# CALCULUS OF SEVERAL VARIABLES

Lecture notes for MA 642

Rudi Weikard

$$\int_\phi d\omega = \int_{\partial\phi} \omega$$

Version of August 3, 2022

# Contents


Preface iii

Chapter 1. Limits and continuity 1
- 1.1. Norms 1
- 1.2. Limits and continuity 2

Chapter 2. Differentiation 5
- 2.1. The total derivative 5
- 2.2. Partial derivatives 6
- 2.3. Taylor's theorem and extrema 8
- 2.4. The inverse and implicit function theorems 9
- 2.5. Extrema under constraints 11

Chapter 3. The multi-dimensional Riemann integral 13

Chapter 4. Integration of differential forms 17
- 4.1. Integration along paths 17
- 4.2. Integration over surfaces 17
- 4.3. The general case 18
- 4.4. Stokes' theorem 22

Appendix A. Vector spaces and linear transformations 25
- A.1. Vector spaces 25
- A.2. Linear operators 26
- A.3. Some facts about spectral theory 28

Appendix B. Miscellaneous 29
- B.1. Algebra 29

List of special symbols 31

Index 33

# Preface

These are notes for a rigorous course on multi-variable calculus, the calculus of differentiation and integration of functions of several variables.

Two excellent books on the subject are the following:

- Walter Rudin. *Principles of mathematical analysis.* McGraw-Hill Book Co., New York-Auckland-Düsseldorf, third edition, 1976. (Chapters 9 & 10)
- Michael Spivak. *Calculus on manifolds. A modern approach to classical theorems of advanced calculus.* W. A. Benjamin, Inc., New York-Amsterdam, 1965.

To a large extent my notes follow one or the other of these books. The notes are terse giving the students an opportunity to devise proofs for themselves.

The notes presuppose a familiarity of the reader with single-variable calculus, topology, and linear algebra. Some results from linear algebra are collected in Appendix A (without proof). Appendix B gathers a few more miscellaneous facts.

Also at the end of the notes the reader may find an index of terms and a list of symbols which refer to the page where they are introduced.

Finally a word on notation: Throughout the notes the symbols $j$, $k$, $\ell$, $m$, and $n$ will refer to elements of $\mathbb{N}$, the set of natural numbers. Also, the symbol $\Omega$ denotes an open set in $\mathbb{R}^n$ unless noted otherwise.

CHAPTER 1

# Limits and continuity

## 1.1. Norms

**1.1.1 The inner product in $\mathbb{R}^n$.** Given $x$ and $y$ in $\mathbb{R}^n$ we define their *inner product* or *scalar product* $x \cdot y$ by

$$x \cdot y = x^\top y = \sum_{j=1}^{n} x_j y_j.$$

Here $x_j$ and $y_j$ denote the components of $x$ and $y$, respectively.

The inner product is a map from $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. Let $x, y, z \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$. Then the following statements hold:

(1) $x \cdot x > 0$ unless $x = 0$.
(2) $x \cdot y = y \cdot x$.
(3) $x \cdot (\alpha y + \beta z) = \alpha(x \cdot y) + \beta(x \cdot z)$.
(4) If $x = 0$ or $y = 0$, then $x \cdot y = 0$.

Note that $x \cdot y$ may be 0 even though neither $x$ nor $y$ is 0. Also, property (4) may be established from properties (1) – (3) without referring to the specific definition of $x \cdot y$.

**1.1.2 Schwarz's inequality.** For any two vectors $x$ and $y$ in $\mathbb{R}^n$ *Schwarz's inequality*

$$|x \cdot y| \leq \sqrt{x \cdot x}\sqrt{y \cdot y}$$

holds. To see this assume $y \neq 0$ and find the minimum of $t \mapsto (x + ty) \cdot (x + ty)$ which cannot be negative.

**1.1.3 The norm on $\mathbb{R}^n$.** Given $x \in \mathbb{R}^n$ we define its *norm* $|x|$ by

$$|x| = \sqrt{x \cdot x}.$$

We are using the same symbol for the norm of a vector in $\mathbb{R}^n$ and the absolute value of a number in $\mathbb{R}$. This cannot cause confusion even when $n = 1$.

The norm is a map from $\mathbb{R}^n$ to $[0, \infty)$. Let $x, y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$. Then the following statements hold:

(1) $|x| > 0$ unless $x = 0$.
(2) $|\alpha x| = |\alpha||x|$.
(3) $|x + y| \leq |x| + |y|$ (the *triangle inequality*).
(4) $|x| = 0$ if and only if $x = 0$.

Again property (4) follows from (1) – (3) directly (without referral to the specific definition of the norm).

**1.1.4 $\mathbb{R}^n$ as a metric space.** Given $x$ and $y$ in $\mathbb{R}^n$ we define their *distance* by $|x - y|$.

The distance function (also called a *metric*) is a map from $\mathbb{R}^n \times \mathbb{R}^n \to [0, \infty)$. Let $x, y, z \in \mathbb{R}^n$. Then the following statements hold:

(1) $|x-y|=0$ if and only if $x=y$.
(2) $|x-y|=|y-x|$.
(3) $|x-y|\le|x-z|+|z-y|$ (the *triangle inequality*).

These properties may be proved without referring to the specific definition of the norm.

A set $S$ equipped with a distance function $X\times X\to[0,\infty)$ satisfying properties (1) – (3) (properly interpreted) is called a *metric space*. In particular, every subset of $\mathbb{R}^n$ is a metric space.

**1.1.5 The norm of a linear operator.** Suppose that $A$ is a linear operator from $\mathbb{R}^n$ to $\mathbb{R}^m$ and let $M=\max\{|A_{j,k}|:1\le j\le m, 1\le k\le n\}$. Then $|Ax|^2\le mnM^2|x|^2$. Hence

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

is a finite number called the *norm* of $A$. In fact $\|A\|\le\sqrt{mn}M$.

Note that $|Ax|\le\|A\|\,|x|$ for all $x\in\mathbb{R}^n$. In fact, $\|A\|=\inf\{C:\forall x\in\mathbb{R}^n:|Ax|\le C|x|\}$.

**1.1.6 Properties of the operator norm.** Let $A,B\in L(\mathbb{R}^n,\mathbb{R}^m)$ and $\alpha\in\mathbb{R}$. Then the following statements hold (justifying the use of the word norm):

(1) $\|A\|>0$ unless $A=0$.
(2) $\|\alpha A\|=|\alpha|\,\|A\|$.
(3) $\|A+B\|\le\|A\|+\|B\|$ (the triangle inequality).

In particular, $L(\mathbb{R}^n,\mathbb{R}^m)$ is a metric space.

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

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

Note, however, that $AB$ may not be defined.

**1.1.7 The invertible linear operators form an open set.** Suppose $A,B\in L(\mathbb{R}^n,\mathbb{R}^n)$ and that $A$ is invertible. If $\gamma=\|B-A\|\,\|A^{-1}\|<1$, then $|Bx|\ge(1-\gamma)|x|/\|A^{-1}\|$ so that $B$ is also invertible. In fact, $\|B^{-1}\|\le\|A^{-1}\|/(1-\gamma)$. Hence the set of invertible linear operators on $\mathbb{R}^n$ is open in the space $L(\mathbb{R}^n,\mathbb{R}^n)$.

## 1.2. Limits and continuity

The concepts of limits for and continuity of functions between metric spaces is a familiar from topology. Nevertheless we review these here for functions between euclidean spaces.

**1.2.1 Limits.** Suppose $f$ is a function from $\Omega$ to $\mathbb{R}^m$ and $x_0$ is a point in $\overline{\Omega}$, the closure of $\Omega$. The vector $L\in\mathbb{R}^m$ is called a *limit* of $f$ at $x_0$, if the following statement holds: for every $\varepsilon>0$ there is a $\delta>0$ such that, for all $x\in\Omega$, we have that $0<|x-x_0|<\delta$ implies $|f(x)-L|<\varepsilon$.

A limit, if it exists, is uniquely determined by $f$ and $x_0$. We denoted it by $\lim_{x_0} f$ or, when convenient, by $\lim_{x\to x_0} f(x)$.

The function $f$ has limit $L$ at $x_0$ if and only if the components $f_k$ have limit $L_k$ for each $k=1,...,m$.

**1.2.2 Continuity.** Suppose $f$ is a function from $\Omega$ to $\mathbb{R}^m$ and $x_0$ is a point in $\Omega$. We say that $f$ is *continuous* at $x_0$, if the following statement holds: for every $\varepsilon>0$ there is a $\delta>0$ such that, for all $x\in\Omega$, we have that $|x-x_0|<\delta$ implies $|f(x)-f(x_0)|<\varepsilon$.

We see immediately that $f$ is continuous at $x_0$, if and only if it has a limit at $x_0$ which coincides with $f(x_0)$.

The function $f$ is called continuous on $\Omega$, if it is continuous at every point of $\Omega$. The set of all continuous functions from $\Omega$ to $\mathbb{R}^m$ is denoted by $C^0(\Omega, \mathbb{R}^m)$.

**1.2.3 Continuity of the norm.** The norm defined in 1.1.3, a function from $\mathbb{R}^n$ to $[0, \infty)$, is continuous.

**1.2.4 Linear operators are continuous.** If $A$ is a linear operator from $\mathbb{R}^n$ to $\mathbb{R}^m$, then the map $x \mapsto Ax$ is continuous.

**1.2.5 Continuity of the operator inverse.** The map $A \mapsto A^{-1}$ defined on the set of all invertible operators on $\mathbb{R}^n$ is continuous.

**1.2.6 Limit rules.** Suppose $f$ and $g$ are functions from $\Omega$ to $\mathbb{R}^m$, $h$ is a function from $\Omega$ to $\mathbb{R}$, and $x_0$ is a point in $\overline{\Omega}$. Also assume that $f$, $g$, and $h$ have limits at $x_0$. Then the following are true:

(1) $\lim_{x_0}(f + g) = \lim_{x_0} f + \lim_{x_0} g$.
(2) $\lim_{x_0} f \cdot g = (\lim_{x_0} f) \cdot (\lim_{x_0} g)$.
(3) $\lim_{x_0} hf = (\lim_{x_0} h)(\lim_{x_0} f)$.

Lastly, suppose that $f : \Omega \to \mathbb{R}^m$ has values in the open set $\Omega'$ and limit $y_0$ at $x_0$ and that $p : \Omega' \to \mathbb{R}^k$ has limit $z_0$ at $y_0$. Then $p \circ f$ has limit $z_0$ at $x_0$.

Since the concepts of limit and continuity are closely related these limit rules imply analogous rules for continuity.

CHAPTER 2

# Differentiation

## 2.1. The total derivative

**2.1.1 Definition.** Suppose $f$ is a function from $\Omega$ to $\mathbb{R}^m$ and $x$ is a point in $\Omega$. If there exists a linear operator $A : \mathbb{R}^n \to \mathbb{R}^m$, i.e., an $m \times n$-matrix $A$, such that

$$\lim_{h\to 0} \frac{|f(x+h) - f(x) - Ah|}{|h|} = 0,$$

we say that $f$ is *differentiable* at $x$ and call $A$ a *total derivative* or just a *derivative* of $f$ at the point $x$.

If $f$ is differentiable at every point of $\Omega$ we say that $f$ is differentiable on $\Omega$.[1]

**2.1.2 Uniqueness of the derivative.** Suppose $f$ is as in 2.1.1. If $A$ and $B$ are total derivatives of $f$ at $x$, then $A = B$.

Therefore it is customary to denote *the* total derivative of $f$ at $x$ by $f'(x)$. If $f$ is differentiable on $\Omega$, the map $x \mapsto f'(x)$ is a function from $\Omega$ to $\mathbb{R}^{m\times n}$.

**2.1.3 Linear approximation.** The function $f : \Omega \to \mathbb{R}^m$ is differentiable at $x_0$ if and only if there exists a linear operator $A : \mathbb{R}^n \to \mathbb{R}^m$ and a function $r : \Omega \to \mathbb{R}^m$ such that (i) $r$ is continuous at $x_0$, (ii) $r(x_0) = 0$, and (iii) the identity

$$f(x) = f(x_0) + A(x - x_0) + |x - x_0|r(x)$$

holds. Of course, $A$ is then equal to $f'(x_0)$.

The function $\mathbb{R}^n \to \mathbb{R}^m : x \mapsto f(x_0) + f'(x_0)(x - x_0)$ is called the *linear approximation* of $f$ at $x_0$.

**2.1.4 Examples.** Suppose $\Omega = \mathbb{R}^n$ and $f(x) = Ax + b$ where $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ and $b \in \mathbb{R}^m$. Then $f'(x) = A$ for every $x \in \mathbb{R}^n$.

Let $f : \mathbb{R}^2 \to \mathbb{R} : \binom{x}{y} \mapsto x^2 y$. Find $f'\binom{2}{3}$.

**2.1.5 Differentiability implies continuity.** If a function is differentiable at a given point, then it is also continuous there.

**2.1.6 Differentiation rules for sums and products.** Suppose $f$ and $g$ are functions from $\Omega$ to $\mathbb{R}^m$, $h$ is a function from $\Omega$ to $\mathbb{R}$, and $x$ is a point in $\Omega$. Also assume that $f$, $g$, and $h$ are differentiable at $x$. Then the following statements hold:

(1) $(f + g)'(x) = f'(x) + g'(x)$.
(2) $(f \cdot g)'(x) = f(x)^\top g'(x) + g(x)^\top f'(x)$.
(3) $(hf)'(x) = h(x)f'(x) + f(x)h'(x)$.

---

[1] Later we need the concept of differentiability on a compact set $K$. A function is called continuously differentiable on $K$, if it may be extended to a continuously differentiable function in some open set containing $K$.

**2.1.7 The chain rule.** Suppose $f : \Omega \to \mathbb{R}^m$ and $g : \Omega' \to \mathbb{R}^k$ where $\Omega'$ is an open set in $\mathbb{R}^m$ containing the range of $f$. If $f$ is differentiable at $x$ and $g$ is differentiable at $f(x)$, then $g \circ f$ is differentiable at $x$ and

$$(g \circ f)'(x) = g'(f(x))f'(x).$$

This is known as the *chain rule*. Its proof uses the linear approximations of $f$ and $g$ at $x$ and $f(x)$, respectively.

**2.1.8 Differentiable functions are locally Lipschitz.** Suppose $B$ is an open ball in $\mathbb{R}^n$, that $f : B \to \mathbb{R}^m$ is differentiable, and that there is a number $M$ such that $\|f'(x)\| \leq M$ for all $x \in B$. Then $f$ satisfies a *Lipschitz condition*, i.e.,

$$|f(x_2) - f(x_1)| \leq M|x_2 - x_1|$$

whenever $x_1, x_2 \in B$.

If $m = 1$ we even get $f(x_2) - f(x_1) = f'(x)(x_2 - x_1)$ for some point $x$ on the line joining $x_1$ and $x_2$.

Sketch of proof: Let $\gamma : [0, 1] \to B : t \mapsto x_1 + t(x_2 - x_1)$. Consider the function $g : [0, 1] \to \mathbb{R} : t \mapsto (f(x_2) - f(x_1)) \cdot f(\gamma(t))$. Product rule, chain rule, and the mean value theorem for one variable imply

$$|f(x_2) - f(x_1)|^2 = g(1) - g(0) = (f(x_2) - f(x_1))^\top f'(\gamma(t_0))\gamma'(t_0)$$

for some $t_0 \in (0, 1)$. □

**2.1.9 Functions with derivative $0$ are constant.** Any function $f$ with $f'(x) = 0$ for all $x$ in its domain must be constant as long as any two points in its domain can be connected by a continuous path, i.e., a continuous function $\gamma$ from $[0, 1]$ to the domain of $f$ such that $\gamma(0)$ and $\gamma(1)$ are the given points.

## 2.2. Partial derivatives

**2.2.1 Partial derivatives.** Recall that $(e_1, ..., e_n)$ is the standard basis in $\mathbb{R}^n$. Let $f = (f_1, ..., f_m)^\top$ be a function from $\Omega$ to $\mathbb{R}^m$, and $x$ a point in $\Omega$. If $1 \leq j \leq n$ and $1 \leq \ell \leq m$, define

$$(D_j f_\ell)(x) = \lim_{t \to 0} \frac{f_\ell(x + te_j) - f_\ell(x)}{t}$$

if the limit exists.

The numbers $(D_j f_\ell)(x)$, $j = 1, ..., n$, $\ell = 1, ..., m$ are called *partial derivatives* of $f$ at $x$.

**2.2.2 Differentiability implies the existence of the partial derivatives.** Suppose $f : \Omega \to \mathbb{R}^m$ is differentiable at $x$. Then the partial derivatives $D_j f_\ell$ exist and

$$f'(x) = \begin{pmatrix} (D_1 f_1)(x) & \cdots & (D_n f_1)(x) \\ \vdots & & \vdots \\ (D_1 f_m)(x) & \cdots & (D_n f_m)(x) \end{pmatrix}.$$

**2.2.3 Continuously differentiable functions.** If $f' : \Omega \to L(\mathbb{R}^n, \mathbb{R}^m)$ is continuous, then $f$ is called continuously differentiable. The set of all continuously differentiable functions from $\Omega$ to $\mathbb{R}^m$ is denoted by $C^1(\Omega, \mathbb{R}^m)$.

Theorem. $f \in C^1(\Omega, \mathbb{R}^m)$ if and only if the partial derivatives $D_j f_\ell : \Omega \to \mathbb{R}$, $j = 1, ..., n$ and $\ell = 1, ..., m$, are continuous.

SKETCH OF PROOF: If $x \mapsto f'(x)$ is continuous, then so are all the partial derivatives. Conversely, if all the partial derivatives are continuous, then so is the matrix $A$ of partial derivatives (given in 2.2.2) as a function of $x$ but we need to show that $A = f'(x)$. To this end assume first $m = 1$ and let $h = x - x_0 = \sum_{j=1}^n h_j e_j$, $v_0 = x_0$, and $v_k = x_0 + \sum_{j=1}^k h_j e_j$. Then, using the mean value theorem for functions of one variable,

$$f(x) - f(x_0) = \sum_{k=1}^n (f(v_k) - f(v_{k-1})) = \sum_{k=1}^n (D_k f)(v_{k-1} + t_k h_k e_k) h_k$$

for $t_k \in (0,1)$. Since $A(x - x_0) = \sum_{k=1}^n h_k (D_k f)(x_0)$ the claim follows using the continuity of $x \mapsto (D_k f)(x)$. □

**2.2.4 Derivatives of higher order.** Let $f$ be a function from $\Omega$ to $\mathbb{R}^m$. The partial derivatives $D_j f_\ell : \Omega \to \mathbb{R}$ may themselves have partial derivatives $D_k(D_j f_\ell)$, $k = 1, ..., n$. These are called partial derivatives of the second order. If they are continuous then $D_j f_\ell \in C^1(\Omega, \mathbb{R})$. If this is the case for all $j = 1, ..., n$ and $\ell = 1, ..., m$ we say that $f$ is twice continuously differentiable and denote the space of such functions by $C^2(\Omega, \mathbb{R}^m)$.

More generally, $C^r(\Omega, \mathbb{R}^m)$ is the space of those functions from $\Omega$ to $\mathbb{R}^m$ for which all partial derivatives of order up to and including $r \in \mathbb{N}$ exist and are continuous.

**2.2.5 Another mean value theorem.** Suppose $\Omega$ is an open subset in $\mathbb{R}^2$ and $f$ a real-valued function on $\Omega$ for which $D_1 f$ and $D_2 D_1 f$ exist everywhere. If $(a, b) \in \Omega$ and if $u$ and $v$ are so small that the rectangle $Q$ with vertices $(a, b)$, $(a+u, b)$, $(a, b+v)$, and $(a+u, b+v)$ is still in $\Omega$, then there is a point $(x, y) \in Q$ such that

$$f(a+u, b+v) - f(a+u, b) - f(a, b+v) + f(a, b) = uv(D_2 D_1 f)(x, y).$$

SKETCH OF PROOF: Let $\phi : [a, a+u] \to \mathbb{R}$ and $\psi : [b, b+v] \to \mathbb{R}$ be given by $\phi(t) = f(t, b+v) - f(t, b)$ and $\psi(t) = (D_1 f)(x, t)$ for a certain $x \in (a, a+u)$. The mean value theorem for functions of one variable applies to both $\phi$ and $\psi$. □

**2.2.6 Mixed partial derivatives commute.** If $f \in C^k(\Omega, \mathbb{R})$, $j_1, ... j_k \in \{1, ..., n\}$, and $\pi$ is a permutation of $\{1, ..., k\}$, then

$$D_{j_1} ... D_{j_k} f = D_{j_{\pi(1)}} ... D_{j_{\pi(k)}} f.$$

SKETCH OF PROOF: This follows from the following statement in which we assume that $k = 2$ and that $n = 2$ so that $\Omega$ is an open subset of $\mathbb{R}^2$. Suppose $f \in C^1(\Omega, \mathbb{R})$ and that $D_2 D_1 f$ exists and is continuous there. Then $D_1 D_2 f$ also exists and equals $D_2 D_1 f$ in $\Omega$.

Given $\varepsilon > 0$ it follows from 2.2.5 that

$$\left| \frac{f(a+u, b+v) - f(a+u, b) - f(a, b+v) + f(a, b)}{uv} - (D_2 D_1 f)(a, b) \right| < \varepsilon/2$$

for all sufficiently small but non-zero $u$ and $v$. Thus, taking $v \to 0$,

$$\left| \frac{(D_2 f)(a+u, b) - (D_2 f)(a, b)}{u} - (D_2 D_1 f)(a, b) \right| \leq \varepsilon/2 < \varepsilon.$$

□

**2.2.7 The gradient.** Suppose that all partial derivatives of $f : \Omega \to \mathbb{R}$ exist at $x \in \Omega$. The column vector

$$(\nabla f)(x) = ((D_1 f)(x), ..., (D_n f)(x))^\top$$

is called the *gradient* of $f$ at $x$.

Thus, if $f$ is differentiable at $x$, then $(\nabla f)(x) = f'(x)^\top$.

**2.2.8 Directional derivatives.** Suppose $f : \Omega \to \mathbb{R}$ is differentiable at $x$ and $u \in \mathbb{R}^n$. Then

$$\lim_{t\to 0} \frac{f(x+tu) - f(x)}{t} = u \cdot (\nabla f)(x).$$

If $u$ is a unit vector this is called the *directional derivative* of $f$ in direction $u$ at $x$.

We now define $(u \cdot \nabla)^0$ to be the identity operator, even if $u = 0$ and, recursively,

$$(u \cdot \nabla)^j f = u \cdot \nabla[(u \cdot \nabla)^{j-1} f]$$

for $j = 1, ..., k$ provided that $f \in C^k(\Omega, \mathbb{R})$.

**2.2.9 The multi-index notation.** A multi-index is an element of $\mathbb{N}_0^n$ for some natural number $n$. If $\alpha$ is such a multi-index we define

$$|\alpha| = \alpha_1 + ... + \alpha_n \quad \text{and} \quad \alpha! = \alpha_1!...\alpha_n!.$$

Furthermore, if $x \in \mathbb{R}^n$, we set

$$x^\alpha = x_1^{\alpha_1}...x_n^{\alpha_n}.$$

Finally,

$$D^\alpha f = D_1^{\alpha_1}...D_n^{\alpha_n} f$$

if $f \in C^{|\alpha|}(\Omega, \mathbb{R})$.

Using this notation and taking into account that mixed partial derivatives commute, as explained in 2.2.6, we obtain by induction and a version of the multinomial theorem (see B.1) that

$$[(u \cdot \nabla)^k f](x) = \sum_{\ell_k=1}^{n} ... \sum_{\ell_1=1}^{n} u_{\ell_k}...u_{\ell_1}(D_{\ell_k}...D_{\ell_1} f)(x) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} u^\alpha (D^\alpha f)(x)$$

for $k = 1, ..., r$ provided that $f \in C^r(\Omega, \mathbb{R})$ and $u \in \mathbb{R}^n$.

## 2.3. Taylor's theorem and extrema

**2.3.1 Taylor's theorem.** Suppose $\Omega$ is convex, $f \in C^r(\Omega, \mathbb{R})$ for some $r \in \mathbb{N}$, and $x_0, x \in \Omega$. Then there exists a number $t \in (0,1)$ such that

$$\begin{aligned} f(x) &= \sum_{k=0}^{r-1} \frac{1}{k!}[((x-x_0)\cdot\nabla)^k f](x_0) + \frac{1}{r!}[((x-x_0)\cdot\nabla)^r f](x_0 + t(x-x_0)) \\ &= \sum_{|\alpha|<r} \frac{(D^\alpha f)(x_0)}{\alpha!}(x-x_0)^\alpha + \sum_{|\alpha|=r} \frac{(D^\alpha f)(x_0 + t(x-x_0))}{\alpha!}(x-x_0)^\alpha. \end{aligned}$$

Sketch of proof: Let $h = x - x_0$ and $g = f \circ \gamma$ where $\gamma : [0,1] \to \Omega : t \mapsto x_0 + th$. Then, by the chain rule, $g'(t) = f'(\gamma(t))h = [(h \cdot \nabla)f](\gamma(t))$. Induction shows that $g^{(k)}(t) = [(h \cdot \nabla)^k f](\gamma(t))$ for $1 \le k \le r$. Now apply Taylor's theorem for one variable. □

**2.3.2 Extrema.** Let $x_0$ be a point in the domain of a real-valued function $f$. If there is a neighborhood $U$ of $x_0$ such that $f(x) \le f(x_0)$ for all $x \in U$, then $x_0$ is called a *relative maximum* of $f$. If the inequality is strict except for $x = x_0$, then $x_0$ is called a *strict relative maximum* of $f$. The terms *relative minimum* and *strict relative minimum* are analogously defined. A (strict) relative *extremum* is a point which is either a (strict) relative maximum or minimum.

**2.3.3 Critical points.** If $f$ is differentiable in its domain and if $f'(x_0) = 0$, then $x_0$ is called a *critical point* of $f$.

If $x_0$ is a relative extremum of a differentiable function $f$, then $x_0$ is a critical point of $f$. Thus we have a necessary condition for a point $x_0$ to be a relative extremum of $f$.

**2.3.4 A sufficient criterion for the presence of an extremum.** The following theorem gives a sufficient condition for a point $x_0$ to be a relative extremum of $f$. If $f$ is twice continuously differentiable the $n^2$ second order partial derivatives $(D_j D_k f)(x)$ form a real symmetric matrix, called the *Hessian* of $f$ at $x$. We will denote it by $H(x)$.

THEOREM. Let $f \in C^2(\Omega, \mathbb{R})$. Suppose that there is a point $x_0$ such that $f'(x_0) = 0$ and $H(x_0)$ is positive (negative) definite. Then $x_0$ is a strict relative minimum (maximum) of $f$.

SKETCH OF PROOF: Suppose $H(x_0)$ is positive definite, i.e., its smallest eigenvalue is positive. For sufficiently small $h$ we obtain from Taylor's theorem the existence of a $t \in (0, 1)$ such that

$$f(x_0 + h) - f(x_0) = \sum_{|\alpha|=2} \frac{(D^\alpha f)(x_0 + th)}{\alpha!} h^\alpha = \frac{1}{2} h^\top H(x_0 + th)h.$$

Since the quadratic form $q$ given by $q(h) = h^\top H(x_0)h$ is positive definite, we obtain that $f(x_0 + h) - f(x_0) > 0$ for sufficiently small but non-zero $h$. □

**2.3.5 A sufficient criterion for the absence of an extremum.** Let $f \in C^2(\Omega, \mathbb{R})$ and suppose that there is a point $x_0$ such that $f'(x_0) = 0$. Let $H$ denote the Hessian of $f$. Then the following two equivalent statements hold:

(1) If $H(x_0)$ is indefinite, then $x_0$ is not an extremum of $f$.
(2) If $x_0$ is an extremum of $f$, then $H(x_0)$ is semi-definite.

## 2.4. The inverse and implicit function theorems

**2.4.1 The geometric series.** Suppose $a$ is a real number and $|a| < 1$. Then

$$\sum_{k=0}^{\infty} a^k = \frac{1}{1-a}.$$

**2.4.2 Contraction mappings.** Let $(M, d)$ be a metric space and $T$ a function from a subset of $M$ to $M$. $T$ is called a *contraction mapping* or a *contraction*, if there is an $\alpha < 1$ such that

$$d(T(x), T(y)) \leq \alpha d(x, y)$$

for all $x, y \in M$.

It is easy to see that every contraction mapping is continuous.

**2.4.3 Fixed points.** Let $M$ be a set and $T$ a function from a subset of $M$ to $M$. A point $x$ in the domain of $T$ for which $T(x) = x$ is called a *fixed point* of $T$.

If $T$ is a contraction, it can have at most one fixed point.

**2.4.4 The contraction mapping theorem.** Let $(M, d)$ be complete metric space and $T : M \to M$ a contraction mapping. Then there is a unique fixed point of $T$.

SKETCH OF PROOF: Uniqueness of the fixed point follows from 2.4.3.

For existence of a fixed point pick $y_0$ and define $y_1 = T(y_0)$, $y_2 = T(y_1)$ and so forth. Then

$$d(y_{m+1}, y_m) \le \alpha d(y_m, y_{m-1}) \le ... \le \alpha^m d(y_1, y_0)$$

and

$$d(y_{m+k}, y_m) \le d(y_1, y_0) \sum_{j=0}^{k-1} \alpha^{m+j} \le \frac{\alpha^m}{1-\alpha} d(y_1, y_0).$$

It follows that $m \mapsto y_m$ is a Cauchy sequence and, using completeness, that it has a limit $y \in M$. Since $T$ is continuous, this limit is a fixed point. □

**2.4.5 The inverse function theorem.** If $f \in C^1(\Omega, \mathbb{R}^n)$ and $f'(x_0)$ is invertible, then there are open sets $U$ and $V$ in $\mathbb{R}^n$ such that $x_0 \in U$, $f'(x)$ is invertible for all $x \in U$, $f(U) = V$, and $f|_U : U \to V$ is bijective. Moreover, the inverse $g$ of $f|_U$ is continuously differentiable on $V$.

Sketch of proof: Let $A = f'(x_0)$.

(a) Since $f'$ is continuous and $\lambda = 1/(2\left\|A^{-1}\right\|) > 0$ there is an open ball $U$ centered at $x_0$ such that $\|f'(x) - f'(x_0)\| < \lambda$ for all $x \in U$. According to 1.1.7, $f'(x)$ is invertible for all such $x$.

(b) Next we show that $f|_U$ is injective. For a fixed $y \in \mathbb{R}^n$ define $\phi : U \to \mathbb{R}^n$ by

$$\phi(x) = x + A^{-1}(y - f(x)).$$

Then $\phi$ is a contraction and therefore has at most one fixed point.

(c) Next we prove that $V = f(U)$ is open. Pick $z \in V$ so that $z = f(x_1)$ for some $x_1 \in U$ and choose $r$ such that $\overline{B(x_1, r)} \subset U$. To show that $B(z, \lambda r) \subset V$ pick a $y \in B(z, \lambda r)$ and consider the associated function $\phi$. Since $\phi$ maps $\overline{B(x_1, r)}$ to itself the contraction mapping theorem 2.4.4 applies and guarantees the existence of a fixed point $x_2$ of $\phi$ and hence $f(x_2) = y$.

(d) Define $g : V \to U$ to be the inverse of $f|_U$. Pick $y, y + v \in V$. Let $x = g(y)$ and $x + u = g(y + v)$. Then $f(x) = y$ and $f(x + u) = y + v$ and hence, letting $B = f'(x)^{-1}$,

$$|g(y+v)-g(y)-Bv| = |B(v-f'(x)u)| = |B(f(x+u)-f(x)-f'(x)u)| \le |u| \, \|B\| \, |r(x+u)| \quad (1)$$

for some function $r$ which is continuous at $x$ and vanishes there. With $y$ we associate, as above, a function $\phi$ and obtain $\phi(x + u) - \phi(x) = u - A^{-1}v$ and, since $\phi$ is a contraction, $|u - A^{-1}v| \le |u|/2$. Hence $|u| \le |v|/\lambda$. This and (1) show that $g$ is differentiable at $y$.

(e) Now we may apply the chain rule to $f(g(y)) = y$ to obtain $g'(y) = f'(g(y))^{-1}$ and conclude that $g'$ is continuous. □

**2.4.6 The implicit function theorem.** Suppose $\Omega$ is an open set in $\mathbb{R}^{n+m}$, $f \in C^1(\Omega, \mathbb{R}^n)$, and $f(\binom{x_0}{y_0}) = 0$ for some $x_0 \in \mathbb{R}^n$ and $y_0 \in \mathbb{R}^m$ such that $\binom{x_0}{y_0} \in \Omega$. Let $f'(\binom{x_0}{y_0}) = (A_1, A_2)$ where $A_1 \in \mathbb{C}^{n\times n}$ and $A_2 \in \mathbb{C}^{n\times m}$. Assume that $A_1$ is invertible.

Then there exist open sets $U \subset \Omega$ and $W \subset \mathbb{R}^m$ and a function $g \in C^1(W, \mathbb{R}^n)$ with the following properties: $y_0 \in W$, $x_0 = g(y_0)$, $\binom{g(y)}{y} \in U$ and $f(\binom{g(y)}{y}) = 0$ for all $y \in W$, and $g'(y_0) = -(A_1)^{-1}A_2$.

Sketch of proof: To simplify notation we will frequently write $(x, y)$ for the vector $\binom{x}{y} \in \mathbb{R}^{n+m}$ when $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$. In particular, we will write $f(x, y)$ in place of $f(\binom{x}{y})$. Moreover, if we write a vector in $\mathbb{R}^{n+m}$ as a pair $(x, y)$, we tacitly assume that $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$. A (rectangular) zero-matrix of any size will be denoted by 0 while an identity matrix of any size will be denoted by $\mathbb{1}$.

(a) Define $F(x,y) = \begin{pmatrix} f(x,y) \\ y \end{pmatrix}$, a function from $\Omega$ to $\mathbb{R}^{n+m}$. If $f' = (P, Q)$ with $P(x,y) \in \mathbb{R}^{n\times n}$ and $Q(x,y) \in \mathbb{R}^{n\times m}$, then $F' = \begin{pmatrix} P & Q \\ 0 & \mathbb{1} \end{pmatrix} \in C^0(\Omega, \mathbb{R}^{n+m})$. Moreover, $F'(x_0,y_0) = \begin{pmatrix} A_1 & A_2 \\ 0 & \mathbb{1} \end{pmatrix}$ and is therefore invertible.

(b) We may now apply the inverse function theorem to $F$. It shows that there are open sets $U, V = F(U) \subset \mathbb{R}^{n+m}$ such that $(x_0, y_0) \in U$ and $G = (F|_U)^{-1} : V \to U$ is bijective and continuously differentiable. Then $W = \{y \in \mathbb{R}^m : (0,y) \in V\}$ is open and contains $y_0$. If $G(0,y) = (x,z) \in U$, then $z = y$ and $x$ is uniquely determined by $y$. Define $g : W \to \mathbb{R}^n : y \mapsto x$.

(c) It is easy to see that $g(y_0) = x_0$. But it remains to show that $g'(y_0) = -(A_1)^{-1}A_2$ and $g \in C^1(W, \mathbb{R}^n)$. To see this note first that, by the inverse function theorem, $F'(x,y)$ is invertible for every $(x,y) \in U$. This entails that $P(x,y)$ is invertible. Thus

$$G' = (F' \circ G)^{-1} = \begin{pmatrix} (P\circ G)^{-1} & -(P\circ G)^{-1}(Q\circ G) \\ 0 & \mathbb{1} \end{pmatrix}.$$

Now, using $G(0,y) = (g(y), y)$, note that $g'(y) = -P(g(y),y)^{-1}Q(g(y),y)$. □

## 2.5. Extrema under constraints

**2.5.1 Extrema under constraints.** Suppose $h \in C^1(\Omega, \mathbb{R})$. Instead of looking for extrema of $h$ in $\Omega$ we will now consider the problem of finding extrema of $h$ in certain non-open subsets of $\Omega$. To be precise, we want to find extrema of $h$ among those points $x$ in $\Omega$ which also satisfy the constraints $f(x) = 0$ where $f \in C^1(\Omega, \mathbb{R}^m)$.

**2.5.2 An example.** Find the points closest to the origin on the parabola $x + y^2 = 3$. Here the function $h$ is given by $h(x,y) = \sqrt{x^2+y^2}$. Since the distance has a minimum if and only if its square has a minimum, we may choose, more simply, $h(x,y) = x^2 + y^2$. The constraint is given by $f(x,y) = x + y^2 - 3$. In this case, any point $(x,y)$ satisfying the constraint satisfies $x = 3 - y^2$. Hence the square of the distance of a point $(x,y)$ on the parabola to the origin is $(3-y^2)^2 + y^2 = y^4 - 5y^2 + 9$. For a minimum we need $4y^3 - 10y = 0$ which gives critical points at $(3,0)$ and $(1, \pm\sqrt{10})/2$. The latter are the minima.

While things are not always so easy, this example gives us an important hint, namely that it is useful to solve the equation $f(x,y) = 0$ for one of the variables, say $x$. This gives us a function $x = g(y)$ so that $f(g(y), y) = 0$ and we want then to minimize $y \mapsto H(y) = h(g(y), y)$. We need to look for critical points of $H$, i.e., for zeros of $H'$. The chain rule gives us

$$h'(g(y),y)\begin{pmatrix} g'(y) \\ 1 \end{pmatrix} = 0.$$

The identity $f(g(y), y) = 0$ gives, in addition,

$$f'(g(y),y)\begin{pmatrix} g'(y) \\ 1 \end{pmatrix} = 0.$$

Taking these equations simultaneously shows that the $2\times 2$-matrix $\begin{pmatrix} h'(g(y),y) \\ f'(g(y),y) \end{pmatrix}$ has 0 as an eigenvalue 0 with eigenvector $\begin{pmatrix} g'(y) \\ 1 \end{pmatrix}$. Therefore the rows are linearly dependent, i.e., $(h' + \lambda f')(g(y), y) = 0$ for some suitable $\lambda$. Studying the set where

$$(h' + \lambda f')(x,y) = 0 \tag{2}$$

often allows for some progress without solving the constraint equation explicitly.

Returning to our example we find that equation (2) becomes

$$(2x + \lambda, 2y + 2\lambda y) = (0,0).$$

The first equation gives $x = -\lambda/2$. The second is satisfied for $y = 0$ or $\lambda = -1$. In the former case the constraint gives $x = 3$. In the latter case we obtain first $x = 1/2$ and then, from the constraint, that $y = \pm\sqrt{10}/2$.

**2.5.3 Lagrange's multiplier method.** Let $h \in C^1(\Omega, \mathbb{R})$ and $f \in C^1(\Omega, \mathbb{R}^m)$ where $m < m + k = n$. Assume that $x_0$ is an extremum of the restriction of $h$ to the set of those points $x \in \Omega$ satisfying $f(x) = 0$ (so that, in particular, $f(x_0) = 0$). Furthermore, assume that $f'(x_0)$ has maximal rank $m$. Then there exists a row $\lambda = (\lambda_1, ..., \lambda_m)$ such that

$$(h + \lambda f)'(x_0) = 0.$$

Sketch of proof: After possibly relabelling the independent variables $x_j$ we may assume that $f'(x_0) = (A_1, A_2)$ where $A_1$ is an invertible $m \times m$-matrix and $A_2$ is an $m \times k$-matrix. We also write $x_0 = \binom{\alpha}{\beta}$ where $\alpha \in \mathbb{R}^m$ and $\beta \in \mathbb{R}^k$. Similarly, $h'(x_0) = (b_1, b_2)$ with $b_1^\top \in \mathbb{R}^m$ and $b_2^\top \in \mathbb{R}^k$.

Since $A_1$ is invertible the equation $b_1 + \lambda A_1 = 0$ has a unique solution $\lambda$ (a row with $m$ components). Hence $(h + \lambda f)'(x_0) = (0, b_2 + \lambda A_2)$. We need to show that $b_2 + \lambda A_2 = 0$.

By the implicit function theorem there exists a neighborhood $W$ of $\beta$ and a function $g \in C^1(W, \mathbb{R}^m)$ such that $g(\beta) = \alpha$ and $f(g(w), w) = 0$ for all $w \in W$. The chain rule gives therefore $A_1 g'(\beta) + A_2 = 0$. Multiplying on the left with $\lambda$ gives

$$-b_1 g'(\beta) + \lambda A_2 = 0. \tag{3}$$

According to our assumption the function $w \mapsto H(w) = h(g(w), w)$ has a relative extremum at $\beta$. Hence

$$0 = H'(\beta) = b_1 g'(\beta) + b_2. \tag{4}$$

Combining equations (3) and (4) shows that indeed $b_2 + \lambda A_2 = 0$. ☐

**2.5.4 Example.** Which points on the ellipse given as the intersection of the plane $x + y + 2z = 2$ and the paraboloid $z = x^2 + y^2$ are farthest from and closest to the origin?

CHAPTER 3

# The multi-dimensional Riemann integral

**3.0.1 $n$-cells.** Given $a, b \in \mathbb{R}^n$ such that $a_k \leq b_k$ we call the set

$$I = \{x \in \mathbb{R}^n : a_k \leq x_k \leq b_k \text{ for } k = 1, ..., n\}$$

a closed *$n$-cell*. $I$ is called an open $n$-cell if

$$I = \{x \in \mathbb{R}^n : a_k < x_k < b_k \text{ for } k = 1, ..., n\}.$$

The quantity

$$|I| = \prod_{k=1}^{n} (b_k - a_k)$$

is called the *volume* of $I$.

**3.0.2 Partitions.** The set $P = \{I_1, ..., I_r\}$ of closed $n$-cells $I_k$ is called a *partition* of the closed $n$-cell $I$ if the union of the $I_k$ is $I$ and if the interiors of the $I_k$ are pairwise disjoint.

If $P^* = \{J_1, ..., J_s\}$ is also a partition of $I$ and if for every $J_k$ there is an $I_\ell$ such that $J_k \subset I_\ell$ then $P^*$ is called a *refinement* of $P$.

For any two partitions $P$ and $P'$ of $I$ there is a partition $P^*$ which is a refinement of both $P$ and $P'$. $P^*$ is called a common refinement of $P$ and $P'$.

**3.0.3 Riemann sums.** Suppose $I$ is a closed $n$-cell, $f : I \to \mathbb{R}$ is a bounded function, and $P = \{I_1, ..., I_r\}$ is a partition of $I$. For every $I_k \in P$ define $M_k = \sup\{f(x) : x \in I_k\}$ and $m_k = \inf\{f(x) : x \in I_k\}$. Define

$$U(P, f) = \sum_{k=1}^{r} M_k |I_k| \quad \text{and} \quad L(P, f) = \sum_{k=1}^{r} m_k |I_k|.$$

Suppose that $P$ and $P'$ are partitions of an $n$-cell $I$ and that $P^*$ is a common refinement of $P$ and $P'$. Then

$$L(P, f) \leq L(P^*, f) \leq U(P^*, f) \leq U(P', f).$$

**3.0.4 The Riemann integral.** Let $I$ be a closed $n$-cell and $f : I \to \mathbb{R}$ a bounded function. The numbers

$$\overline{\int_I} f = \inf\{U(P, f) : P \text{ is a partition of } I\}$$

and

$$\underline{\int_I} f = \sup\{L(P, f) : P \text{ is a partition of } I\}$$

are called the upper and lower Riemann integral of $f$ over $I$.

If upper and lower Riemann integral of $f$ over $I$ coincide, then we say that $f$ is Riemann integrable over $I$ and we define

$$\int_I f = \overline{\int_I} f = \underline{\int_I} f,$$

the *Riemann integral* of $f$ over $I$.

**3.0.5 A criterion for integrability.** Suppose $I$ is a closed $n$-cell and $f : I \to \mathbb{R}$ is a bounded function. Then $f$ is Riemann integrable if and only if, for every positive $\varepsilon$, there is a partition $P$ such that

$$U(P, f) - L(P, f) < \varepsilon.$$

**3.0.6 Continuous functions are Riemann integrable.** If $f$ is a continuous real-valued function on the closed $n$-cell $I$, then $f$ is Riemann integrable over $I$.

Sketch of proof: Since $f$ is uniformly continuous on $I$ one may construct an appropriate partition. □

**3.0.7 Sets of measure zero.** A set $E \subset \mathbb{R}^n$ is said to have measure zero if, for every $\varepsilon > 0$, there are countably many open $n$-cells $U_j$, $j \in \mathbb{N}$, such that $E \subset \bigcup_{j=1}^{\infty} U_j$ and $\sum_{j=1}^{\infty} |U_j| < \varepsilon$.

Any set with countably many elements has measure zero. Moreover, if each of the countably many sets $E_n$, $n \in \mathbb{N}$, has measure zero than so does the set $\bigcup_{n=1}^{\infty} E_n$.

**3.0.8 Examples.** Let $I = [a_1, b_1] \times ... \times [a_n, b_n]$ be an $n$-cell and fix $j \in \{1, ..., n\}$. The sets $\{x \in I : x_j = a_j\}$ and $\{x \in I : x_j = b_j\}$ are called *faces* of the cell. Each face has measure zero.

Let $E$ be a closed $(n-1)$-cell and $f$ a continuous, real-valued function on $E$. Then the graph of $f$, i.e., the set $\{\left(\begin{smallmatrix} x \\ f(x) \end{smallmatrix}\right) : x \in E\} \subset \mathbb{R}^n$, has measure zero. To see this let $\varepsilon > 0$ be given and let $\{R_1, ..., R_k\}$ be a collection of intervals of length $\epsilon$ partitioning a closed interval containing $f(E)$. If $\{E_1, .., E_N\}$ is a partition of $E$ so that, for $j = 1, ..., N$, we have $M_j - m_j < \varepsilon$, then the graph of $f|_{E_j}$ lies in at most two of the sets $E_j \times R_k$. From this the conclusion follows.

**3.0.9 Oscillation.** Suppose $E \subset \mathbb{R}^n$ and $f$ a bounded function from $E$ to $\mathbb{R}$. For each $x_0 \in E$ and $\delta > 0$ define

$$M(x_0, \delta) = \sup\{f(x) : x \in E, |x - x_0| < \delta\} \text{ and } m(x_0, \delta) = \inf\{f(x) : x \in E, |x - x_0| < \delta\}.$$

Then

$$\operatorname{osc}(x_0) = \lim_{\delta \to 0} (M(x_0, \delta) - m(x_0, \delta))$$

exists for all $x_0 \in E$. It is called the *oscillation* of $f$ at $x_0$.

The function $f$ is continuous at $x_0$ if and only if $\operatorname{osc}(x_0) = 0$.

**3.0.10 Riemann integrable functions are nearly continuous.** A bounded real-valued function $f$ defined on a closed $n$-cell is Riemann integrable if and only if the set of points where it is not continuous has measure zero.

Sketch of proof: Denote the domain of $f$ by $I$ and define $B_k = \{x \in I : \operatorname{osc}(x) \geq 1/k\}$, $B = \bigcup_{k=1}^{\infty} B_k$, and $C = \sup\{|f(x)| : x \in I\}$.

Assume that $f$ is integrable. Let $P = \{I_1, ..., I_N\}$ be a partition of $I$ such that $U(P, f) - L(P, f) < \varepsilon/k$ and assume that $\{I_1, ... I_\ell\}$ is the set of those cells in $P$ whose interiors intersect $B_k$. Then

$$\frac{1}{k} \sum_{j=1}^{\ell} |I_j| \leq \sum_{j=1}^{\ell} |I_j|(M_j - m_j) \leq \sum_{j=1}^{N} |I_j|(M_j - m_j) = U(P, f) - L(P, f) < \varepsilon/k.$$

Since the faces of the $n$-cells have measure zero it follows that $B_k$, and hence $B$, has measure zero.

Assume $B$ has measure zero. Then there are open cells $U_j$, $j \in \mathbb{N}$, such that $B \subset \bigcup_{j=1}^{\infty} U_j$ and $\sum_{j=1}^{\infty} |U_j| < \varepsilon$. Moreover, if $x_0 \in I \setminus B$, then $f$ is continuous at $x_0$. Hence there is an open $n$-cell $V_{x_0}$ such that $\sup\{f(x) : x \in \overline{V_{x_0}}\} - \inf\{f(x) : x \in \overline{V_{x_0}}\} < \varepsilon$. Since $I$ is compact, it is covered by a finite collection of the $U_j$ and the $V_{x_k}$. There is a partition $P = \{I_1, ..., I_r\}$ of $I$ such that each of the first $\ell$ cells is in one of the $\overline{U_j}$ while each of the remaining ones is in one of the $\overline{V_{x_k}}$. For this partition we have $U(P, f) - L(P, f) < 2C\varepsilon + |I|\varepsilon$. □

**3.0.11 Integrals over bounded sets.** Let $E$ be a bounded subset of $\mathbb{R}^n$ and $f$ a bounded function from $E$ to $\mathbb{R}$. If $I$ is a closed $n$-cell containing $E$ we extend $f$ to a function defined on $I$ by setting it equal to 0 on $I \setminus E$. Denoting the extension by $f_e$ we define $\int_E f = \int_I f_e$, if the latter exists. While $E$ is contained in many $n$-cells, this definition does not depend on the choice of such a cell. We then say that $f$ is Riemann integrable over $E$.

In particular, if the boundary $\partial E$ of $E$ has measure zero and $f : E \to \mathbb{R}$ is continuous, then $f$ is Riemann integrable over $E$.

**3.0.12 Properties of the Riemann integral.** Let $E$ be bounded a subset of $\mathbb{R}^n$ and assume that $f$ and $g$ are Riemann integrable over $E$. Then the following statements hold:

(1) $\int_E (f + g) = \int_E f + \int_E g$.
(2) $\int_E (cf) = c \int_E f$ whenever $c \in \mathbb{R}$.
(3) $fg$ is Riemann integrable over $E$.
(4) If $f \geq 0$ then $\int_E f \geq 0$.
(5) $|f|$ is Riemann integrable over $E$ and $|\int_E f| \leq \int_E |f|$.
(6) If $E_1 \cup E_2 = E$ and $E_1 \cap E_2 = \emptyset$, then $\int_E f = \int_{E_1} f + \int_{E_2} f$.

**3.0.13 Iterated integrals.** Suppose $A \subset \mathbb{R}^n$ is a closed $n$-cell, $B \subset \mathbb{R}^m$ is a closed $m$-cell, and $f : A \times B \to \mathbb{R}$ is Riemann integrable over $A \times B$. Then $x \mapsto \varphi(x) = \underline{\int_B} f(x, \cdot)$ and $x \mapsto \psi(x) = \overline{\int_B} f(x, \cdot)$ are Riemann integrable. Moreover,

$$\int_{A \times B} f = \int_A \varphi = \int_A \psi. \tag{5}$$

Sketch of proof: Let $P = (A_1, ..., A_N)$ be a partition of $A$ and $Q = (B_1, ..., B_M)$ a partition of $B$ giving rise to a partition $R$ of $A \times B$ consisting of the cells $A_j \times B_k$. If $m_k(x) = \inf\{f(x, y) : y \in B_k\}$ we get $\varphi(x) \geq \sum_{k=1}^M m_k(x)|B_k|$. If $x \in A_j$ we have $m_k(x) \geq m_{j,k} = \inf\{f(x, y) : x \in A_j, y \in B_k\}$ and hence

$$\inf\{\varphi(x) : x \in A_j\} \geq \sum_{k=1}^M m_{j,k}|B_k|.$$

Thus

$$L(R, f) = \sum_{j=1}^N \sum_{k=1}^M m_{j,k}|A_j||B_k| \leq L(P, \varphi) \leq U(P, \varphi).$$

Similarly, $U(R, f) \geq U(P, \psi) \geq L(P, \psi)$. Also, since $\varphi \leq \psi$ and $f$ is Riemann integrable, we get that both $\varphi$ and $\psi$ are Riemann integrable and that (5) holds. □

**3.0.14 Differentiating an integral.** Let $f : [x_0, x_1] \times [y_0, y_1] \to \mathbb{R}$ be a continuous function such that $D_2 f$ is also continuous. Define $F(y) = \int_{[a,b]} f(\cdot, y)$ for $y \in [y_0, y_1]$. Then

$$F'(y) = \int_{[a,b]} (D_2 f)(\cdot, y).$$

Sketch of proof: For $y > y_0$ we have $F(y) - F(y_0) = \int_{[a,b]} \int_{[y_0,y]} (D_2 f)(x,u) du dx$. Hence, given $\varepsilon > 0$,

$$\left| \frac{F(y) - F(y_0)}{y - y_0} - \int_{[a,b]} (D_2 f)(\cdot, y) \right| < \varepsilon(b-a)$$

when $|y - y_0|$ is sufficiently small. This uses that $D_2 f$ is, in fact, uniformly continuous. $\square$

CHAPTER 4

# Integration of differential forms

Recall that $\Omega$ always denotes an open subset of $\mathbb{R}^n$.

## 4.1. Integration along paths

**4.1.1 Smooth paths.** A smooth path in $\Omega$ is a continuously differentiable function from $Q^1 = [0, 1]$ to $\Omega$.

**4.1.2 Integration along a smooth path.** Given a smooth path $\gamma$ in $\Omega$, we may integrate a list $\omega = (\omega_1, ..., \omega_n)$ of continuous real-valued functions defined on $\Omega$ along $\gamma$ by defining

$$\int_\gamma \omega = \int_{[0,1]} \sum_{j=1}^{n} \omega_j(\gamma(t))\gamma_j'(t)dt.$$

For example, if $\omega(x) = (x_2, x_1, 0)$ and $\gamma(t) = (2t^3, 3t, t^2)^\top$, then

$$\int_\gamma \omega = 6.$$

If $\omega(x) = (0, x_1)$ and $\gamma(t) = (a\cos(2\pi t), b\sin(2\pi t))^\top$, then

$$\int_\gamma \omega = \pi ab.$$

## 4.2. Integration over surfaces

**4.2.1 Smooth surfaces.** A smooth surface in $\Omega$ is a continuously differentiable function from $Q^2 = \{(x, y) : 0 \le x, 0 \le y, x + y \le 1\}$ to $\Omega$.

**4.2.2 Integration over a smooth surface.** Given a smooth surface $\phi$ in $\Omega$ we define, for $\alpha = (\alpha_1, \alpha_2) \in \{1, ..., n\}^2$, the Jacobian determinants

$$J(\phi, \alpha) = \det \begin{pmatrix} D_1\phi_{\alpha_1} & D_2\phi_{\alpha_1} \\ D_1\phi_{\alpha_2} & D_2\phi_{\alpha_2} \end{pmatrix}.$$

These are continuous functions from $Q^2$ to $\mathbb{R}$. Note that $J(\phi, (k, k)) = 0$ and $J(\phi, (k, \ell)) = -J(\phi, (\ell, k))$ for $\ell, k = 1, ..., n$.

Now we define the integral of an array of continuous real-valued functions $\omega_{j,k}$, $j, k = 1, ..., n$, defined on $\Omega$ by

$$\int_\phi \omega = \int_{Q^2} \sum_{j=1}^{n} \sum_{k=1}^{n} \omega_{j,k} \circ \phi \ J(\phi, (j, k)) = \int_{Q^2} \sum_{1 \le j < k \le n} (\omega_{j,k} - \omega_{k,j}) \circ \phi \ J(\phi, (j, k)).$$

For example, if

$$\omega(x) = \begin{pmatrix} x_1 & x_3 & 0 \\ 0 & x_2 & x_1 \\ x_2 & 0 & x_3 \end{pmatrix} \quad \text{and} \quad \phi(s,t) = \begin{pmatrix} \sin(\pi s)\cos(2\pi t) \\ \sin(\pi s)\sin(2\pi t) \\ \cos(\pi s) \end{pmatrix},$$

then

$$\int_\gamma \omega = 2\pi.$$

## 4.3. The general case

**4.3.1 The standard $k$-simplex.** Suppose $k \in \mathbb{N}$. Then the set $Q^k = \{x \in \mathbb{R}^k : 0 \le x_j, x_1 + ... + x_k \le 1\}$ is called the *standard $k$-simplex* in $\mathbb{R}^k$. We also define the standard 0-simplex to be $Q^0 = \mathbb{R}^0 = \{0\}$.

**4.3.2 $k$-surfaces.** If $k \in \mathbb{N}$ we define a $k$-surface in $\Omega$ to be a function $\phi \in C^1(Q^k, \Omega)$. A 0-surface in $\Omega$ is a point in $\Omega$. $Q^k$ is called the parameter domain of $\phi$.

**4.3.3 Multi-indices.** We have introduced the concept of a multi-index in 2.2.9. In this chapter we need a slightly different kind of object. Henceforth, given $n, k \in \mathbb{N}$, we call a list of $k$ elements from $\{1, ..., n\}$ a $k$-index of type $n$. The set of $k$-indices of type $n$ is denoted by $N_k^n$. It has precisely $n^k$ elements.

A $k$-index $\beta$ is called a *basic* $k$-index if $\beta_1 < \beta_2 < ... < \beta_k$. There are $\binom{n}{k}$ basic $k$-indices in $N_k^n$ if $k \le n$ and none if $k > n$. The set of all basic $k$-indices of type $n$ is denoted by $I_k^n$.

**4.3.4 Jacobians.** Given $n$ continuously differentiable real-valued functions $\phi_1, ..., \phi_n$ defined on $Q^k$ and a $k$-index $\alpha = (\alpha_1, ..., \alpha_k)$ of type $n$ we define the Jacobian

$$J(\phi, \alpha) = \det \begin{pmatrix} D_1\phi_{\alpha_1} & \cdots & D_k\phi_{\alpha_1} \\ \vdots & & \vdots \\ D_1\phi_{\alpha_k} & \cdots & D_k\phi_{\alpha_k} \end{pmatrix}$$

which is a continuous function on $Q^k$.

**4.3.5 The vector space $W_k^n(\Omega)$.** The real vector space of functions $\omega : N_k^n \to C^0(\Omega, \mathbb{R})$ is denoted by $W_k^n(\Omega)$. In other words, an element of $W_k^n(\Omega)$ assigns to each $k$-index $\alpha$ of type $n$ a function $\omega_\alpha \in C^0(\Omega, \mathbb{R})$. We also define $W_0^n(\Omega) = C^0(\Omega, \mathbb{R})$.

The set of functions from $N_k^n$ to $\mathbb{R}$ is an $n^k$-dimensional vector space. It has a standard basis $e_\alpha$, $\alpha \in N_k^n$, defined by $\alpha(\beta) = 1$ if $\alpha = \beta$ and $\alpha(\beta) = 0$ if $\alpha \ne \beta$.[1]

Treating the $e_\alpha$ as (constant) functions on $\Omega$ we may now represent an element of $W_k^n(\Omega)$ by $\omega = \sum_{\alpha \in N_k^n} \omega_\alpha e_\alpha$.

**4.3.6 Integration over a $k$-surface.** Suppose $\phi$ is a $k$-surface in $\Omega$ and $\omega \in W_k^n(\Omega)$. Then we define

$$\int_\phi \omega = \int_{Q^k} \sum_{\alpha \in N_k^n} \omega_\alpha \circ \phi \; J(\phi, \alpha)$$

if $k > 0$. If $k = 0$ we set $\int_\phi \omega = \omega(\phi(0))$.

[1] This construction is exactly parallel to the construction of the standard base in $\mathbb{R}^n$ if one considers it as the space of real-valued functions on $\{1, ..., n\}$.

**4.3.7 Differential $k$-forms.** We call two functions $\omega_1, \omega_2 \in W_k^n(\Omega)$ equivalent, if $\int_\phi \omega_1 = \int_\phi \omega_2$ for all $k$-surfaces $\phi$ in $\Omega$. This relation is an equivalence relation and will be denoted by $\omega_1 \sim \omega_2$.

DEFINITION. Suppose $k \in \mathbb{N}$. A *differential form* $\omega$ of order $k$ in $\Omega$, or simply a *$k$-form* in $\Omega$, is an equivalence class of functions $\tilde{\omega} \in W_k^n(\Omega)$. A differential form of order 0 in $\Omega$, or simply a 0-form in $\Omega$ is a continuous real-valued function on $\Omega$.

A $k$-form in $\Omega$ assigns to each $k$-surface in $\Omega$ a real number. If $\tilde{\omega}$ is a representative of $\omega$ and $\phi$ is a $k$-surface, we shall write $\omega(\phi) = \int_\phi \tilde{\omega}$ or $\int_\phi \omega = \int_\phi \tilde{\omega}$.

The differential forms of order $k$ form a real vector space.

**4.3.8 Elementary properties of $k$-forms.** If $\alpha$ is a $k$-index and $\pi$ is a permutation of $\{1, ..., k\}$ we define $\alpha_\pi = (\alpha_{\pi(1)}, ..., \alpha_{\pi(k)})$.

Suppose $b \in C^0(\Omega, \mathbb{R})$, $\alpha \in N_k^n$, and $\pi$ is a transposition of $\{1, ..., k\}$. Then $be_\alpha \sim -be_{\alpha_\pi}$. If $\alpha_j = \alpha_\ell$ for some $j \neq \ell$ then $be_\alpha \sim 0$. If $\tilde{\omega} \in W_k^n(\Omega)$ with $k > n$, then $\tilde{\omega} \sim 0$.

Suppose $\omega_1$ and $\omega_2$ are $k$-forms in $\Omega$ and $\phi$ is a $k$-surface in $\Omega$. Let $c$ be real number. Then the following statements are true.

(1) $\int_\phi (\omega_1 + \omega_2) = \int_\phi \omega_1 + \int_\phi \omega_2$.

(2) $\int_\phi c\omega = c \int_\phi \omega$.

**4.3.9 Basic representatives of $k$-forms.** Suppose $k \leq n$. If the entries of the $k$-index $\alpha$ are pairwise distinct, there is a permutation $\pi$ of $\{1, ..., k\}$ such that $\beta = \alpha_\pi$ is a basic $k$-index. Since $e_\alpha \sim (-1)^\pi e_\beta$ we can, for any $k$-form $\omega$, choose a representative $\tilde{\omega}$ such that $\tilde{\omega}_\alpha = 0$ unless $\alpha$ is a basic $k$-index. Such a representative is called a basic representative of $\omega$. If $k > n$ and $\omega$ is a $k$-form, then $\omega = 0$. In this case the basic representative has $\omega_\alpha = 0$ for all $\alpha$.

THEOREM. Suppose $k \leq n$. Let $\omega$ be a $k$-form in $\Omega$ and $\tilde{\omega}$ a basic representative of $\omega$. Then $\omega = 0$ if and only if $\tilde{\omega}_\alpha = 0$ for every $\alpha \in N_k^n$. In other words, the equivalence class of representatives of a $k$-form contains precisely one basic representative.

SKETCH OF PROOF. Suppose, by way of contradiction, that $\tilde{\omega}_\alpha(x_0) > 0$ for some $x_0 \in \Omega$ and some basic $k$-index $\alpha$. Construct a $k$-surface $\phi$ in a sufficiently small neighborhood of $x_0$ such that $J(\phi, \alpha) = 1$ and $J(\phi, \beta) = 0$ for all basic $k$-indices $\beta \neq \alpha$. Then $\int_\phi \omega = \int_{Q^k} \tilde{\omega}_\alpha \circ \phi \; J(\phi, \alpha) > 0$, the desired contradiction. □

**4.3.10 The wedge product of differential forms.** Suppose $p, q \in \mathbb{N}$, $\omega$ is a $p$-form, and $\lambda$ is a $q$-form in $\Omega$. Let $\tilde{\omega}$ be the basic representative of $\omega$ and $\tilde{\lambda}$ the basic representative of $\lambda$, i.e,[2]

$$\tilde{\omega} = \sum_{\alpha \in I_p^n} \tilde{\omega}_\alpha e_\alpha \quad \text{and} \quad \tilde{\lambda} = \sum_{\beta \in I_q^n} \lambda_\beta e_\beta.$$

Then we define the $(p+q)$-form $\omega \wedge \lambda$ to be the form represented by

$$\sum_{\alpha \in I_p^n} \sum_{\beta \in I_q^n} \tilde{\omega}_\alpha \tilde{\lambda}_\beta e_{(\alpha,\beta)}.$$

We also define the product of 0-forms with $k$-forms: If $\omega$ and $\lambda$ are both 0-forms then $\omega \wedge \lambda$ is the 0-form given by the product of the continuous functions $\omega$ and $\lambda$. If $\omega$ is a 0-form and

[2] Customarily, an expression of the form $\sum_{a \in A} b(a)$ is set equal to 0 if $A = \emptyset$.

$\lambda$ is a $q$-form, then

$$\omega \wedge \lambda = \sum_{\beta \in I_q^n} \omega \lambda_\beta e_\beta.$$

Similarly, if $\omega$ is a $p$-form and $\lambda$ is a 0-form, then

$$\omega \wedge \lambda = \sum_{\alpha \in I_p^n} \omega_\alpha \lambda e_\alpha.$$

Note that $\omega \wedge \lambda = 0$ if $p + q > n$.

The wedge product of differential forms is associative and left and right distributive but not commutative. In fact, $\omega \wedge \lambda = (-1)^{pq}\, \lambda \wedge \omega$.

**4.3.11 Differentiation of differential forms.** We say a differential form of order $k$ is of class $C^r$ if the functions $\omega_\alpha$ in its basic representation $\sum_{\alpha \in I_k^n} \omega_\alpha e_\alpha$ are elements of $C^r(\Omega, \mathbb{R})$. A form of class $C^r$ is called an $r$ times continuously *differentiable* form.

We will now define an operator $d$ which maps $k$-forms of class $C^r$ to $(k+1)$-forms of class $C^{r-1}$.

If $f$ is a 0-form of class $C^1$ in $\Omega$ we define $df$ to be the 1-form with basic representative

$$\sum_{j=1}^{n} (D_j f) e_j$$

using that $I_1^n = \{1, ..., n\}$.

If $k \geq 1$ and $\omega$ is a $k$-form of class $C^1$ with basic representative

$$\tilde{\omega} = \sum_{\alpha \in I_k^n} \tilde{\omega}_\alpha e_\alpha,$$

we define $d\omega$ to be the $(k+1)$-form represented by

$$\sum_{\alpha \in I_k^n} \sum_{j=1}^{n} (D_j \tilde{\omega}_\alpha) e_{(j,\alpha)}.$$

**4.3.12 Examples.** The following are important examples.

(1) Consider $x \mapsto x_j$ as a 0-form. Then $dx_j := d(x \mapsto x_j)$ has representative $e_j$. It is therefore customary to write $dx_j$ for $e_j$ and representatives of general 1-forms in $\Omega$ may be written as $\sum_{j=1}^{n} \omega_j dx_j$, where the $\omega_j$ are continuous real-valued functions on $\Omega$.
(2) $dx_j \wedge dx_\ell$ is represented by $e_{j,\ell}$.
(3) $d^2 x_j = d(d(x \mapsto x_j) = 0$.
(4) Let $\phi$ be a 1-surface and $f$ a 0-form of class $C^1$. Then

$$\int_\phi df = \int_{[0,1]} \sum_{j=1}^{n} (D_j f)(\phi)\phi_j' = \int_{[0,1]} (f \circ \phi)' = f(\phi(1)) - f(\phi(0)).$$

(5) Let $\omega$ be the 1-form with basic representative $x_p e_q$, $1 \leq p, q \leq n$. Then $d\omega$ is represented by $e_{p,q} = dx_p \wedge dx_q$. In particular, $d\omega = 0$ if $p = q$.

**4.3.13 Differentiation rules.** Suppose $\omega$ is a differentiable $p$-form and $\lambda$ is a differentiable $q$-form in $\Omega$. Then the following statements hold:

(1) If $p = q$, then $d(\omega + \lambda) = d\omega + d\lambda$.
(2) If $c \in \mathbb{R}$, then $d(c\omega) = cd\omega$.
(3) $d(\omega \wedge \lambda) = (d\omega) \wedge \lambda + (-1)^p \omega \wedge (d\lambda)$.

Moreover, $d^2 = 0$ on twice continuously differentiable forms.

**4.3.14 Change of variables.** Suppose $T \in C^1(\Omega, \Omega')$ where $\Omega'$ is an open set in $\mathbb{R}^m$. If $k \in \mathbb{N}$ and $\omega$ is a $k$-form in $\Omega'$ with basic representative

$$\tilde{\omega} = \sum_{\alpha \in I_k^m} \tilde{\omega}_\alpha e_\alpha^{(m)}$$

we define a $k$-form $\omega_T$ in $\Omega$ by setting

$$\tilde{\omega}_T = \sum_{\alpha \in I_k^m} (\tilde{\omega}_\alpha \circ T) dT_{\alpha_1} \wedge \ldots \wedge dT_{\alpha_k} = \sum_{\beta \in I_k^n} \Big( \sum_{\alpha \in I_k^m} (\tilde{\omega}_\alpha \circ T) t_{\alpha,\beta} \Big) e_\beta^{(n)}$$

where

$$t_{\alpha,\beta} = \det \begin{pmatrix} D_{\beta_1} T_{\alpha_1} & \cdots & D_{\beta_k} T_{\alpha_1} \\ \vdots & & \vdots \\ D_{\beta_1} T_{\alpha_k} & \cdots & D_{\beta_k} T_{\alpha_k} \end{pmatrix}.$$

If $\omega$ is a 0-form we set $\omega_T = \omega \circ T$.

Example: Suppose $n = 2$, $m = 3$, $k = 2$, $T(x_1, x_2) = (x_1^2 + x_2^2, x_1 x_2, x_2)^\top$, and $\omega(y) = y_2 y_3 e_{1,2}^{(3)} + y_1\ e_{2,3}^{(3)}$. Then $\omega_T = (2x_1^3 x_2^2 - 2x_1 x_2^4 + x_1^2 x_2 + x_2^3) e_{1,2}^{(2)}$.

**4.3.15 Basic properties of variable changes.** Let $\Omega$ and $\Omega'$ be open sets in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Suppose that $T \in C^1(\Omega, \Omega')$ and that $\omega$ is a $p$-form and $\lambda$ a $q$-form in $\Omega'$. Then

(1) If $p = q$ then $(\omega + \lambda)_T = \omega_T + \lambda_T$.
(2) $(\omega \wedge \lambda)_T = \omega_T \wedge \lambda_T$.
(3) If $\omega$ is of class $C^1$ and $T \in C^2(\Omega, \Omega')$ then $\omega_T$ is of class $C^1$ and $d(\omega_T) = (d\omega)_T$.

**4.3.16 Compositions of variable changes.** Suppose $\Omega$, $\Omega'$, and $\Omega''$ are open sets in $\mathbb{R}^n$, $\mathbb{R}^m$, and $\mathbb{R}^p$, respectively. Let $T \in C^1(\Omega, \Omega')$ and $S \in C^1(\Omega', \Omega'')$ and $\omega$ a $k$-form in $\Omega''$. Then $ST = S \circ T \in C^1(\Omega, \Omega'')$ and

$$\omega_{ST} = (\omega_S)_T.$$

Sketch of proof: Show this first for the case when $k = 0$ and $k = 1$. Then use 4.3.15. □

**4.3.17 Variable changes and integration.** Suppose $\Omega$ and $\Omega'$ are open sets in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. If $\omega$ is a $k$-form in $\Omega'$, $\phi$ is a $k$-surface in $\Omega$ and $T \in C^1(\Omega, \Omega')$, then $\omega_T$ is a $k$-form in $\Omega$, $T \circ \phi$ is a $k$-surface in $\Omega'$ and

$$\int_{T \circ \phi} \omega = \int_\phi \omega_T.$$

Sketch of proof: First show that this true when $k = n$ and $\phi$ is the identity on $Q^k$. Then note that

$$\int_{T \circ \phi} \omega = \int_{\mathbb{1}} \omega_{T \circ \phi} = \int_{\mathbb{1}} (\omega_T)_\phi = \int_\phi \omega_T.$$

□

## 4.4. Stokes' theorem

**4.4.1 Chains.** Let $\mathcal{S}_k(\Omega)$ denote the set of $k$-surfaces in $\Omega$. A $k$-chain in $\Omega$ is a function $f:\mathcal{S}_k(\Omega)\to\mathbb{Z}$ such that $f(s)=0$ for all but finitely many $s\in\mathcal{S}_k(\Omega)$. We define the sum of two $k$-chains $f$ and $g$ by $(f+g)(s)=f(s)+g(s)$ and an integer multiple of a $k$-chain by $(rf)(s)=rf(s)$ when $r\in\mathbb{Z}$. Then $f+g$ and $rf$ are again $k$-chains in $\Omega$.

Defining $f(s_0)=1$ and $f(s)=0$ for $s\neq s_0$ shows that we may consider a $k$-surface as a $k$-chain. A function $f$ defined this way will be denoted by $[s_0]$. We may now represent $k$-chains as $n_1[s_1]+...+n_\ell[s_\ell]$ with integers $n_j$ and $k$-surfaces $s_j$, $j=1,...,\ell$.

The set of $k$-chains $\mathcal{S}_k(\Omega)\to\mathbb{Z}$ is denoted by $\mathcal{C}_k(\Omega)$.

A $k$-chain $n_1[s_1]+...+n_\ell[s_\ell]$ is of class $C^r$, if each of the $k$-surfaces $s_1$, ..., $s_\ell$ is of class $C^r$ for some $r$ in $\mathbb{N}$.

**4.4.2 Boundaries.** Suppose $\ell\le k$. The points $e_0^{(\ell)}=0\in\mathbb{R}^\ell$ and $e_j^{(\ell)}$, $j=1,...,\ell$, are the vertices of $Q^\ell$. Let $p_0$, $p_1$, ..., $p_\ell$ be points in $\mathbb{R}^k$. The $\ell$-surface

$$Q^\ell\to\mathbb{R}^k:(u_1,...,u_\ell)\mapsto p_0+\sum_{j=1}^{\ell}u_j(p_j-p_0)$$

is called an *affine $\ell$-simplex* in $\mathbb{R}^k$ which we denote by $\langle p_0,...,p_\ell\rangle$. Note that $p_j$ is the image of $e_j^{(\ell)}$. For the associated chain we will write $[p_0,...,p_\ell]$ instead of $[\langle p_0,...,p_\ell\rangle]$.

An affine $\ell$-simplex $\langle p_0,...,p_\ell\rangle$ has $\ell+1$ *faces* $\langle p_0,...,\not{p_j},...,p_\ell\rangle$, $j=0,...,\ell$. Each of these faces is an affine $(\ell-1)$-simplex and the chain

$$\sum_{j=0}^{k}(-1)^j[p_0,...,\not{p_j},...,p_\ell]$$

is called the *boundary* of $[p_0,...,p_\ell]$.

More generally, if $\phi$ is a $k$-surface $\in\mathbb{R}^n$, we define

$$\partial[\phi]=\sum_{j=0}^{k}(-1)^j[\phi\circ\langle e_0^{(k)},...,\not{e_j^{(k)}},...,e_k^{(k)}\rangle]$$

noting that $\langle e_0^{(k)},...,\not{e_j^{(k)}},...,e_k^{(k)}\rangle$ maps $Q^{k-1}$ to $Q^k$.

Finally, if $\psi=\sum_{j=1}^{\ell}n_j[\phi_j]$ is any chain in $\mathcal{C}_k(\Omega)$ we define its boundary as

$$\partial\psi=\sum_{j=1}^{\ell}n_j\partial[\phi_j].$$

**4.4.3 Examples.** Suppose $k=3$ and $\ell=2$. Let $p_0=(0,0,0)^\top$, $p_1=(1,1,1)^\top$, and $p_2=(0,1,1)^\top$. The 2-surface $\langle p_0,p_1,p_2\rangle$ represents a triangle in $\mathbb{R}^3$. The boundary $\partial[p_0,p_1,p_2]$ consists of the three edges of the triangle. Also, $\partial(\partial[p_0,p_1,p_2])=0$.

Find the boundary of the 2-surface

$$\phi(s,t)=\begin{pmatrix}\sin(\pi s)\cos(2\pi t)\\ \sin(\pi s)\sin(2\pi t)\\ \cos(\pi s)\end{pmatrix},\quad (s,t)\in Q^2.$$

Plot the surface and its boundary.

**4.4.4 $\partial^2=0$.** For any $k$-chain $\sigma$ we have $\partial^2\sigma=0$.

**4.4.5 Integration over chains.** Let $\gamma$ be a $k$-chain in $\Omega$ and $\omega$ a $k$-form in $\Omega$. Since $\gamma = \sum_{j=1}^{\ell} n_j[\phi_j]$ with $k$-surfaces $\phi_j$ we define

$$\omega(\gamma) = \int_\gamma \omega = \sum_{j=1}^{\ell} n_j \int_{\phi_j} \omega.$$

**4.4.6 The fundamental theorem of calculus.** Suppose $n = 1$, $\Omega = \mathbb{R}$ and $a, b \in \mathbb{R}$ with $a < b$. Let $\phi$ be the 1-chain $[a, b]$ given by the 1-surface $\langle a, b\rangle$, i.e., by $[0, 1] \to \mathbb{R} : t \mapsto a + (b - a)t$. Also, let $f$ be a continuously differentiable 0-form in $\mathbb{R}$. Then $df$ is a 1-form represented by $f'e_1 = f'dx$ and the following identity, called the fundamental theorem of calculus, holds:

$$\int_\phi df = \int_{\partial\phi} f.$$

Sketch of proof: If we use the letter $\phi$ also for the function $\langle a, b\rangle$ well-known results from calculus show

$$\int_\phi df = \int_{[0,1]} f'(\phi)\phi' = \int_{[0,1]} (f \circ \phi)' = f(\phi(1)) - f(\phi(0))$$

and, since $\partial\phi = [\phi(1)] - [\phi(0)]$

$$\int_{\partial\phi} f = \int_{\phi(1)} f - \int_{\phi(0)} f = f(\phi(1)) - f(\phi(0)).$$

□

**4.4.7 Stokes' theorem.** Let $\Omega$ be an open subset of $\mathbb{R}^n$ and $k \in \mathbb{N}$. If $\phi$ is a $k$-chain of class $C^2$ in $\Omega$ and $\omega$ is a $(k-1)$-form of class $C^1$ in $\Omega$, then

$$\int_\phi d\omega = \int_{\partial\phi} \omega.$$

Sketch of proof: It is sufficient to prove that

$$\int_\phi (D_r f) e_{r,\alpha} = \int_{\partial\phi} f e_\alpha$$

for $r \in \{1, ..., k\}$, $\alpha = (1, ..., \not{r}, ..., k)$, $\phi = [e_0, ..., e_k]$, and a continuously differentiable function $f : Q^k \to \mathbb{R}$. To see this recall that $\langle e_0, ..., e_k\rangle$ is the identity map, which we denote by $\mathbb{1}$, on $Q^k$.

Note that $J(\mathbb{1}, (r, \alpha)) = (-1)^{r+1}$ so that $\int_\phi (D_r f) e_{r,\alpha} = (-1)^{r+1} \int_{Q^k} D_r f$. To evaluate this we use iterated integrals and integrate first over the $r$-th component $t_r$ of $Q^k$ from 0 to $s = 1 - t_1 - ... - \not{t_r} - ... - t_k$. This gives

$$\int_\phi (D_r f) e_{r,\alpha} = (-1)^{r+1} \int_{Q^{k-1}} (f(t_1, ..., t_{r-1}, s, t_{r+1}, ...t_k) - f(t_1, ..., t_{r-1}, 0, t_{r+1}, ...t_k)).$$

Defining $\sigma^{(j)} = \langle e_0, ..., \not{e_j}, ..., e_k\rangle$ the boundary of $\phi$ is given by $\sum_{j=0}^{k} (-1)^j [\sigma^{(j)}]$. Then $J(\sigma^{(j)}, \alpha) = 0$ unless $j = 0$ or $j = r$. In fact, $J(\sigma^{(r)}, \alpha) = 1$ and $J(\sigma^{(0)}, \alpha) = (-1)^{r+1}$. Hence

$$\int_{\partial\phi} f e_\alpha = \sum_{j=0}^{k} (-1)^j \int_{\sigma^{(j)}} f e_\alpha = (-1)^r \int_{Q^{k-1}} (f \circ \sigma^{(r)} - f \circ \sigma^{(0)}).$$

It remains to prove that $\int_{Q^{k-1}}(f(t_1,...,t_{r-1},s,t_{r+1},...t_k) = \int_{Q^{k-1}} f\circ\sigma^{(0)}$ which follows from 4.3.17 for a simple transformation $T : Q^r \to Q^r$. □

**4.4.8 Green's theorem.** Let $n = 2$. If $\phi$ is 2-chain in $\Omega \subset \mathbb{R}^2$ and if $\omega = fe_1 + ge_2$ is a 1-form of class $C^1$, then $d\omega = (D_1 g - D_2 f)e_{1,2}$. Stokes' theorem is called Green's theorem in this case

$$\int_\phi (D_1 g - D_2 f)e_{1,2} = \int_{\partial\phi}(fe_1 + ge_2).$$

In particular, when $f(x_1, x_2) = -x_2/2$ and $g(x_1, x_2) = x_1/2$ we get

$$\int_\phi e_{1,2} = \frac{1}{2}\int_{\partial\phi}(x_1e_2 - x_2e_1).$$

This quantity is called the area of (the range of ) $\phi$. Use Green's theorem to find the area of the triangle with vertices $p_0 = (0, 0)^\top$, $p_1 = (a, 0)^\top$, and $p_2 = (b, c)^\top$ where $a, b, c > 0$.

APPENDIX A

# Vector spaces and linear transformations

## A.1. Vector spaces

**A.1.1 Euclidean vector spaces.** $\mathbb{R}^n$ is the set of all ordered lists of $n$ real numbers. Its elements are called *vectors*, real numbers themselves are sometimes called *scalars*. The entries of a list defining a vector are called *components* or *coordinates*. We will usually think of the lists as columns rather than rows. For typographical reasons we shall often use the notation $(a_1, ...., a_n)^\top$ for the column whose components are $a_1$, ..., $a_n$.

Two elements of $\mathbb{R}^n$ may be *added* componentwise, i.e.,

$$(a_1, ...., a_n)^\top + (b_1, ..., b_n)^\top = (a_1 + b_1, ..., a_n + b_n)^\top.$$

If $\alpha$ is a scalar and $a$ is a vector, we define

$$\alpha(a_1, ...., a_n)^\top = (\alpha a_1, ..., \alpha a_n)^\top.$$

This is called the *scalar multiplication* of $a$ by $\alpha$. With these operations $\mathbb{R}^n$ is a real vector space in the sense of Linear Algebra.

There is also a canonical *inner product* (or *scalar product*) associated with $\mathbb{R}^n$:

$$x \cdot y = \sum_{j=1}^{n} x_j y_j$$

when $x = (x_1, ..., x_n)^\top$ and $x = (y_1, ..., y_n)^\top$.

Equipped with vector addition, scalar multiplication, and inner product as just defined $\mathbb{R}^n$ is called the *euclidean vector space* of dimension $n$.

**A.1.2 Linear combinations.** If $x_1, ..., x_n \in \mathbb{R}^n$ and $\alpha_1, ..., \alpha_n \in \mathbb{R}$, the vector

$$\alpha_1 x_1 + ... + \alpha_n x_n$$

is called a *linear combination* of $x_1, ..., x_n$.

**A.1.3 Linearly independence.** The vectors $x_1, ..., x_n \in \mathbb{R}^n$ are called *linearly independent* if $\alpha_1 x_1 + ... + \alpha_n x_n = 0$ implies that $\alpha_1 = ... = \alpha_n = 0$. Otherwise, they are called *linearly dependent*.

A set is called linearly independent, if any finite number of its elements are linearly independent.

**A.1.4 Subspaces.** A nonempty subset $S$ of $\mathbb{R}^n$ is called a *subspace* of $\mathbb{R}^n$ if $\alpha x + \beta y \in S$ whenever $x, y \in S$ and $\alpha, \beta \in \mathbb{R}$. A subspace is a vector space with respect to the operations of vector addition and scalar multiplication defined in A.1.1.

**A.1.5 Spans.** Let $A$ be a nonempty subset of $\mathbb{R}^n$. The set of all linear combinations of finitely many elements of $A$ is called the *span* of $A$. The span of $A$, denoted by $\operatorname{span} A$, is a subspace of $\mathbb{R}^n$. If $W = \operatorname{span} A$ we also say that $W$ is spanned by $A$ or that $A$ spans $W$.

If $A = \emptyset$ we define span $A = \{0\}$, the trivial vector space. Here we wrote, as is customary, 0 for the vector $(0, ..., 0)^\top$.

**A.1.6 Bases and dimension.** Suppose $V$ is a subspace of $\mathbb{R}^n$. A set $B \subset V$ is called a *basis* of $V$, if it is linearly independent and spans $V$. The empty set is a basis of the trivial vector space $\{0\}$. Every basis of $V$ has the same number of elements. This number is called the dimension of $V$.

We call $(v_1, ..., v_n) \in V^n$ an *ordered basis* of $V$, if $v_1, ..., v_n$ are pairwise distinct and form a basis of $V$.

The vectors $e_1 = (1, 0, ..., 0)^\top$, $e_2 = (0, 1, 0, ..., 0)^\top$, ... $e_n = (0, ..., 0, 1)^\top$ form a basis of $\mathbb{R}^n$. The ordered basis $(e_1, ..., e_n)$ is called the *standard basis* of $\mathbb{R}^n$. Sometimes we may want to emphasize the dimension of the space to which a standard basis element belongs. Then we use $e_j^{(n)}$ instead of $e_j$.

## A.2. Linear operators

**A.2.1 Linear operators.** Let $V$ and $W$ be two vector spaces over $\mathbb{R}$. The function $F : V \to W$ is called a *linear operator* or a *linear transformation*, if

$$F(\alpha x + \beta y) = \alpha F(x) + \beta F(y)$$

for all $\alpha, \beta \in \mathbb{R}$ and all $x, y \in V$.

If $F$ is a linear operator we have $F(0) = 0$ and $F(-x) = -F(x)$.

It is customary to write $Fx$ in place of $F(x)$.

**A.2.2 Kernel and range.** The *kernel* of a linear operator $F : V \to W$ is the set $\ker F = \{x \in V : F(x) = 0\}$. The *range* of a linear transformation $F : V \to W$ is the set $\operatorname{ran} F = F(V) = \{F(x) : x \in V\}$ of all images of $F$.

Kernel and range of $F$ are subspaces of $V$ and $W$, respectively. The dimension of $\ker F$ is called the *nullity* of $F$ while the dimension of $\operatorname{ran} F$ is called the *rank* of $F$.

**A.2.3 The vector space of linear operators.** The set of all linear operators from the vector space $V$ to the vector space $W$ is denoted by $L(V, W)$. We define an addition and a scalar multiplication of linear operators by $(F+G)(x) = F(x)+G(x)$ and $(\alpha F)(x) = \alpha F(x)$ when $F$ and $G$ are linear operators and $\alpha$ a real number. One may then show that $L(V, W)$ is a real vector space.

**A.2.4 The fundamental theorem of Linear Algebra.** Suppose $V$ and $W$ are finite-dimensional vector spaces and $T \in L(V, W)$. Then

$$\dim(\ker T) + \dim(\operatorname{ran} T) = \dim V.$$

This is also known as the *rank-nullity theorem*.

**A.2.5 Compositions of linear operators.** Suppose $U$, $V$, and $W$ are finite-dimensional vector spaces. If $F : U \to V$ and $G : V \to W$ are linear operators we define

$$(G \circ F)(x) = G(F(x))$$

for all $x \in U$. Then $G \circ F$, the composition of $G$ and $F$, is a linear transformation from $U$ to $W$. Note that it makes no sense to define $F \circ G$ unless $W \subset U$.

For simplicity one often writes $GF$ in place of $G \circ F$ and $F^2$ in place of $F \circ F$.

**A.2.6 Matrices and linear operators between euclidean vector spaces.** Let $T$ be a linear operator from $\mathbb{R}^n$ to $\mathbb{R}^m$. Then

$$Te_j^{(n)} = \sum_{\ell=1}^{m} M_{\ell,j} e_\ell^{(m)}$$

where the $M_{\ell,j}$ are appropriate real numbers. These are customarily arranged in a rectangular grid $M$ with $m$ rows and $n$ columns, i.e.,

$$M = \begin{pmatrix} M_{1,1} & \cdots & M_{1,n} \\ \vdots & & \vdots \\ M_{m,1} & \cdots & M_{m,n} \end{pmatrix}.$$

$M$ is called an $m \times n$-*matrix*. The set of all $m \times n$-matrix with real entries is denoted by $\mathbb{R}^{m\times n}$.

Of course, an $m \times n$-matrix $M$ determines a linear operator from $\mathbb{R}^n$ to $\mathbb{R}^m$. Thus, assuming standard bases in both domain and range, it is sensible to identify linear operators from $\mathbb{R}^n$ to $\mathbb{R}^m$ with the corresponding $m \times n$ matrices.

**A.2.7 Matrix algebra.** The operations of addition, scalar multiplication, and composition of linear operators between euclidean vector spaces are reflected in corresponding algebraic operations on matrices. Specifically, addition and scalar multiplication are represented by

$$M + N \begin{pmatrix} M_{1,1} + N_{1,1} & \cdots & M_{1,n} + N_{1,n} \\ \vdots & & \vdots \\ M_{m,1} + N_{m,1} & \cdots & M_{m,n} + N_{m,n} \end{pmatrix} \quad \text{and} \quad \alpha M = \begin{pmatrix} \alpha M_{1,1} & \cdots & \alpha M_{1,n} \\ \vdots & & \vdots \\ \alpha M_{m,1} & \cdots & \alpha M_{m,n} \end{pmatrix}$$

when $M$ and $N$ are $m \times n$ matrices

The composition of linear transformations turns into a multiplication of matrices, if we define the product of an $\ell \times m$-matrix $M$ and an $m \times n$-matrix $N$ by

$$(MN)_{j,k} = \sum_{s=1}^{m} M_{j,s} N_{s,k}, j = 1, ..., \ell, k = 1, ..., n.$$

Note that it is necessary that the number of columns of $M$ equals the number of rows of $N$ in order to form the product $MN$. This reflects the fact that the range of the operator associated with $N$ has to be in the domain of the operator associated with $M$. Thus matrix multiplication is not commutative (but it is associative).

**A.2.8 Distributive laws in matrix algebra.** We have the following distributive laws for matrices $A, B, C$ whenever it makes sense to form the sums and products in question: $(A + B)C = AC + BC$, $A(B + C) = AB + AC$, and $\alpha(AB) = (\alpha A)B = A(\alpha B)$.

**A.2.9 Square matrices.** A matrix is called a *square matrix* if it has as many columns as it has rows. The elements $M_{1,1}$, ..., $M_{n,n}$ of an $n \times n$-matrix are called *diagonal elements* and together they form the *main diagonal* of the matrix. A matrix is called a *diagonal matrix*, if its only non-zero entries are on the main diagonal.

The *identity transformation* $F(x) = x$ defined on an $n$-dimensional vector space is represented by the *identity matrix* $\mathbb{1}$ which is an $n \times n$-matrix all of whose entries are 0 save for the ones on the main diagonal which are 1.

**A.2.10 Inverses.** A linear operator $T$ from $\mathbb{R}^n$ to $\mathbb{R}^n$ as well as the associated matrix is called *invertible*, if it is bijective. Since here domain and co-domain have the same dimension, the rank-nullity theorem guarantees that $T$ is injective if and only if it surjective. Thus $T$ is invertible if and only if $\ker T = \{0\}$.

**A.2.11 Determinants.** We define the determinant of an $n\times n$-matrix recursively. If $n=1$ we set $\det A = A$. If $n>1$ we define the minor $M_{j,k}$ to be the $(n-1)\times(n-1)$-matrix obtained from $A$ by deleting row $j$ and column $k$. Then we define

$$\det A = \sum_{j=1}^{n} (-1)^{j+n} \det M_{j,n} A_{j,n}.$$

The determinant has the following properties: (i) $\det A = 0$ if the rows or the columns are linearly dependent. (ii) If $B$ is obtained by switching two rows or two columns of $A$, then $\det B = -\det A$. (iii) If $B$ is obtained by multiplying a row or a column of $A$ by the number $c$, then $\det B = c \det A$.

## A.3. Some facts about spectral theory

**A.3.1 Eigenvalues and eigenvectors.** Suppose $A \in L(V,V)$ where $V$ is a real vector space. If there exists a non-zero element $x \in V$ and a real number $\lambda$ such that $Ax = \lambda x$, then $\lambda$ is called an *eigenvalue* of $A$ and $x$ an *eigenvector* associated with $\lambda$.

The kernel of $A - \lambda \mathbb{1}$ is called the *geometric eigenspace* of $A$ associated with $\lambda$.

**A.3.2 Symmetric matrices.** A matrix $M \in \mathbb{R}^{n\times n}$ is called *symmetric*, if $M_{j,k} = M_{k,j}$. Recall that $M$ represents a linear operator from $\mathbb{R}^n$ to $\mathbb{R}^n$.

If $M$ is a symmetric matrix in $\mathbb{R}^{n\times n}$, then all its eigenvalues are real. Moreover, $\mathbb{R}^n$ has an orthonormal basis consisting of eigenvectors.

**A.3.3 Quadratic forms.** A homogeneous quadratic polynomial in $n$ variables with real coefficients is called a *quadratic form* over $\mathbb{R}$. Such a quadratic form is given by

$$q(x) = \sum_{j=1}^{n} \sum_{k=1}^{n} Q_{j,k} x_j x_k = x^\top Q x$$

where $Q \in \mathbb{R}^{n\times n}$. Note that the matrix $Q$ may always be chosen to be symmetric.

A quadratic form $q$ is called positive (or negative) *semi-definite*, if $q(x) \geq 0$ (or $q(x) \leq 0$) whenever $x \in \mathbb{R}^n$. It is called positive or negative *definite* if the inequalities are strict when $x \neq 0$. A quadratic form which is not semi-definite is called *indefinite*. These expressions are also used to characterize real symmetric matrices.

The following statements are true if $Q$ is chosen symmetric:

(1) $q$ is positive (negative) definite if and only if all eigenvalues of $Q$ are positive (negative).
(2) $q$ is positive (negative) semi-definite if and only if none of the eigenvalues of $Q$ are negative (positive).

APPENDIX B

# Miscellaneous

## B.1. Algebra

**B.1.1 The multinomial theorem.** Let $n$ and $k$ be a natural numbers and $x_1$, ..., $x_n$ real numbers. Then

$$(x_1 + ... + x_n)^k = \sum \frac{k!}{\alpha_1!...\alpha_n!} x_1^{\alpha_1}...x_n^{\alpha_n}$$

where the sum is over all choices of non-negative integers $\alpha_j$, $j = 1, ..., n$, such that $\alpha_1 + ... + \alpha_n = k$.

**B.1.2 Permutations.** A *permutation* of a finite set $X$ is a bijection from $X$ to itself. The set of such permutations is a group under composition. A permutation $\tau$ is called a *transposition* if there are distinct elements $x, y \in X$ such that $\tau(x) = y$ and $\tau(y) = x$ while $\tau(z) = z$ whenever $z \in X \setminus \{x, y\}$. Every permutation is a composition of transpositions. Such factorizations of permutations are not unique. However, if one factorization of a permutation $\pi$ has an even number of factors then this is true for all factorizations of $\pi$. One defines therefore the *parity* of a permutation $\pi$, denoted by $(-1)^\pi$, to be $(-1)^\ell = \pm 1$, if it has a factorization consisting of $\ell$ transpositions. If $\ell$ is even $\pi$ is called an even permutation and otherwise an odd permutation.

# List of special symbols

$\overline{\Omega}$: the closure of $\Omega$, 2

$\mathbb{1}$: the identity transformation or identity matrix, 27

$J(\phi, \alpha)$: the Jacobian determinant associated with the $k$-index $\alpha$, 18

$\ker F$: the kernel of $F$, 26

$L(V, W)$: the space of linear operators from $V$ to $W$, 26

$N_k^n$: the set of $k$-indices of type $n$, 18
$I_k^n$: the set of basic$k$-indices of type $n$, 18
$W_k^n$: a function space giving rise to differential forms, 18

$\|A\|$: the norm of the operator $A$, 2
$|x|$: the norm of the vector $x$, 1

$\mathbb{R}^{m \times n}$: the set of real $m \times n$-matrices, 27
$\operatorname{ran} F$: the range of $F$, 26

$Q^k$: the standard $k$-simplex, 18
span: the span of a set, 25
$e_k$ or $e_k^{(n)}$: the $k$-th member of the standard basis of $\mathbb{R}^n$, 26
$e_\alpha$: an element of the standard basis of $\mathbb{R}^{N_k^n}$, 18

# Index


$k$-form, 19
$k$-index, 18
  basic, 18
$n$-cell, 13

basis, 26
  ordered, 26
  standard, 26
boundary, 22

chain rule, 6
component, 25
continuity, 2
contraction, 9
coordinate, 25
critical point, 9

definite, 28
derivative, 5
  partial, 6
  total, 5
diagonal
  element, 27
  main, 27
diagonal matrix, 27
differentiable
  form, 20
  function, 5
differential form, 19
directional derivative, 8
distance, 1

eigenspace
  geometric, 28
eigenvalue, 28
eigenvector, 28
euclidean vector space, 25
extremum, 8

face of a cell, 14
face of an affine simplex, 22
fixed point, 9

gradient, 7

Hessian, 9

identity matrix, 27
identity transformation, 27
indefinite, 28
inner product, 1, 25
invertible, 27

kernel, 26

limit, 2
linear approximation, 5
linear combination, 25
linear independence, 25
linear operator, 26
linearly dependence, 25
Lipschitz condition, 6

main diagonal, 27
matrix, 27
  diagonal, 27
  square , 27
maximum
  relative, 8
  strict relative, 8
metric, 1
metric space, 2
minimum
  relative, 8
  strict relative, 8

norm, 1
  of a linear operator, 2

nullity, 26

oscillation, 14

parity, 29
partition of an $n$-cell, 13
permutation
  of a set, 29

quadratic form, 28

range, 26
rank, 26
rank-nullity theorem, 26
refinement, 13
Riemann integral, 13

scalar, 25
scalar multiplication, 25
scalar product, 1, 25
Schwarz's inequality, 1
semi-definite, 28
simplex
  affine, 22
  standard, 18
span, 25
subspace, 25
symmetric matrix, 28

total derivative, 5
transformation
  linear, 26
transposition, 29
triangle inequality
  for metric spaces, 2
  for normed spaces, 1

vector, 25
vector addition in $\mathbb{R}^n$, 25
volume of an $n$-cell, 13