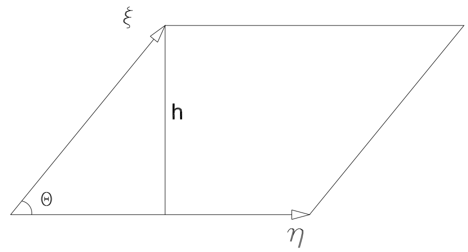

Figure 4.3:

Example 4.1.6 *Let $\mathbb{R}^3$ be Euclidean space and consider $\xi = \xi_1 e_1 + \xi_2 e_2 + \xi_3 e_3$, $\eta = \eta_1 e_1 + \eta_2 e_2 + \eta_3 e_3$ in $\mathbb{R}^3$. Define*

$$\omega^2 : \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}, \qquad (\xi, \eta) \longmapsto \omega^2(\xi, \eta) = \begin{vmatrix} \xi_1 & \xi_2 \\ \eta_1 & \eta_2 \end{vmatrix}.$$

In view of the previous example, ω^2 is a 2-form on $\mathbb{R}^3$. From a geometric point of view, the modulus of $\omega^2(\xi, \eta)$ is the area of the projection on the $x_1 x_2$-plane of the parallelogram generated by the vectors ξ and η.

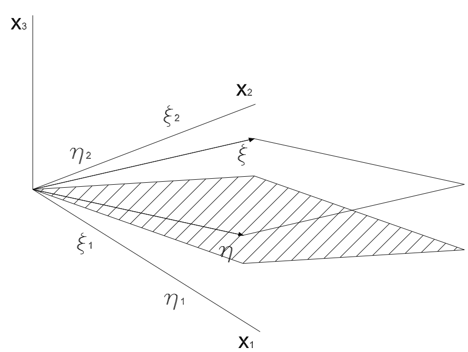

Figure 4.4:

Example 4.1.7 *Let $\mathbb{R}^3$ be Euclidean space and let v be a uniform velocity field. Define*

$$\omega^2 : \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}, \qquad (\xi, \eta) \longmapsto \omega^2(\xi, \eta) = [v, \xi, \eta] = \langle v, \xi \wedge \eta \rangle.$$

Indeed, writing $v = v_1 e_1 + v_2 e_2 + v_3 e_3$, $\xi = \xi_1 e_1 + \xi_2 e_2 + \xi_3 e_3$, $\eta = \eta_1 e_1 + \eta_2 e_2 + \eta_3 e_3$, we obtain

$$\omega^2(\xi, \eta) = \begin{vmatrix} v_1 & v_2 & v_3 \\ \xi_1 & \xi_2 & \xi_3 \\ \eta_1 & \eta_2 & \eta_3 \end{vmatrix}.$$

It is not difficult to show that ω^2 is a bilinear, antisymmetric form and hence a 2-form. From a physical point of view, $\omega^2(\xi, \eta)$ represents the flow of the fluid crossing the area of the parallelogram generated by the vectors ξ and η.

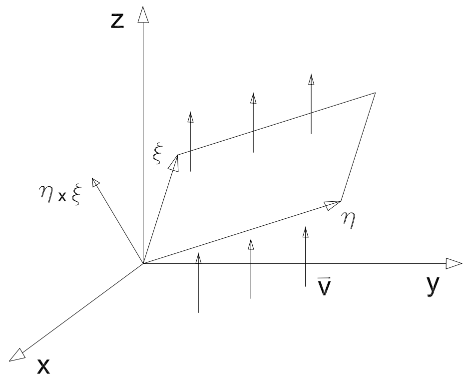

Figure 4.5:

The set of 2-forms $\omega^2 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, which we denote by $\mathcal{A}^2(\mathbb{R}^n)$, just as in the case of $(\mathbb{R}^n)^*$, is a vector space endowed with the operations

$$(\omega_1^2 + \omega_2^2)(\xi_1, \xi_2) = \omega_1^2(\xi_1, \xi_2) + \omega_2^2(\xi_1, \xi_2),$$

$$(\lambda\omega_1^2)(\xi_1, \xi_2) = \lambda\, \omega_1^2(\xi_1, \xi_2).$$

As we know, if $\mathbb{R}^n$ is a vector space of dimension n and $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a bilinear form, then, given a basis $\beta = \{e_1, \ldots, e_n\}$ of $\mathbb{R}^n$, we associate to B a matrix $[B]_\beta^\beta$, called the matrix of the bilinear form B in the basis β, as follows:

If $\xi = a_1e_1 + \cdots + a_ne_n$ and $\eta = b_1e_1 + \cdots + b_ne_n$, then

$$[B(\xi, \eta)] = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \begin{bmatrix} B(e_1, e_1) & \cdots & B(e_1, e_n) \\ \vdots & \ddots & \vdots \\ B(e_n, e_1) & \cdots & B(e_n, e_n) \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

This correspondence is bijective. Moreover, there is an isomorphism between bilinear forms and their corresponding matrices $[B]_\beta^\beta$. If $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is bilinear and symmetric, then $[B]_\beta^\beta$ is a symmetric matrix and vice versa. Similarly, the same holds for antisymmetric forms.

From the above, if $\omega^2 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a bilinear, antisymmetric form, then $[\omega^2(\xi, \eta)]$ is of the type

$$\begin{bmatrix} 0 & B_{12} & B_{13} & \cdots & B_{1(n-1)} & B_{1n} \\ -B_{12} & 0 & B_{23} & \cdots & B_{2(n-1)} & B_{2n} \\ -B_{13} & -B_{23} & 0 & \cdots & B_{3(n-1)} & B_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -B_{1(n-1)} & -B_{2(n-1)} & -B_{3(n-1)} & \cdots & 0 & B_{(n-1)n} \\ -B_{1n} & -B_{2n} & -B_{3n} & \cdots & -B_{(n-1)n} & 0 \end{bmatrix}.$$

A basis for the set of all matrices of this type is given by matrices of the form

$$\begin{bmatrix} 0 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & -1 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \vdots & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix},$$

where the entries 1 and -1 occupy positions which are symmetric with respect to the main diagonal. A natural question then arises:

How many matrices of this type are there?

It is easy to see that the number of such matrices is related to the number of positions that can be occupied by the entry 1 in the 'upper triangle' (or equivalently by -1 in the 'lower triangle'). Clearly, this number is given by the sum of the arithmetic progression

$$(n-1, n-2, \ldots, 1),$$

that is,

$$\frac{n(n-1)}{2} = \binom{n}{2}.$$

Therefore, in view of the isomorphism described above, the dimension of $\mathcal{A}^2(\mathbb{R}^n)$ is

$$\binom{n}{2}.$$

Later we shall exhibit a basis for this space whose number of vectors is precisely

$$\binom{n}{2}.$$

Definition 4.1.8 *A k-**form** is a map*

$$\omega^k : \mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n \longrightarrow \mathbb{R}$$

which is k-linear and antisymmetric. In other words:

$$\text{i)}\quad \omega^k(\lambda_1\xi_1 + \lambda_1'\xi_1', \xi_2, \dots, \xi_k) = \lambda_1\omega^k(\xi_1, \xi_2, \dots, \xi_k) + \lambda_1'\omega^k(\xi_1', \xi_2, \dots, \xi_k),$$

$$\text{ii)}\quad \omega^k(\xi_{i_1}, \xi_{i_2}, \dots, \xi_{i_k}) = (-1)^\nu\, \omega^k(\xi_1, \xi_2, \dots, \xi_k),$$

where $(i_1, i_2, \dots, i_k)$ is a permutation of $(1, 2, \dots, k)$ and

$$\nu = \begin{cases} 0, & \text{if the permutation is even,} \\ 1, & \text{if the permutation is odd.} \end{cases}$$

Generalising the previous examples, we can consider as an example of a k-form on $\mathbb{R}^n$ $(k \le n)$ the map

$$\omega^k : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \longrightarrow \mathbb{R}, \qquad (\xi_1, \dots, \xi_k) \longmapsto \omega^k(\xi_1, \dots, \xi_k) = \begin{vmatrix} \xi_{11} & \xi_{12} & \cdots & \xi_{1k} \\ \xi_{21} & \xi_{22} & \cdots & \xi_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \xi_{k1} & \xi_{k2} & \cdots & \xi_{kk} \end{vmatrix},$$

where

$$\begin{cases} \xi_1 = \xi_{11}e_1 + \xi_{12}e_2 + \cdots + \xi_{1k}e_k + \cdots + \xi_{1n}e_n, \\ \xi_2 = \xi_{21}e_1 + \xi_{22}e_2 + \cdots + \xi_{2k}e_k + \cdots + \xi_{2n}e_n, \\ \quad\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \\ \xi_k = \xi_{k1}e_1 + \xi_{k2}e_2 + \cdots + \xi_{kk}e_k + \cdots + \xi_{kn}e_n. \end{cases}$$

Note that there are

$$\binom{n}{k}$$

ways of choosing examples of this kind, it being enough to choose k among the n directions $e_1, \dots, e_n$ in $\mathbb{R}^n$.

In fact, $\omega^k(\xi_1, \dots, \xi_k)$ represents the 'oriented volume' of the projected parallelotope with edges $\xi_1, \dots, \xi_k$ (oriented because it carries a sign).

Denoting by $\mathcal{A}^k(\mathbb{R}^n)$ the set of all k-forms on $\mathbb{R}^n$, and using an argument analogous to that used for 2-forms, we find that the dimension of $\mathcal{A}^k(\mathbb{R}^n)$ is

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

4.2 Exterior Product

Definition 4.2.1 *Let* $\omega_1, \omega_2 : \mathbb{R}^n \to \mathbb{R}$ *be 1â€“forms and consider the map*

$$\omega_1 \wedge \omega_2 : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R},$$

$$(\xi_1, \xi_2) \longmapsto (\omega_1 \wedge \omega_2)(\xi_1, \xi_2) = \begin{vmatrix} \omega_1(\xi_1) & \omega_2(\xi_1) \\ \omega_1(\xi_2) & \omega_2(\xi_2) \end{vmatrix}.$$

This map is called the ***exterior product of the 1-forms*** ω_1 ***and*** ω_2.

Now, setting

$$\omega : \mathbb{R}^n \longrightarrow \mathbb{R} \times \mathbb{R}, \qquad \xi \longmapsto \omega(\xi) = (\omega_1(\xi), \omega_2(\xi)),$$

we have, in particular,

$$\omega(\xi_1) = (\omega_1(\xi_1), \omega_2(\xi_1)), \qquad \omega(\xi_2) = (\omega_1(\xi_2), \omega_2(\xi_2)),$$

and therefore $(\omega_1 \wedge \omega_2)(\xi_1, \xi_2)$ is the area of the parallelogram with sides $\omega(\xi_1)$ and $\omega(\xi_2)$ in $\mathbb{R}^2$.

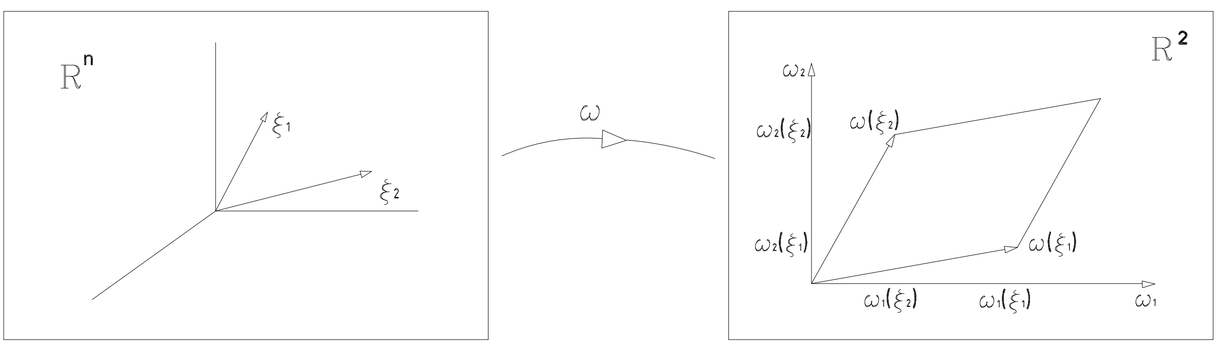

Figure 4.6:

We observe that the exterior product $\omega_1 \wedge \omega_2$ defined above is a 2-form. Indeed, linearity follows from the fact that ω_1 and ω_2 are 1-forms (and hence linear maps), and from the fact that the determinant is linear in each of its row vectors when the others are kept fixed. Antisymmetry also follows from the determinant, as an intrinsic property which can be found in standard analysis textbooks.

From what we have seen, it makes sense to define the map

$$\phi : (\mathbb{R}^n)^* \times (\mathbb{R}^n)^* \longrightarrow \mathcal{A}^2(\mathbb{R}^n), \qquad (\omega_1, \omega_2) \longmapsto \phi(\omega_1, \omega_2) = \omega_1 \wedge \omega_2.$$

Since the determinant of a 2×2 matrix is a bilinear, antisymmetric form in its row vectors, we see that ϕ is a bilinear, antisymmetric map on $(\mathbb{R}^n)^*$.

Let $X_1, \ldots, X_n$ be basic 1-forms, that is, for each $i = 1, \ldots, n$ define

$$X_i : \mathbb{R}^n \to \mathbb{R}, \qquad a_1 e_1 + \cdots + a_n e_n \longmapsto a_i,$$

as in the previous section.

Given $i, j \in \{1, \ldots, n\}$, what does the exterior product $X_i \wedge X_j$ represent?

By definition, for $\xi_1, \xi_2 \in \mathbb{R}^n$ we have

$$(X_i \wedge X_j)(\xi_1, \xi_2) = \begin{vmatrix} X_i(\xi_1) & X_j(\xi_1) \\ X_i(\xi_2) & X_j(\xi_2) \end{vmatrix},$$

which is the area of the parallelogram with sides $\omega(\xi_1)$ and $\omega(\xi_2)$.

However,

$$\omega(\xi_1) = (X_i(\xi_1), X_j(\xi_1)), \qquad \omega(\xi_2) = (X_i(\xi_2), X_j(\xi_2)).$$

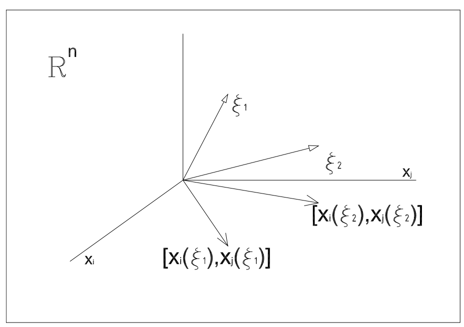

Figure 4.7:

Thus $(X_i \wedge X_j)(\xi_1, \xi_2)$ represents the 'oriented area' of the parallelogram generated by the vectors ξ_1 and ξ_2.

Note that

$$\begin{aligned} (X_i \wedge X_j)(\xi_1, \xi_2) &= \begin{vmatrix} X_i(\xi_1) & X_j(\xi_1) \\ X_i(\xi_2) & X_j(\xi_2) \end{vmatrix} \\ &= X_i(\xi_1) X_j(\xi_2) - X_i(\xi_2) X_j(\xi_1) \\ &= -[X_i(\xi_2) X_j(\xi_1) - X_i(\xi_1) X_j(\xi_2)] \\ &= -\begin{vmatrix} X_i(\xi_2) & X_j(\xi_2) \\ X_i(\xi_1) & X_j(\xi_1) \end{vmatrix} \\ &= -(X_j \wedge X_i)(\xi_1, \xi_2), \qquad \forall\, (\xi_1, \xi_2) \in \mathbb{R}^n \times \mathbb{R}^n. \end{aligned}$$

In particular, $(X_i \wedge X_i)(\xi_1, \xi_2) = -(X_i \wedge X_i)(\xi_1, \xi_2)$, which implies $(X_i \wedge X_i) = 0$, for all $(\xi_1, \xi_2) \in \mathbb{R}^n \times \mathbb{R}^n$.

Thus

$$(1) \quad X_i \wedge X_j = -(X_j \wedge X_i), \qquad (2) \quad X_i \wedge X_i \equiv 0.$$

We claim that the exterior forms $(X_i \wedge X_j)$ with $i < j$ are linearly independent, and there are $\frac{n(n-1)}{2}$ of them.

Indeed, suppose $\sum_{i<j} a_{ij}(X_i \wedge X_j) = 0$ (where 0 denotes the zero 2-form). Then, for $k, l \in \{1, \ldots, n\}$ with $k < l$, we have

$$\sum_{i<j} a_{ij}(X_i \wedge X_j)(e_k, e_l) = 0. \tag{1}$$

On the other hand,

$$(X_i \wedge X_j)(e_k, e_l) = \begin{vmatrix} X_i(e_k) & X_j(e_k) \\ X_i(e_l) & X_j(e_l) \end{vmatrix}.$$

Let $I = \{i, j\}$ and $J = \{k, l\}$. There are two cases:

(1) $I = J$.

In this case, if $i = k$ then necessarily $j = l$, and if $i = l$ then necessarily $j = k$. However, since $i < j$ and $k < l$, the only possible case would be $i = l$ and $j = k$, which would give $l = i < j = k$, a contradiction. Therefore

$$\begin{aligned} (X_i \wedge X_j)(e_k, e_l) &= (X_i \wedge X_j)(e_i, e_j) \\ &= \begin{vmatrix} X_i(e_i) & X_j(e_i) \\ X_i(e_j) & X_j(e_j) \end{vmatrix} \\ &= \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1. \end{aligned}$$

(2) $I \neq J$.

In this case there exists $i_k \in I \setminus J$. Thus i_k differs from every $j \in J$, and hence $X_{i_k}(e_m) = 0$ for all $m \in J$. Say $i_k = i$. Then

$$(X_i \wedge X_j)(e_k, e_l) = \begin{vmatrix} X_i(e_k) & X_j(e_k) \\ X_i(e_l) & X_j(e_l) \end{vmatrix} = \begin{vmatrix} X_{i_k}(e_k) & X_j(e_k) \\ X_{i_k}(e_l) & X_j(e_l) \end{vmatrix} = 0,$$

because $X_{i_k}(e_k) = X_{i_k}(e_l) = 0$.

Returning to (1), we obtain

$$\sum_{i<j} a_{ij}(X_i \wedge X_j)(e_i, e_j) = 0,$$

which implies $a_{ij} = 0$ for all i, j with $i < j$, proving that these vectors are linearly independent.

When we consider the exterior forms $X_i \wedge X_j$ with $i < j$, how many such forms are there?

There are as many as there are sets $I = \{i, j\}$ with $i < j$, and there are

$$\binom{n}{2}$$

such sets.

$$\begin{bmatrix} 0 & X_1 \wedge X_2 & X_1 \wedge X_3 & \dots & X_1 \wedge X_n \\ X_2 \wedge X_1 & 0 & X_2 \wedge X_3 & \dots & X_2 \wedge X_n \\ \vdots & \vdots & 0 & \dots & \vdots \\ X_n \wedge X_1 & X_n \wedge X_2 & X_n \wedge X_3 & \dots & 0 \end{bmatrix}.$$

We shall now prove that the exterior products $\{X_i \wedge X_j\}_{i<j}$ span $\mathcal{A}^2(\mathbb{R}^n)$, that is,

$$[X_i \wedge X_j]_{i<j} = \mathcal{A}^2(\mathbb{R}^n).$$

Indeed, let $\omega^2 \in \mathcal{A}^2(\mathbb{R}^n)$ and $\xi, \eta \in \mathbb{R}^n$. Then

$$\begin{aligned} \omega^2(\xi, \eta) &= \omega^2\left(\sum_{i=1}^n a_i e_i, \sum_{j=1}^n b_j e_j\right) \\ &= \omega^2(a_1 e_1 + \cdots + a_n e_n,\ b_1 e_1 + \cdots + b_n e_n) \\ &= a_1 b_1 \omega^2(e_1, e_1) + \cdots + a_1 b_n \omega^2(e_1, e_n) \\ &+ a_2 b_1 \omega^2(e_2, e_1) + \cdots + a_2 b_n \omega^2(e_2, e_n) \\ &+ \qquad \vdots \qquad \cdots \qquad \vdots \\ &+ a_n b_1 \omega^2(e_n, e_1) + \cdots + a_n b_n \omega^2(e_n, e_n) \\ &= \sum_i \sum_j a_i b_j \omega^2(e_i, e_j) \\ &= \sum_i \sum_j X_i(\xi) X_j(\eta) \omega^2(e_i, e_j) \\ &= \sum_{i<j} X_i(\xi) X_j(\eta) \omega^2(e_i, e_j) + \sum_{i>j} X_i(\xi) X_j(\eta) \omega^2(e_i, e_j) \\ &= \sum_{i<j} X_i(\xi) X_j(\eta) \omega^2(e_i, e_j) + \sum_{i>j} -X_i(\xi) X_j(\eta) \omega^2(e_j, e_i). \end{aligned}$$

Renaming indices $i = k$, $j = l$ in the second sum, we obtain

$$\omega^2(\xi, \eta) = \sum_{i<j} X_i(\xi) X_j(\eta) \omega^2(e_i, e_j) + \sum_{l<k} -X_k(\xi) X_l(\eta) \omega^2(e_l, e_k).$$

Setting $l = i$, $k = j$ in the second sum, we have

$$\begin{aligned}
\omega^2(\xi,\eta) &= \sum_{i<j} X_i(\xi)X_j(\eta)\omega^2(e_i,e_j) + \sum_{i<j} -X_j(\xi)X_i(\eta)\omega^2(e_i,e_j) \\
&= \sum_{i<j}[X_i(\xi)X_j(\eta) - X_j(\xi)X_i(\eta)]\omega^2(e_i,e_j) \\
&= \sum_{i<j}(X_i \wedge X_j)(\xi,\eta)\,\omega^2(e_i,e_j) \\
&= \left(\sum_{i<j}\omega^2(e_i,e_j)\,X_i \wedge X_j\right)(\xi,\eta),
\end{aligned}$$

which proves the claim.

Definition 4.2.2 *Let $\omega_1,\ldots,\omega_k$ $(k \le n)$ be 1-forms on $\mathbb{R}^n$ and consider the map*

$$\begin{aligned}
\omega_1 \wedge \cdots \wedge \omega_k : \mathbb{R}^n \times \cdots \times \mathbb{R}^n &\longrightarrow \mathbb{R} \\
(\xi_1,\xi_2,\ldots,\xi_k) &\longmapsto (\omega_1 \wedge \cdots \wedge \omega_k)(\xi_1,\xi_2,\ldots,\xi_k),
\end{aligned}$$

where

$$(\omega_1 \wedge \cdots \wedge \omega_k)((\xi_1,\xi_2,\ldots,\xi_k)) = \begin{vmatrix} \omega_1(\xi_1) & \cdots & \omega_k(\xi_1) \\ \vdots & \ddots & \vdots \\ \omega_1(\xi_k) & \cdots & \omega_k(\xi_k) \end{vmatrix}.$$

Such application is called the exterior product of the k-forms.

Now, setting

$$\omega : \mathbb{R}^n \longrightarrow \mathbb{R} \times \ldots \times \mathbb{R}, \qquad \xi \longmapsto (\omega_1(\xi),\ldots,\omega_k(\xi)),$$

we have, in particular,

$$\begin{aligned}
\omega(\xi_1) &= (\omega_1(\xi_1),\ldots,\omega_k(\xi_1)), \\
&\vdots \\
\omega(\xi_k) &= (\omega_1(\xi_k),\ldots,\omega_k(\xi_k)).
\end{aligned}$$

It follows that

$(\omega_1 \wedge \ldots \wedge \omega_k)(\xi_1,\ldots,\xi_k) =$ oriented volume of the parallelotope defined by $\omega(\xi_1),\ldots,\omega(\xi_k)$ in $\mathbb{R}^k$,

which is precisely the determinant above.

Note that the exterior product $(\omega_1 \wedge \ldots \wedge \omega_k)$ defined above is a k-form, thanks to the linearity of the maps involved and to the antisymmetry of the determinant.

In this way, the map

$$\begin{aligned} \phi : (\mathbb{R}^n)^* \times \ldots \times (\mathbb{R}^n)^* &\longrightarrow \mathcal{A}^k(\mathbb{R}^n) \\ (\omega_1, \omega_2, \ldots, \omega_k) &\longmapsto \omega_1 \wedge \ldots \wedge \omega_k \end{aligned}$$

is well defined and is multilinear and antisymmetric.

Let $X_1, \ldots, X_k$ be basic 1-forms on $\mathbb{R}^n$, and let $\{i_1, i_2, \ldots, i_k\}$ be a permutation of $\{1, 2, \ldots, k\}$. Then, analogously to the case of $\mathcal{A}^2(\mathbb{R}^n)$, $(X_{i_1} \wedge X_{i_2} \wedge \ldots \wedge X_{i_k})(\xi_1, \xi_2, \ldots, \xi_k)$ represents the oriented volume of the parallelotope generated by the vectors $(\xi_1, \xi_2, \ldots, \xi_k)$ projected onto the subspace $X_{i_1} X_{i_2} \ldots X_{i_k}$.

Our aim from now on is to determine a basis for $\mathcal{A}^k(\mathbb{R}^n)$. We claim that the exterior forms $X_{i_1} \wedge X_{i_2} \wedge \ldots \wedge X_{i_k}$, with $1 \leq i_1 < i_2 < \ldots < i_k \leq n$, are linearly independent, that they are $\begin{pmatrix} n \\ k \end{pmatrix}$ in number and, moreover, that they span $\mathcal{A}^k(\mathbb{R}^n)$, that is, they form a basis of this space. Before that, however, we need some preliminary results, as we shall see below.

Lemma 4.2.3 *Let $\varphi, \psi : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be k-linear maps and let G be a generating set of the vector space $\mathbb{R}^n$. If $\varphi(\xi_1, \xi_2, \ldots, \xi_k) = \psi(\xi_1, \xi_2, \ldots, \xi_k)$ for every k-tuple $(\xi_1, \xi_2, \ldots, \xi_k)$ of elements of G, then $\varphi = \psi$.*

Proof:

We use induction on k.

If $k = 1$, then $\varphi, \psi : \mathbb{R}^n \to \mathbb{R}$ are linear maps. For every $x \in \mathbb{R}^n$ we have

$$x = \sum \alpha_i \xi_i, \qquad \xi_i \in G,$$

since G generates $\mathbb{R}^n$.

Therefore

$$\begin{aligned} \varphi(x) &= \varphi\left(\sum \alpha_i \xi_i\right) = \sum \alpha_i \varphi(\xi_i) \\ &= \sum \alpha_i \psi(\xi_i) = \psi\left(\sum \alpha_i \xi_i\right) = \psi(x). \end{aligned}$$

Assume now that the statement holds for $(k-1)$ and let us prove it for k. Indeed:

For each $\xi \in G$, define $(k-1)$-linear maps

$$\varphi_\xi : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \longrightarrow \mathbb{R}, \qquad \psi_\xi : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \longrightarrow \mathbb{R}$$

by

$$\varphi_\xi(\eta_1, \ldots, \eta_{k-1}) = \varphi(\eta_1, \ldots, \eta_{k-1}, \xi),$$
$$\psi_\xi(\eta_1, \ldots, \eta_{k-1}) = \psi(\eta_1, \ldots, \eta_{k-1}, \xi).$$

By the hypothesis of the lemma, φ_ξ and ψ_ξ take the same values on all $(k-1)$-tuples of elements of G. By the induction hypothesis we conclude that $\varphi_\xi = \psi_\xi$, that is,

$$\varphi(\eta_1, \ldots, \eta_{k-1}, \xi) = \psi(\eta_1, \ldots, \eta_{k-1}, \xi),$$

for all $\eta_1, \ldots, \eta_{k-1} \in \mathbb{R}^n$ and $\xi \in G$.

On the other hand, every element $\eta_k \in \mathbb{R}^n$ is a linear combination of elements of G. Thus, for arbitrary $\eta_1, \ldots, \eta_{k-1} \in \mathbb{R}^n$, we have

$$\begin{aligned} \varphi(\eta_1, \ldots, \eta_k) &= \varphi(\eta_1, \ldots, \eta_{k-1}, \sum_i \alpha_i \xi_i) \\ &= \sum_i \alpha_i \varphi(\eta_1, \ldots, \eta_{k-1}, \xi_i) \\ &= \sum_i \alpha_i \psi(\eta_1, \ldots, \eta_{k-1}, \xi_i) \\ &= \psi(\eta_1, \ldots, \eta_k), \end{aligned}$$

as required. □

Lemma 4.2.4 *Let $\varphi, \psi : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be k-forms and let $\{e_1, e_2, \ldots, e_n\}$ be a basis of $\mathbb{R}^n$. If for every increasing sequence $i_1 < \ldots < i_k$ of k integers between 1 and n we have*

$$\varphi(e_{i_1}, \ldots, e_{i_k}) = \psi(e_{i_1}, \ldots, e_{i_k}),$$

then $\varphi = \psi$.

Proof:

Let $(j_1, j_2, \ldots, j_k)$ be an arbitrary k-tuple of integers between 1 and n. If there are repeated indices in this list, then

$$\varphi(e_{j_1}, \ldots, e_{j_k}) = \psi(e_{j_1}, \ldots, e_{j_k}) = 0.$$

Indeed, suppose there exist $j_s \neq j_t$ with $e_{j_s} = e_{j_t} = e_{j_0}$. Then

$$\varphi(e_{j_1}, \ldots, e_{j_s}, \ldots, e_{j_t}, \ldots, e_{j_k}) = \psi(e_{j_1}, \ldots, e_{j_s}, \ldots, e_{j_t}, \ldots, e_{j_k}),$$

since φ is antisymmetric. Consequently

$$2\,\varphi(e_{j_1}, \ldots, e_{j_0}, \ldots, e_{j_0}, \ldots, e_{j_k}) = 0, \qquad \forall\, e_{j_1}, \ldots, e_{j_k}.$$

Analogously, we obtain the same conclusion for ψ.

If, however, all indices in this list are distinct, then by means of successive transpositions we can rearrange the numbers $j_1, \ldots, j_k$ into increasing order $i_1 < \ldots < i_k$. If ν transpositions are needed, the antisymmetry of φ and ψ, together with the hypothesis of the lemma, give

$$\begin{aligned} \varphi(e_{j_1}, \ldots, e_{j_k}) &= (-1)^{\nu} \varphi(e_{i_1}, \ldots, e_{i_k}) \\ &= (-1)^{\nu} \psi(e_{i_1}, \ldots, e_{i_k}) \\ &= \psi(e_{j_1}, \ldots, e_{j_k}). \end{aligned}$$

Thus, the k-linear maps φ and ψ satisfy the hypothesis of Lemma 1, and are therefore equal. □

Proposition 4.2.5 *Let $\{X_1, X_2, \ldots, X_n\}$ be a basis of $(\mathbb{R}^n)^*$. The k-forms*

$$X_I = X_{i_1} \wedge X_{i_2} \wedge \ldots \wedge X_{i_k},$$

where $I = \{i_1 < \ldots < i_k\}$ runs over the subsets of $\{1, \ldots, n\}$ with k elements, form a basis of $\mathcal{A}^k(\mathbb{R}^n)$. In particular,

$$\dim \mathcal{A}^k(\mathbb{R}) = \begin{pmatrix} n \\ k \end{pmatrix}.$$

Proof:

Let $\omega^k \in \mathcal{A}^k(\mathbb{R})$. For each $I = \{i_1 < \ldots < i_k\}$ set

$$\alpha_I = \omega^k(e_{i_1}, \ldots, e_{i_k}),$$

where $\{e_{i_1}, e_{i_2}, \ldots, e_{i_k}\} \subset \mathbb{R}^n$ is the basis corresponding to the basis of the dual space $(\mathbb{R})^*$, that is, $\{X_{i_1}, X_{i_2}, \ldots, X_{i_k}\} \subset (\mathbb{R}^n)^*$.

The kâ€“form

$$\begin{aligned}\varphi &= \sum_I \alpha_I X_I \\ &= \sum_{1\le i_1<\ldots<i_k\le n} \omega^k(e_{i_1},\ldots,e_{i_k})\, X_{i_1}\wedge X_{i_2}\wedge\ldots\wedge X_{i_k}\end{aligned}$$

is such that, for every increasing sequence $J=\{j_1<\ldots<j_k\}$ of integers between 1 and n, we have

$$\begin{aligned}\varphi(e_{j_1},\ldots,e_{j_k}) &= \sum_I \alpha_I X_I(e_{j_1},\ldots,e_{j_k}) \\ &= \sum_{1\le i_1<\ldots<i_k\le n} \big(\omega^k(e_{i_1},\ldots,e_{i_k})\, X_{i_1}\wedge X_{i_2}\wedge\ldots\wedge X_{i_k}\big)(e_{j_1},\ldots,e_{j_k}) \\ &= \sum_{1\le i_1<\ldots<i_k\le n} \omega^k(e_{i_1},\ldots,e_{i_k})\, (X_{i_1}\wedge X_{i_2}\wedge\ldots\wedge X_{i_k})(e_{j_1},\ldots,e_{j_k}).\end{aligned}$$

On the other hand,

$$(X_{i_1}\wedge X_{i_2}\wedge\ldots\wedge X_{i_k})(e_{j_1},\ldots,e_{j_k}) = \begin{vmatrix} X_{i_1}(e_{j_1}) & X_{i_2}(e_{j_1}) & \cdots & X_{i_k}(e_{j_1}) \\ X_{i_1}(e_{j_2}) & X_{i_2}(e_{j_2}) & \cdots & X_{i_k}(e_{j_2}) \\ \vdots & \vdots & \ddots & \vdots \\ X_{i_1}(e_{j_k}) & X_{i_2}(e_{j_k}) & \cdots & X_{i_k}(e_{j_k}) \end{vmatrix}.$$

There are two cases to consider.

1) $I = J$.

In this case, $i_s = j_s$ for all $s\in\{1,\ldots,k\}$. Indeed, we argue by induction on k. The case $k=2$ has already been proved for 2-forms. Suppose it holds for $(k-1)$ and let us prove it for k.

Since $i_s = j_s$ for all $s\in\{1,\ldots,k-1\}$, we must have $i_k = j_k$. Otherwise, if $i_k\ne j_k$, then $i_k = j_{r_0}$ for some $r_0\in\{1,\ldots,k-1\}$, and also $j_k = i_{s_0}$ for some $s_0\in\{1,\ldots,k-1\}$. Hence

$$j_k = i_{s_0}\le i_{k-1} < i_k, \qquad i_k = j_{r_0}\le j_{k-1} < j_k,$$

which is impossible. It follows that $i_s = j_s$ for all $s\in\{1,\ldots,k\}$, and therefore

$$(X_{i_1}\wedge X_{i_2}\wedge\ldots\wedge X_{i_k})(e_{j_1},\ldots,e_{j_k}) = \begin{vmatrix} X_{i_1}(e_{j_1}) & X_{i_2}(e_{j_1}) & \cdots & X_{i_k}(e_{j_1}) \\ X_{i_1}(e_{j_2}) & X_{i_2}(e_{j_2}) & \cdots & X_{i_k}(e_{j_2}) \\ \vdots & \vdots & \ddots & \vdots \\ X_{i_1}(e_{j_k}) & X_{i_2}(e_{j_k}) & \cdots & X_{i_k}(e_{j_k}) \end{vmatrix} = 1.$$

2) $I \neq J$.

In this case there exists $i_{k_0} \in I$ such that $i_{k_0} \notin J$, more precisely, i_{k_0} is different from every element $j \in J$. Hence

$$X_{i_{k_0}}(e_j) = 0, \qquad \forall j \in J.$$

Thus $(X_{i_1} \wedge X_{i_2} \wedge \ldots \wedge X_{i_k})(e_{j_1}, \ldots, e_{j_k}) = 0$, since it is the determinant of a matrix whose kâ€“th column is zero.

Therefore, going back to the computation of $\varphi(e_{j_1}, \ldots, e_{j_k})$, we obtain

$$\varphi(e_{j_1}, \ldots, e_{j_k}) = \omega^k(e_{j_1}, \ldots, e_{j_k}).$$

By Lemma (2.1.4) it follows that $\varphi = \omega^k$, that is, $\omega^k = \sum_I \alpha_I X_I$. This proves that the k-forms $X_I = X_{i_1} \wedge X_{i_2} \wedge \ldots \wedge X_{i_k}$ span $\mathcal{A}^k(\mathbb{R}^n)$. Moreover, these forms are linearly independent, because from any linear combination $\varphi = \sum_I \alpha_I X_I = 0$ we deduce, for every $J = \{j_1 < \ldots < j_k\}$, that

$$0 = \varphi(e_{j_1}, \ldots, e_{j_k}) = \sum_I \alpha_I X_I(e_{j_1}, \ldots, e_{j_k}) = \alpha_J.$$

□

Remark 4.2.6 *The previous proposition is the most important fact about k-forms. An important special case occurs when* $k = n$. *Then* $\dim \mathcal{A}^k(\mathbb{R}^n) = 1$. *This means that, up to a constant factor, there is only one antisymmetric form of degree* n *on an* n*-dimensional vector space.*

Proposition 4.2.7 *Let* $\varphi : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *be a* k*â€“linear, antisymmetric map. If* $\xi_1, \ldots, \xi_k \in \mathbb{R}^n$ *are linearly dependent, then*

$$\varphi(\xi_1, \xi_2, \ldots, \xi_n) = 0.$$

Proof:

One of the vectors $\xi_1, \ldots, \xi_k$ is a linear combination of the others. Say

$$\xi_1 = a_2\xi_2 + \ldots + a_k\xi_k.$$

Then

$$\begin{aligned} \varphi(k_1, k_2, \ldots, k_k) &= \varphi\left(\sum_{i=2}^{k} a_i\xi_i, \xi_2, \ldots, \xi_k\right) \\ &= \sum_{i=2}^{k} a_i\varphi(\xi_i, \xi_2, \ldots, \xi_k) = 0, \end{aligned}$$

because

$$\varphi(\xi_2, \xi_2, \xi_3, \ldots, \xi_k) = \varphi(\xi_3, \xi_2, \xi_3, \ldots, \xi_k) = \ldots = \varphi(\xi_k, \xi_2, \xi_3, \ldots, \xi_k) = 0,$$

since φ is antisymmetric. □

Corollary 4.2.8 *The exterior product $\omega_1 \wedge \ldots \wedge \omega_k$ is a nonzero kâ€“form if and only if $\omega_1, \ldots, \omega_k$ are linearly independent in $(\mathbb{R}^n)^*$.*

Proof:

Since the map $\phi : (\mathbb{R}^n)^* \times \ldots \times (\mathbb{R}^n)^* \longrightarrow \mathcal{A}^k(\mathbb{R}^n)$ defined by $(\omega_1, \omega_2, \ldots, \omega_k) \mapsto \omega_1 \wedge \ldots \wedge \omega_k$ is kâ€“linear and antisymmetric, it follows from Proposition (1.2.7) that if $\omega_1 \wedge \ldots \wedge \omega_k \neq 0$, then $\omega_1, \ldots, \omega_k$ are linearly independent. Conversely, if these functionals are linearly independent, we may extend them to a basis of $(\mathbb{R}^n)^*$. Let $\{e_1, e_2, \ldots, e_n\} \subset \mathbb{R}^n$ be the basis corresponding to this dual basis. Then for all i, j between 1 and k we have

$$\omega_i(e_j) = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

Hence $(\omega_i(e_j))_{i,j}$ is the $k \times k$ identity matrix, and it follows that

$$(\omega_1 \wedge \ldots \wedge \omega_k)(e_1, e_2, \ldots, e_k) = 1.$$

In particular, $\omega_1 \wedge \ldots \wedge \omega_k \neq 0$. □

Corollary 4.2.9 *If $k > n$, then $\mathcal{A}^k(\mathbb{R}^n) = \{0\}$.*

Proof:

Indeed, in this case any k vectors in $\mathbb{R}^n$ are linearly dependent. □

Remark 4.2.10 *k-forms of the type $(\omega_1 \wedge \ldots \wedge \omega_k)$, where $\omega_1, \ldots, \omega_k \in (\mathbb{R}^n)^*$, are called **decomposable**. It also follows from Proposition 1 that every k-form can be written (in a nonâ€“unique way) as a sum of decomposable k-forms.*

In fact, not every kâ€“form is decomposable, but Proposition 1 shows that every element of $\mathcal{A}^k(\mathbb{R}^n)$ can be written as the sum of decomposable k-forms.

Take a basis $\{X_1, X_2, \ldots, X_n\}$ of $(\mathbb{R}^n)^*$, dual to the basis $\{e_1, e_2, \ldots, e_n\}$ of $\mathbb{R}^n$. Suppose that the linear functionals $\omega_1, \ldots, \omega_k \in (\mathbb{R}^n)^*$ are expressed in terms of the basis $\{X_i\}_{1\leq i\leq n}$ as

$$\omega_i = \sum_{s=1}^{n} a_{is} X_s, \qquad (i = 1, \ldots, k).$$

What are the coordinates of the exterior product $\omega^k = \omega_1 \wedge \ldots \wedge \omega_k$ with respect to the basis $\{X_J\}$ $(J = \{j_1, \ldots < j_k\})$ of the space $\mathcal{A}^k(\mathbb{R}^n)$?

We know that there is a unique expression

$$\omega_1 \wedge \ldots \wedge \omega_k = \sum_J \alpha_J X_J$$

with

$$\alpha_J = \omega^k(e_{j_1}, \ldots, e_{j_k}) = (\omega_1 \wedge \ldots \wedge \omega_k)(e_{j_1}, \ldots, e_{j_k}) = \det(\omega_i(e_{j_s}))_{1\leq i,s\leq k}$$

for every $J = \{j_1 < \ldots < j_k\} \subset \{1, \ldots, n\}$.

Now, the matrix $A = (a_{ij})$, with k rows and n columns, determined by the coordinates of the functionals ω_i relative to the basis $\{X_j\}$, is characterised by

$$a_{ij} = \omega_i(e_j), \qquad 1 \leq i \leq k,\ 1 \leq j \leq n.$$

Indeed, since $\omega_i = \sum_{s=1}^{n} a_{is} X_s$ for $i = 1, \ldots, k$, we have

$$\omega_i(e_j) = a_{i1}X_1(e_j) + \ldots + a_{ij}X_j(e_j) + \ldots + a_{in}X_n(e_j) = a_{ij}.$$

For every subset $J \subset \{1, \ldots, n\}$ with k elements, the matrix A has a $k \times k$ submatrix, denoted by A_J, obtained by selecting the k columns of $A = (a_{ij})$ whose indices j belong to J. Then

$$\alpha_J = \det(A_J) = \det(\omega_i(e_{j_s})).$$

Thus

$$(\omega_1 \wedge \ldots \wedge \omega_k) = \sum_J \alpha_J X_J = \sum_J \det(A_J)\, X_J,$$

the sum being taken over all subsets $J \subset \{1, \ldots, n\}$ with k elements.

More explicitly,

$$\omega_1 \wedge \ldots \wedge \omega_k = \sum_{1\leq j_1<j_2<\ldots<j_k\leq n} \det(\omega_i(e_{j_s}))_{1\leq i,s\leq k}\, (X_{j_1} \wedge X_{j_2} \wedge \ldots \wedge X_{j_k}).$$

In particular, if $k = n$ we obtain

$$\omega_1 \wedge \ldots \wedge \omega_n = \det(A)(X_1 \wedge \ldots \wedge X_n),$$

where $A = (a_{ij})$ is the changeâ€“ofâ€“basis matrix from $\{X_1, X_2, \ldots, X_n\}$ to $\{\omega_1, \omega_2, \ldots, \omega_n\}$.

We now use this result to establish a useful identity for determinants.

Lagrange's Identity

Let $A = (a_{ij})$ be an $n \times k$ matrix, $n \geq k$, and A^* its transpose. Then

$$\det(A^*A) = \sum_J [\det(A_J)]^2,$$

where J runs over all subsets of $\{1, \ldots, n\}$ with k elements and A_J is the matrix obtained from A by choosing the k rows whose indices belong to J.

Indeed, we have

$$A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} \end{vmatrix}, \qquad A^* = \begin{vmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1k} & a_{2k} & \cdots & a_{nk} \end{vmatrix}.$$

Let $v_1, \ldots, v_k \in \mathbb{R}^n$ be the column vectors of the matrix A. For each $j = 1, \ldots, k$ we have $v_j = (a_{1j}, \ldots, a_{nj})$. If $\{X_1, X_2, \ldots, X_n\} \subset (\mathbb{R}^n)^*$ is the canonical basis of the dual space, then

$$X_i(v_j) = X_i\left(\sum_{r=1}^{n} a_{rj} e_r\right) = \sum_{r=1}^{n} a_{rj} X_i(e_r) = a_{ij}.$$

As we saw earlier, for every subset $J = \{j_1 < \ldots < j_k\} \subset \{1, \ldots, n\}$ we have

$$\begin{aligned} X_J(v_1, v_2, \ldots, v_k) &= (X_{j_1} \wedge \ldots \wedge X_{j_k})(v_1, \ldots, v_k) \\ &= \begin{vmatrix} X_{j_1}(v_1) & X_{j_2}(v_1) & \cdots & X_{j_k}(v_1) \\ X_{j_1}(v_2) & X_{j_2}(v_2) & \cdots & X_{j_k}(v_2) \\ \vdots & \vdots & \ddots & \vdots \\ X_{j_1}(v_k) & X_{j_2}(v_k) & \cdots & X_{j_k}(v_k) \end{vmatrix} \\ &= \begin{vmatrix} a_{j_1 1} & a_{j_2 1} & \cdots & a_{j_k 1} \\ a_{j_1 2} & a_{j_2 2} & \cdots & a_{j_k 2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{j_1 k} & a_{j_2 k} & \cdots & a_{j_k k} \end{vmatrix} \\ &= \begin{vmatrix} a_{j_1 1} & a_{j_1 2} & \cdots & a_{j_1 k} \\ a_{j_2 1} & a_{j_2 2} & \cdots & a_{j_2 k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{j_k 1} & a_{j_k 2} & \cdots & a_{j_k k} \end{vmatrix} = \det(A_J). \end{aligned}$$

Now consider the functionals $\omega_1, \ldots, \omega_k \in (\mathbb{R}^n)^*$ given by

$$\omega_i = \sum_{r=1}^{n} a_{ri} X_r.$$

Then, for any $1 \leq i, j \leq k$,

$$\omega_i(v_j) = \sum_{r=1}^{n} a_{ri} X_r(v_j) = \sum_{r=1}^{n} a_{ri} a_{rj},$$

so $\omega_i(v_j)$ is the (i, j)â€“entry of the product matrix $A^* A$.

It follows from the expression obtained earlier,

$$\omega_1 \wedge \ldots \wedge \omega_k = \sum_J \det(A_J) X_J,$$

that

$$\begin{aligned} \det(A^* A) &= \det \left(\left(\sum_{r=1}^{n} a_{ri} a_{rj} \right)_{1 \leq i,j \leq k} \right) \\ &= (\omega_1 \wedge \ldots \wedge \omega_k)(v_1, v_2, \ldots, v_k) \\ &= \sum_J \det(A_J)\, X_J(v_1, v_2, \ldots, v_k) \\ &= \sum_J \det(A_J)\, \det(A_J) \\ &= \sum_J [det(A_J)]^2. \end{aligned}$$

4.3 Exterior Product of Monomials

Let $\omega_1, ..., \omega_k, \omega_{k+1}, ..., \omega_{k+l}$ be 1-forms on $\mathbb{R}^n$. Then $(\omega_1 \wedge ... \wedge \omega_k)$ and $(\omega_{k+1} \wedge ... \wedge \omega_{k+l})$ are, respectively, a decomposable k-form and a decomposable l-form, which we shall simply call **binomials**.

We now define the exterior product of a k-form by an lâ€“form, or, in other words, of two binomials $(\omega_1 \wedge ... \wedge \omega_k)$ and $(\omega_{k+1} \wedge ... \wedge \omega_{k+l})$, obtaining as a result a $(k+l)$-form of the type

$$\omega_1 \wedge ... \wedge \omega_k \wedge \omega_{k+1} \wedge ... \wedge \omega_{k+l}.$$

More precisely, we wish to obtain a bilinear map

$$\varphi : \mathcal{A}^k(\mathbb{R}^n) \times \mathcal{A}^l(\mathbb{R}^n) \longrightarrow \mathcal{A}^{k+l}(\mathbb{R}^n)$$

such that, in particular, for decomposable forms we have

$$\varphi(\omega_1 \wedge ... \wedge \omega_k, \omega_{k+1} \wedge ... \wedge \omega_{k+l}) = \omega_1 \wedge ... \wedge \omega_k \wedge \omega_{k+1} ... \wedge \omega_{k+l}, \tag{1}$$

for any 1-forms $\omega_1, ..., \omega_k, \omega_{k+1}, ..., \omega_{k+l}$.

Before proceeding, let us consider two preliminary results.

Lemma 4.3.1 *Let $\varphi, \psi : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^k$ be bilinear maps. Let G_1 and G_2 be generating sets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. If $\varphi(v, w) = \psi(v, w)$ for all $v \in G_1$ and $w \in G_2$, then $\varphi = \psi$.*

Proof:

Let $G_1 = \{x_1, \ldots, x_p\}$ and $G_2 = \{y_1, \ldots, y_q\}$. Take $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$. Then

$$x = \sum_{i=1}^{p} a_i x_i, \qquad y = \sum_{j=1}^{q} b_j y_j.$$

Thus

$$\begin{aligned}
\varphi(x, y) &= \varphi\left(\sum_{i=1}^{p} a_i x_i, \sum_{j=1}^{q} b_j y_j\right) \\
&= \sum_{i=1}^{p} \sum_{j=1}^{q} a_i b_j \varphi(x_i, y_j) \\
&= \sum_{i=1}^{p} \sum_{j=1}^{q} a_i b_j \psi(x_i, y_j) \\
&= \psi\left(\sum_{i=1}^{p} a_i x_i, \sum_{j=1}^{q} b_j y_j\right) \\
&= \psi(x, y).
\end{aligned}$$

□

Lemma 4.3.2 *Let $\{e_1, e_2, \ldots, e_n\}$ and $\{\overline{e}_1, \ldots, \overline{e}_m\}$ be bases of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. For each pair (i, j) of integers with $1 \leq i \leq n$ and $1 \leq j \leq m$, suppose we are given a vector $\omega_{ij} \in \mathbb{R}^k$. Then there exists a unique bilinear map $\varphi : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^k$ such that $\varphi(e_i, \overline{e}_j) = \omega_{ij}$ for every pair (i, j).*

Proof:

Any two vectors $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ can be written as

$$x = \sum_{i=1}^{n} \alpha_i e_i, \qquad y = \sum_{j=1}^{m} \beta_j \overline{e}_j,$$

where the coefficients α_i and β_j are uniquely determined.

Define the map $\varphi : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^k$ by

$$\varphi(x, y) = \sum_i \sum_j \alpha_i \beta_j \omega_{ij}.$$

Clearly, φ is bilinear and satisfies the condition $\varphi(e_i, \overline{e}_j) = \omega_{ij}$, since the coordinates of e_i in the basis $\{e_k\}$ are 0 when $i \neq k$ and 1 when $i = k$, and likewise the coordinates of $\overline{e}_j$ in the basis $\{\overline{e}_k\}$ are 0 when $j \neq k$ and 1 when $j = k$. Uniqueness follows from Lemma (4.3.2). □

Returning now to our problem, note that since the vector spaces $\mathcal{A}^k(\mathbb{R}^n)$ and $\mathcal{A}^l(\mathbb{R}^n)$ are generated by decomposable forms, Lemma (4.3.2) implies that if there exists a bilinear map

$$\varphi : \mathcal{A}^k(\mathbb{R}^n) \times \mathcal{A}^l(\mathbb{R}^n) \longrightarrow \mathcal{A}^{k+l}(\mathbb{R}^n)$$

satisfying (1), then it is unique.

Take an arbitrary but fixed basis $\{X_1, X_2, \ldots, X_n\}$ of $(\mathbb{R}^n)^*$. For each $I = \{i_1 < \ldots < i_k\}$ and $J = \{i_{k+1}, \ldots, i_{k+l}\}$ contained in $\{1, \ldots, n\}$, set

$$\varphi(X_I, X_J) = X_{i_1} \wedge X_{i_2} \wedge \ldots \wedge X_{i_k} \wedge X_{i_{k+1}} \wedge \ldots \wedge X_{i_{k+l}}, \tag{2}$$

where $X_I = X_{i_1} \wedge X_{i_2} \wedge \ldots \wedge X_{i_k}$ and $X_J = X_{i_{k+1}} \wedge \ldots \wedge X_{i_{k+l}}$.

Recall that the k-forms $X_I = X_{i_1} \wedge X_{i_2} \wedge \ldots \wedge X_{i_k}$, where $I = \{i_1 < \ldots < i_k\}$ runs over the subsets of $\{1, \ldots, n\}$ with k elements, form a basis of $\mathcal{A}^k(\mathbb{R}^n)$.

Moreover, since there are $\begin{pmatrix} n \\ k \end{pmatrix}$ ways to choose k elements among the n elements of $\{1, \ldots, n\}$, we can enumerate these choices by defining, for each $i \in \{1, \ldots, \begin{pmatrix} n \\ k \end{pmatrix}\}$, a set I_i collecting the elements of the iâ€“th choice. Thus there is a bijective correspondence between $(i, j) \in \{1, \ldots, \begin{pmatrix} n \\ k \end{pmatrix}\} \times \left\{1, \ldots, \begin{pmatrix} n \\ l \end{pmatrix}\right\}$ and the pairs (I_i, J_j). Hence, by (2) and Lemma (4.3.2), we can extend φ to a bilinear map from $\mathcal{A}^k(\mathbb{R}^n) \times \mathcal{A}^l(\mathbb{R}^n)$ into $\mathcal{A}^{k+l}(\mathbb{R}^n)$.

Consider the diagram in Figure 4.8:

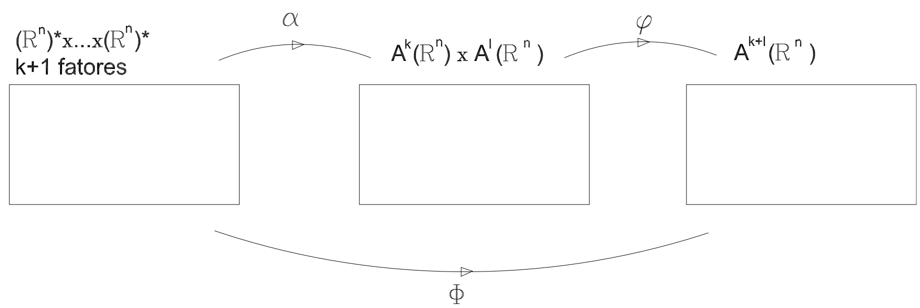

Figure 4.8:

where

$$\begin{aligned}\Phi(\omega_1, \ldots, \omega_k, \omega_{k+1}, \ldots, \omega_{k+l}) &= \omega_1 \wedge \cdots \wedge \omega_k \wedge \omega_{k+1} \wedge \cdots \wedge \omega_{k+l}, \\ \alpha(\omega_1, \ldots, \omega_k, \omega_{k+1}, \ldots, \omega_{k+l}) &= (\omega_1 \wedge \cdots \wedge \omega_k, \omega_{k+1} \wedge \cdots \wedge \omega_{k+l}).\end{aligned}$$

We shall prove that $\varphi \circ \alpha = \Phi$.

According to Lemma (4.2.3) of the previous section, it suffices that these maps coincide on any $(k+l)$-tuple of elements of $\{X_1, X_2, \ldots, X_n\}$. Thus it is enough to prove that

$$(\varphi \circ \alpha)(X_{i_1}, \ldots, X_{i_k}, X_{i_{k+1}}, \ldots, X_{i_{k+l}}) = \Phi(X_{i_1}, \ldots, X_{i_k}, X_{i_{k+1}}, \ldots, X_{i_{k+l}})$$

or, equivalently, that

$$(\varphi \circ \alpha)(X_{i_1} \wedge \ldots \wedge X_{i_k}, X_{i_{k+1}} \wedge \ldots \wedge X_{i_{k+l}}) = \Phi(X_{i_1} \wedge \ldots \wedge X_{i_k}, X_{i_{k+1}} \wedge \ldots \wedge X_{i_{k+l}}) \quad (3)$$

for any $(k+l)$-tuple $(X_{i_1}, \ldots, X_{i_k}, X_{i_{k+1}}, \ldots, X_{i_{k+l}})$ of basis elements.

The equality in (3) is obvious when one of the sequences $(i_1, \ldots, i_k)$ or $(i_{k+1}, \ldots, i_{k+l})$ has repetitions, since both sides are then equal to zero. If neither sequence has repetitions and, moreover, $i_1 < \ldots < i_k$ and $i_{k+1} < \ldots < i_{k+l}$, then the equality in (3) is precisely the definition of φ. Finally, if the sequences have no repetitions but at least one of them is not in increasing order, we may rearrange both into increasing order by successive transpositions. Since each transposition of two indices changes the sign of both sides of (3), we conclude that the equality in question holds in all cases.

Given $\omega^k \in \mathcal{A}^k(\mathbb{R}^n)$ and $\omega^l \in \mathcal{A}^l(\mathbb{R}^n)$, we usually write $\omega^k \overline{\wedge} \omega^l \in \mathcal{A}^{k+l}(\mathbb{R}^n)$ instead of $\varphi(\omega^k, \omega^l)$. In particular, if

$$\omega^k = \omega_1 \wedge \ldots \wedge \omega_k \quad \text{and} \quad \omega^l = \omega_{k+1} \wedge \ldots \wedge \omega_{k+l},$$

then

$$\begin{aligned} \omega^k \overline{\wedge} \omega^l &= \varphi(\omega^k, \omega^l) \\ &= \omega_1 \wedge \cdots \wedge \omega_k \wedge \omega_{k+1} \wedge \cdots \wedge \omega_{k+l} \\ &= \omega^k \wedge \omega^l. \end{aligned}$$

We claim that if $\omega_1^k = \omega_1^1 \wedge \ldots \wedge \omega_k^1$, $\omega_2^k = \omega_1^2 \wedge \ldots \wedge \omega_k^2$ and $\omega^l = \omega_{k+1} \wedge \ldots \wedge \omega_{k+l}$, then

$$(\lambda_1 \omega_1^k + \lambda_2 \omega_2^k) \overline{\wedge} \lambda \omega^l = \lambda_1 \lambda (\omega_1^k \wedge \omega^l) + \lambda_2 \lambda (\omega_2^k \wedge \omega^l). \tag{4}$$

Indeed, by the bilinearity of φ we have

$$\begin{aligned} (\lambda_1 \omega_1^k + \lambda_2 \omega_2^k) \overline{\wedge} \lambda \omega^l &= \varphi(\lambda_1 \omega_1^k + \lambda_2 \omega_2^k, \lambda \omega^l) \\ &= \lambda_1 \lambda \varphi(\omega_1^k, \omega^l) + \lambda_2 \lambda \varphi(\omega_2^k, \omega^l) \\ &= \lambda_1 \lambda (\omega_1^k \overline{\wedge} \omega^l) + \lambda_2 \lambda (\omega_2^k \overline{\wedge} \omega^l) \\ &= \lambda_1 \lambda (\omega_1^k \wedge \omega^l) + \lambda_2 \lambda (\omega_2^k \wedge \omega^l). \end{aligned}$$

Thus, for decomposable forms, the distributive property (4) holds.

The exterior product of decomposable forms enjoys the following properties.

(5) **Anticommutativity:**

$$\omega^k \overline{\wedge} \omega^l = (-1)^{kl} \, \omega^l \overline{\wedge} \omega^k.$$

Indeed, let $\omega^k = \omega_1 \wedge \ldots \wedge \omega_k$ and $\omega^l = \omega_{k+1} \wedge \ldots \wedge \omega_{k+l}$. Then

$$\begin{aligned} \omega^k \overline{\wedge} \omega^l &= \varphi(\omega^k, \omega^l) \\ &= \omega_1 \wedge \ldots \wedge \omega_k \wedge \omega_{k+1} \wedge \ldots \wedge \omega_{k+l} \\ &= (-1)^k (\omega_{k+1} \wedge \omega_1 \wedge \ldots \wedge \omega_k \wedge \omega_{k+2} \wedge \ldots \wedge \omega_{k+l}) \\ &= (-1)^k (-1)^k \ldots (-1)^k (\omega_{k+1} \wedge \omega_{k+2} \wedge \ldots \wedge \omega_{k+l} \wedge \omega_1 \wedge \ldots \wedge \omega_k) \\ &= (-1)^{kl} (\omega^l \overline{\wedge} \omega^k). \end{aligned}$$

(6) **Associativity:**

$$(\omega^k \overline{\wedge} \omega^l) \overline{\wedge} \omega^m = \omega^k \overline{\wedge} (\omega^l \overline{\wedge} \omega^m).$$

This is immediate.

Now, given $\omega^k \in \mathcal{A}^k(\mathbb{R}^n)$ and $\omega^l \in \mathcal{A}^l(\mathbb{R}^n)$, how can we characterise their exterior product $\omega^k \overline{\wedge} \omega^l$?

According to Proposition (1.2.5) of the previous section, we have

$$\omega^k = \sum_I \alpha_I X_I \quad \text{and} \quad \omega^l = \sum_J \alpha_J X_J,$$

where

(i) $I = \{i_1 < \ldots < i_k\}$ and $J = \{j_1 < \ldots < j_l\}$ run over the subsets of $\{1, \ldots, n\}$ with k and l elements, respectively;

(ii) $\alpha_I = \omega^k(e_{i_1}, \ldots, e_{i_k})$ and $\alpha_J = \omega^l(e_{j_1}, \ldots, e_{j_k})$ (where e_j is the basis of $\mathbb{R}^n$);

(iii) $X_I = X_{i_1} \wedge X_{i_2} \wedge \ldots \wedge X_{i_k}$ and $X_J = X_{j_1} \wedge X_{j_2} \wedge \ldots \wedge X_{j_l}$.

Thus, by the bilinearity of φ, we have

$$\begin{aligned} \omega^k \overline{\wedge} \omega^l &= \varphi\left(\sum_I \alpha_I X_I, \sum_J \alpha_J X_J\right) \\ &= \sum_I \sum_J \alpha_I \alpha_J \varphi(X_I, X_J) \\ &= \sum_I \sum_J \alpha_I \alpha_J (X_I \overline{\wedge} X_J) \\ &= \sum_I \sum_J \alpha_I \alpha_J (X_I \wedge X_J). \end{aligned}$$

Therefore, for $(\xi_1, \ldots, \xi_k, \xi_{k+1}, \ldots, \xi_{k+l}) \in \mathbb{R}^n \times \ldots \times \mathbb{R}^n$,

$$\omega^k \overline{\wedge} \omega^l(\xi_1, \ldots, \xi_k, \xi_{k+1}, \ldots, \xi_{k+l}) = \sum_I \sum_J \alpha_I \alpha_J (X_I \wedge X_J)(\xi_1, \ldots, \xi_k, \xi_{k+1}, \ldots, \xi_{k+l}).$$

More explicitly,

$$\begin{aligned} &\omega^k \overline{\wedge} \omega^l(\xi_1, \ldots, \xi_k, \xi_{k+1}, \ldots, \xi_{k+l}) \\ &= \sum_{1 \le i_1 < \ldots < i_k \le n} \sum_{1 \le j_1 < \ldots < j_l \le n} \omega^k(e_{i_1}, \ldots, e_{i_k})\, \omega^l(e_{j_1}, \ldots, e_{j_k})\, (X_{i_1} \wedge \ldots \wedge X_{i_k} \wedge X_{j_1} \wedge \ldots \wedge X_{j_l})(\xi_1, \ldots, \xi_k, \xi_k \end{aligned}$$

Equivalently,

$$\begin{aligned} \omega^k \overline{\wedge} \omega^l(\xi_1, \ldots, \xi_k, \xi_{k+1}, \ldots, \xi_{k+l}) &= \sum_{1 \le i_1 < \ldots < i_k \le n} \sum_{1 \le j_1 < \ldots < j_l \le n} (-1)^\nu \omega^k(e_{i_1}, \ldots, e_{i_k})\, \omega^l(e_{j_1}, \ldots, e_{j_k}) \\ &\cdot (X_{i_1} \wedge \ldots \wedge X_{i_k} \wedge X_{j_1} \wedge \ldots \wedge X_{j_l})(\xi_1, \ldots, \xi_k, \xi_{k+1}, \ldots, \xi_{k+l}), \end{aligned}$$

where

$$\nu = \begin{cases} 0, & \text{if the permutation } \{i_1, \dots, i_k, j_1, \dots, j_l\} \text{ is even,} \\ 1, & \text{if it is odd.} \end{cases}$$

The exterior product $\omega^k \overline{\wedge} \omega^l$ enjoys the properties

(1) **Anticommutativity:**

$$\omega^k \overline{\wedge} \omega^l = (-1)^{kl} (\omega^l \overline{\wedge} \omega^k);$$

(2) **Associativity:**

$$(\omega^k \overline{\wedge} \omega^l) \overline{\wedge} \omega^m = \omega^k \overline{\wedge} (\omega^l \overline{\wedge} \omega^m);$$

(3) **Distributivity:**

$$(\lambda_1 \omega_1^k + \lambda_2 \omega_2^k) \overline{\wedge} \lambda \omega^l = \lambda_1 \lambda (\omega_1^k \overline{\wedge} \omega^l) + \lambda_2 \lambda (\omega_2^k \overline{\wedge} \omega^l).$$

Indeed, it suffices to prove these properties for decomposable forms, since the general case, as we have seen, reduces to sums of decomposable forms. However, this has already been done above.

4.4 Induced Forms

Definition 4.4.1 *Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, $v \mapsto T.v$, be a linear map and consider its transpose*

$$T^* : (\mathbb{R}^m)^* \longrightarrow (\mathbb{R}^n)^*, \qquad w \longmapsto T^*.w,$$

where

$$\begin{array}{rcl} T^*.w : \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ v & \longmapsto & (T^*.w)(v) = w(T.v). \end{array}$$

Thus T^* is well defined, thanks to the linearity of T and w, and moreover the relation

$$(T^*.w)(v) = w(T.v), \quad \forall w \in (\mathbb{R}^m)^*, \ \forall v \in \mathbb{R}^n \tag{1}$$

holds.

Remark 4.4.2 *If $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ and $S : \mathbb{R}^m \longrightarrow \mathbb{R}^p$ are linear maps, then*

$$(S \circ T)^* = T^* \circ S^*.$$

Indeed, let $w \in (\mathbb{R}^p)^*$ and $v \in \mathbb{R}^n$. Then

$$\begin{aligned}
\big[(S \circ T)^* w\big](v) &= w\big[(S \circ T)(v)\big] \\
&= w\big(S(Tv)\big) \\
&= (S^* w)(Tv) \\
&= \big[T^*(S^* w)\big](v) \\
&= \big[(T^* \circ S^*)w\big](v).
\end{aligned}$$

This notion can be generalised. For every k, the linear map $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ determines a new linear map

$$\begin{aligned}
T^* : \mathcal{A}^k(\mathbb{R}^m) &\longrightarrow \mathcal{A}^k(\mathbb{R}^n) \\
w &\longmapsto T^* w
\end{aligned}$$

defined by

$$(T^* w)(v_1, \ldots, v_k) = w(Tv_1, \ldots, Tv_k), \quad \forall w \in \mathcal{A}^k(\mathbb{R}^m),\ \forall v_1, \ldots, v_k \in \mathbb{R}^n. \tag{2}$$

The transformation T^* is said to be *induced by T on forms of degree k*. The kâ€“form $T^* w$ is called the *form induced by T on $\mathbb{R}^n$*, or the *pullâ€“back* of the kâ€“form w to the space $\mathbb{R}^n$.

Proposition 4.4.3 *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map. Then:*

(a) $T^*(w_1 + w_2) = T^*(w_1) + T^*(w_2)$, for all $w_1, w_2 \in \mathcal{A}^k(\mathbb{R}^m)$.

(b) $T^*(\alpha w) = \alpha\, T^*(w)$, for all $\alpha \in \mathbb{R}$ and all $w \in \mathcal{A}^k(\mathbb{R}^m)$.

(c) $T^*(w_1 \wedge w_2) = T^*(w_1) \wedge T^*(w_2)$, for all $w_1, w_2 \in (\mathbb{R}^m)^*$.

Proof:

(a) Let $v_1, \ldots, v_k \in \mathbb{R}^n$ and $w_1, w_2 \in \mathcal{A}^k(\mathbb{R}^m)$. Then

$$\begin{aligned}
[T^*(w_1 + w_2)](v_1, v_2, \ldots, v_k) &= (w_1 + w_2)(Tv_1, Tv_2, \ldots, Tv_k) \\
&= w_1(Tv_1, Tv_2, \ldots, Tv_k) + w_2(Tv_1, Tv_2, \ldots, Tv_k) \\
&= (T^* w_1)(v_1, v_2, \ldots, v_k) + (T^* w_2)(v_1, v_2, \ldots, v_k).
\end{aligned}$$

(b) Let $\alpha \in \mathbb{R}$ and $w \in \mathcal{A}^k(\mathbb{R}^m)$. Then

$$\begin{aligned}
[T^*(\alpha w)](v_1, v_2, \ldots, v_k) &= (\alpha w)(Tv_1, Tv_2, \ldots, Tv_k) \\
&= \alpha\, w(Tv_1, Tv_2, \ldots, Tv_k) \\
&= \alpha [T^*(w)](v_1, v_2, \ldots, v_k).
\end{aligned}$$

(c) Let $w_1, w_2 \in (\mathbb{R}^m)^*$ and $v_1, v_2 \in \mathbb{R}^n$. Then

$$\begin{aligned}
[T^*(w_1 \wedge w_2)](v_1, v_2) &= (w_1 \wedge w_2)(Tv_1, Tv_2) \\
&= \begin{vmatrix} w_1(Tv_1) & w_2(Tv_1) \\ w_1(Tv_2) & w_2(Tv_2) \end{vmatrix} \\
&= \begin{vmatrix} (T^*w_1)(v_1) & (T^*w_2)(v_1) \\ (T^*w_1)(v_2) & (T^*w_2)(v_2) \end{vmatrix} \\
&= \big[(T^*w_1) \wedge (T^*w_2)\big](v_1, v_2).
\end{aligned}$$

□

Remark 4.4.4 *Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear transformation and denote by*

$$T^* : \mathcal{A}^r(\mathbb{R}^m) \to \mathcal{A}^r(\mathbb{R}^n)$$

its induced map (without specifying the degree r in the notation). Then, for $w^k \in \mathcal{A}^k(\mathbb{R}^m)$ and $w^l \in \mathcal{A}^l(\mathbb{R}^m)$, we have

$$T^*(w^k \wedge w^l) = (T^*w^k) \wedge (T^*w^l).$$

Indeed, this relation holds, as we have seen, when w^k and w^l are decomposable forms. Now, for general w^k and w^l, we can write

$$w^k = \sum_I \alpha_I X_I, \qquad w^l = \sum_J \alpha_J X_J.$$

Hence

$$w^k \wedge w^l = \sum_I \sum_J \alpha_I \alpha_J (X_I \wedge X_J).$$

Consequently,

$$\begin{aligned} T^*(w^k \wedge w^l) &= \sum_I \sum_J \alpha_I \alpha_J \, T^*(X_I \wedge X_J) \\ &= \sum_I \sum_J \alpha_I \alpha_J \big(T^* X_I \wedge T^* X_J\big) \\ &= \left(\sum_I \alpha_I \, T^* X_I\right) \wedge \left(\sum_J \alpha_J \, T^* X_J\right) \\ &= T^* \left(\sum_I \alpha_I X_I\right) \wedge T^* \left(\sum_J \alpha_J X_J\right) \\ &= T^*(w^k) \wedge T^*(w^l). \end{aligned}$$

Let $\{X_1, X_2, \ldots, X_n\} \subset (\mathbb{R}^n)^*$ and $\{Y_1, \ldots, Y_m\} \subset (\mathbb{R}^m)^*$ be dual bases of the bases $\{e_1, \ldots, e_n\}$ and $\{f_1, \ldots, f_m\}$, respectively. Let $T = (t_{ij})$ be the matrix (with m rows and n columns) of the linear transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ with respect to these bases, that is,

$$T e_j = \sum_{i=1}^{m} t_{ij} f_i, \qquad j = 1, \ldots, n. \tag{3}$$

Then the matrix of $T^* : (\mathbb{R}^m)^* \longrightarrow (\mathbb{R}^n)^*$ is the transpose of the matrix T.

Indeed, if

$$T^* Y_i = \sum_{j=1}^{n} \bar{t}_{ij} X_j, \qquad i = 1, \ldots, m,$$

then, from (1), for $k \in \{1, \ldots, n\}$ we have

$$(T^* Y_i)(e_k) = Y_i\big(T(e_k)\big), \qquad i = 1, \ldots, m.$$

Thus

$$\left(\sum_{j=1}^{n} \bar{t}_{ij} X_j\right)(e_k) = Y_i \left(\sum_{l=1}^{m} t_{lk} f_l\right).$$

Equivalently,

$$\sum_{j=1}^{n} \bar{t}_{ij} X_j(e_k) = \sum_{l=1}^{m} t_{lk} Y_i(f_l).$$

Hence

$$\bar{t}_{ik} = t_{ik}, \qquad \forall i \in \{1, \ldots, m\},\ \forall k \in \{1, \ldots, n\}.$$

Thus

$$T^* Y_i = \sum_{j=1}^{n} t_{ij} X_j, \qquad i = 1, \ldots, m. \tag{4}$$

We now determine the matrix of the induced linear transformation

$$T^* : \mathcal{A}^k(\mathbb{R}^m) \longrightarrow \mathcal{A}^k(\mathbb{R}^n)$$

with respect to the bases $(Y_I) = (Y_{i_1} \wedge \ldots \wedge Y_{i_k})$ and $(X_J) = (X_{j_1} \wedge \ldots \wedge X_{j_k})$, where the sets $I = \{i_1 < \ldots < i_k\}$ and $J = \{j_1 < \ldots < j_k\}$ run, respectively, over the subsets of $\{1, \ldots, m\}$ and $\{1, \ldots, n\}$ with k elements.

We have $T^* Y_I \in \mathcal{A}^k(\mathbb{R}^n)$ and therefore

$$T^* Y_I = \sum_J \alpha_{IJ} \, X_J,$$

when $J = \{j_1 < \ldots < j_k\}$. Moreover,

$$X_J(e_{j_1}, \ldots, e_{j_k}) = \begin{vmatrix} X_{j_1}(e_{j_1}) & \cdots & X_{j_k}(e_{j_1}) \\ \vdots & \ddots & \vdots \\ X_{j_1}(e_{j_k}) & \cdots & X_{j_k}(e_{j_k}) \end{vmatrix} = 1.$$

Consequently

$$(T^* Y_I)(e_{j_1}, \ldots, e_{j_k}) = \alpha_{IJ}.$$

Thus, from (2) and (3) we obtain

$$\begin{aligned}
\alpha_{IJ} &= (T^* Y_I)(e_{j_1}, \ldots, e_{j_k}) \\
&= Y_I(Te_{j_1}, \ldots, Te_{j_k}) \\
&= \begin{vmatrix} Y_{i_1}(Te_{j_1}) & \cdots & Y_{i_k}(Te_{j_1}) \\ \vdots & \ddots & \vdots \\ Y_{i_1}(Te_{j_k}) & \cdots & Y_{i_k}(Te_{j_k}) \end{vmatrix} \\
&= \begin{vmatrix} Y_{i_1}\left(\sum_{r=1}^m t_{rj_1} f_r\right) & \cdots & Y_{i_k}\left(\sum_{r=1}^m t_{rj_1} f_r\right) \\ \vdots & \ddots & \vdots \\ Y_{i_1}\left(\sum_{r=1}^m t_{rj_k} f_r\right) & \cdots & Y_{i_k}\left(\sum_{r=1}^m t_{rj_k} f_r\right) \end{vmatrix} \\
&= \begin{vmatrix} \sum_{r=1}^m t_{rj_1} Y_{i_1}(f_r) & \cdots & \sum_{r=1}^m t_{rj_1} Y_{i_k}(f_r) \\ \vdots & \ddots & \vdots \\ \sum_{r=1}^m t_{rj_k} Y_{i_1}(f_r) & \cdots & \sum_{r=1}^m t_{rj_k} Y_{i_k}(f_r) \end{vmatrix} \\
&= \begin{vmatrix} t_{i_1 j_1} & \cdots & t_{i_k j_1} \\ \vdots & \ddots & \vdots \\ t_{i_1 j_k} & \cdots & t_{i_k j_k} \end{vmatrix} \\
&= \det(t_{i_\mu j_\nu}),
\end{aligned}$$

with $1 \leq \mu \leq k$ and $1 \leq \nu \leq k$.

Denoting by (T_{IJ}) the $k \times k$ submatrix obtained from the matrix $T = (t_{ij})$ by selecting all entries t_{ij} with $i \in I$ and $j \in J$, we have

$$T^* Y_I = \sum_J \det(T_{IJ})\, X_J. \tag{4}$$

Thus, for each $I = \{i_1 < \ldots < i_k\} \subset \{1, \ldots, m\}$,

$$T^* Y_I = \sum_J \det(T_{IJ})\, X_J, \tag{5}$$

where $J = \{j_1 < \ldots < j_k\}$ runs over the subsets of $\{1, \ldots, n\}$ with k elements. There are

$$\binom{m}{k}$$

sets of type I and

$$\binom{n}{k}$$

sets of type J. Consequently, the matrix of T^* has

$$\binom{n}{k}$$

rows and

$$\binom{m}{k}$$

columns. It is the transpose of the matrix whose entries $\alpha_{IJ} = \det(T_{IJ})$ have as row indices the subsets $I = \{i_1 < \ldots < i_k\} \subset \{1, \ldots, m\}$.

In particular, if $m = n = k$, then the linear map

$$T^* : \mathcal{A}^n(\mathbb{R}^n) \longrightarrow \mathcal{A}^n(\mathbb{R}^n)$$

satisfies

$$T^*(Y_1 \wedge \ldots \wedge Y_n) = \det(T)(X_1 \wedge \ldots \wedge X_n), \tag{6}$$

where $T = (t_{ij})$ is the matrix of $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ as above.

More particularly, if the bases in $(\mathbb{R}^n)^*$ coincide, then

$$T^*(X_1 \wedge \ldots \wedge X_n) = \det(T)(X_1 \wedge \ldots \wedge X_n).$$

We now examine how the coordinates of a form $w^k \in \mathcal{A}^k(\mathbb{R}^n)$ change when we perform a change of basis in $\mathbb{R}^n$.

Let $\{e_1, e_2, \ldots, e_n\}$ and $\{\overline{e}_1, \overline{e}_2, \ldots, \overline{e}_n\}$ be bases in $\mathbb{R}^n$ related by

$$e_j = \sum_{i=1}^{n} a_{ij}\overline{e}_i, \qquad j = 1, \ldots, n. \tag{8}$$

Their dual bases $\{X_1, X_2, \ldots, X_n\}$ and $\{\overline{X}_1, \overline{X}_2, \ldots, \overline{X}_n\}$ in $(\mathbb{R}^n)^*$ satisfy

$$\overline{X}_i = \sum_{j=1}^{n} a_{ij}X_j, \qquad i = 1, \ldots, n. \tag{9}$$

Indeed, if for each $i \in \{1, \ldots, n\}$ we have $\overline{X}_i = \sum_{j=1}^{n} b_{ij}X_j$, then, for all $i, k \in \{1, \ldots, n\}$, we obtain:

$$\text{(i)} \quad \overline{X}_i(e_k) = \sum_{j=1}^{n} b_{ij}X_j(e_k) = b_{ik},$$

$$\text{(ii)} \quad \overline{X}_i(e_k) = \overline{X}_i\Big(\sum_{l=1}^{n} a_{lk}\overline{e}_l\Big) = \sum_{l=1}^{n} a_{lk}\overline{X}_i(\overline{e}_l) = a_{ik}.$$

Thus, from (i) and (ii) we get $a_{ik} = b_{ik}$, which proves (9).

For any subsets $I, J \subset \{1, \ldots, n\}$ with k elements we denote by A_{IJ} the $k \times k$ submatrix of $A = (a_{ij})$ formed by the entries a_{ij} with $i \in I$ and $j \in J$.

From the above it follows that

$$\overline{X}_I = \sum_J \det(A_{IJ})X_J.$$

Indeed, $\overline{X}_I \in \mathcal{A}^k(\mathbb{R}^n)$ and therefore $\overline{X}_I = \sum_J a_{IJ}X_J$, where $a_{IJ} = \overline{X}_I(e_J)$. Hence $\overline{X}_I = \sum_J \overline{X}_I(e_J)X_J$. If $I = \{i_1 < \ldots < i_k\}$ and $J = \{j_1 < \ldots < j_k\}$, then, by (9),

$$\begin{aligned}
\overline{X}_I(e_J) &= \begin{vmatrix} \overline{X}_{i_1}(e_{j_1}) & \cdots & \overline{X}_{i_k}(e_{j_1}) \\ \overline{X}_{i_1}(e_{j_2}) & \cdots & \overline{X}_{i_k}(e_{j_2}) \\ \vdots & \ddots & \vdots \\ \overline{X}_{i_1}(e_{j_k}) & \cdots & \overline{X}_{i_k}(e_{j_k}) \end{vmatrix} \\
&= \begin{vmatrix} \sum_{r=1}^{n} a_{i_1 r}X_r(e_{j_1}) & \cdots & \sum_{r=1}^{n} a_{i_k r}X_r(e_{j_1}) \\ \vdots & \ddots & \vdots \\ \sum_{r=1}^{n} a_{i_1 r}X_r(e_{j_k}) & \cdots & \sum_{r=1}^{n} a_{i_k r}X_r(e_{j_k}) \end{vmatrix} \\
&= \begin{vmatrix} a_{i_1 j_1} & \cdots & a_{i_k j_1} \\ \vdots & \ddots & \vdots \\ a_{i_1 j_k} & \cdots & a_{i_k j_k} \end{vmatrix} \\
&= \det(A_{IJ}).
\end{aligned}$$

Thus, if a form $w^k \in \mathcal{A}^k(\mathbb{R}^n)$ admits the expressions

$$w^k = \sum_J \alpha_J X_J \quad \text{and} \quad w^k = \sum_I \beta_I \overline{X}_I$$

with respect to the bases (X_J) and $(\overline{X}_I)$, we have

$$\begin{aligned} w^k &= \sum_J \alpha_J X_J \\ &= \sum_I \beta_I \sum_J \det(A_{IJ}) X_J \\ &= \sum_I \sum_J \beta_I \det(A_{IJ}) X_J \\ &= \sum_J \left[\sum_I \det(A_{IJ}) \beta_I \right] X_J. \end{aligned}$$

Comparing the coefficients of X_J, we obtain

$$\alpha_J = \sum_I \det(A_{IJ}) \beta_I.$$

Remark 4.4.5 *In the language of classical tensor calculus, a k-form is described, in each basis of $\mathbb{R}^n$, by its coordinates α_J, so that a change of basis in $\mathbb{R}^n$ induces a change of coordinates for the form according to the expression above.*

It is worthwhile to note the particular case in which $w \in \mathcal{A}^n(\mathbb{R}^n)$. We then have the bases $\{X\}$ and $\{\overline{X}\}$ of $\mathcal{A}^n(\mathbb{R}^n)$, with $X = X_1 \wedge \ldots \wedge X_n$ and $\overline{X} = \overline{X}_1 \wedge \ldots \wedge \overline{X}_n$, and $\overline{X} = (\det A) X$, where $A = (a_{ij})$ is the changeâ€“ofâ€“basis matrix. An arbitrary n-form $w \in \mathcal{A}^n(\mathbb{R}^n)$ can be written as

$$w = \alpha X = \beta \overline{X},$$

with $\alpha = (\det A)\beta$, keeping in mind the relations

$$\overline{X}_i = \sum_{j=1}^n a_{ij} X_j,$$

which define the matrix A.

4.5 The Volume Element

Let $\mathbb{R}^n$ be an n-dimensional oriented inner product space. Recall that to orient a vector space is to choose a basis, call it positive, and declare positive every other basis whose change-of-basis matrix has determinant greater than zero.

We now define an n-form called the volume element on $\mathbb{R}^n$. First, choose a positive orthonormal basis $\{e_1, e_2, \ldots, e_n\}$ in $\mathbb{R}^n$. Given a sequence of vectors $\xi_1, \ldots, \xi_n \in \mathbb{R}^n$, for each $j = 1, \ldots, n$ we can write

$$\xi_j = \sum_{i=1}^{n} a_{ij} e_i.$$

Let (A_{ij}) be the resulting $n \times n$ matrix, and define the map

$$\begin{aligned} \omega : \mathbb{R}^n \times \cdots \times \mathbb{R}^n &\longrightarrow \mathbb{R} \\ (\xi_1, \ldots, \xi_n) &\longmapsto \omega(\xi_1, \ldots, \xi_n) = \det(A). \end{aligned}$$

Clearly $\omega \in \mathcal{A}^k(\mathbb{R}^n)$ is alternating and n-linear. We claim that ω does not depend on the choice of basis we made. To prove this, we introduce the *Gram matrix* $G = (\langle \xi_i, \xi_j \rangle)$, whose entry in the i-th row and j-th column is the inner product $\langle \xi_i, \xi_j \rangle$. Indeed,

$$\begin{aligned} \langle \xi_i, \xi_j \rangle &= \left\langle \sum_{k=1}^{n} a_{ki} e_k, \sum_{s=1}^{n} a_{sj} e_s \right\rangle \\ &= \sum_{k=1}^{n} \sum_{s=1}^{n} a_{ki} a_{sj} \langle e_k, e_s \rangle \\ &= \sum_{k=1}^{n} a_{ki} a_{kj}. \end{aligned}$$

It follows that $G = A^* A$, where A^* is the transpose of A. In fact,

$$\begin{aligned} AA^* &= \begin{vmatrix} \sum_k a_{k1} a_{k1} & \cdots & \sum_k a_{k1} a_{kn} \\ \vdots & \ddots & \vdots \\ \sum_k a_{kn} a_{k1} & \cdots & \sum_k a_{kn} a_{kn} \end{vmatrix} \\ &= \begin{vmatrix} \langle \xi_1, \xi_1 \rangle & \cdots & \langle \xi_1, \xi_n \rangle \\ \vdots & \ddots & \vdots \\ \langle \xi_n, \xi_1 \rangle & \cdots & \langle \xi_n, \xi_n \rangle \end{vmatrix}. \end{aligned}$$

Thus

$$\det G = \det(AA^*) = \det A \cdot \det A^* = (\det A)^2.$$

In particular, $\det G \geq 0$, and $\det G = 0$ if and only if the vectors $\xi_1, \ldots, \xi_n$ are linearly dependent.

We conclude that

$$\omega(\xi_1, \ldots, \xi_n) = \pm\sqrt{\det\big(\langle \xi_i, \xi_j \rangle\big)},$$

where the sign $+$ or $-$ is the sign of $\det A$. The equality above shows that the definition of ω is independent of the choice of basis. Thus $\omega(\xi_1, \ldots, \xi_n) > 0$ when the vectors $\xi_1, \ldots, \xi_n$ form a positive basis, and if $\xi_1, \ldots, \xi_n$ are linearly dependent then

$$\omega(\xi_1, \ldots, \xi_n) = 0.$$

In the case where $\mathbb{R}^n$ is the n-dimensional Euclidean space, $|\det(A)|$ is the volume of the parallelepiped with edges $\xi_1, \ldots, \xi_n$, so that $\omega(\xi_1, \ldots, \xi_n)$ is the *oriented volume*, as mentioned in Section 1.

4.6 The Cross Product

Let $\mathbb{R}^3$ be three-dimensional Euclidean space.

The cross product

$$\times : \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}^3, \quad (\xi, \eta) \longmapsto \xi \times \eta,$$

is the bilinear map defined as follows.

Consider the canonical basis $\{e_1, e_2, e_3\} \subset \mathbb{R}^3$. Set

$$e_1 \times e_1 = e_2 \times e_2 = e_3 \times e_3 = 0,$$

$$e_1 \times e_2 = -e_2 \times e_1 = e_3,$$

$$e_2 \times e_3 = -e_3 \times e_2 = e_1,$$

$$e_3 \times e_1 = -e_1 \times e_3 = e_2.$$

Now, given arbitrary vectors $\xi = (x_1, x_2, x_3)$ and $\eta = (y_1, y_2, y_3)$, we have

$$\begin{aligned} \xi \times \eta &= (x_1 e_1 + x_2 e_2 + x_3 e_3) \times (y_1 e_1 + y_2 e_2 + y_3 e_3) \\ &= (x_2 y_3 - x_3 y_2) e_1 + (x_3 y_1 - x_1 y_3) e_2 + (x_1 y_2 - x_2 y_1) e_3. \end{aligned}$$

Clearly $\xi \times \xi = 0$. Thus the cross product is a bilinear and antisymmetric map. Note that the inner products $\langle \xi \times \eta, \xi \rangle$ and $\langle \xi \times \eta, \eta \rangle$ are zero, and consequently the cross product $\xi \times \eta$ is perpendicular to the plane spanned by these vectors (if ξ and η are linearly dependent, they do not span a plane, and in this case we have $\xi \times \eta = 0$).

For each $\eta \in \mathbb{R}^3$, define $\omega_\eta^1 : \mathbb{R}^3 \to \mathbb{R}$ by

$$\omega_\eta^1(\xi) = \langle \eta, \xi \rangle.$$

Then ω_η^1 is a 1-form. Consider the map

$$\begin{array}{rcl} \varphi : \mathbb{R}^3 & \longrightarrow & (\mathbb{R}^3)^* \\ \eta & \longmapsto & \omega_\eta^1. \end{array}$$

(i) φ is linear:

Let $\eta_1, \eta_2 \in \mathbb{R}^3$ and $\alpha, \beta \in \mathbb{R}$. Then

$$\begin{array}{rcl} \omega^1_{\alpha\eta_1+\beta\eta_2}(\xi) & = & \langle \alpha\eta_1 + \beta\eta_2, \xi \rangle \\ & = & \langle \alpha\eta_1, \xi \rangle + \langle \beta\eta_2, \xi \rangle \\ & = & \omega^1_{\alpha\eta_1}(\xi) + \omega^1_{\beta\eta_2}(\xi), \quad \forall \xi \in \mathbb{R}^3. \end{array}$$

(ii) φ is one-to-one.

If $\omega^1_{\eta_1} = \omega^1_{\eta_2}$, then $\omega^1_{\eta_1}(\xi) = \omega^1_{\eta_2}(\xi)$ for all $\xi \in \mathbb{R}^3$. Thus

$$\langle \eta_1 - \eta_2, \xi \rangle = 0, \quad \forall \xi \in \mathbb{R}^3.$$

In particular, taking $\xi = \eta_1 - \eta_2$, we obtain

$$\langle \eta_1 - \eta_2, \eta_1 - \eta_2 \rangle = 0,$$

which implies $\eta_1 = \eta_2$.

Since φ is surjective by construction, it follows that φ is an isomorphism. Hence $\mathbb{R}^3$ and $(\mathbb{R}^3)^*$ are isomorphic vector spaces.

Now, for each $\eta \in \mathbb{R}^3$, define

$$\begin{array}{rcl} \omega_\eta^2 : \mathbb{R}^3 \times \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ (\xi_1, \xi_2) & \longmapsto & \omega_\eta^2(\xi_1, \xi_2) = \langle \eta, \xi_1 \times \xi_2 \rangle. \end{array}$$

The map ω^2_η is a 2-form, which motivates the definition of

$$\begin{aligned} \psi : \mathbb{R}^3 &\longrightarrow \mathcal{A}^2(\mathbb{R}^3) \\ \eta &\longmapsto \omega^2_\eta. \end{aligned}$$

Arguing as we did for the map φ, we obtain an isomorphism between $\mathbb{R}^3$ and $\mathcal{A}^2(\mathbb{R}^3)$. In fact, $n = 3$ is the only dimension in which this happens, since

$$\dim \mathcal{A}^2(\mathbb{R}^n) = n \iff n = 3.$$

Indeed,

$$\dim \mathcal{A}^2(\mathbb{R}^n) = \begin{pmatrix} n \\ 2 \end{pmatrix} = \frac{n!}{2!(n-2)!}.$$

Thus,

$$\frac{n!}{2!(n-2)!} = n \iff \frac{n(n-1)(n-2)!}{2!(n-2)!} = n \iff n^2 - 3n = 0 \iff n = 3.$$

Given $\eta = (x_1, x_2, x_3) \in \mathbb{R}^3$ and the projections $X_i : \mathbb{R}^3 \to \mathbb{R}$ $(i = 1, 2, 3)$, we have

$$\omega^1_\eta = x_1 X_1 + x_2 X_2 + x_3 X_3.$$

Indeed, take $\xi = (y_1, y_2, y_3) \in \mathbb{R}^3$. Then

$$\begin{aligned} \big(x_1X_1 + x_2X_2 + x_3X_3\big)(\xi) &= x_1X_1(\xi) + x_2X_2(\xi) + x_3X_3(\xi) \\ &= x_1y_1 + x_2y_2 + x_3y_3 \\ &= \langle \eta, \xi \rangle = \omega^1_\eta(\xi). \end{aligned}$$

Thus we have the isomorphism

$$x_1e_1 + x_2e_2 + x_3e_3 \in \mathbb{R}^3 \ \approx \ x_1X_1 + x_2X_2 + x_3X_3 \in (\mathbb{R}^3)^*.$$

For $\eta = (x_1, x_2, x_3)$ we claim that

$$\omega^2_\eta = x_1(X_2 \wedge X_3) + x_2(X_1 \wedge X_3) + x_3(X_1 \wedge X_2).$$

Indeed, with $\xi_1 = (y_1, y_2, y_3)$ and $\xi_2 = (z_1, z_2, z_3)$,

$$\begin{aligned}
&\big(x_1(X_2 \wedge X_3) + x_2(X_1 \wedge X_3) + x_3(X_1 \wedge X_2)\big)(\xi_1, \xi_2) \\
&\quad = x_1 \begin{vmatrix} X_2(\xi_1) & X_3(\xi_1) \\ X_2(\xi_2) & X_3(\xi_2) \end{vmatrix} + x_2 \begin{vmatrix} X_1(\xi_1) & X_3(\xi_1) \\ X_1(\xi_2) & X_3(\xi_2) \end{vmatrix} + x_3 \begin{vmatrix} X_1(\xi_1) & X_2(\xi_1) \\ X_1(\xi_2) & X_2(\xi_2) \end{vmatrix} \\
&\quad = x_1 \begin{vmatrix} y_2 & y_3 \\ z_2 & z_3 \end{vmatrix} + x_2 \begin{vmatrix} y_1 & y_3 \\ z_1 & z_3 \end{vmatrix} + x_3 \begin{vmatrix} y_1 & y_2 \\ z_1 & z_2 \end{vmatrix} \\
&\quad = x_1(y_2 z_3 - z_2 y_3) + x_2(y_1 z_3 - z_1 y_3) + x_3(y_1 z_2 - z_1 y_2) \\
&\quad = \langle \eta, \xi_1 \times \xi_2 \rangle \\
&\quad = \omega^2_\eta(\xi_1, \xi_2).
\end{aligned}$$

Therefore we have the isomorphism

$$x_1 e_1 + x_2 e_2 + x_3 e_3 \in \mathbb{R}^3 \;\approx\; x_1(X_2 \wedge X_3) + x_2(X_1 \wedge X_3) + x_3(X_1 \wedge X_2) \in \mathcal{A}^2(\mathbb{R}^3).$$

On the other hand, given $\eta_1 = (y_1, y_2, y_3)$ and $\eta_2 = (z_1, z_2, z_3)$,

$$\begin{aligned}
\omega^1_{\eta_1} \wedge \omega^1_{\eta_2} &= (y_1 X_1 + y_2 X_2 + y_3 X_3) \wedge (z_1 X_1 + z_2 X_2 + z_3 X_3) \\
&= y_1 z_2 (X_1 \wedge X_2) + y_1 z_3 (X_1 \wedge X_3) + y_2 z_1 (X_2 \wedge X_1) \\
&\quad + y_2 z_3 (X_2 \wedge X_3) + y_3 z_1 (X_3 \wedge X_1) + y_3 z_2 (X_3 \wedge X_2) \\
&= (y_1 z_2 - z_1 y_2)(X_1 \wedge X_2) + (y_2 z_3 - z_2 y_3)(X_2 \wedge X_3) \\
&\quad + (y_1 z_3 - z_1 y_3)(X_1 \wedge X_3) \\
&= \omega^2_{\eta_1 \times \eta_2}.
\end{aligned}$$

Thus

$$\omega^1_{\eta_1} \wedge \omega^1_{\eta_2} = \omega^2_{\eta_1 \times \eta_2}.$$

From now on, our aim is to generalise the cross product.

We define the 'cross product'

$$\xi_1 \times \cdots \times \xi_n$$

of n vectors in $\mathbb{R}^{n+1}$ as the vector $\xi \in \mathbb{R}^{n+1}$ such that, for every $\eta \in \mathbb{R}^{n+1}$,

$$\langle \eta, \xi \rangle = \det(\eta, \xi_1, \ldots, \xi_n),$$

where $\det(\eta, \xi_1, \ldots, \xi_n)$ is the determinant of the $(n+1) \times (n+1)$ matrix whose columns are the vectors $\eta, \xi_1, \ldots, \xi_n$, in this order. That is, if

$$\eta = (y_1, \ldots, y_{n+1}),$$

$$\xi_1 = (x_{11}, x_{21}, \ldots, x_{(n+1)1}),$$
$$\xi_2 = (x_{12}, x_{22}, \ldots, x_{(n+1)2}),$$
$$\vdots \quad \vdots \qquad \vdots$$
$$\xi_n = (x_{1n}, x_{2n}, \ldots, x_{(n+1)n}),$$

then the cross product $\xi_1 \times \cdots \times \xi_n$ is the vector $\xi \in \mathbb{R}^{n+1}$ such that, for every $\eta \in \mathbb{R}^{n+1}$,

$$\langle \eta, \xi \rangle = \begin{vmatrix} y_1 & x_{11} & \cdots & x_{1n} \\ y_2 & x_{21} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n+1} & x_{(n+1)1} & \cdots & x_{(n+1)n} \end{vmatrix}_{(n+1)\times(n+1)}.$$

The coordinates of the vectors

$$\xi_j = \sum_{i=1}^{n+1} x_{ij} e_i, \quad j = 1, \ldots, n,$$

with respect to the canonical basis of $\mathbb{R}^{n+1}$ form a matrix $M = (x_{ij})$ with $(n+1)$ rows and n columns. We denote by $M_{(i)}$ the $n \times n$ matrix obtained from M by omitting its i-th row.

By the definition of the cross product, for each $i = 1, \ldots, n+1$ we have

$$\begin{aligned} \langle e_i, \xi_1 \times \cdots \times \xi_n \rangle &= \det(e_i, \xi_1, \ldots, \xi_n) \\ &= \begin{vmatrix} 0 & x_{11} & x_{12} & \cdots & x_{1n} \\ 0 & x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{i1} & x_{i2} & \cdots & x_{in} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & x_{(n+1)1} & x_{(n+1)2} & \cdots & x_{(n+1)n} \end{vmatrix} \\ &= (-1)^{i+1} \det(M_{(i)}), \end{aligned}$$

where the last equality follows from expanding the determinant along the first column. Therefore

$$\xi_1 \times \cdots \times \xi_n = \sum_{i=1}^{n+1} (-1)^{i+1} \det(M_{(i)}) e_i$$

is the expression of the vector $\xi_1 \times \cdots \times \xi_n$ in the canonical basis of $\mathbb{R}^{n+1}$.

From this, or directly from the definition, we see that the cross product is an n-linear antisymmetric map from $\mathbb{R}^{n+1} \times \cdots \times \mathbb{R}^{n+1}$ to $\mathbb{R}^{n+1}$.

The cross product enjoys the following properties:

(1) $\xi_1 \times \cdots \times \xi_n = 0$ whenever the vectors $\xi_1, \ldots, \xi_n$ are linearly dependent.

Indeed, this follows from the fact that the determinant in the definition is alternating (and hence vanishes when two of the columns coincide).

(2) $\xi_1 \times \cdots \times \xi_n$ is perpendicular to each ξ_j.

In fact, by definition

$$\langle \xi_j, \xi_1 \times \cdots \times \xi_n \rangle = \det(\xi_j, \xi_1, \ldots, \xi_n) = 0,$$

since this determinant has two equal columns.

(3) $\|\xi_1 \times \cdots \times \xi_n\|$ is the volume of the parallelepiped generated by the vectors $\xi_1, \ldots, \xi_n$.

We know that the volume of the parallelepiped $[\xi_1, \ldots, \xi_n]$ is $\sqrt{\det G}$, where the Gram matrix G has as its (i,j)-entry the inner product

$$\langle \xi_i, \xi_j \rangle = \sum_{k=1}^{n+1} x_{ki} x_{kj}.$$

Hence

$$\begin{aligned} M^* M &= \begin{vmatrix} \sum_{k=1}^{n+1} x_{k1}x_{k1} & \cdots & \sum_{k=1}^{n+1} x_{k1}x_{kn} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{n+1} x_{kn}x_{k1} & \cdots & \sum_{k=1}^{n+1} x_{kn}x_{kn} \end{vmatrix} \\ &= \begin{vmatrix} \langle \xi_1, \xi_1 \rangle & \cdots & \langle \xi_1, \xi_n \rangle \\ \vdots & \ddots & \vdots \\ \langle \xi_n, \xi_1 \rangle & \cdots & \langle \xi_n, \xi_n \rangle \end{vmatrix} = G. \end{aligned}$$

Thus, by Lagrange's identity we can write

$$\begin{aligned} \text{vol}[\xi_1, \ldots, \xi_n] &= \sqrt{\det G} \\ &= \sqrt{\det(M^* M)} \\ &= \sqrt{\sum_{i=1}^{n+1} [\det(M_{(i)})]^2} \\ &= \|\xi_1 \times \cdots \times \xi_n\|. \end{aligned}$$

(4) $\det(\xi_1 \times \cdots \times \xi_n, \xi_1, \ldots, \xi_n) > 0$ whenever the vectors $\xi_1, \ldots, \xi_n$ are linearly independent.

Indeed, observe that by definition (in particular when $\eta = \xi = \xi_1 \times \cdots \times \xi_n$),

$$\begin{aligned}\det(\xi_1 \times \cdots \times \xi_n, \xi_1, \ldots, \xi_n) &= \langle \xi_1 \times \cdots \times \xi_n, \xi_1 \times \cdots \times \xi_n \rangle \\ &= \|\xi_1 \times \cdots \times \xi_n\|^2 = \det G,\end{aligned}$$

and $\det G \neq 0$ if and only if $\{\xi_1, \ldots, \xi_n\}$ is linearly independent.

We now show that the properties above characterise the cross product. Indeed:

By property (1) it suffices to consider the case where $\xi_1, \ldots, \xi_n$ are linearly independent. Property (2) determines the direction of the vector $\xi_1 \times \cdots \times \xi_n$, that is, the line through the origin containing it. Property (3) gives its length, while property (4) tells us which of the two non-zero vectors with that length is the cross product, namely the one for which $(\xi_1 \times \cdots \times \xi_n, \xi_1, \ldots, \xi_n)$ is a positive basis of $\mathbb{R}^{n+1}$; in other words, it fixes the 'sense' of this vector along that direction.

Thus, from all the above, the cross product $\xi_1 \times \cdots \times \xi_n$ does not depend on the positive orthonormal basis chosen to define it. The cross product is an nâ€“linear antisymmetric map

$$\begin{aligned}\times : \mathbb{R}^{n+1} \times \cdots \times \mathbb{R}^{n+1} &\longrightarrow \mathbb{R}^{n+1} \\ (\xi_1, \ldots, \xi_n) &\longmapsto \xi_1 \times \cdots \times \xi_n = \sum_{i=1}^{n+1} (-1)^{i+1} \det(M_{(i)}) e_i.\end{aligned}$$

We can generalise the isomorphisms constructed at the beginning of this section by defining the following maps.

For each $\eta \in \mathbb{R}^{n+1}$, define

$$\begin{aligned}\omega_\eta^1 : \mathbb{R}^{n+1} &\longrightarrow \mathbb{R} \\ \xi &\longmapsto \omega_\eta^1(\xi) = \langle \eta, \xi \rangle\end{aligned}$$

and the map

$$\begin{aligned}\varphi : \mathbb{R}^{n+1} &\longrightarrow (\mathbb{R}^{n+1})^* \\ \eta &\longmapsto \omega_\eta^1.\end{aligned}$$

This is an isomorphism between $\mathbb{R}^{n+1}$ and $(\mathbb{R}^{n+1})^*$.

Now, for each $\eta \in \mathbb{R}^{n+1}$ define

$$\begin{aligned}\omega_\eta^n : \mathbb{R}^{n+1} \times \cdots \times \mathbb{R}^{n+1} &\longrightarrow \mathbb{R} \\ (\xi_1, \ldots, \xi_n) &\longmapsto \omega_\eta^n(\xi_1, \ldots, \xi_n) = \langle \eta, \xi_1 \times \cdots \times \xi_n \rangle\end{aligned}$$

and the map

$$\begin{aligned} \psi : \mathbb{R}^{n+1} &\longrightarrow \mathcal{A}^n(\mathbb{R}^{n+1}) \\ \eta &\longmapsto \omega_\eta^n. \end{aligned}$$

This is an isomorphism between $\mathbb{R}^{n+1}$ and $\mathcal{A}^n(\mathbb{R}^{n+1})$.

In an analogous way, given $\eta = (x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1}$ and the projections

$$X_i : \mathbb{R}^{n+1} \to \mathbb{R}, \quad i = 1, \dots, n+1,$$

we have

$$\omega_\eta^1 = \sum_{i=1}^{n+1} x_i X_i.$$

Note that we have

$$\binom{n+1}{n} = n+1$$

subsets of the form $I = \{i_1 < \cdots < i_n\}$ running through the set $\{1, \dots, n+1\}$. Thus there are as many such subsets as there are coordinates of the vector $\eta = (x_1, \dots, x_{n+1})$. Therefore, for each $i \in \{1, \dots, n+1\}$ we choose one of the I_i in such a way that I_i does not contain the element i. Then

$$X_\eta^n = \sum_i x_i X_{I_i}.$$

More precisely,

$$X_\eta^n = x_1(X_2 \wedge X_3 \wedge \cdots \wedge X_{n+1}) + x_2(X_1 \wedge X_3 \wedge \cdots \wedge X_{n+1}) + \cdots + x_{n+1}(X_1 \wedge X_2 \wedge \cdots \wedge X_n).$$

Finally, given $\eta_1, \dots, \eta_n \in \mathbb{R}^{n+1}$ we also have

$$\omega_{\eta_1}^1 \wedge \cdots \wedge \omega_{\eta_n}^1 = \omega_{\eta_1 \times \cdots \times \eta_n}^n.$$

We leave the proof of these results to the reader.

Chapter 5

Differential Forms in $\mathbb{R}^n$

5.1 Differential Forms in $\mathbb{R}^n$

Definition 5.1.1 *Let $p \in \mathbb{R}^n$ be an arbitrary but fixed point. The set of vectors attached at p is called the tangent space of $\mathbb{R}^n$ at p, denoted by $\mathbb{R}^n_p$. More precisely,*

$$\mathbb{R}^n_p = \{(p, v);\ v \in \mathbb{R}^n\}.$$

We sometimes use the notation $v(p)$ or simply v_p to represent the elements $(p, v) \in \mathbb{R}^n_p$. To fix ideas, consider the diagram below:

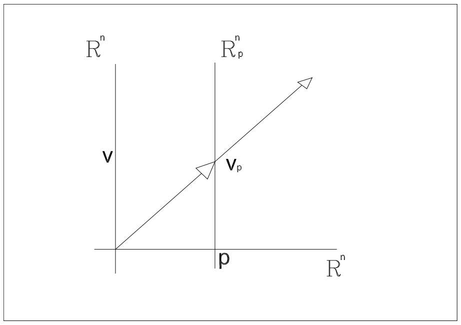

Figure 5.1:

$\mathbb{R}^n_p$ is a vector space endowed with the operations

$$(p, v_1) + (p, v_2) = (p, v_1 + v_2),$$

$$\alpha(p, v_1) = (p, \alpha v_1), \quad \forall v_1, v_2 \in \mathbb{R}^n \text{ and } \forall \alpha \in \mathbb{R},$$

that is, the natural addition and scalar multiplication.

In fact, $\mathbb{R}^n_p$ is a vector space isomorphic to $\mathbb{R}^n$. Indeed, consider the linear bijection from $\mathbb{R}^n$ onto $\mathbb{R}^n_p$ given by $v \mapsto (p, v) = v_p$. Through this isomorphism the vectors of the canonical basis $\{e_1, \ldots, e_n\}$ of $\mathbb{R}^n$ can be 'identified' with their translates $\{(e_1)_p, \ldots, (e_n)_p\}$ at the point p.

Definition 5.1.2 *We define an inner product on $\mathbb{R}^n_p$ by setting, for $v_p, w_p \in \mathbb{R}^n_p$,*

$$\langle v_p, w_p \rangle_p = \langle v, w \rangle.$$

Definition 5.1.3 *A vector field on $\mathbb{R}^n$ is a map $v : \mathbb{R}^n \to \mathbb{R}^n_p$ which associates to each point $p \in \mathbb{R}^n$ a vector $v(p) \in \mathbb{R}^n_p$.*

In view of the identification above, v can be written in the form

$$v(p) = \sum_{i=1}^{n} a_i(p) e_i.$$

The vector field v is said to be differentiable when the functions $a_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \ldots, n$ are differentiable.

Definition 5.1.4 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable map. We denote by df_p the differential of the function f at the point p in $\mathbb{R}^n_p$ and define*

$$\begin{array}{rcl} df_p : \mathbb{R}^n_p & \longrightarrow & \mathbb{R} \\ (p, v) = v_p & \longmapsto & (df)(p)(v_p) = Df(p)(v), \end{array}$$

where $Df(p) : \mathbb{R}^n \to \mathbb{R}$ is the (Fréchet) differential of f at the point p in $\mathbb{R}^n$. Clearly, $(df)(p)$ is a linear map and therefore $df_p \in (\mathbb{R}^n_p)^*$.

For each tangent space $\mathbb{R}^n_p$, consider the dual space $(\mathbb{R}^n_p)^*$. A basis for $(\mathbb{R}^n_p)^*$ is obtained by taking

$$(dX_i)(p), \quad i = 1, \ldots, n,$$

where $X_i : \mathbb{R}^n \to \mathbb{R}$ is the i-th coordinate projection.

Indeed, if $v_p \in \mathbb{R}^n_p$ then $v_p = (p, v)$ and $v = \sum_{i=1}^n v_i e_i$. Thus, by the previous definition we have

$$(dX_i)(p)(v_p) = (DX_i)(p)(v) = \sum_{j=1}^{n} \frac{\partial X_i}{\partial x_j}(p) v_j.$$

However,

$$\frac{\partial X_i}{\partial x_j}(p) = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

Hence

$$(dX_i)(p)(v_p) = v_i. \tag{1}$$

In particular, for $v_p = (e_j)_p$ we get

$$(dX_i)(p)(e_j)_p = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

Thus $\{(dX_i)_p;\ 1 \leq i \leq n\}$ is the 'dual basis' of $\{(e_i)_p;\ 1 \leq i \leq n\}$, that is, the dual basis of $\mathbb{R}^n_p$.

Proposition 5.1.5 *If $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable, then*

$$df = \frac{\partial f}{\partial x_1} dX_1 + \cdots + \frac{\partial f}{\partial x_n} dX_n.$$

Proof: For every $v_p = (p, v)$ with $v = \sum_{i=1}^n v_i e_i$, by Definition (2.1.4) and (1),

$$\begin{aligned} (df)(p)(v_p) &= Df(p)(v) \\ &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(p)\, v_i \\ &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(p)(dX_i)(p)(v_p) \\ &= \sum_{i=1}^n \big(\tfrac{\partial f}{\partial x_i} dX_i\big)(p)(v_p). \end{aligned}$$

By the arbitrariness of $v_p \in \mathbb{R}^n_p$ we obtain

$$(df)(p) = \sum_{i=1}^n \big(\tfrac{\partial f}{\partial x_i} dX_i\big)(p),$$

and, by the arbitrariness of $p \in \mathbb{R}^n$, we conclude

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dX_i.$$

□

Remark 5.1.6 *The differential df can be understood as the map which to each $p \in \mathbb{R}^n$ associates $df_p \in (\mathbb{R}^n_p)^*$, where*

$$df_p : \mathbb{R}^n_p \longrightarrow \mathbb{R},$$
$$v_p \longmapsto (df)(p)(v_p) = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i} dX_i \right)(p)(v_p) = Df(p)(v),$$

and which, by a slight abuse of terminology, we also call the differential of f.

Let $\mathcal{A}^k(\mathbb{R}^n_p)$ be the vector space of k-forms

$$\omega^k : \mathbb{R}^n_p \times \cdots \times \mathbb{R}^n_p \to \mathbb{R}. \tag{2}$$

If $\omega_1, \ldots, \omega_k$ are 1-forms, we can obtain from them a kâ€“form as in (2),

$$\omega_1 \wedge \cdots \wedge \omega_k,$$

(called decomposable) defined by

$$\begin{array}{rcl} \omega_1 \wedge \cdots \wedge \omega_k : \mathbb{R}^n_p \times \cdots \times \mathbb{R}^n_p & \longrightarrow & \mathbb{R} \\ (v_1, \ldots, v_k) & \longmapsto & (\omega_1 \wedge \cdots \wedge \omega_k)(v_1, \ldots, v_k), \end{array}$$

where

$$(\omega_1 \wedge \cdots \wedge \omega_k)(v_1, \ldots, v_k) = \det\big(\omega_i(v_j)\big)_{1 \le i,j \le k}. \tag{3}$$

As seen in the previous chapter, $\omega_1 \wedge \cdots \wedge \omega_k$ is a k-form. In particular,

$$(dx_{i_1})_p \wedge \cdots \wedge (dx_{i_k})_p \in \mathcal{A}^k(\mathbb{R}^n_p).$$

We denote this element by

$$(dx_{i_1} \wedge \cdots \wedge dx_{i_k})_p.$$

In accordance with Proposition (4.2.5) we have the analogous result:

Proposition 5.1.7 *The set*

$$\big\{(dx_{i_1} \wedge \cdots \wedge dx_{i_k})_p\big\}, \quad i_1 < \cdots < i_k,$$

where $i_j \in \{1, \ldots, n\}$, forms a basis of $\mathcal{A}^k(\mathbb{R}^n_p)$.

Definition 5.1.8 *An exterior kâ€“form on $\mathbb{R}^n$ is a map which to each $p \in \mathbb{R}^n$ associates $\omega(p) \in \mathcal{A}^k(\mathbb{R}^n_p)$.*

In view of Proposition (4.1.7) we can write

$$\omega(p) = \sum_{i_1<i_2<\cdots<i_k} a_{i_1\cdots i_k}(p)\,(dx_{i_1} \wedge \cdots \wedge dx_{i_k}), \quad i_j \in \{1, 2, \ldots, n\}, \tag{4}$$

where $a_{i_1\cdots i_k}$ are functions from $\mathbb{R}^n$ into $\mathbb{R}$. If these functions are differentiable, ω is called a differentiable k-form.

As before, we denote by I the k-tuple

$$\vec{i} = (i_1, \ldots, i_k), \quad i_1 < \cdots < i_k, \quad i_j \in \{1, \ldots, n\},$$

and, for simplicity, we use the notation

$$\omega(p) = \sum_I a_I(p)\,dX_I(p),$$

or simply

$$\omega = \sum_I a_I dX_I,$$

when there is no risk of confusion.

By convention, a differential 0-form on $\mathbb{R}^n$ is a differentiable function.

Example 5.1.9 *In $\mathbb{R}^4$ we have the following types of exterior forms:*

0*â€“forms: functions on* $\mathbb{R}^4$.

1*â€“forms:*

$$a_1 dX_1 + a_2 dX_2 + a_3 dX_3 + a_4 dX_4.$$

2*â€“forms:*

$$a_{12}(dX_1 \wedge dX_2) + a_{13}(dX_1 \wedge dX_3) + a_{14}(dX_1 \wedge dX_4) + a_{23}(dX_2 \wedge dX_3) + a_{24}(dX_2 \wedge dX_4) + a_{34}(dX_3 \wedge dX_4).$$

3*â€“forms:*

$$a_{123}(dX_1 \wedge dX_2 \wedge dX_3) + a_{124}(dX_1 \wedge dX_2 \wedge dX_4) + a_{134}(dX_1 \wedge dX_3 \wedge dX_4) + a_{234}(dX_2 \wedge dX_3 \wedge dX_4).$$

4*â€“forms:*

$$a_{1234}(dX_1 \wedge dX_2 \wedge dX_3 \wedge dX_4).$$

Unless otherwise stated, from now on we shall deal only with differentiable k-forms.

If ω_1^k and ω_2^k are two k-forms:

$$\omega_1^k = \sum_I a_I dX_I \quad \text{and} \quad \omega_2^k = \sum_I b_I dX_I, \quad I = (i_1, \ldots, i_k),$$

with $i_1 < \cdots < i_k$, we define the sum $\omega_1^k + \omega_2^k$ by

$$\omega_1^k + \omega_2^k = \sum_I (a_I + b_I) dX_I. \tag{5}$$

If ω^k is a kâ€“form and ω^l is an ℓ-form, we can also, as in Section 3 of the previous chapter, define the exterior product $\omega^k \wedge \omega^l$ as follows:

If $\omega^k = \sum_I a_I dX_I$, $I = (i_1, \ldots, i_k)$, $i_1 < \cdots < i_k$, and $\omega^l = \sum_J b_J dX_J$, $J = (j_1, \ldots, j_l)$, $j_1 < \cdots < j_l$, we set

$$\omega_1^k \wedge \omega_2^l = \sum_{I,J} a_I b_J (dX_I \wedge dX_J). \tag{6}$$

The exterior product, as in Section 3 of the previous chapter, enjoys the following properties:

(1) $(\omega^k \wedge \omega^l) = (-1)^{kl} (\omega^l \wedge \omega^k)$ (anticommutativity);

(2) $(\omega^k \wedge \omega^l) \wedge \omega^m = \omega^k \wedge (\omega^l \wedge \omega^m)$ (associativity);

(3) $(\omega_1^k + \omega_2^k) \wedge \lambda\omega^l = (\omega_1^k \wedge \omega^l) + (\omega_2^k \wedge \omega^l)$ (distributivity).

We also recall that the exterior product defined in (6) has the crucial property that if $\omega_1, \ldots, \omega_k$ are 1-forms, then the exterior product $(\omega_1 \wedge \cdots \wedge \omega_k)$ (which is a decomposable kâ€“form) coincides with the form defined in (3).

Definition 5.1.10 *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a differentiable map. We denote by df_p the derivative of the function f at the point p in $\mathbb{R}^n_p$, defined by*

$$\begin{aligned} df_p : \mathbb{R}^n_p &\longrightarrow \mathbb{R}^m_{f(p)} \\ v_p &\longmapsto (df)(p)(v_p) = (f'(p) \cdot v)_{f(p)}, \end{aligned}$$

where $f'(p) : \mathbb{R}^n \to \mathbb{R}^m$ is the derivative of f at p.

Sometimes, instead of df_p, we also denote this map by f_*. Such a map is clearly linear, because $f'(p)$ is linear. Therefore $df_p \in \mathcal{L}(\mathbb{R}^n_p, \mathbb{R}^m_{f(p)})$ and we obtain a map

$$p \in \mathbb{R}^n \longmapsto df_p \in \mathcal{L}(\mathbb{R}^n_p, \mathbb{R}^m_{f(p)}).$$

Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a differentiable map. Given an exterior kâ€“form ω on $\mathbb{R}^m$, we define an exterior k-form $f^*\omega$ on $\mathbb{R}^n$ by setting, for each $p \in \mathbb{R}^n$ and each list of vectors $v_1, \ldots, v_k \in \mathbb{R}^n_p$,

$$\big[(f^*\omega)(p)\big](v_1, \ldots, v_k) = \omega\big(f(p)\big)\big(df_p(v_1), \ldots, df_p(v_k)\big). \tag{7}$$

In this way we obtain a map $f^*\omega$ which to each $p \in \mathbb{R}^n$ associates $(f^*\omega)(p) \in \mathcal{A}^k(\mathbb{R}^n_p)$ defined as in (7), that is, an exterior kâ€“form on $\mathbb{R}^n$.

On the other hand, as we know, the derivative $df_p : \mathbb{R}^n_p \longrightarrow \mathbb{R}^m_{f(p)}$, being linear, induces a linear transformation

$$[df_p]^* : \mathcal{A}^k(\mathbb{R}^m_{f(p)}) \longrightarrow \mathcal{A}^k(\mathbb{R}^n_p) \tag{8}$$

given by

$$\omega \longmapsto [df_p]^*(\omega),$$

where

$$[df_p]^*(\omega) : \mathbb{R}^n_p \times \cdots \times \mathbb{R}^n_p \longrightarrow \mathbb{R} \tag{9}$$

is defined by

$$\vec{v} \longmapsto [df_p]^*(\omega)(\vec{v}) = \omega\big(df_p(v_1), \ldots, df_p(v_k)\big),$$

as seen in Section 4 of Chapter 1.

Thus, from (7) and (9) we obtain, in particular for $\omega(f(p)) \in \mathcal{A}^k(\mathbb{R}^m_{f(p)})$,

$$(f^*\omega)(p) = [df_p]^*(\omega(f(p))). \tag{10}$$

Conclusion

Every differentiable map $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ induces a linear transformation f^* which sends exterior forms on $\mathbb{R}^m$ to exterior forms on $\mathbb{R}^n$. This transformation is one of the main reasons why differential forms are so useful for studying maps between surfaces, as we shall see later.

In the case of a 0â€“form, that is, a map $g : \mathbb{R}^m \longrightarrow \mathbb{R}$, we set $f^*g = g \circ f$, which is clearly a differentiable 0-form on $\mathbb{R}^n$, i.e. an exterior 0â€“form.

Proposition 5.1.11 *If $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is differentiable, then:*

(a) $f^*(\omega_1 + \omega_2) = f^*(\omega_1) + f^*(\omega_2)$, where ω_1, ω_2 are k-forms on $\mathbb{R}^m$.

(b) $f^*(g\,\omega) = f^*(g)\,f^*(\omega)$*, where* g *is a* 0*â€“form and* ω *is a* k*â€“form on* $\mathbb{R}^m$*.*

(c) $f^*(\omega_1 \wedge \omega_2) = f^*(\omega_1) \wedge f^*(\omega_2)$*, where* ω_1, ω_2 *are differential forms on* $\mathbb{R}^m$ *(of degrees adding up to the degree of the wedge product).*

Proof:

Items (a) and (c) have already been proved in the general case in Section 4 of Chapter 1. We now prove item (b).

Let $p \in \mathbb{R}^n$ and $v_1, \ldots, v_k \in \mathbb{R}^n_p$. Then, by (10),

$$\begin{aligned} f^*(g\,\omega)(p)(\vec{v}) &= [df_p]^*\big(g\,\omega\big)(f(p))(v_1, \ldots, v_k) \\ &= (g\,\omega)(f(p))\big(df_p(v_1), \ldots, df_p(v_k)\big) \\ &= g(f(p))\,\omega(f(p))\big(df_p(v_1), \ldots, df_p(v_k)\big) \\ &= (g \circ f)(p)\,[df_p]^*(\omega(f(p)))(v_1, \ldots, v_k) \\ &= (g \circ f)(p)\,(f^*\omega)(p)(v_1, \ldots, v_k) \\ &= f^*(g)(p)\,(f^*\omega)(p)(v_1, \ldots, v_k). \end{aligned}$$

□

Proposition 5.1.12 *If* $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *is a differentiable map, then*

$$f^*(dX_i) = df_i = \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}\,dX_j.$$

Proof:

Let $p \in \mathbb{R}^n$ and $v_p \in \mathbb{R}^n_p$. Then

$$\begin{aligned} [f^*(dX_i)(p)](v_p) &= dX_i(f(p))\big(df_p(v_p)\big) \\ &= dX_i(f(p))\big(f'(p)\,v\big)_{f(p)} \\ &= dX_i(f(p))\big(Df_1(p)v, \ldots, Df_m(p)v\big)_{f(p)} \\ &= Df_i(p)v \\ &= \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(p)\,v_j \\ &= \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(p)\,dX_j(p)(v_p) \\ &= \Big[\sum_{j=1}^{n} \Big(\frac{\partial f_i}{\partial x_j}\,dX_j\Big)(p)\Big](v_p). \end{aligned}$$

□

Remark 5.1.13 *In fact, the map f^* (which sends exterior k-forms on $\mathbb{R}^m$ to exterior k-forms on $\mathbb{R}^n$) is equivalent to a change of variables.*

Indeed, let $\omega = \sum_I a_I dY_I$ be a k-form on $\mathbb{R}^m$. Using the properties in Proposition (2.1.11), we obtain:

$$\begin{aligned} f^*\omega &= f^*\left(\sum_I a_I dY_I\right) \\ &= \sum_I f^*(a_I)\, f^*(dY_I) \\ &= \sum_I (a_I \circ f)\, f^*(dY_{i_1} \wedge \cdots \wedge dY_{i_k}) \\ &= \sum_I (a_I \circ f)\, f^*(dY_{i_1}) \wedge \cdots \wedge f^*(dY_{i_k}). \end{aligned} \tag{11}$$

On the other hand, by Proposition (5.1.12),

$$f^*(dY_{i_j}) = df_{i_j} = \sum_{k=1}^{m} \frac{\partial f_{i_j}}{\partial y_k}\, dY_k.$$

Hence, from (11) we obtain

$$f^*(\omega) = \sum_I (a_I \circ f)\, df_{i_1} \wedge \cdots \wedge df_{i_k}.$$

More explicitly,

$$f^*(\omega) = \sum_I a_I\big(f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n)\big)\, df_{i_1} \wedge \cdots \wedge df_{i_k},$$

where the f_i and df_i are functions of the variables x_j.

Therefore, to apply f^* to ω is equivalent to substituting, in ω, the variables Y_i and their differentials dY_i by the functions of x_k and dx_k given by

$$\begin{cases} y_1 &= f_1(x_1, \ldots, x_n), \\ \vdots & \vdots \\ y_m &= f_m(x_1, \ldots, x_n). \end{cases}$$

5.2 The Exterior Differential

We now define an operation on differentiable exterior k-forms which generalises the usual differentiation of functions.

If $f : \mathbb{R}^n \to \mathbb{R}$ is a 0-form (a C^2 differentiable function), its differential

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\, dX_i,$$

which is an identity in $(\mathbb{R}^n_p)^*$, is a differentiable exterior 1-form. Recall that df is the map

$$p \in \mathbb{R}^n \longmapsto df_p \in (\mathbb{R}^n_p)^*,$$

where

$$\begin{aligned} df_p : \mathbb{R}^n_p &\longrightarrow \mathbb{R} \\ v_p &\longmapsto df_p(v_p) = Df(p)(v) = \sum_{i=1}^{n} \big(\tfrac{\partial f}{\partial x_i}\, dX_i\big)(p)(v_p). \end{aligned}$$

We wish to define, by analogy, an operation which sends a differentiable exterior k-form to a differentiable exterior $(k+1)$-form.

Definition 5.2.1 *Let ω be a differentiable exterior k-form of class C^k, that is, a map*

$$\omega : p \in \mathbb{R}^n \longmapsto \omega(p) \in \mathcal{A}^k(\mathbb{R}^n_p)$$

such that

$$\omega(p) = \sum_I a_I(p)\, dX_I(p),$$

with functions $a_I \in C^k(\mathbb{R}^n)$.

We define the exterior differential of ω to be the differentiable exterior $(k+1)$-form of class C^{k-1}, that is, the map

$$d\omega : p \in \mathbb{R}^n \longmapsto d\omega(p) \in \mathcal{A}^{k+1}(\mathbb{R}^n_p)$$

such that

$$d\omega(p) = \sum_I da_I(p) \wedge dX_I(p) = \sum_{j,I} \frac{\partial a_I(p)}{\partial x_j}\, dX_j(p) \wedge dX_I(p).$$

Remark 5.2.2 *(i) If ω is a form of degree zero, that is a function*

$$\omega = f : \mathbb{R}^n \to \mathbb{R},$$

then

$$d\omega = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\, dX_i.$$

(ii) If ω is a 1-form, then for each $p \in \mathbb{R}^n$ we can write

$$\omega = \sum_i a_i \, dX_i,$$

and therefore

$$\begin{aligned}
d\omega &= \sum_i da_i \wedge dX_i \\
&= \sum_{i,j} \frac{\partial a_i}{\partial x_j} \, dX_j \wedge dX_i \\
&= \sum_{j<i} \frac{\partial a_i}{\partial x_j} \, dX_j \wedge dX_i + \sum_{i<j} \frac{\partial a_i}{\partial x_j} \, dX_j \wedge dX_i \\
&= \sum_{i<j} \frac{\partial a_j}{\partial x_i} \, dX_i \wedge dX_j - \sum_{i<j} \frac{\partial a_i}{\partial x_j} \, dX_i \wedge dX_j \\
&= \sum_{i<j} \left(\frac{\partial a_j}{\partial x_i} - \frac{\partial a_i}{\partial x_j} \right) dX_i \wedge dX_j,
\end{aligned}$$

where all quantities above are evaluated at the point $p \in \mathbb{R}^n$.

For simplicity, from now on we shall omit the point p in the notation.

In particular:

If $\omega = M\,dx + N\,dy$, then

$$d\omega = \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} \right) dx \wedge dy.$$

If $\omega = P\,dx + Q\,dy + R\,dz$, then

$$\begin{aligned}
d\omega = &\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx \wedge dy + \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) dy \wedge dz \\
&+ \left(\frac{\partial R}{\partial x} - \frac{\partial P}{\partial z} \right) dx \wedge dz.
\end{aligned}$$

Proposition 5.2.3 *Let ω_1, ω_2 be exterior differential forms of class C^1 and let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a map of class C^2 (a 0-form of class C^2). Then:*

(a) If $\omega : \mathbb{R}^n \to \mathbb{R}$ is a differentiable 0â€“form, then $d\omega$ is the usual differential of a function.

(b) $d(\omega_1 + \omega_2) = d\omega_1 + d\omega_2$, for ω_1, ω_2 kâ€“forms.

(c) If ω is of class C^2, then $d(d\omega) = d^2\omega = 0$.

(d) $d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^k \omega_1 \wedge d\omega_2$, *where* ω_1 *is a* k*â€“form and* ω_2 *is an* ℓ*â€“form.*

(e) $d(f^*\omega) = f^*(d\omega)$.

Proof:

(a) Obvious.

(b) Let ω_1, ω_2 be two k- forms. Then, for each $p \in \mathbb{R}^n$,
$\omega_1 = \sum_I a_I dX_I$ and $\omega_2 = \sum_I b_I dX_I$. Thus $\omega_1 + \omega_2 = \sum_I (a_I + b_I) dX_I$ and hence

$$\begin{aligned} d(\omega_1 + \omega_2) &= \sum_I d(a_I + b_I) \wedge dX_I \\ &= \sum_I da_I \wedge dX_I + \sum_I db_I \wedge dX_I \\ &= d\omega_1 + d\omega_2. \end{aligned}$$

As seen in earlier sections, to establish the remaining properties it suffices to treat the case in which, for each $p \in \mathbb{R}^n$, $\omega_1 = a\, dX_I$ and $\omega_2 = b\, dX_J$ are monomials.

Let us proceed.

(c) If ω is as above and of class C^2, then for each p we have $\omega = a\, dX_I$. Hence

$$d\omega = \sum_{j=1}^{n} \frac{\partial a}{\partial x_j}\, dX_j \wedge dX_I,$$

and therefore

$$\begin{aligned} d(d\omega) &= \left[\sum_{k,j=1}^{n} \frac{\partial^2 a}{\partial x_k \partial x_j}\, dX_k \wedge dX_j \right] \wedge dX_I \\ &= \left[\sum_{j<k} \left(\frac{\partial^2 a}{\partial x_j \partial x_k} - \frac{\partial^2 a}{\partial x_k \partial x_j} \right) dX_j \wedge dX_k \right] \wedge dX_I = 0, \end{aligned}$$

by Schwarz's theorem on the equality of mixed partial derivatives.

(d) For each $p \in \mathbb{R}^n$ we have $\omega_1 = a\, dX_I$ and $\omega_2 = b\, dX_J$. Thus

$$\omega_1 \wedge \omega_2 = ab\, (dX_I \wedge dX_J).$$

Hence

$$\begin{aligned} d(\omega_1 \wedge \omega_2) &= d(ab) \wedge dX_I \wedge dX_J \\ &= (da \cdot b + a \cdot db) \wedge (dX_I \wedge dX_J) \\ &= b\, da \wedge (dX_I \wedge dX_J) + a\, db \wedge (dX_I \wedge dX_J) \\ &= b\, (da \wedge dX_I) \wedge dX_J + (-1)^k a\, (dX_I \wedge db) \wedge dX_J. \qquad (1) \end{aligned}$$

On the other hand, $d\omega_1 = da \wedge dX_I$. Thus

$$d\omega_1 \wedge \omega_2 = (da \wedge dX_I) \wedge b\, dX_J = b\,(da \wedge dX_I) \wedge dX_J. \tag{2}$$

Similarly, $d\omega_2 = db \wedge dX_J$, so

$$\omega_1 \wedge d\omega_2 = a\, dX_I \wedge db \wedge dX_J. \tag{3}$$

Combining (1), (2) and (3), we obtain

$$d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^k\, \omega_1 \wedge d\omega_2.$$

(e) Finally, to prove the last property, we begin with the case where the form ω reduces to a function $g : \mathbb{R}^m \to \mathbb{R}$. Then, by the chain rule, at every point $p \in \mathbb{R}^n$ we have

$$dg(f(p)) \cdot f'(p) = d(g \circ f)(p).$$

Therefore, for any $v \in \mathbb{R}^n$,

$$f^*(dg)(p)(v) = dg(f(p))\big(df_p(v)\big) = dg(f(p))\big(f'(p)\, v\big) = d(g \circ f)(p)\, v.$$

Hence

$$f^*(dg) = d(g \circ f) = d(f^*g).$$

Now consider a form $\omega = a\, dX_I = a(dx_{i_1} \wedge \cdots \wedge dx_{i_k})$ of arbitrary degree k. From (d), (c) and a straightforward induction one shows that if
$a : \mathbb{R}^m \to \mathbb{R}$ is of class C^1 and $g_1, \dots, g_k$ are of class C^2, then

$$d\big(a(dg_1 \wedge \cdots \wedge dg_k)\big) = da \wedge dg_1 \wedge \cdots \wedge dg_k.$$

We also recall that $f^*(\alpha \wedge \beta) = f^*(\alpha) \wedge f^*(\beta)$. Thus

$$f^*\omega = f^*a \cdot \big(f^*dX_{i_1} \wedge \cdots \wedge f^*dX_{i_k}\big) = f^*a \cdot d(X_{i_1} \circ f) \wedge \cdots \wedge d(X_{i_k} \circ f).$$

Therefore

$$\begin{aligned} d(f^*\omega) &= d(f^*a) \wedge d(X_{i_1} \circ f) \wedge \cdots \wedge d(X_{i_k} \circ f) \\ &= f^*\big(da \wedge dX_{i_1} \wedge \cdots \wedge dX_{i_k}\big) = f^*(d\omega), \end{aligned}$$

as claimed. □

Chapter 6

Differentiable Surfaces

6.1 Differential Forms on Surfaces

Although differential forms were introduced in the previous chapter on $\mathbb{R}^n$, they, like everything related to differentiability, naturally live on differentiable surfaces (manifolds). Let us recall this concept.

A surface of class C^k and dimension $m < n$ in $\mathbb{R}^n$ is a subset $S \subset \mathbb{R}^n$ such that for each $p \in S$ there exists a neighbourhood V of p in $\mathbb{R}^n$ and a map $f : U \subset \mathbb{R}^m \to V \cap S$ of class C^k on the open set U such that

i) f is a differentiable diffeomorphism;

ii) the differential $Df(q) : \mathbb{R}^m \to \mathbb{R}^n$ is injective for every $q \in U$.

The map $f : U \subset \mathbb{R}^m \to S$ is called a parametrisation of S. To fix ideas, consider the diagram of a surface S in $\mathbb{R}^3$:

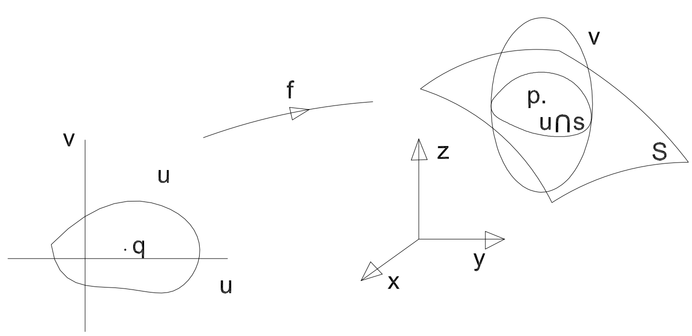

Figure 6.1:

Given a surface $S \subset \mathbb{R}^n$ of dimension m and class C^k $(k \geq 1)$ and a point $p \in S$, the tangent space to S at p in $\mathbb{R}^n$ is the vector subspace $T_p(S)$. We can describe it in the following ways:

i) Take a parametrisation $\varphi : V_0 \to V \cap S$ of class C^k of a neighbourhood V of p with $\varphi(q) = p$. We then set

$$T_p(S) = \varphi'(q)(\mathbb{R}^m).$$

ii) Consider all curves $\lambda : (-\varepsilon, \varepsilon) \to S$ with $\lambda(0) = p$, differentiable at 0. Then $T_p(S)$ is defined as the set of velocity vectors $\lambda'(0)$ of these curves.

By the first definition we obtain $T_p(S)$ as an m-dimensional vector subspace of $\mathbb{R}^n$, but this description depends on the parametrisation φ. The second definition does not depend on the choice of parametrisation, although it is not immediately obvious from it that $T_p(S)$ is a vector space. However, one proves that the two definitions are equivalent and therefore

$$T_p(S) = \varphi'(q)(\mathbb{R}^m) = \{\lambda'(0) \in \mathbb{R}^n;\ \lambda : (-\varepsilon, \varepsilon) \to S \cap V,\ \lambda(0) = p\},$$

with λ differentiable at 0.

The most important fact that follows from the definition of a surface is that the change of parameters is a diffeomorphism of class C^k.

More precisely, if $\varphi : U_0 \to \varphi(U_0)$ and $\psi : W_0 \to \psi(W_0)$ are two parametrisations such that

$$\varphi(U_0) \cap \psi(W_0) = W \neq \emptyset,$$

that is, they both contain the point p, then the maps

$$\psi^{-1} \circ \varphi : \varphi^{-1}(W) \to \mathbb{R}^m \quad \text{and} \quad \varphi^{-1} \circ \psi : \psi^{-1}(W) \to \mathbb{R}^m$$

are diffeomorphisms of class C^k.

As a consequence, we can introduce the concept of differentiable maps between surfaces. Indeed, a map

$$f : S \longrightarrow \mathbb{R}^k$$

is said to be differentiable at a point $p \in S$ if there exists a parametrisation φ of class C^k, $\varphi : V_0 \to V$, of a neighbourhood V of p in S such that

$$f \circ \varphi : V_0 \to \mathbb{R}^k$$

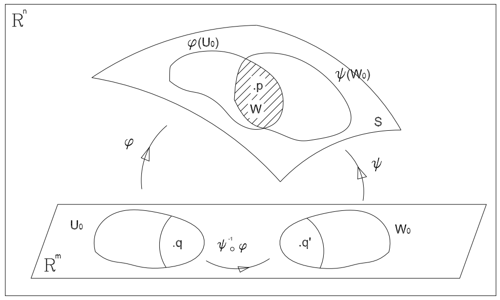

Figure 6.2:

(V_0 open in $\mathbb{R}^m$) is differentiable at the point $q = \varphi^{-1}(p)$. If we take another parametrisation $\psi : W_0 \to W$ of class C^k of a neighbourhood W of p in S, then $f \circ \varphi$ is differentiable at $\varphi^{-1}(p)$ if and only if $f \circ \psi$ is differentiable at $\psi^{-1}(p)$, since

$$f \circ \psi = (f \circ \varphi) \circ (\varphi^{-1} \circ \psi),$$

and

$$\varphi^{-1} \circ \psi : \psi^{-1}(V \cap W) \longrightarrow \varphi^{-1}(V \cap W)$$

is a diffeomorphism of class C^k.

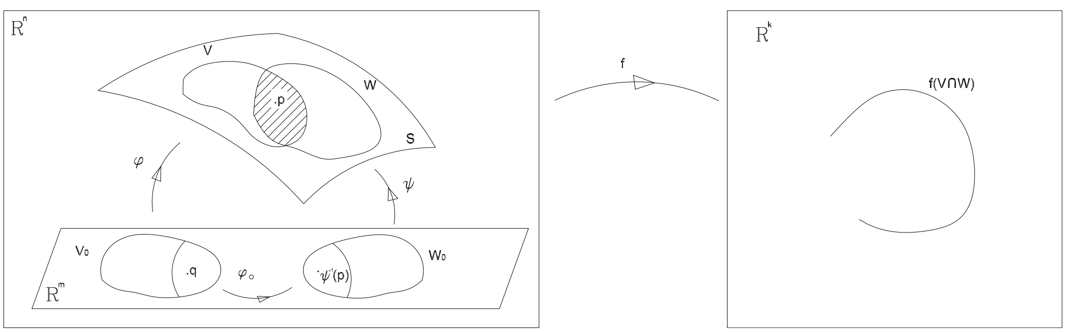

Figure 6.3:

If $f : S \to \mathbb{R}^k$ is differentiable at $p \in S$, its derivative at p is the linear transformation

$$f'(p) : T_p(S) \longrightarrow \mathbb{R}^k \tag{1}$$

given by

$$v \longmapsto f'(p)v$$

defined as follows:

Take a parametrisation $\varphi : V_0 \to V$ such that $\varphi(q) = p$. Given a vector $v \in T_p(S)$, we have $v = \varphi'(q)v_0$ for some $v_0 \in \mathbb{R}^m$. We then set

$$f'(p)v = (f \circ \varphi)'(q)v_0.$$

The linear map (1) is well defined, because if

$$\psi : W_0 \to W$$

is another parametrisation with $p = \psi(q')$ and $v = \psi'(q')w_0$ for some $w_0 \in \mathbb{R}^m$, then

$$(f \circ \varphi)'(q)v_0 = (f \circ \psi)'(q')w_0.$$

Indeed, we know that $\psi = \varphi \circ \xi$, where

$$\xi = \varphi^{-1} \circ \psi : \psi^{-1}(V \cap W) \to \varphi^{-1}(V \cap W)$$

is a diffeomorphism of class C^k with $\xi(q') = q$. Then

$$\varphi'(q)v_0 = v = \psi'(q')w_0 = (\varphi \circ \xi)'(q')w_0 = \varphi'(q)\,\xi'(q')w_0.$$

Since $\varphi'(q)$ is injective, it follows that $\xi'(q')w_0 = v_0$, and therefore

$$\begin{aligned}(f \circ \psi)'(q')w_0 &= (f \circ \varphi \circ \xi)'(q')w_0 = (f \circ \varphi)'(q)\,\xi'(q')w_0 \\ &= (f \circ \varphi)'(q)v_0.\end{aligned}$$

Any velocity vector $v \in T_pS$ is the velocity vector $v = \lambda'(0)$ of a curve $\lambda : (-\varepsilon, \varepsilon) \to S$ with $\lambda(0) = p$. Then the image

$$f'(p)v = (f \circ \lambda)'(0)$$

is the velocity vector at 0 of the curve $(f \circ \lambda) : (-\varepsilon, \varepsilon) \to \mathbb{R}^k$.

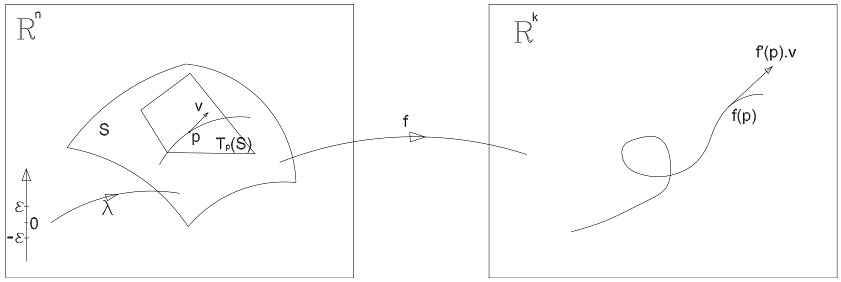

Figure 6.4:

If $\overline{S} \subset \mathbb{R}^k$ is another C^k-surface and the map

$$f : S \longrightarrow \mathbb{R}^k$$

is differentiable at p and satisfies $f(S) \subset \overline{S}$, we shall say that

$$f : S \to \overline{S} \tag{2}$$

is differentiable at p. The observation we have just made about $f'(p)v$ as the velocity vector of a curve shows that if (2) is differentiable at $p \in S$, then the derivative $f'(p)$ is a linear map from $T_p(S)$ to $T_{f(p)}(\overline{S})$:

$$\begin{aligned} f'(p) : T_p(S) &\longrightarrow T_{f(p)}(\overline{S}) \\ v &\longmapsto f'(p)v. \end{aligned}$$

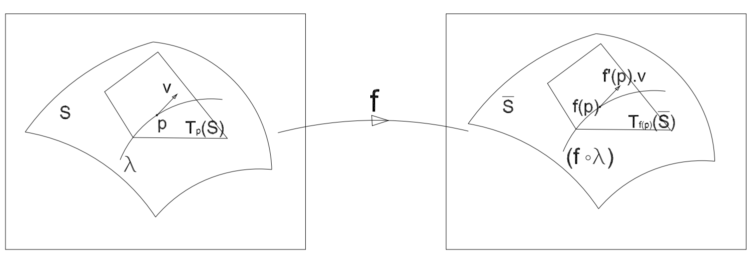

Figure 6.5:

Note that the chain rule holds: if $f : S \to \overline{S}$ is differentiable at $p \in S$ and $g : \overline{S} \to \mathbb{R}^s$ is differentiable at $f(p)$, then

$$g \circ f : S \longrightarrow \mathbb{R}^s$$

is differentiable at p and

$$(g \circ f)'(p) = g'(f(p)) \circ f'(p).$$

In complete analogy with the previous setting, we may define an exterior k-form on a surface S of dimension m in $\mathbb{R}^n$ as a map

$$\omega : p \in S \longrightarrow \omega(p) \in \mathcal{A}^k(T_pS). \tag{3}$$

If $k = 0$, an exterior 0â€“form on S is simply a real-valued function $\omega : S \to \mathbb{R}$.

Let $\varphi : U_0 \to U$ be a parametrisation of an open set $U \subset S$. At each point $p = \varphi(q) \in U$ we have the basis

$$\left\{ \frac{\partial \varphi(q)}{\partial u_1}, \ldots, \frac{\partial \varphi(q)}{\partial u_m} \right\} \subset T_pS. \tag{4}$$

Indeed,

$$\frac{\partial \varphi(q)}{\partial u_1} = \varphi'(q)e_1, \ldots, \frac{\partial \varphi(q)}{\partial u_m} = \varphi'(q)e_m.$$

Since the derivative of the parametrisation

$$\begin{aligned} \varphi'(q) : \mathbb{R}^m &\longrightarrow \mathbb{R}^n \\ v &\longmapsto \varphi'(q)v \end{aligned}$$

is linear and injective (by definition), the family in (4) is indeed a basis of $T_p(S)$. Now, if $\psi : V_0 \to V$ is another parametrisation of S such that $U \cap V \neq \emptyset$ and $p = \varphi(q) = \psi(q')$, then we also have the basis

$$\left\{ \frac{\partial \psi}{\partial v_1}(q'), \ldots, \frac{\partial \psi}{\partial v_m}(q') \right\} \subset T_pS.$$

There is, of course, a relation of the form

$$\frac{\partial \varphi}{\partial u_j}(q) = \sum_{i=1}^{m} a_{ij} \frac{\partial \psi}{\partial v_i}(q').$$

To determine the coefficients a_{ij}, we again use the diffeomorphism (see Figure 6.6)

$$\xi = \psi^{-1} \circ \varphi : \varphi^{-1}(U \cap V) \longrightarrow \psi^{-1}(U \cap V).$$

If $\xi_1, \ldots, \xi_m$ are the coordinate functions of ξ, the equality $\varphi = \psi \circ \xi$ yields

$$\varphi'(x) = \psi'(\xi(x))\, \xi'(x) \Rightarrow \varphi'(q) = \psi'(q')\, \xi'(q).$$

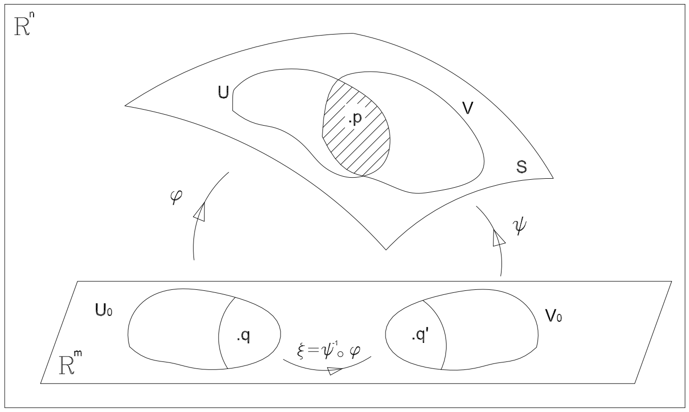

Figure 6.6:

Consequently,

$$\begin{pmatrix} \frac{\partial \varphi_1}{\partial u_1} & \cdots & \frac{\partial \varphi_1}{\partial u_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial \varphi_n}{\partial u_1} & \cdots & \frac{\partial \varphi_n}{\partial u_m} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^m \frac{\partial \psi_1}{\partial v_i}\frac{\partial \xi_i}{\partial u_1} & \cdots & \sum_{i=1}^m \frac{\partial \psi_1}{\partial v_i}\frac{\partial \xi_i}{\partial u_m} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^m \frac{\partial \psi_n}{\partial v_i}\frac{\partial \xi_i}{\partial u_1} & \cdots & \sum_{i=1}^m \frac{\partial \psi_n}{\partial v_i}\frac{\partial \xi_i}{\partial u_m} \end{pmatrix}.$$

Thus

$$\begin{aligned} \frac{\partial \varphi}{\partial u_j}(q) &= \left(\frac{\partial \varphi_1}{\partial u_j}, \ldots, \frac{\partial \varphi_n}{\partial u_j}\right) \\ &= \left(\sum_{i=1}^m \frac{\partial \psi_1}{\partial v_i}\frac{\partial \xi_i}{\partial u_j}, \ldots, \sum_{i=1}^m \frac{\partial \psi_n}{\partial v_i}\frac{\partial \xi_i}{\partial u_j}\right) \\ &= \sum_{i=1}^m \frac{\partial \xi_i}{\partial u_j}(q) \left(\frac{\partial \psi_1}{\partial v_i}, \ldots, \frac{\partial \psi_n}{\partial v_i}\right)(q') \\ &= \sum_{i=1}^m \frac{\partial \xi_i}{\partial u_j}(q)\, \frac{\partial \psi}{\partial v_i}(q'). \end{aligned}$$

Hence

$$\frac{\partial \varphi}{\partial u_j}(q) = \sum_{i=1}^m \frac{\partial \psi}{\partial v_i}(q')\, \frac{\partial \xi_i}{\partial u_j}(q),$$

and therefore

$$a_{ij} = \frac{\partial \xi_i}{\partial u_j}(q).$$

Note that the change-of-basis matrix (a_{ij}) from the basis $\left(\frac{\partial \psi}{\partial v_i}(q')\right)$ to the basis $\left(\frac{\partial \varphi}{\partial u_i}(q)\right)$ in the vector space $T_p(S)$ is precisely the Jacobian matrix $J(\xi(q))$, where $\xi = \psi^{-1} \circ \varphi$ is the diffeomorphism mentioned above.

We use the notation

$$\{du_1, \ldots, du_m\} \subset (T_pS)^*$$

for the dual basis of

$$\left\{\frac{\partial\varphi(q)}{\partial u_1}, \ldots, \frac{\partial\varphi(q)}{\partial u_m}\right\}.$$

Indeed, $du_1, \ldots, du_m$ are exterior 1-forms on U; that is, for each $i \in \{1, \ldots, m\}$, du_i is the map

$$du_i : p \in U \longmapsto du_i(p) \in T_p(S)^*.$$

For each $p \in U$ we shall write du_i instead of $du_i(p)$ when there is no risk of confusion. Thus, at each point $p = \varphi(q) \in U$ the kâ€“forms

$$du_I = du_{i_1} \wedge \cdots \wedge du_{i_k}, \quad I = \{i_1 < \cdots < i_k\} \subset \{1, 2, \cdots, m\},$$

form a basis of $\mathcal{A}^k(T_p(S))$.

Given an exterior kâ€“form ω on S, we can write, at each point $p = \varphi(q) \in U$,

$$\omega(p) = \omega(\varphi(q)) = \sum_I a_I(q)\, du_I. \tag{6}$$

Thus the exterior form determines, for each parametrisation $\varphi : U_0 \to U$ in S, a family of functions $a_I : U_0 \to \mathbb{R}$, in number

$$\binom{m}{k},$$

called the coordinates of the form ω with respect to the parametrisation φ. Indeed, according to Proposition (1.2.5),

$$a_I(q) = \omega(\varphi(q))\Big(\frac{\partial\varphi}{\partial u_{i_1}}(q), \ldots, \frac{\partial\varphi}{\partial u_{i_k}}(q)\Big), \quad \forall q \in U_0. \tag{7}$$

Now let $\psi : V_0 \to V$ be another parametrisation of S, with $U \cap V \neq \emptyset$. For each $p = \varphi(q) = \psi(q') \in U \cap V$ we have the dual bases

$$\left\{\frac{\partial\psi}{\partial v_1}(q'), \ldots, \frac{\partial\psi}{\partial v_m}(q')\right\} \subset T_p(S) \quad \text{and} \quad \{dv_1, \ldots, dv_m\} \subset (T_p(S))^*,$$

which are related to the bases determined by φ as follows:

$$\frac{\partial\varphi}{\partial u_j}(q) = \sum_{i=1}^m \frac{\partial v_i}{\partial u_j}\frac{\partial\psi}{\partial v_i}(q') \quad \text{and} \quad dv_i = \sum_{j=1}^m \frac{\partial v_i}{\partial u_j}\, du_j. \tag{8}$$

The first identity follows from (4) and (5), and the second from what we saw in Section 4 of Chapter 1.

In these equalities, $(\frac{\partial v_i}{\partial u_j})$ is the Jacobian matrix of the change of parameters $\psi^{-1} \circ \varphi$, evaluated at q; the derivative $\frac{\partial \varphi}{\partial u_j}$ is taken at q, and $\frac{\partial \psi}{\partial v_i}$ at $q' = \psi^{-1}(\varphi(q))$.

The parametrisation ψ determines in $\mathcal{A}^k(T_pS)$ the basis consisting of the k-forms $dv_I = dv_{i_1} \wedge \cdots \wedge dv_{i_k}$.

From Chapter 1 we know that if $p = \varphi(q) = \psi(q') \in U \cap V$ and, in addition,

$$\omega(p) = \sum_J a_J(q)\, du_J = \sum_I b_I(q')\, dv_I, \tag{9}$$

then

$$a_J(q) = \sum_I \det\left(\frac{\partial v_I}{\partial u_J}\right) b_I(q'), \tag{10}$$

where $\left(\frac{\partial v_I}{\partial u_J}\right)$ is the $k \times k$ matrix formed by the entries $\frac{\partial v_i}{\partial u_j}(q)$ such that $i \in I$ and $j \in J$.

In terms of classical tensor calculus, an exterior k-form on a surface S may be thought of as an assignment which, to each parametrisation $\varphi : U_0 \to U$ in S, associates the $\binom{m}{k}$ functions $a_J : U_0 \to \mathbb{R}$, called the coordinates of the form with respect to φ, in such a way that if to another parametrisation $\psi : V_0 \to V$ correspond the functions $b_I : V_0 \to \mathbb{R}$ and $\varphi(q) = \psi(q')$, then the coordinate change formulas (10) hold.

It is worth highlighting the important particular case $k = m$, that is, when the degree of the form equals the dimension of the surface. In this case the form has only one coordinate in each parametrisation. Thus:

For every point $p = \varphi(q) = \psi(q') \in U \cap V$, we have

$$\omega(p) = a(q)\, du_1 \wedge \cdots \wedge du_m = b(q')\, dv_1 \wedge \cdots \wedge dv_m, \tag{11}$$

where the functions $a : U_0 \to \mathbb{R}$ and $b : V_0 \to \mathbb{R}$ satisfy

$$a(q) = \det\left(\frac{\partial v_i}{\partial u_j}\right) b(q'), \tag{12}$$

where $q \in \varphi^{-1}(U \cap V)$, $q' = (\psi^{-1} \circ \varphi)(q)$, and $\det\left(\frac{\partial v_i}{\partial u_j}\right)$ is the Jacobian determinant of the diffeomorphism $(\psi^{-1} \circ \varphi)$ evaluated at q.

Let S be a surface of class C^m. An exterior kâ€“form on S is said to be of class C^k $(k < m)$ if S can be covered by images U of C^mâ€“parametrisations $\varphi : U_0 \to U$, with respect to which $\omega = \sum a_I\, du_I$, where all functions $a_I : U_0 \to \mathbb{R}$ are of class C^k.

The coordinate change formulas (10) show that if the coordinates of ω in a parametrisation $\psi \in C^m$ are functions of class C^k $(k < m)$, then they remain of class C^k in any other parametrisation $\varphi \in C^m$ with $\mathrm{Im}(\varphi) \cap \mathrm{Im}(\psi) \neq \emptyset$. An exterior 0-form of class C^k on S is simply a function $f : S \to \mathbb{R}$ of class C^k.

The coordinate change formulas express the 'invariance of differential forms'. The simplest case, given in (12), ensures, for instance, that the integral of an exterior mâ€“form is well defined on an m- dimensional surface, as we shall see next.

6.2 Integration of Differential Forms

We begin this section by defining the support of a top-degree form on a surface.

Definition 6.2.1 *The* **support** *of an exterior form ω on a surface S is the closure (relative to S) of the set of points $p \in S$ such that $\omega(p) \not\equiv 0$. Denoting the support of ω by* supp(ω), *we set*

$$\mathrm{supp}(\omega) = \overline{\{p \in S;\ \omega(p) \not\equiv 0\}}^{S}.$$

Equivalently,

$$\mathrm{supp}(\omega) = \overline{\{p \in S;\ \omega(p) \not\equiv 0\}}^{\mathbb{R}^n} \cap S.$$

Thus, $p \in \mathrm{supp}(\omega)$ means that every neighbourhood of p contains points p' such that $\omega(p') \not\equiv 0$. Observe that, by definition, the support of ω is always a closed subset of S. Hence, if the form

$$\omega : p \in S \longmapsto \omega(p) \in \mathcal{A}^k(T_pS)$$

is continuous and $\omega(p) \neq 0$, then $\omega \neq 0$ in some neighbourhood of p. Thus $p \in \mathrm{int}(\mathrm{supp}(\omega))$ (relative to S). In other words, if $p \in S$ and $p \notin \mathrm{int}(\mathrm{supp}(\omega))$, then $\omega(p) \equiv 0$.

It follows that if $\omega \in C^0$ and $p \in S$ is a boundary point of supp(ω), then $\omega(p) = 0$, although $p \in \mathrm{supp}(\omega)$.

We now define the integral of a continuous exterior mâ€“form ω with compact support on an oriented mâ€“dimensional surface S, in the particular case where supp(ω) is contained in the image of a positive parametrisation $\varphi : U_0 \to U$.

Definition 6.2.2 *Let S be an oriented surface of class C^1 and dimension m, and let $\omega : p \in S \mapsto \omega(p) \in \mathcal{A}^k(T_pS)$ be a continuous exterior mâ€“form. Suppose that* supp(ω)

is a compact set, contained in the image of a positive parametrisation $\varphi : U_0 \to U$. *In terms of this parametrisation we can write*

$$\omega(p) = a(q)\, du_1 \wedge \cdots \wedge du_m$$

for every $p = \varphi(q) \in U$, *where the continuous function*

$$a : U_0 \longrightarrow \mathbb{R}$$

has compact support equal to $\varphi^{-1}(\mathrm{supp}(\omega))$. *By definition,*

$$\int_S \omega = \int_K a(u)\, du, \tag{13}$$

where $K \subset \mathbb{R}^m$ *is any compact Jordan-measurable subset contained in* U_0 *and containing* $\mathrm{supp}(a)$.

We now make some remarks concerning this definition:

i) One may drop the requirement that $K \subset U_0$ (still assuming that $\mathrm{supp}(a) \subset K$), provided we regard a as a continuous function defined on K, vanishing at the points of $K \setminus U_0$. In this way a ceases to be continuous only on a set of measure zero in $\mathbb{R}^m$, which does not affect the value of the integral in (13).

ii) Since S is orientable, there exists a family of parametrisations $\varphi : U_0 \to U$ (of class C^1) such that any two of them are always coherent.

More precisely, if $\varphi : U_0 \to U$ and $\psi : V_0 \to V$ are two parametrisations of this family, then either $U \cap V = \emptyset$ or, if $U \cap V \neq \emptyset$, the Jacobian determinant of the diffeomorphism between the two parametrisations is positive at every point
$q \in \varphi^{-1}(U \cap V)$. The parametrisations in this family are called positive.

We must still show that $\int_S \omega$, as defined above, is independent of the choice of positive parametrisation φ.

Indeed, let $\psi : V_0 \to V$ be another positive parametrisation of S, with $\mathrm{supp}(\omega) \subset V$ and

$$\omega(p) = b(q')\, dv_1 \wedge \cdots \wedge dv_m$$

for every $p = \psi(q') \in V$. The function $b : V_0 \to \mathbb{R}$ is continuous, its support is equal to $\psi^{-1}(\mathrm{supp}(\omega))$, and for every $q \in \varphi^{-1}(U \cap V)$ we have

$$a(q) = J(q)\, b(q'), \tag{14}$$

where $J(q)$ is the Jacobian determinant of $(\psi^{-1} \circ \varphi)$ at q, and $q' = \psi^{-1}(\varphi(q))$. Note that $J(q) > 0$ for all $q \in \varphi^{-1}(U \cap V)$.

To compute $\int_S \omega$ in terms of the parametrisation φ, let K be a compact Jordanâ€œmeasurable set such that

$$\varphi^{-1}(\text{supp}(\omega)) \subset K \subset \varphi^{-1}(U \cap V),$$

and to compute $\int_S \omega$ in terms of ψ, let $L = \psi^{-1}(\varphi(K))$ be the compact Jordanâ€œmeasurable set (its boundary has measure zero).

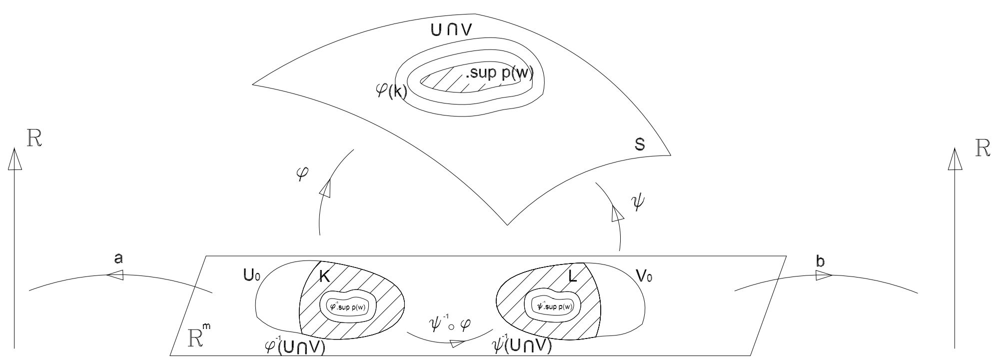

Figure 6.7:

By an ordinary change of variables we obtain

$$\int_L b(v)\, dv = \int_K b(\psi^{-1}\varphi(u))\, J(u)\, du = \int_K a(u)\, du.$$

Thus the surface integral $\int_S \omega$ is well defined whenever ω is a continuous exterior m-form with compact support contained in some parameterised neighbourhood of an oriented m-dimensional surface. Later we shall define the integral $\int_S \omega$ under more general assumptions on the form ω.

Now consider a continuous k-form ω whose support K is not contained in a single coordinate neighbourhood, and let us define the surface integral $\int_S \omega$.

Roughly speaking, we take a covering $\{V_\alpha\}$ of the oriented surface S by coordinate neighbourhoods, and a smooth partition of unity subordinate to this covering $\{V_\alpha\}$, that is, a family of differentiable functions $\varphi_1, \ldots, \varphi_m : S \to \mathbb{R}$ such that

i) $\sum_{i=1}^{m} \varphi_i = 1$;

ii) $0 \leq \varphi_i \leq 1$ and the support of φ_i is contained in some $V_{\alpha_i} = V_i$;

and we then define the integral of the k-form ω on S.

For each i, let $\omega_i = \varphi_i \, \omega$. Then

$$\sum_{i=1}^{m} \omega_i = \sum_{i=1}^{m} \varphi_i \, \omega = \omega \sum_{i=1}^{m} \varphi_i = \omega.$$

Moreover, the support of each form ω_i is contained in the parameterised neighbourhood V_i associated with φ_i. Indeed, if $\varphi_i(a) \neq 0$ and $\omega(a) \neq 0$, then $a \in \mathrm{supp}(\varphi_i) \cap \mathrm{supp}(\omega)$, so that $\mathrm{supp}(\omega_i) \subset \mathrm{supp}(\varphi_i) \subset V_i$. Since $\mathrm{supp}(\omega_i)$ is a closed subset of $\mathrm{supp}(\varphi_i)$, it is compact. Hence, by the previous definition, $\int_S \omega_i$ makes sense.

In view of these considerations, we set

$$\int_S \omega = \sum_{i=1}^{m} \int_S \varphi_i \, \omega.$$

It remains to show that this definition is independent of the particular partition of unity chosen.

In fact, let $\{W_\beta\}$ be another covering of S that induces on S the same orientation as $\{V_\alpha\}$ and let $\{\psi_j\}_{j=1}^{s}$ be the corresponding partition of unity. Thus $\{V_\alpha \cap W_\beta\}$ is a covering of S, and the family of functions $\varphi_i \psi_j$ is a partition of unity subordinated to this covering. In this case, we have

$$\begin{aligned} \sum_{i=1}^{m} \int_S \varphi_i \omega &= \sum_{i=1}^{m} \int_S \varphi_i \Big(\sum_{j=1}^{s} \psi_j \Big) \omega \\ &= \sum_{i,j} \int_S \varphi_i \psi_j \omega. \end{aligned}$$

Similarly,

$$\sum_{j=1}^{s} \int_S \psi_j \omega = \sum_{j=1}^{s} \int_S \psi_j \Big(\sum_{i=1}^{m} \varphi_i \Big) \omega = \sum_{i,j} \int_S \varphi_i \psi_j \omega,$$

and this shows the desired independence.

In summary, the integral of a differential form with compact support reduces to a multiple integral.

6.3 Surfaces with boundary

From now on, we shall enlarge the concept of a surface, so that it will come to include, for example, closed balls in the Euclidean space. For this, we shall allow parametrisations to have open subsets of subspaces as domains.

Definition 6.3.1 *A half-space in* $\mathbb{R}^n$ *is a set of the form*

$$H = \{x \in \mathbb{R}^n;\ \alpha(x) \leq 0\},$$

where $\alpha \in (\mathbb{R}^n)^*$ *is a non-zero linear functional. The boundary of the half-space is the set* $\partial H = \{x \in \mathbb{R}^n;\ \alpha(x) = 0\}$.

If $A \subset H$ is open in the half-space $H \subset \mathbb{R}^n$, we say that $f : A \to \mathbb{R}^m$ is differentiable if there exists a function $\overline{f} : U \to \mathbb{R}^m$, where $U \subset \mathbb{R}^n$ is open, $U \supset A$, such that the restriction of $\overline{f}$ to A coincides with f.

Observe that, if $f : A \to \mathbb{R}^m$ is differentiable, then for each $x \in A$, the derivative $f'(x) : \mathbb{R}^n \to \mathbb{R}^m$ is well defined. In fact, if $x \in \operatorname{int}(H)$ there is nothing to prove. Now let $x \in A \cap \partial H$, and let $\{v_1, \ldots, v_n\}$ be a basis of $\mathbb{R}^n$ such that $v_i \in H$ for $1 \leq i \leq n$. Note that such a basis exists, since for any $v \in \mathbb{R}^n$ we have $v \in H$ or $-v \in H$. Thus, if $x \in \partial H$, then for every $t \leq 0$ we have $x + tv_i \in H$, because, as

$$H = \{y \in \mathbb{R}^n;\ \alpha(y) \leq 0\},$$

we obtain

$$\alpha(x + tv_i) = \alpha(x) + t\alpha(v_i) \leq 0, \qquad \forall t \leq 0.$$

In particular, letting $t \to 0$ through positive values,

$$\begin{aligned} \overline{f}'(x)v_i &= \lim_{t \to 0^+} \frac{\overline{f}(x + tv_i) - \overline{f}(x)}{t} \\ &= \lim_{t \to 0^+} \frac{f(x + tv_i) - f(x)}{t}. \end{aligned}$$

In view of the above, we now enlarge the notion of parametrisation.

Definition 6.3.2 *A parametrisation (of class* C^k *and dimension* m*) of a set* $U \subset \mathbb{R}^n$ *is a homeomorphism* $\varphi : U_0 \to U$ *of class* C^k*, defined on an open set* U_0 *of a half-space of* $\mathbb{R}^m$*, such that* $\varphi'(u) : \mathbb{R}^m \to \mathbb{R}^n$ *is an injective linear map for each* $u \in U_0$.

Definition 6.3.3 *A set* $S \subset \mathbb{R}^n$ *is called a surface with boundary (of dimension* m *and class* C^k*) if every* $x \in S$ *belongs to an open set* $U \subset \mathbb{R}^n$ *which is the image of a parametrisation* $\varphi : U_0 \to U$*, of class* C^k*, defined on an open set* U_0 *of some half-space of* $\mathbb{R}^m$*. If* S *is a surface with boundary, the boundary of* S *is the set consisting of those points* $x \in S$ *such that, for every parametrisation* $\varphi : U_0 \to U$ *of class* C^k *onto some open* $U \subset S$ *with* $x \in \varphi(U_0)$*, one has* $x \in \partial U_0$*. Moreover, this definition does not depend on the parametrisation chosen.*

Given $x \in S$, it suffices that there exists a parametrisation of class C^k of an open set $U \subset S$ with $x = \varphi(u)$ and $u \in \partial U_0$, in order to have $x \in \partial S$. Indeed, let $\varphi : U_0 \to U$ be a parametrisation of class C^k such that $x = \varphi(u)$ and $u \in \partial U_0$. Suppose, by contradiction, that there exists another parametrisation $\psi : V_0 \to V$ of class C^k such that $x = \psi(v)$ but $v \notin \partial V_0$.

Consider $W = U \cap V$, which is clearly non-empty since $x \in U \cap V$, and the map

$$\varphi^{-1} \circ \psi : \psi^{-1}(W) \to \varphi^{-1}(W),$$

which is a diffeomorphism. Since $v \notin \partial V_0$, there exists a neighbourhood B of v such that $B \subset \psi^{-1}(W)$ and

$$B \cap \partial H = \emptyset,$$

that is, B does not intersect the hyperplane $\alpha = 0$. Say that

$$B \subset H = \{x \in \mathbb{R}^m;\ \alpha(x) \leq 0\}.$$

Restricting $\varphi^{-1} \circ \psi$ to B, we have a differentiable map

$$\varphi^{-1} \circ \psi : B \to H,$$

with Jacobian non-zero at some point $q_2 \in U$. By the Inverse Function Theorem, there exists a diffeomorphism between a neighbourhood $G \subset B$ of v and a neighbourhood of $\varphi^{-1} \circ \psi(v)$ in $\mathbb{R}^m$. But, since $u \in \partial U_0$, for any neighbourhood V of u we have $V \cap \partial H \neq \emptyset$, and in particular,

$$\varphi^{-1} \circ \psi(G) \cap \partial H \neq \emptyset,$$

which is a contradiction, because

$$\varphi^{-1} \circ \psi(B) \subset \varphi^{-1}(W) \subset H.$$

This proves the claim.

We now examine the relation between the dimension of a surface with boundary and the dimension of its boundary, and how these objects are related.

Proposition 6.3.4 *If S is a surface with boundary of class C^k and dimension m $(m < n)$, then its boundary ∂S is a (boundaryless) surface of class C^k and dimension $m - 1$.*

Proof:

Indeed, the parametrisations that characterise ∂S as a surface are the restrictions to the boundary $\partial U_0 = U_0 \cap \partial H$ of the parametrisations $\varphi : U_0 \to U$, of class C^k on S, whose image is an open set $U \subset \mathbb{R}^n$ and such that $U \cap \partial S \neq \emptyset$. To verify that the dimension of ∂S is $m-1$, note that another way to parametrise ∂U is the following: write the elements of $\mathbb{R}^m$ as $u = (u_0, \ldots, u_{m-1})$, set $H_0 = \{u \in \mathbb{R}^m;\ u_0 \leq 0\}$ and identify ∂H_0 with $\mathbb{R}^{m-1}$ via the correspondence

$$(0, u_1, \ldots, u_{m-1}) \mapsto (u_1, \ldots, u_{m-1}).$$

Next, we standardise the parametrisations of class C^k on S by considering only those defined on open subsets of the half-space H_0. If $\varphi : U_0 \to U$ is standard and $U \cap \partial S \neq \emptyset$, the restriction $\varphi|_{\partial U_0} : \partial U_0 \to \partial U$ is a parametrisation of the surface ∂S, defined on an open subset $\partial U_0 \subset \mathbb{R}^{m-1}$. In this way, $\dim(\partial S) = m - 1$. □

As in the case of regular surfaces, for surfaces with boundary one also has the notion of 'tangent plane', a notion which is likewise local.

Definition 6.3.5 *Let $S \subset \mathbb{R}^n$ be a surface with boundary of class C^1 and dimension m $(m < n)$. To each point $x \in S$ we associate a vector subspace $T_xS \subset \mathbb{R}^n$ of dimension m, called the tangent space to S at x, defined as the image $\varphi'(u)(\mathbb{R}^m)$, where $\varphi : U_0 \to U$ is any parametrisation of class C^k of an open set $U \subset S$ such that $x = \varphi(u)$.*

If $x \in \partial S$, then U_0 is open in a half-space $H \subset \mathbb{R}^m$ with $u = \varphi^{-1}(x) \in \partial H$. The image $\varphi'(u)(\partial H) = T_x(\partial S)$ is the tangent space to the boundary ∂S at x. Obviously, $T_x(\partial S) \subset T_xS$, and it is a subspace of dimension $m-1$. As seen earlier, the tangent space $T_xS = \varphi'(u)(\mathbb{R}^m)$ does not depend on the parametrisation used to define it.

Definition 6.3.6 *Let $S \subset \mathbb{R}^n$ be a surface with boundary and let $x \in \partial S$. We say that a vector $w \in T_xS$ points outwards from the surface S if there exists a parametrisation $\varphi : U_0 \to U$ of class C^1 on an open set U_0 of a half-space $H \subset \mathbb{R}^m$, with image an open set $U \subset S$, such that $x = \varphi(u) \in U$ and $w = \varphi'(u)w_0$, where $w_0 \in \mathbb{R}^m$ points outwards from the half-space H. Moreover, this concept does not depend on the parametrisation chosen.*

For $x \in \partial S$, the tangent space T_xS contains not only the distinguished subspace $T_x(\partial S)$ but also a half-space, formed by the vectors that point outwards from the surface S.

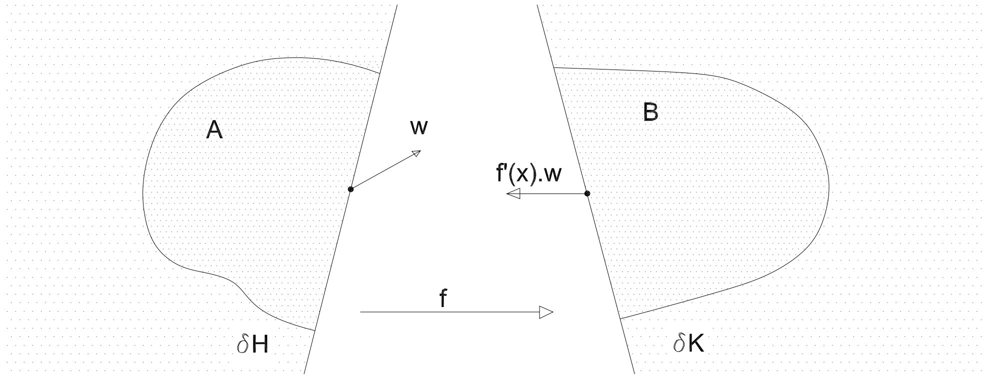

Figure 6.8:

At each point $x \in \partial S$, the vectors tangent to ∂S together with the vectors that point outwards from S at x form a half-space of $T_x S$. Among the vectors that point outwards from S at x, there is a unique one of length 1 which is normal to $T_x(\partial S)$. We denote this vector by $\nu(x)$. In this way, we obtain a field of unit vectors $\nu : \partial S \to \mathbb{R}^n$ normal to ∂S.

Indeed, if $\varphi : U_0 \to U$ is a parametrisation of class C^k defined on the open set U_0 of the half-space $H \subset \mathbb{R}^m$, choose a basis $\{v_0, \ldots, v_{m-1}\} \subset \mathbb{R}^m$ such that v_0 points outwards from H and $\{v_1, \ldots, v_{m-1}\} \subset \partial H$. Then:

$$\nu(x) = \frac{\varphi'(u)v_1 \times \cdots \times \varphi'(u)v_{m-1}}{\|\varphi'(u)v_1 \times \cdots \times \varphi'(u)v_{m-1}\|},$$

for every $x = \varphi(u) \in \partial U = \partial S \cap U$ (here we are assuming that the basis $\{\varphi'(u)v_1, \ldots, \varphi'(u)v_{m-1}\}$ is positive).

Thus, if S is a surface with boundary of class C^k and dimension m in $\mathbb{R}^n$, then its boundary ∂S is an orientable surface.

Definition 6.3.7 *A surface with boundary is said to be orientable if it admits a coherent atlas of class C^1, that is, given parametrisations φ, ψ of S, the change of parametrisation has positive Jacobian at each point of its domain.*

From the above, we conclude that the boundary of a surface is endowed with a natural orientation. Furthermore, if the surface S is an oriented surface, then it induces an orientation on its boundary.

Proposition 6.3.8 *If S is an orientable surface, then its boundary is also orientable.*

Proof:

Let A be the set of parametrisations $\varphi : U_0 \to U$ of S, of class C^1, with the following properties:

i) U_0 is connected;

ii) U_0 is open in the half-space $H_0 = \{u = (u_0, \dots, u_{m-1}) \in \mathbb{R}^m;\ u_0 \leq 0\}$;

iii) φ is positive with respect to the orientation of S.

Note that A, under the above conditions, is an atlas on S. Identifying, as before, $\mathbb{R}^{m-1}$ with ∂H_0, let A_0 be the set of restrictions $\varphi_0 = \varphi_{|\partial U_0}$ of the parametrisations $\varphi \in A$ such that $\partial U_0 = U_0 \cap \mathbb{R}^{m-1} \neq \emptyset$. Note that A_0 is a C^1 atlas on ∂S, and moreover it is coherent. In fact, if $\varphi_0 : \partial U_0 \to U$ and $\psi_0 : \partial V_0 \to V$ belong to A_0, with $\partial U \cap \partial V \neq \emptyset$, then the change of parametrisation $\psi_0^{-1} \circ \varphi_0$ is the restriction to the boundary of the diffeomorphism $\psi^{-1} \circ \varphi$ on its domain. Let $u \in \partial U \cap \partial V$ and $A = (\psi^{-1} \circ \varphi)'(u) : \mathbb{R}^m \to \mathbb{R}^m$. Since A is coherent, we have $\det A > 0$. As $\psi^{-1} \circ \varphi$ is a diffeomorphism from the open set $\varphi^{-1}(U \cap V) \subset H_0$ onto the open set $\psi^{-1}(U \cap V) \subset H_0$, it follows that $A(\partial H_0) = \partial H_0$, i.e., $Av_i = (0, a_{1i}, \dots, a_{m-1,i})$ for every $i = 1, 2, \dots, m-1$. Since $e_0 = (1, 0, \dots, 0)$ points outwards from H_0, we have $Ae_0 = (a_{00}, a_{10}, \dots, a_{m-1,0})$ also pointing outwards from H_0, hence $a_{00} > 0$. Thus, the matrix of A has the form

$$\begin{pmatrix} a_{00} & 0 & \cdots & 0 \\ a_{10} & a_{11} & \cdots & a_{1,m-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m-1,0} & a_{m-1,1} & \cdots & a_{m-1,m-1} \end{pmatrix}$$

with $a_{00} > 0$, and therefore

$$\det A = a_{00} \det A_0,$$

where $A_0 = A_{|\mathbb{R}^{m-1}}$ is the Jacobian of $\psi^{-1} \circ \varphi$ at the point u. Hence $\det A_0 > 0$, so A_0 is coherent. □

The orientation defined on ∂S by the atlas A_0 is said to be induced by the orientation of S.

With respect to the orientation induced by S on ∂S, a basis $\{w_1, \dots, w_{m-1}\} \subset T_x(\partial S)$ is positive if and only if, for any vector w_0 that points outwards from S, the set $\{w_0, w_1, \dots, w_{m-1}\}$ is a positive basis of T_xS.

In particular, if $\nu(x) \in T_xS$ is the unit vector tangent to S and normal to ∂S at x, pointing outwards from S, then $\{w_1, \dots, w_{m-1}\} \subset T_x(\partial S)$ is a positive basis if and only if the basis $\{\nu(x), w_1, \ldots, w_{m-1}\} \subset T_x S$ is positive.

Indeed, we have $\{w_1, \ldots, w_{m-1}\} \subset T_x(\partial S)$ positive if and only if

$$w_j = \sum_{i=1}^{m} a_{ij} \frac{\partial \varphi}{\partial u_i}(u), \quad j = 1, \ldots, m-1,$$

where the $(m-1) \times (m-1)$ matrix $A_0 = (a_{ij})$ has positive determinant and $\varphi : U_0 \to U$ is a parametrisation defined on an open set U_0 of the half-space

$$H_0 = \{(u_0, \ldots, u_{m-1}) \in \mathbb{R}^m;\ u_0 \leq 0\},$$

with $u \in \partial H_0$ and $\varphi(u) = x$.

Since $e_0 = (1, 0, \ldots, 0)$ points outwards from H_0,

$$\frac{\partial \varphi}{\partial u_0}(u) = \varphi'(u) e_0 \in T_x S$$

points outwards from the surface S. Hence, if $w_0 \in T_x S$ is any vector pointing outwards from S, then

$$w_0 = a_{00} \frac{\partial \varphi}{\partial u_0}(u) + a_{10} \frac{\partial \varphi}{\partial u_1}(u) + \cdots + a_{m-1,0} \frac{\partial \varphi}{\partial u_{m-1}}(u), \quad a_{00} > 0,$$

and for $j = 1, 2, \ldots, m-1$ we have

$$w_j = 0 \cdot \frac{\partial \varphi}{\partial u_0}(u) + a_{1j} \frac{\partial \varphi}{\partial u_1}(u) + \cdots + a_{m-1,j} \frac{\partial \varphi}{\partial u_{m-1}}(u).$$

Thus, the matrix A of the change of basis from

$$\left\{ \frac{\partial \varphi}{\partial u_0}(u), \frac{\partial \varphi}{\partial u_1}(u), \ldots, \frac{\partial \varphi}{\partial u_{m-1}}(u) \right\}$$

to the basis

$$\{w_0, \ldots, w_{m-1}\}$$

has the form

$$A = \begin{pmatrix} a_{00} & 0 & \cdots & 0 \\ a_{10} & \vdots & & \vdots \\ \vdots & \ddots & & A_0 \\ a_{m-1,0} & \ddots & & \vdots \end{pmatrix}.$$

Hence $\det A = a_{00} \det A_0$, that is, $\det A > 0 \Leftrightarrow \det A_0 > 0$.

This means that, when $w_0 \in T_x S$ points outwards from S, the set $\{w_1, \ldots, w_{m-1}\} \subset T_x(\partial S)$ is a positive basis if and only if $\{w_0, w_1, \ldots, w_{m-1}\} \subset T_x S$ is a positive basis.

6.4 Stokes's Theorem and applications

We are now in a position to state Stokes's Theorem, which allows the computation of surface integrals, and to reformulate it in the classical cases.

Teorema 6.4.1 (Stokes's Theorem) *Let ω be a differential form of class C^1, of degree m, with compact support on an oriented surface S of dimension $m+1$, whose boundary ∂S is endowed with the induced orientation. Then:*

$$\int_S d\omega = \int_{\partial S} \omega.$$

Proof:

Let K be the support of ω. If $\omega = \omega_1 + \cdots + \omega_k$ and the theorem holds for each term ω_i, then it holds for the sum ω, since $d\omega = \sum d\omega_i$ and

$$\int_S d\omega = \sum \int_S d\omega_i = \sum_{i=1}^{k} \int_{\partial S} \omega_i = \int_{\partial S} \omega.$$

We consider the following cases:

(i) K is contained in the image of a positive parametrisation $\varphi : U_0 \to U$ and $U \cap \partial S = \emptyset$.

In this case, for every $x = \varphi(u) \in U$ we have

$$\omega(x) = \sum_{i=0}^{m} a_i(u)\, du_0 \wedge \cdots \wedge du_{i-1} \wedge du_{i+1} \wedge \cdots \wedge du_{m+1},$$

with $a_i(u)$ differentiable functions. Hence

$$d\omega(x) = \left[\sum_{i=0}^{m} (-1)^i \frac{\partial a_i}{\partial u_i}\right] du_0 \wedge \cdots \wedge du_m.$$

Since $U \cap \partial S = \emptyset$, ω vanishes on ∂S, and therefore $i^*\omega = 0$, so that

$$\int_{\partial S} i^*\omega = 0.$$

We now show that

$$\int_S d\omega = \int_{U_0} \left(\sum_{i=0}^{m} (-1)^i \frac{\partial a_i}{\partial u_i}\right) du_0 \wedge \cdots \wedge du_m = 0.$$

Indeed, extend each function a_i to H_0 by setting

$$a_i(u) = \begin{cases} a_i(u), & u \in U_0, \\ 0, & u \in H_0 \setminus U_0, \end{cases}$$

so that, as $\varphi^{-1}(K) \subset U_0$, the functions a_i thus defined are differentiable on H_0. Let $Q \subset H_0$ be the m-dimensional parallelepiped given by $u_i^1 \leq u_i \leq u_i^0$, $0 \leq i \leq m$, containing $\varphi^{-1}(K)$ in its interior. Then

$$\begin{aligned} \int_S d\omega &= \sum_{i=0}^{m} \int_{U_0} (-1)^i \frac{\partial a_i}{\partial u_i}(u)\, du_0 \wedge \cdots \wedge du_m \\ &= \sum_{i=0}^{m} (-1)^i \int_{U_0} \frac{\partial a_i}{\partial u_i}(u)\, du_0 \wedge \cdots \wedge du_m \\ &= \sum_{i=0}^{m} (-1)^i \int_Q \left[\frac{\partial a_i}{\partial u_i}(u)\, du_i\right] du_0 \wedge \cdots \wedge du_{i-1} \wedge du_{i+1} \wedge \cdots \wedge du_m \\ &= \sum_{i=0}^{m} (-1)^i \int_Q \big[a_i(u_0, \ldots, u_{i-1}, u_i^0, u_{i+1}, \ldots, u_m) \\ &\qquad - a_i(u_0, \ldots, u_{i-1}, u_i^1, u_{i+1}, \ldots, u_m)\big]\, du_0 \cdots du_{i-1} du_{i+1} \cdots du_m = 0, \end{aligned}$$

because $a_i(u_0, \ldots, u_i^0, \ldots, u_m) = a_i(u_0, \ldots, u_i^1, \ldots, u_m) = 0$.

(ii) Now suppose that $U \cap \partial S \neq \emptyset$. Then the inclusion map is written as

$$i(u_0) = 0, \qquad i(u_j) = u_j,\ j = 1, 2, \ldots, m.$$

By the definition of the induced orientation, the restriction of φ to $\partial U_0 = U_0 \cap H_0$ is a positive parametrisation of ∂S, whose image is ∂U. As in case *(i)*, consider the extension of the functions a_i by

$$a_i(u) = a_i(u), \quad u \in U_0,$$

$$a_i(u) = 0, \quad u \in H_0 \setminus U_0,$$

which are differentiable. Let $K = \prod_{i=0}^{m} [\alpha_i, \beta_i]$ be a rectangle containing $\varphi^{-1}(K)$ with $\beta_0 = 0$, so that $K \subset H_0$. For each $i = 0, 1, \ldots, m$ let K_i be the Cartesian product of the intervals $[\alpha_j, \beta_j]$, $j \neq i$. In particular, $K_0 = \prod_{i=1}^{m} [\alpha_i, \beta_i]$ is a rectangle in ∂H_0 containing $\varphi^{-1}(\mathrm{supp}(i^*\omega))$. For every $x = \varphi(u) = \varphi(0, u_1, \ldots, u_m) \in \partial U$, clearly (by the induced orientation)

$$(i^*\omega)(x) = a_0(0, u_1, \ldots, u_m)\, du_1 \wedge \cdots \wedge du_m,$$

and hence

$$\int_{\partial S} \omega = \int_{\partial S} i^*\omega = \int_{K_0} a_0(0, u_1, \ldots, u_m)\, du_1 \wedge \cdots \wedge du_m.$$

It follows that

$$\begin{aligned}\int_S d\omega &= \sum_{i=0}^{m} \int_K \frac{\partial a_i}{\partial u_i}(u)\, du_0 du_1 \cdots du_m \\ &= \sum_{i=0}^{m} \int_{K_i} \left[\int_{\alpha_i}^{\beta_i} \frac{\partial a_i}{\partial u_i}(u)\, du_i\right] du_0 \cdots du_{i-1} du_{i+1} \cdots du_m \\ &= \sum_{i=0}^{m} \int_{K_i} \left[a_i(u_0, \dots, \beta_i, \dots, u_m) - a_i(u_0, \dots, \alpha_i, \dots, u_m)\right] du_0 \cdots du_{i-1} du_{i+1} \cdots du_m.\end{aligned}$$

Since

$$a_i(u_0, \dots, \beta_i, \dots, u_m) = a_i(u_0, \dots, \alpha_i, \dots, u_m) = 0, \quad i = 1, 2, \dots, m,$$

and

$$a_0(u_0, u_1, \dots, u_m) = 0$$

for $u_0 < 0$, we obtain

$$\int_S d\omega = \int_{K_0} a_0(0, u_1, \dots, u_m)\, du_1 \cdots du_m = \int_{\partial S} \omega.$$

(iii) Finally, consider the general case. Let $\{V_\alpha\}$ be a covering of S by coordinate neighbourhoods, and let $\varphi_1, \dots, \varphi_m$ be a differentiable partition of unity subordinated to $\{V_\alpha\}$. The forms $\omega_j = \varphi_j \omega$, $j = 1, 2, \dots, m$, satisfy the hypotheses of either case *(i)* or case *(ii)*. Moreover, since

$$\sum_j d\varphi_j = 0,$$

we have $\sum \omega_j = \omega$ and $\sum d\omega_j = d\omega$. Therefore

$$\begin{aligned}\int_S d\omega &= \sum_{j=1}^{m} \int_S d\omega_j \\ &= \sum_{j=1}^{m} \int_{\partial S} i^* \omega_j \\ &= \int_{\partial S} \sum_{j=1}^{m} i^* \omega_j \\ &= \int_{\partial S} i^* \Big(\sum_{j=1}^{m} \omega_j\Big) \\ &= \int_{\partial S} i^* \omega.\end{aligned}$$

□

From now on, our aim is to reformulate Stokes's Theorem in its classical forms.

6.4.1 The divergence theorem

Let $X = (a_0, \ldots, a_m)$ be a vector field of class C^k, defined on an open set $U \subset \mathbb{R}^{m+1}$ by its $m+1$ coordinate functions $a_i : U \to \mathbb{R}$. To the field X we associate the differential form

$$\alpha_X = \sum_{i=0}^{m} (-1)^i a_i \, dx_0 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_m,$$

of class C^k on the open set U. Expansion of a determinant along the first column shows that, for $x \in U$ and $w_1, \ldots, w_m \in \mathbb{R}^{m+1}$,

$$\alpha_X(x)(w_1, \ldots, w_m) = \det(X(x), w_1, \ldots, w_m).$$

Indeed,

$$\begin{aligned}
\alpha_X(x)(w_1, \ldots, w_m) &= \sum_{i=0}^{m} (-1)^i a_i(x) \, (dx_0 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_m) \\
&\quad (w_1, \ldots, w_m) \\
&= \sum_{i=0}^{m} (-1)^i a_i(x) \det \begin{pmatrix} dx_0(w_1) & \cdots & dx_0(w_i) & \cdots & dx_0(w_m) \\ dx_1(w_1) & \vdots & \cdots & \vdots & \vdots \\ dx_{i-1}(w_1) & \vdots & \ddots & \vdots & \vdots \\ dx_{i+1}(w_1) & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ dx_m(w_1) & \vdots & \ldots & \vdots & dx_m(w_m) \end{pmatrix} \\
&= \det \begin{pmatrix} a_0(x) & w_{01} & \cdots & w_{0m} \\ a_1(x) & w_{11} & \cdots & w_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ a_m(x) & w_{m1} & \cdots & w_{mm} \end{pmatrix} \\
&= \det(X(x), w_1, \ldots, w_m).
\end{aligned}$$

If $M \subset U$ is an oriented surface and the vector field X has compact support, we define $\int_M X$ as $\int_M \langle X, \nu \rangle \omega$, where ω is the m-form element of volume on M and ν is the unit normal vector field determining the orientation of M. Observe that at each point $x \in M$ this form coincides with the previously defined $\alpha_X(x)$. In fact, given any positive basis $\{w_1, \ldots, w_m\} \subset T_x M$, we have:

$$\begin{aligned}
\alpha_X(x)(w_1, \ldots, w_m) &= \det(X(x), w_1, \ldots, w_m) \\
&= \langle X(x), w_1 \times \cdots \times w_m \rangle \\
&= \langle X(x), \nu(x) \rangle \| w_1 \times \cdots \times w_m \| \\
&= \langle X(x), \nu(x) \rangle \, \omega(x)(w_1, \ldots, w_m).
\end{aligned}$$

Hence, $\alpha_X = \langle X, \nu \rangle \omega$.

The volume element on a surface is always denoted by dM. Thus, if $X = (a_0, \ldots, a_m)$ is a vector field with compact support, continuous on the open set $U \subset \mathbb{R}^{m+1}$ and $M \subset U$ is an oriented surface, then

$$\int_M \langle X, \nu \rangle \, dM = \int_M \sum_{i=0}^{m} a_i \, dx_0 \cdots dx_{i-1} dx_{i+1} \cdots dx_m.$$

When X is a vector field of class C^1 on the open set $U \subset \mathbb{R}^{m+1}$ and $\Omega \subset U$ is a compact domain with regular boundary of class C^k $(k \geq 1)$ (that is, Ω is a compact surface with boundary, of dimension $m+1$ and class C^k, contained in U), the interior of Ω is a bounded open subset of $\mathbb{R}^{m+1}$ and its boundary $\partial\Omega$ is a compact oriented surface in $\mathbb{R}^{m+1}$. The differential of the form α_X is

$$d\alpha_X = \left(\sum_{i=0}^{m} (-1)^i \frac{\partial a_i}{\partial x_i} \right) dx_0 \wedge \cdots \wedge dx_m.$$

We define the function $\operatorname{div} X : U \to \mathbb{R}$ by

$$\operatorname{div} X(x) = \frac{\partial a_0}{\partial x_0}(x) + \cdots + \frac{\partial a_m}{\partial x_m}(x),$$

the divergence of the vector field X. Stokes's Theorem allows us to conclude that, under these conditions, if $M = \partial\Omega$ then:

$$\int_M \langle X, \nu \rangle dM = \int_\Omega (\operatorname{div} X) \, dx,$$

where, in the second member, we have the integral of the continuous function $\operatorname{div} X$ on the Jordan-measurable compact set $\Omega \subset \mathbb{R}^{m+1}$.

6.4.2 Stokes's Theorem (vector form)

Let $X = (a, b, c)$ be a vector field of class C^1 on the open set $U \subset \mathbb{R}^3$ containing the compact oriented surface M (of dimension 2), whose boundary C is endowed with the induced orientation. To the field X we associate the 1-form $\beta_X = a \, dx + b \, dy + c \, dz$ of class C^1 on U. Stokes's Theorem can be written explicitly as

$$\int_M \left(\frac{\partial c}{\partial y} - \frac{\partial b}{\partial z} \right) dy \wedge dz + \left(\frac{\partial a}{\partial z} - \frac{\partial c}{\partial x} \right) dz \wedge dx + \left(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y} \right) dx \wedge dy = \int_C a \, dx + b \, dy + c \, dz.$$

We want to express the above identity in vector notation. To this end, consider the curl of X, which is the field $\operatorname{rot} X : U \to \mathbb{R}^3$ given by

$$\operatorname{rot} X = \left(\frac{\partial c}{\partial y} - \frac{\partial b}{\partial z}, \frac{\partial a}{\partial z} - \frac{\partial c}{\partial x}, \frac{\partial b}{\partial x} - \frac{\partial a}{\partial y} \right).$$

Then the first member of the previous equality can be written as

$$\int_M \langle \operatorname{rot} X, \nu \rangle \, dM,$$

where ν is the unit normal vector field to M and dM is the area element. For the second member, let ds denote the arc length element of C. For $x \in C$ and $v \in T_xC$, we have $ds(v) = \pm|v|$ (depending on the direction of v). If $\tau(x) \in T_xC$ is the unit tangent vector pointing in the positive direction of C, then $\tau(x) = \pm\frac{v}{|v|}$ for all $v \neq 0$ in T_xC. Thus, if $0 \neq v \in T_xC$, we obtain

$$\beta_X(v) = \langle X, v \rangle = \left\langle X, \frac{v}{|v|} \right\rangle |v| = \langle X, \pm\tau \rangle \pm ds(v) = \langle X, \tau \rangle \, ds(v).$$

Hence $\beta_X = \langle X, \tau \rangle \, ds$, and therefore

$$\int_M \langle \operatorname{rot} X, \nu \rangle \, dM = \int_C \langle X, \tau \rangle \, ds.$$

The first member represents the flux of the field $\operatorname{rot} X$ across the surface M, and the second member is the circulation of the field X along the boundary $C = \partial M$.

6.4.3 Green's Theorem

Green's Theorem concerns a compact surface with boundary, of class C^1 and dimension 2 in $\mathbb{R}^2$, that is, a compact domain $M \subset \mathbb{R}^2$ with regular boundary of class C^1. The domain M has the natural orientation of $\mathbb{R}^2$, and its boundary ∂M is endowed with the induced orientation. Let $f, g : M \to \mathbb{R}$ be functions of class C^1. Green's Theorem states that

$$\int_M \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} \right) dx \, dy = \int_{\partial M} f \, dx + g \, dy.$$

This is precisely Stokes's Theorem applied to the 1-form $\beta = f \, dx + g \, dy$, defined on the surface with boundary M.

Bibliography

[1] BATLE, R.G. *Elementos de Análise Real.* Editora Campus, Rio de Janeiro, 1983

[2] CURANT, R. *Diferential and Integral Calculus, vol II.* Interscience, New York ,1937

[3] DIEUDONNE, J. *Foundations of Modern Analysis.* Academis Press ,1960.

[4] DO CARMO, M.P *Formas Diferenciais e Aplicações.* IMPA, Rio de Janeiro, 1983

[5] LIMA, E. L. *EspaÃ§os Métricos.* Projeto Euclides, IMPA - Rio de Janeiro, 1977.

[6] LIMA, E. L. *Curso de Análise, vol 1 e 2.* Projeto Euclides, IMPA - Rio de Janeiro, 1981.

[7] RUDIN, W. *Principles of Mathematical Analysis.* McGRAW-HILL International Book Company, 1964.

[8] SPIVAK, M. *Calculus on Manifolds.* Bernjamin, New York ,1965.