CALCULUS OF SEVERAL VARIABLES

Lecture notes for MA 642

Rudi Weikard

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

Version of July 30, 2024

Contents

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, ℓ, m, and n will refer to elements of $\mathbb{N}$, the set of natural numbers. Also, the symbol Ω denotes an open set in $\mathbb{R}^n$ unless noted otherwise.

Hints and comments for the instructor are in blue.

CHAPTER 1

Limits and continuity

1.1. Inner products, norms, and metrics

1.1.1 The inner product on $\mathbb{R}^n$. Let X be a real vector space and assume that for every pair $(x, y) \in X \times X$ we have a number $\langle x, y\rangle \in \mathbb{R}$ such that the following properties hold when $x, y, z \in X$ and $\alpha, \beta \in \mathbb{R}$:

(1) $\langle x, x\rangle > 0$ unless $x = 0$.
(2) $\langle x, y\rangle = \langle y, x\rangle$.
(3) $\langle x, \alpha y + \beta z\rangle = \alpha\langle x, y\rangle + \beta\langle x, z\rangle$.

Then the map $(x, y) \mapsto \langle x, y\rangle$ is called an *inner product* or *scalar product* on X and X, equipped with an inner product, is called an inner product space.

We necessarily have that $\langle x, y\rangle = 0$ when $x = 0$ or $y = 0$.

On $\mathbb{R}^n$ an inner product is given by

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

Here x_j and y_j denote the components of x and y, respectively. Note that $x \cdot y$ may be 0 even though neither x nor y is 0. We will frequently use the notation $x \cdot y$ for the inner product on $\mathbb{R}^n$.

1.1.2 Schwarz's inequality. For any two vectors x and y in an inner product space *Schwarz's inequality*

$$|\langle x, y\rangle| \leq \sqrt{\langle x, x\rangle}\sqrt{\langle y, y\rangle}$$

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

1.1.3 The norm on $\mathbb{R}^n$. Let X be a real vector space and assume that for every element $x \in X$ we have a number $\|x\|$ such that the following properties hold when $x, y \in X$ and $\alpha \in \mathbb{R}$:

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

Then the map $x \mapsto \|x\|$ is called a *norm* on X and X, equipped with a norm, is called a normed (vector) space.

Every inner product space is a normed space with the norm given by

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

In particular, $\|x\| = 0$ if and only if $x = 0$.

Since $\mathbb{R}^n$ is an inner product space it is also a normed space. We have

$$|x|^2 = \sum_{j=1}^{n} x_j^2.$$

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 causes no harm even when $n = 1$.

1.1.4 $\mathbb{R}^n$ as a metric space. Let X be a set and assume that for every pair $(x, y) \in X \times X$ we have a number $d(x, y) \geq 0$ such that the following properties hold when $x, y, z \in X$:

(1) $d(x, y) = 0$ if and only if $x = y$.
(2) $d(x, y) = d(y, x)$.
(3) $d(x, y) \leq d(x, z) + d(z, y)$ (the *triangle inequality*).

Then d is called a *distance function* or a *metric*, the number $d(x, y)$ is called the *distance* of x and y, and X, equipped with a distance function is called *metric space*.

Every normed space is a metric space with the distance function given by $d(x, y) = |x - y|$. In particular, $\mathbb{R}^n$ is a metric space.

Recall that metric spaces are topological spaces when we use the open balls $B(x, r) = \{y \in X : d(x, y) < r\}$ as a base for the topology.

1.1.5 The norm of a linear operator. Suppose that A is a linear operator from $\mathbb{R}^n$ to $\mathbb{R}^m$. Denote the entries of the matrix associated with A using the standard bases in $\mathbb{R}^n$ and $\mathbb{R}^m$ by $A_{j,k}$, i.e., $A_{j,k} = e_j^{(m)} \cdot Ae_k^{(n)}$ and let $M = \max\{|A_{j,k}| : 1 \leq j \leq m, 1 \leq k \leq n\}$. Then $|Ax|^2 \leq mnM^2|x|^2$ by Cauchy-Schwarz $\sum_{k=1}^{n} |x_j| \leq \sqrt{n}|x|$. Hence

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

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

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

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

$$\|BA\| \leq \|B\|\|A\|.$$

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 $|Bx| \geq |Ax| - |(B - A)x| \geq (1 - \gamma)|x|/\|A^{-1}\|$ so that B is also invertible. In fact, $\|B^{-1}\| \leq \|A^{-1}\|/(1 - \gamma)$. Choose $x = B^{-1}y$. 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 Ω to $\mathbb{R}^m$ and x_0 is a point in $\overline{\Omega}$, the closure of Ω. 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 Ω to $\mathbb{R}^m$ and x_0 is a point in Ω. 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 Ω, if it is continuous at every point of Ω. The set of all continuous functions from Ω 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. $A^{-1} - A_0^{-1} = A^{-1}(A_0 - A)A_0^{-1}$.

1.2.6 Limit rules. Suppose f and g are functions from Ω to $\mathbb{R}^m$, h is a function from Ω 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 Ω' 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 Ω to $\mathbb{R}^m$ and x_0 is a point in Ω. If there exists a linear operator $A : \mathbb{R}^n \to \mathbb{R}^m$, i.e., an $m \times n$-matrix A, such that

$$\lim_{x \to x_0} \frac{|f(x) - f(x_0) - A(x - x_0)|}{|x - x_0|} = 0,$$

we say that f is *differentiable* at x_0 and call A a *total derivative* or just a *derivative* of f at the point x_0.

If f is differentiable at every point of Ω we say that f is differentiable on Ω.[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 $A = B$. Suppose $0 \neq h \in \mathbb{R}^n$ and $t > 0$. Then $|(A - B)h| = \frac{1}{|t|}|(A - B)(th)| \to 0$ as $t \to 0$.

Therefore it is customary to denote *the* total derivative of f at x by $f'(x)$. If f is differentiable on Ω, the map $x \mapsto f'(x)$ is a function from Ω 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 $f'\binom{2}{3} = (12, 4)$.

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 Ω to $\mathbb{R}^m$, h is a function from Ω to $\mathbb{R}$, and x is a point in Ω. 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.

Consider the dot product and fix x_0. We have continuous vector-valued functions r_f and r_g which vanish at x_0. Define r by

$$|h|r(x) = f(x) \cdot g(x) - f(x_0) \cdot g(x_0) - (f(x_0)^\top g'(x_0) + g(x_0)^\top f'(x_0))h$$

where $h = x - x_0 \neq 0$. Then, using Schwarz's inequality and the operator norm, we get

$$|r(x)| \leq |h|(\|f'(x_0)\| + |r_f(x)|)(\|g'(x_0)\| + |r_g(x)|) + |f(x_0)||r_g(x)| + |g(x_0)||r_f(x)|$$

which tends to 0 as x tends to x_0.

2.1.7 The chain rule. Suppose $f : \Omega \to \mathbb{R}^m$ and $g : \Omega' \to \mathbb{R}^k$ where Ω' 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 segment 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 γ 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 Ω to $\mathbb{R}^m$, and x a point in Ω. 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 at x and

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

Suppose $f : \mathbb{R}^2 \to \mathbb{R}$ is defined by $f(0,0) = 0$ and $f(x,y) = xy/(x^2+y^2)$ when $(x,y) \neq (0,0)$. Then the partial derivatives of f exist at $(0,0)$ but the function is not continuous and hence not differentiable there.

This works for $y^3/(x^2+y^2)$ which is continuous.

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 Ω 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.

We have, in fact, f' exists and is continuous at a fixed x_0 iff the partial derivatives exist and are continuous at x_0 as the proof actually shows. Therefore split the topic: A sufficient condition for differentiability & Cont. diff. functions.

However, $f'(x_0)$ may exist even if the partial derivatives are not continuous at x_0, as seen in the example $f(x) = x^2 \sin(1/x)$ for $x \neq 0$ and $f(0) = 0$.

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 $A(x_0) = f'(x_0)$ for any $x_0 \in \Omega$. 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 $A(x_0)(x - x_0) = \sum_{k=1}^n h_k A(x_0) e_k = \sum_{k=1}^n h_k (D_k f)(x_0)$ 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 Ω 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 Ω 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 Ω is an open subset in $\mathbb{R}^2$ and f a real-valued function on Ω 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 Ω, 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 ϕ and ψ. □

2.2.6 Mixed partial derivatives commute. If $f \in C^k(\Omega, \mathbb{R})$, $j_1, ... j_k \in \{1, ..., n\}$, and π 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 Ω 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. These assumptions are somewhat stronger than Rudin's to make the statement shorter. Then $D_1 D_2 f$ also exists and equals $D_2 D_1 f$ in Ω.

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 $v \to 0$,

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

□

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 α 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

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

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 Ω 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 $1 \leq k \leq r$. Now apply Taylor's theorem for one variable. □

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

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

If the differentiable function f has a local extremum at x_0, then x_0 is a critical point of f. Consider $g(t) = f(x_0 + te_k)$ for $k = 1, ..., n$. Then g has a local extremum at 0 and is differentiable at 0, and $(D_k f)(x_0) = g'(0) = 0$. For the latter note that $(g(t) - g(0))/t$ is non-negative for $t > 0$ and non-positive for $t < 0$ in case of a minimum. If $g'(0)$ exists it must be 0. Thus we have a necessary condition for a point x_0 to be a local extremum of f.

2.3.4 The Hessian. 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 $H(x)$. Taylor's theorem shows now the existence of a $t \in (0, 1)$ such that

$$\begin{aligned} f(x_0 + h) - f(x_0) &= \sum_{|\alpha|=2} \frac{(D^\alpha f)(x_0 + th)}{\alpha!} h^\alpha = \frac{1}{2} h^\top H(x_0 + th) h \\ &= \frac{1}{2} h^\top H(x_0) h + \frac{1}{2} h^\top \big(H(x_0 + t\phi) - H(x_0)\big) h. \end{aligned}$$

Since H is continuous we have the following claim: If $\epsilon > 0$ there is a $\delta > 0$ such that

$$\|H(x_0 + h) - H(x_0)\| < \epsilon$$

whenever $h \in \mathbb{R}^n$ and $|h| < \delta$. From this we obtain the inequalities

$$f(x_0 + h) - f(x_0) \geq \frac{1}{2} h^\top H(x_0) h - \frac{\epsilon}{2} |h|^2 \tag{1}$$

and

$$f(x_0 + h) - f(x_0) \leq \frac{1}{2} h^\top H(x_0) h + \frac{\epsilon}{2} |h|^2 \tag{2}$$

It is known from Linear Algebra that the eigenvalues of $H(x)$ are all real and the corresponding eigenvectors may be chosen to form an orthonormal basis of $\mathbb{R}^n$.

2.3.5 A sufficient criterion for the presence of an extremum. The following theorem gives a sufficient condition for a point x_0 to be a local extremum of f.

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 local minimum (maximum) of f.

Sketch of proof: Suppose $H(x_0)$ is positive definite, i.e., its smallest eigenvalue λ is positive. Then $h^\top H(x_0) h \geq \lambda |h|^2$ for all $h \in \mathbb{R}^n$. Using this in (1) and choosing there $\epsilon = \lambda/2$ gives

$$f(x_0 + h) - f(x_0) \geq \frac{1}{2} \lambda |h|^2 - \frac{\lambda}{4} |h|^2 = \frac{1}{4} \lambda |h|^2 > 0$$

for sufficiently small h unless $h = 0$. Therefore f has a strict local minimum at x_0. One proves in a similar fashion, using (2), that f has a strict local maximum at x_0 if $H(x_0)$ is negative definite. $\square$

2.3.6 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.

Sketch of proof: The second statement is the contrapositive of the first. To prove the first statement assume that $H(x_0)$ has both a positive eigenvalue λ_+ and a negative eigenvalue λ_- with corresponding normalized eigenvectors ϕ_+ and ϕ_-. Thus

$$\phi_+^\top H(x_0) \phi_+ = \lambda_+ \quad \text{and} \quad \phi_-^\top H(x_0) \phi_- = \lambda_-.$$

Now choose $h = s\phi_+$ and $\epsilon = \lambda_+/2$ in (1) to find

$$f(x_0 + s\phi_+) - f(x_0) \geq \frac{s^2}{4} \lambda_+$$

showing that f can not have a local minimum at x_0. Similarly, using $h = s\phi_-$ and $\epsilon = |\lambda_-|/2$ in (2), gives

$$f(x_0 + s\phi_+) - f(x_0) \leq \frac{s^2}{4} \lambda_-$$

showing that f can not have a local maximum at x_0. $\square$

2.4. The inverse and implicit function theorems

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

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

Either from induction or the identity $as_n = s_{n+1} - 1 = s_n + a^{n+1} - 1$ we may see that $s_n = \sum_{k=0}^n a^k = (a^{n+1}-1)/(a-1)$.

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

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

and

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

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 $\lambda = 1/(2\|A^{-1}\|) > 0$ 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 ϕ is a contraction ($\|\phi'(x)\| \leq 1/2$ and 2.1.8) and therefore has at most one fixed point (choose $y = f(x) = f(x')$).

(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 ϕ. Since ϕ maps $\overline{B(x_1, r)}$ to itself $|\phi(x) - x_1| \leq$

$|\phi(x)-\phi(x_1)|+|\phi(x_1)-x_1|$ the contraction mapping theorem 2.4.4 applies and guarantees the existence of a fixed point x_2 of ϕ 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 $B=f'(x)^{-1}$,

$$|g(y+v)-g(y)-Bv|=|B(v-f'(x)u)|=|B(f(x+u)-f(x)-f'(x)u)|\le|u|\|B\||r(x+u)| \quad (3)$$

for some function r which is continuous at x and vanishes there. With y we associate, as above, a function ϕ and obtain $\phi(x+u)-\phi(x)=u-A^{-1}v$ and, since ϕ is a contraction, $|u-A^{-1}v|\le|u|/2$. Hence $|u|\le|v|/\lambda$. This and (3) 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 (use 1.2.5). ☐

Discuss $f(x)=x+2x^2\sin(1/x)$ to show that continuity of the derivative is "necessary" for local injectivity.

2.4.6 The implicit function theorem. Let Ω be an open subset of $\mathbb{R}^n$ and let $f\in C^1(\Omega,\mathbb{R}^m)$ where $m<n$. Define $k=n-m\ge1$. Suppose

(1) $f(\binom{x_0}{y_0})=0$ for some $\binom{x_0}{y_0}\in\Omega$ with $x_0\in\mathbb{R}^k$ and $y_0\in\mathbb{R}^m$ and
(2) $f'(\binom{x_0}{y_0})=(A_1,A_2)$ where $A_1\in\mathbb{C}^{m\times k}$, $A_2\in\mathbb{C}^{m\times m}$ and A_2 is invertible.

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

Sketch of proof: To simplify notation we will frequently write $f(x,y)$ in place of $f(\binom{x}{y})$. Moreover, if we write a vector in $\mathbb{R}^n$ as a pair $(x,y)^\top$, we tacitly assume that $x\in\mathbb{R}^k$ 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 $F(x,y)=\binom{x}{f(x,y)}$, a function from Ω to $\mathbb{R}^n$. If $f'=(P,Q)$ with $P(x,y)\in\mathbb{R}^{m\times k}$ and $Q(x,y)\in\mathbb{R}^{m\times m}$, then $F'=\begin{pmatrix}\mathbb{1}&0\\P&Q\end{pmatrix}\in C^0(\Omega,\mathbb{R}^n)$. Moreover, $F'(x_0,y_0)=\begin{pmatrix}\mathbb{1}&0\\A_1&A_2\end{pmatrix}$ and is therefore invertible.

(b) We may now apply the inverse function theorem to F. It shows that there are open sets U, $V=F(U)\subset\mathbb{R}^n$ such that $\binom{x_0}{y_0}\in U$ and $\binom{x_0}{0}=F(\binom{x_0}{y_0})\in V$. Moreover, $G=(F|_U)^{-1}:V\to U$ is bijective and continuously differentiable.

Now define $W=\{x\in\mathbb{R}^k:\binom{x}{0}\in V\}$. Then $x_0\in W$ and W is open. If $x\in W$, then $\binom{x}{0}\in V$ and we define

$$G(\tbinom{x}{0})=\tbinom{h(x)}{g(x)}\in U$$

with $h(x)\in\mathbb{R}^k$ and $g(x)\in\mathbb{R}^m$. Then $\binom{x}{0}=F(\binom{h(x)}{g(x)})=\binom{h(x)}{f(h(x),g(x))}$. This implies that $h(x)=x$ and $f(x,g(x))=0$. We have, in particular, $F(\binom{x_0}{g(x_0)})=\binom{x_0}{0}=F(\binom{x_0}{y_0})$. Since F is injective it follows that $g(x_0)=y_0$.

(c) It remains to show that $g\in C^1(W,\mathbb{R}^m)$ and $g'(x_0)=-(A_2)^{-1}A_1$. 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 $Q(x,y)$ is invertible. Thus

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

Now, using $G((\begin{smallmatrix} x \\ 0 \end{smallmatrix})) = (\begin{smallmatrix} x \\ g(x) \end{smallmatrix})$, note that $g'(x) = -Q(x, g(x))^{-1} P(x, g(x))$. Fix x and let $B = Q(x, g(x))^{-1}$ and $A = -Q(x, g(x))^{-1} P(x, g(x))$. Then

$$\frac{|g(x+h) - g(x) - Ah|}{|h|} = \frac{|(\begin{smallmatrix} x+h \\ g(x+h) \end{smallmatrix}) - (\begin{smallmatrix} x \\ g(x) \end{smallmatrix}) - (\begin{smallmatrix} \mathbb{1} \\ A \end{smallmatrix}) h|}{|h|}$$
$$= \frac{|G((\begin{smallmatrix} x+h \\ 0 \end{smallmatrix})) - G((\begin{smallmatrix} x \\ 0 \end{smallmatrix})) - (\begin{smallmatrix} \mathbb{1} & 0 \\ A & B \end{smallmatrix})(\begin{smallmatrix} h \\ 0 \end{smallmatrix})|}{|h|}.$$

The right-hand side tends to 0 as $|h| \to 0$ since $(\begin{smallmatrix} \mathbb{1} & 0 \\ A & B \end{smallmatrix}) = G'((\begin{smallmatrix} x \\ 0 \end{smallmatrix}))$. □

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 Ω we will now consider the problem of finding extrema of h in certain non-open subsets of Ω. To be precise, we want to find extrema of h among those points x in Ω 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

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

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

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

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

$$(h' + \lambda f')(x, y) = 0 \tag{4}$$

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

Returning to our example we find that equation (4) 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 $m \times k$-matrix and A_2 is an invertible $m \times m$-matrix. We also write $x_0 = \binom{\alpha}{\beta}$ where $\alpha \in \mathbb{R}^k$ and $\beta \in \mathbb{R}^m$. Similarly, $h'(x_0) = (b_1, b_2)$ with $b_1^\top \in \mathbb{R}^k$ and $b_2^\top \in \mathbb{R}^m$.

Since A_2 is invertible the equation $b_2 + \lambda A_2 = 0$ has a unique solution $\lambda = -b_2 A_2^{-1}$ (a row with m components). Hence $(h + \lambda f)'(x_0) = (b_1 + \lambda A_1, 0)$. We need to show that $b_1 + \lambda A_1 = 0$.

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

$$\lambda A_1 - b_2 g'(\alpha) = 0. \tag{5}$$

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

$$0 = H'(\alpha) = b_1 + b_2 g'(\alpha). \tag{6}$$

Combining equations (5) and (6) shows that indeed $b_1 + \lambda A_1 = 0$. □

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?

$$(h + \lambda f)' = (2x + \lambda_1 + 2\lambda_2 x, 2y + \lambda_1 + 2\lambda_2 y, 2z + 2\lambda_1 - \lambda_2).$$

Hence $\lambda_1 = -2x(1 + \lambda_2) = -2y(1 + \lambda_2)$.

The case $\lambda_2 = -1$ gives $\lambda_1 = 0$ and $z = -1/2$ which is void.

The case $\lambda_2 \neq -1$ gives $x = y$ and, using the constraints, $z = 1 - x$ and $z = 2x^2$. Hence $x = 1/2$ or $x = -1$. The former gives the closest and the latter the farthest point on the ellipse.

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

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

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

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

The quantity

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

is called the *volume* of I.

3.0.2 Partitions. Recall that a partition of a closed interval $[a, b]$ is a finite subset of $[a, b]$ which contains both a and b. If the number of elements of the partition P_k of $[a_k, b_k]$ is $m + 1$ we label them so that

$$a_k = p_{k,0} < p_{k,1} < \ldots < p_{k,m-1} < p_{k,m} = b_k.$$

Now suppose we have a partition P_k of each of the intervals $[a_k, b_k]$ defining the n-cell $I = \bigtimes_{k=1}^{n} [a_k, b_k]$. Then $P = P_1 \times ... \times P_n$ is called a partition of I. If $(p_{1,j_1}, ..., p_{n,j_n}) \in P$ with $p_{k,j_k} < b_k$ then the n-cell $\bigtimes_{k=1}^{n} [p_{k,j_k}, p_{k,j_{k+1}}]$ is included in I and therefore called a sub-cell of I. The union of all these sub-cells is I and the intersection of the interiors of any two distinct sub-cells is empty.

If $P^* = P_1^* \times ... \times P_n^*$ is also a partition of I and if $P_k \subset P_k^*$ for $k = 1, ..., n$, 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

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

Every sub-cell of P is divided into 2^ℓ sub-cells of P^* where $\ell \in \{0, ..., n\}$.

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 ε, 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_j, $j \in \mathbb{N}$, has measure zero than so does the set $\bigcup_{j=1}^\infty E_j$.

Needs work!

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 $\{\binom{x}{f(x)} : x \in E\} \subset \mathbb{R}^n$, has measure zero. To see this let $\varepsilon > 0$ be given and let $\{R_1, ..., R_k\}$ be a collection of intervals of length ϵ 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 and that ε is given. Then there are open cells U_j, $j \in \mathbb{N}$, such that

$$B \subset \bigcup_{j=1}^{\infty} U_j \subset \bigcup_{j=1}^{\infty} \overline{U_j} \quad \text{and} \quad \sum_{j=1}^{\infty}|\overline{U_j}| < \varepsilon.$$

Moreover, if $x_0 \in I \setminus B$, then f is continuous at x_0. Thus, there is a positive δ such that $|x - x_0| < \delta$ implies $|f(x) - f(x_0)| < \varepsilon/4$. Choose an open n-cell V_{x_0} such that $x_0 \in V_{x_0}$ and the diameter of V_{x_0} is less than δ. It follows that

$$\sup\{f(x) : x \in V_{x_0}\} - \inf\{f(x) : x \in V_{x_0}\} < \varepsilon.$$

Also V_{x_0} includes an open n-cell W_{x_0} such that $x_0 \in W_{x_0}$ and $\overline{W_{x_0}} \subset V_{x_0}$.

We now have that I is covered by the collection of all U_j and W_{x_0}. Since I is compact, it is, in fact, covered by a finite collection of the U_j, $j = 1, ..., J$ and the W_{x_ℓ}, $\ell = 1, ..., L$ and hence by their corresponding closures. Thus

$$I = \left(\bigcup_{j=1}^{J}(\overline{U_j} \cap I)\right) \cup \left(\bigcup_{\ell=1}^{L}(\overline{W_{x_\ell}} \cap I)\right).$$

Collecting all the k-th components of these cells for each $k = 1, ..., n$ gives rise to a partition P of I. The corresponding subcells lie either in one of the U_j or in one of the W_{x_ℓ} (or perhaps in both). The former are small in total volume but $M_j - m_j$ maybe as large as $2C$. For the latter $M_k - m_k < \epsilon$ and their total volume does not exceed $|I|$. For this partition we have therefore $U(P, f) - L(P, f) < 2C\varepsilon + |I|\varepsilon$. ☐

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. If $E \subset I_1$ and $E \subset I_2$, then $E \subset I_1 \cap I_2 = I$. Note that $I \setminus I_j$ is a finite union of cells where $f = 0$. We then say that f is Riemann integrable over E.

In particular, if the boundary ∂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) $\int_E(f+g) = \int_E f + \int_E g$. $m_k(f)+m_k(g) \leq m_k(f+g)$, $M_k(f)+M_k(g) \geq M_k(f+g)$.
(2) $\int_E(cf) = c\int_E f$ whenever $c \in \mathbb{R}$.
(3) fg is Riemann integrable over E. fg is almost everywhere continuous.
(4) If $f \geq 0$ then $\int_E f \geq 0$.

(5) $|f|$ is Riemann integrable over E and $|\int_E f| \le \int_E |f|$. $|\int f| = s\int f = \int sf$ where $s = \pm 1$. But $sf \le |f|$.

(6) If $E_1 \cup E_2 = E$ and $E_1 \cap E_2 = \emptyset$, then $\int_E f = \int_{E_1} f + \int_{E_2} f$. $\chi_E = \chi_{E_1} + \chi_{E_2}$.

Needs work!

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,

$$\int_{A\times B} f = \int_A \varphi = \int_A \psi. \tag{7}$$

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 $\varphi(x) = \underline{\int_B} f(x, \cdot) \ge L(Q, f(x, \cdot)) = \sum_{k=1}^M m_k(x)|B_k|$. If $x \in A_j$ we have $m_k(x) \ge m_{j,k} = \inf\{f(x, y) : x \in A_j, y \in B_k\}$ and hence

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

Thus

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

Similarly, $U(R, f) \ge U(P, \psi) \ge L(P, \psi)$. Also, since $\varphi \le \psi$ and f is Riemann integrable, we get that both φ and ψ are Riemann integrable and that (7) holds. $\square$

3.0.14 The change of variables formula. Suppose $T : \Omega \to \mathbb{R}^n$ is an injective, continuously differentiable function. Assume further that T' is everywhere invertible so that $\det T'(x) \ne 0$ for all $x \in \Omega$. If $f : T(\Omega) \to \mathbb{R}$ or $f \circ T\, |\det T'| : \Omega \to \mathbb{R}$ is Riemann integrable, then so is the other one and

$$\int_{T(\Omega)} f = \int_\Omega f \circ T\, |\det T'|.$$

3.0.15 Differentiating an integral. Let $f : [a, b] \times [c, d] \to \mathbb{R}$ be a continuous function such that $D_2 f$ is also continuous. Define $F(y) = \int_{[a,b]} f(\cdot, y)$ for $y \in (c, d)$. 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 $\varepsilon > 0$,

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

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

$$\int_{[a,b]} \int_{[y_0,y]} |(D_2 f)(x, u) - D_2 f(x, y_0)|\, du dx < \varepsilon(b - a)(y - y_0)$$

when $|u - y_0| \le |y - y_0| < \delta$. $\square$

CHAPTER 4

Integration of differential forms

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

4.1. Integration along paths

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

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

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

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

$$\int_\gamma \omega = \int_{[0,1]} ((3t)(6t^2) + (2t^3)3 + 0(2t))dt = 6.$$

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

$$\int_\gamma \omega = \int_{[0,1]} 2\pi ab\cos(2\pi t)^2 dt = \pi ab.$$

4.2. Integration over surfaces

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

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

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

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

Now we define the integral of an array of n^2 continuous real-valued functions $\omega_{j,k}$, $j, k = 1, ..., n$, defined on Ω 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 the non-trivial Jacobians are $J((1,2)) = 2\pi^2 \sin(\pi s)\cos(\pi s)$, $J((3,1)) = 2\pi^2 \sin(\pi s)^2 \sin(2\pi t)$, and $J((2,3)) = 2\pi^2 \sin(\pi s)^2 \cos(2\pi t)$. Thus

$$\int_\phi \omega = \int_{[0,1]} \int_{[0,1-s]} 2\pi^2 \sin(\pi s) dt\, ds = 2\pi.$$

This is half of the surface area of the unit sphere. To see why let

$$N = (J(2,3), J(3,1), J(1,2))^\top$$

and note that $N = 2\pi^2 \sin(\pi s)\phi$ and $|\phi| = 1$. Hence $\int_\gamma \omega = \int_{Q^2} \phi \cdot N = \int_{Q^2} |N|$ as claimed in calculus books.

4.3. The general case

4.3.1 The standard k-simplex. Suppose $k \in \mathbb{N}$. Then the set $Q^k = \{x \in \mathbb{R}^k : 0 \leq x_j, x_1 + ... + x_k \leq 1\}$ 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 Ω to be a function $\phi \in C^1(Q^k, \Omega)$. A 0-surface in Ω is a point in Ω. Q^k is called the parameter domain of ϕ.

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 $\{1, ..., n\}$, written as a row, a k-index of type n. The set of k-indices of type n is denoted by V_n^k. It has precisely n^k elements.

The set of functions from V_n^k to $\mathbb{R}$ is an n^k-dimensional vector space. It has a standard basis $\mathbf{e}_\alpha$, $\alpha \in V_n^k$, defined by $\mathbf{e}_\alpha(\beta) = 1$ if $\alpha = \beta$ and $\mathbf{e}_\alpha(\beta) = 0$ if $\alpha \neq \beta$. Check linear independence and spanning: Suppose $0 = \sum_{\alpha \in V_n^k} c_\alpha \mathbf{e}_\alpha$. Evaluate at β to see that $c_\beta = 0$, showing linear independence. To show that the $\mathbf{e}_\alpha$ span, let r be an arbitrary function from V_n^k to $\mathbb{R}$. Then $r = \sum_{\alpha \in V_n^k} r(\alpha)\mathbf{e}_\alpha$. This construction is similar to the construction of the standard base in $\mathbb{R}^n = \mathbb{R}^{\{1,...,n\}}$.

A k-index β is called a *basic* k-index if $\beta_1 < \beta_2 < ... < \beta_k$. There are $\binom{n}{k}$ basic k-indices in V_n^k if $k \leq n$ and none if $k > n$. Choose k distinct numbers and order them. The set of all basic k-indices of type n is denoted by I_n^k. When $k = 1$ we have $V_n^1 = I_n^1$.

4.3.4 The vector space $W_n^k(\Omega)$. The functions $\omega : V_n^k \to C^0(\Omega, \mathbb{R})$ form a real vector space upon using the standard definition of sums and constant multiples of functions. This vector space will be denoted by $W_n^k(\Omega)$. In other words, an element of $W_n^k(\Omega)$ assigns to each k-index α of type n a function $\omega_\alpha \in C^0(\Omega, \mathbb{R})$. Clearly, we may represent an element of $W_n^k(\Omega)$ by $\omega = \sum_{\alpha \in V_n^k} \omega_\alpha \mathbf{e}_\alpha$.

We also define $W_n^0(\Omega) = C^0(\Omega, \mathbb{R})$.

4.3.5 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.

4.3.6 Integration over a k-surface. Suppose ϕ is a k-surface in Ω and $\omega \in W_n^k(\Omega)$. Then we define

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

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

Integration over k-surfaces is linear, i.e.,

$$\int_\phi (c_1\omega_1 + c_2\omega_2) = c_1 \int_\phi \omega_1 + c_2 \int_\phi \omega_2$$

when $c_1, c_2 \in \mathbb{R}$ and $\omega_1, \omega_2 \in W_n^k(\Omega)$.

4.3.7 Differential k-forms. Discuss $\omega = \left(\begin{smallmatrix} f & 0 \\ 0 & g \end{smallmatrix}\right)$, $\hat\omega = \left(\begin{smallmatrix} 0 & f \\ f & 0 \end{smallmatrix}\right)$, and $\tilde\omega = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$. We call two functions $\omega_1, \omega_2 \in W_n^k(\Omega)$ equivalent, if $\int_\phi \omega_1 = \int_\phi \omega_2$ for all k-surfaces ϕ in Ω. 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* ω of order k in Ω, or simply a *k-form* in Ω, is an equivalence class of functions $\tilde\omega \in W_n^k(\Omega)$. A differential form of order 0 in Ω, or simply a 0-form in Ω is a continuous real-valued function on Ω.

A k-form in Ω assigns to each k-surface in Ω a real number. If $\tilde\omega$ is a representative of ω and ϕ 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. One needs to show addition and scalar multiplication are well-defined: Let $[\check\omega] = [\hat\omega]$ and $[\check\psi] = [\hat\psi]$. By linearity

$$\int_\Phi (\check\psi + \check\omega) = \int_\Phi (\hat\psi + \hat\omega).$$

Then go through all nine axioms: associative, commutative, zero, negatives, two distributive laws, $(c_1c_2)\omega = c_1(c_2\omega)$, $1\omega = \omega$.

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

Suppose $f \in C^0(\Omega, \mathbb{R})$, $\alpha \in V_n^k$, and π is a transposition of $\{1, ..., k\}$. Then $f\mathbf{e}_\alpha \sim -f\mathbf{e}_{\alpha_\pi}$. If $\alpha_j = \alpha_\ell$ for some $j \neq \ell$ then $f\mathbf{e}_\alpha \sim 0$. If $\tilde\omega \in W_n^k(\Omega)$ with $k > n$, then $\tilde\omega \sim 0$.

Suppose ω_1 and ω_2 are k-forms in Ω and ϕ is a k-surface in Ω. 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 α are pairwise distinct, there is a permutation π of $\{1, ..., k\}$ such that $\beta = \alpha_\pi$ is a basic k-index. Since $\mathbf{e}_\alpha \sim (-1)^\pi \mathbf{e}_\beta$ we can, for any k-form ω, choose a representative $\tilde\omega$ such that $\tilde\omega_\alpha = 0$ unless α is a basic k-index. Such a representative is called a basic representative of

ω. If $k > n$ and ω is a k-form, then $\omega = 0$. We will use the notation $[\alpha]$ for α_π when this is the basic k index associated with α.

THEOREM. Suppose $k \leq n$. Let ω be a k-form in Ω and $\tilde{\omega}$ a basic representative of ω. Then $\omega = 0$ if and only if $\tilde{\omega}_\alpha = 0$ for every $\alpha \in I_n^k$ (and, indeed, in V_n^k). 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 α. Then there is a $\delta > 0$ such that $|\tilde{\omega}_\alpha(x) - \tilde{\omega}_\alpha(x_0)| < \tilde{\omega}_\alpha(x_0)/2$ whenever $|x - x_0| < \delta$. Hence $\tilde{\omega}_\alpha(x) > \tilde{\omega}_\alpha(x_0)/2 > 0$ for $x \in B(x_0, \delta)$. Construct a k-surface ϕ 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. Let $\phi : Q^k \to \Omega$ be given by $\phi(u) = x_0 + \nu \sum_{j=1}^k u_j \mathbf{e}_{\alpha_j}$. Note that $|\phi(u) - x_0| \leq \nu k$. Hence, for $\nu = \delta/k$ and $u \in Q^k$ we have $\tilde{\omega}_\alpha(\phi(u)) > 0$. □

In the following $\wedge$, d and variable changes are defined using basic representatives. Can we extend to arbitrary representatives?

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

$$\tilde{\omega} = \sum_{\alpha \in I_n^p} \tilde{\omega}_\alpha \mathbf{e}_\alpha \quad \text{and} \quad \tilde{\lambda} = \sum_{\beta \in I_n^q} \lambda_\beta \mathbf{e}_\beta.$$

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

$$\sum_{\alpha \in I_n^p} \sum_{\beta \in I_n^q} \tilde{\omega}_\alpha \tilde{\lambda}_\beta \mathbf{e}_{(\alpha, \beta)}.$$

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

$$\omega \wedge \lambda = \sum_{\beta \in I_n^q} \omega \lambda_\beta \mathbf{e}_\beta.$$

Similarly, if ω is a p-form and λ is a 0-form, then

$$\omega \wedge \lambda = \sum_{\alpha \in I_n^p} \omega_\alpha \lambda \mathbf{e}_\alpha.$$

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.To prove associativity one has to first write $\omega \wedge \lambda$ and $\lambda \wedge \tau$ in basic form, which introduces parities into the sums. These can then be undone after looking at both $(\omega \wedge \lambda) \wedge \tau$ and $\omega \wedge (\lambda \wedge \tau)$. In fact, $\omega \wedge \lambda = (-1)^{pq} \lambda \wedge \omega$. After p transpositions of β_q with each α_j we have $(\beta_1, \ldots \beta_{q-1}, \alpha, \beta_q)$. Repeat $q - 1$ times for a total of pq transpositions to arrive at (α, β).

Note that, if $\alpha \in I_n^k$, then

$$\mathbf{e}_\alpha = \mathbf{e}_{\alpha_1} \wedge \ldots \wedge \mathbf{e}_{\alpha_k}.$$

4.3.11 Differentiation of differential forms. We say a differential form of order k is of class C^r if the functions ω_α in its basic representation $\sum_{\alpha \in I_n^k} \omega_\alpha \mathbf{e}_\alpha$ 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^r in Ω we define df to be the 1-form with basic representative

$$\sum_{j=1}^{n} (D_j f)\mathbf{e}_j$$

using that $I_n^1 = \{1, ..., n\}$. df is of class C^{r-1}.

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

$$\tilde{\omega} = \sum_{\alpha \in I_n^k} \tilde{\omega}_\alpha \mathbf{e}_\alpha,$$

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

$$\sum_{\alpha \in I_n^k} \sum_{j=1}^{n} (D_j \tilde{\omega}_\alpha)\mathbf{e}_{(j,\alpha)}.$$

4.3.12 Examples. The following are important examples.

(1) Consider $x \mapsto x_j$ as a 0-form. Then $d(x \mapsto x_j)$ has representative $\mathbf{e}_j$. It is therefore customary to write dx_j for $\mathbf{e}_j$ and representatives of general 1-forms in Ω are often written as $\sum_{j=1}^{n} \omega_j dx_j$, where the ω_j are continuous real-valued functions on Ω. This convention makes no distinction between a variable x_j and the function $x \mapsto x_j$. We will not use it in these notes.

(2) $d^2(x \mapsto x_j) = d(d(x \mapsto x_j) = 0$.

(3) Let ϕ be a 1-surface and f a 0-form of class C^1. Then

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

This is called the fundamental theorem of line integrals.

(4) Let ω be the 1-form with basic representative $x \mapsto x_p \mathbf{e}_q$, $1 \leq p, q \leq n$. Then $d\omega$ is represented by $\mathbf{e}_{p,q}$. In particular, $d\omega = 0$ if $p = q$.

4.3.13 Differentiation rules. Suppose ω is a differentiable p-form and λ is a differentiable q-form in Ω. 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) $d(\omega \wedge \lambda) = (d\omega) \wedge \lambda + (-1)^p \omega \wedge (d\lambda) = (d\omega) \wedge \lambda + (-1)^{pq}(d\lambda) \wedge \omega$.

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

4.3.14 Products of 1-forms and determinants. Suppose $T \in C^1(\Omega, \Omega')$ where Ω' is an open set in $\mathbb{R}^m$, $k \leq m$, and $\alpha \in I_m^k$. Then dT_1, ..., dT_m are 1-forms in Ω and

$$dT_{\alpha_1} \wedge ... \wedge dT_{\alpha_k} = \sum_{\beta \in I_n^k} J(T, \alpha, \beta)\mathbf{e}_\beta^{(n)}$$

where

$$J(T, \alpha, \beta) = \det \begin{pmatrix} D_{\beta_1} T_{\alpha_1} & \cdots & D_{\beta_k} T_{\alpha_1} \\ \vdots & & \vdots \\ D_{\beta_1} T_{\alpha_k} & \cdots & D_{\beta_k} T_{\alpha_k} \end{pmatrix},$$

an extension of our notation introduced in 4.3.5.

Sketch of proof: If $k > n$ we have 0 on both sides of our identity. Hence assume $k \le n$ in the following and suppose $1 \le p \le k$ and $1 \le q \le n$. Then one proves by induction that

$$dT_{\alpha_1} \wedge \ldots \wedge dT_{\alpha_k} = \sum_{m_1=1}^{n} \ldots \sum_{m_k=1}^{n} (D_{m_1} T_{\alpha_1}) \ldots (D_{m_k} T_{\alpha_k})\ \mathbf{e}^{(n)}_{m_1} \wedge \ldots \wedge \mathbf{e}^{(n)}_{m_k}.$$

If the elements of $(m_1, \ldots, m_k)$ are not pairwise distinct the corresponding summand is 0. Each of the remaining lists $(m_1, \ldots, m_k)$ is obtained by a permutation π of some $\beta \in I^k_n$, i.e.,

$$(m_1, \ldots, m_k) = (\pi(\beta_1), \ldots, \pi(\beta_k)).$$

Therefore

$$dT_{\alpha_1} \wedge \ldots \wedge dT_{\alpha_k} = \sum_{\beta \in I^k_n} \sum_{\pi \in S_\beta} (-1)^\pi (D_{\pi(\beta_1)} T_{\alpha_1}) \ldots (D_{\pi(\beta_k)} T_{\alpha_k}) \mathbf{e}^{(n)}_\beta$$

where S_β is the group of all permutations on $\beta = (\beta_1, \ldots, \beta_k)$. Now recall a standard definition of the determinant:

$$\det A = \sum_{\pi \in S_{\{1,\ldots,k\}}} (-1)^\pi a_{1,\pi(1)} \ldots a_{k,\pi(k)}.$$

□

4.3.15 Changing variables. Suppose $T \in C^1(\Omega, \Omega')$ where Ω' is an open set in $\mathbb{R}^m$. If ω is a 0-form we set $\omega_T = \omega \circ T$. If $k \in \mathbb{N}$ and ω is a k-form in Ω' with basic representative

$$\tilde{\omega} = \sum_{\alpha \in I^k_m} \tilde{\omega}_\alpha \mathbf{e}^{(m)}_\alpha$$

we define a k-form ω_T in Ω by setting

$$\tilde{\omega}_T = \sum_{\alpha \in I^k_m} (\tilde{\omega}_\alpha \circ T) dT_{\alpha_1} \wedge \ldots \wedge dT_{\alpha_k}.$$

According to 4.3.14 we have

$$\tilde{\omega}_T = \sum_{\beta \in I^k_n} \Big(\sum_{\alpha \in I^k_m} (\tilde{\omega}_\alpha \circ T) J(T, \alpha, \beta) \Big) \mathbf{e}^{(n)}_\beta.$$

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 $\omega(y) = y_2 y_3 \mathbf{e}^{(3)}_{1,2} + y_1\ \mathbf{e}^{(3)}_{2,3}$. $dT_1 \wedge dT_2 = 2(x_1^2 - x_2^2)\mathbf{e}^{(2)}_{1,2}$, $dT_1 \wedge dT_3 = 2x_1 \mathbf{e}^{(2)}_{1,2}$, and $dT_2 \wedge dT_3 = x_2 \mathbf{e}^{(2)}_{1,2}$. Then $\omega_T = (2x_1^3 x_2^2 - 2x_1 x_2^4 + x_1^2 x_2 + x_2^3)\mathbf{e}^{(2)}_{1,2}$.

4.3.16 Basic properties of variable changes. Let Ω and Ω' be open sets in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Suppose that $T \in C^1(\Omega, \Omega')$ and that ω is a p-form and λ a q-form in Ω'. 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 ω is of class C^1 and $T \in C^2(\Omega, \Omega')$ then ω_T is of class C^1 and $d(\omega_T) = (d\omega)_T$.

Sketch of proof: Let ω and λ have the basic representatives $\sum_{\alpha \in I^p_m} \omega_\alpha \mathbf{e}^{(m)}_\alpha$ and $\sum_{\beta \in I^q_m} \lambda_\beta \mathbf{e}^{(m)}_\beta$, respectively.

To prove the first claim is simple. For the second note that

$$\omega \wedge \lambda = \sum_{\alpha \in I^p_m} \sum_{\beta \in I^q_m} \omega_\alpha \lambda_\beta (-1)^{\pi_{\alpha,\beta}} \mathbf{e}^{(m)}_\gamma$$

where $\gamma = [\alpha, \beta]$. Therefore

$$(\omega \wedge \lambda)_T = \sum_{\alpha \in I_m^p} \sum_{\beta \in I_m^q} (\omega_\alpha \circ T)(\lambda_\beta \circ T)(-1)^{\pi_{\alpha,\beta}} dT_{\gamma_1} \wedge \dots \wedge dT_{\gamma_{p+q}}$$

$$= \sum_{\alpha \in I_m^p} \sum_{\beta \in I_m^q} (\omega_\alpha \circ T)(\lambda_\beta \circ T) dT_{\alpha_1} \wedge \dots \wedge dT_{\beta_q}.$$

On the other hand, using the distributive laws and the fact that $dT \wedge f = f \wedge dT$ for any 0-form f and 1-form dT,

$$\omega_T \wedge \lambda_T = \sum_{\alpha \in I_m^p} (\omega_\alpha \circ T)\ dT_{\alpha_1} \wedge \dots \wedge dT_{\alpha_p} \wedge \lambda_T$$

$$= \sum_{\alpha \in I_m^p} \sum_{\beta \in I_m^q} (\omega_\alpha \circ T)(\lambda_\beta \circ T) dT_{\alpha_1} \wedge \dots \wedge dT_{\alpha_p} \wedge dT_{\beta_1} \wedge \dots \wedge dT_{\beta_q}.$$

First prove the third claim when $\omega = f$ is a 0-form with the help of the chain rule. For the general case we only have to consider $\omega = f\mathbf{e}_\alpha^{(m)}$. Then $\omega_T = (f \circ T) dT_{\alpha_1} \wedge \dots \wedge dT_{\alpha_p}$ and $d\omega = \sum_{j=1}^m (D_j f)(-1)^{\pi_{j,\alpha}} \mathbf{e}_{[j,\alpha]}^{(m)}$. Now use the product rule from 4.3.13 and $d^2 = 0$ to obtain

$$d(\omega_T) = d(f \circ T) \wedge dT_{\alpha_1} \wedge \dots \wedge dT_{\alpha_p}$$

and undoing the ordering

$$(d\omega)_T = \sum_{j=1}^m ((D_j f) \circ T) dT_j \wedge dT_{\alpha_1} \wedge \dots \wedge dT_{\alpha_p}.$$

As above the chain rule gives that $\sum_{j=1}^m ((D_j f) \circ T) dT_j = d(f \circ T)$ which completes the proof. $\square$

4.3.17 Compositions of variable changes. Suppose Ω, Ω', and Ω'' 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 ω a k-form in Ω''. Then $ST = S \circ T \in C^1(\Omega, \Omega'')$ and

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

Sketch of proof: This is trivial when $k = 0$. If $\omega = \mathbf{e}_q^{(p)}$ where $q \in \{1, \dots, p\}$, then $\omega_S = \sum_{j=1}^m (D_j S_q) \mathbf{e}_j^{(m)}$ and

$$(\omega_S)_T = \sum_{j=1}^m ((D_j S_q) \circ T) dT_j = \sum_{j=1}^m ((D_j S_q) \circ T) \sum_{\ell=1}^n D_\ell T_j \mathbf{e}_\ell^{(n)}$$

$$= \sum_{\ell=1}^n \Big(\sum_{j=1}^m ((D_j S_q) \circ T) D_\ell T_j \Big) \mathbf{e}_\ell^{(n)} = \sum_{\ell=1}^n D_\ell (S_q \circ T) \mathbf{e}_\ell^{(n)} = \omega_{ST}.$$

The general case follows now with the aid of 4.3.16: In

$$\omega = \sum_{\alpha \in I_p^k} \omega_\alpha \mathbf{e}_\alpha^{(p)} = \sum_{\alpha \in I_p^k} \omega_\alpha \wedge \mathbf{e}_{\alpha_1}^{(p)} \wedge \dots \wedge \mathbf{e}_{\alpha_k}^{(p)}$$

we may treat each summand and each factor separately. $\square$

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

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

Sketch of proof: First suppose $n = k$, that $\phi = \mathbb{1}$, the identity on Q^k, and that $\omega = f\mathbf{e}_\alpha^{(m)}$. Then $T \circ \phi = T|_{Q^k}$ is a k surface in Ω' and

$$\int_{\mathbb{1}} \omega_T = \int_{\mathbb{1}} (f \circ T) J(T, \alpha, (1, ..., k)) \mathbf{e}_{1,...,k}^{(k)} = \int_{Q_k} (f \circ T) J(T, \alpha, (1, ..., k)) = \int_{T\circ\mathbb{1}} \omega.$$

Using the linearity of the integral we have that $\int_{\mathbb{1}} \omega_T = \int_T \omega$ for all k-forms in Ω'. Now note that

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

$\square$

4.4. Stokes' theorem

4.4.1 Chains. Let $\mathcal{S}_k(\Omega)$ denote the set of k-surfaces in Ω. A k-chain in Ω 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 Ω.

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_ℓ is of class C^r for some r in $\mathbb{N}$.

4.4.2 Integration over chains. Let γ be a k-chain in Ω and ω a k-form in Ω. Since $\gamma = \sum_{j=1}^\ell n_j[\phi_j]$ with k-surfaces ϕ_j we define

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

4.4.3 Boundaries of affine simplices. Suppose $1 \leq \ell \leq k$. Recall that the points $e_j^{(\ell)}$, $j = 1, ..., \ell$ form the standard basis of $\mathbb{R}^\ell$. We define also $e_0^{(\ell)}$ to be the zero vector in $\mathbb{R}^\ell$. Thus the points $e_j^{(\ell)}$, $j = 0, ..., \ell$ are the vertices of Q^ℓ.

Let p_0, p_1, ..., p_ℓ be points in $\mathbb{R}^k$. The ℓ-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 ℓ-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 ℓ-simplex $\langle p_0, ..., p_\ell\rangle$ has $\ell+1$ *faces* $\langle p_0, ..., \cancel{p_j}, ..., p_\ell\rangle$, $j = 0, ..., \ell$. Each of these faces is an affine $(\ell-1)$-simplex and the chain

$$\sum_{j=0}^{k}(-1)^j[p_0, ..., \cancel{p_j}, ..., p_\ell]$$

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

In particular, $\mathbb{1}^{(\ell)} = \langle e_0^{(\ell)}, ..., e_\ell^{(\ell)}\rangle$ is the identity function on Q^ℓ and its boundary is

$$\partial\mathbb{1}^{(\ell)} = \sum_{j=0}^{k}(-1)^j[\sigma_j^{(\ell)}]$$

where $\sigma_j^{(\ell)} = \langle e_0^{(\ell)}, ..., \cancel{e_j^{(\ell)}}, ..., e_\ell^{(\ell)}\rangle$. Note that $\sigma_j^{(\ell)}$ maps Q^{k-1} into (but not onto) Q^k.

4.4.4 Boundaries of surfaces and chains. If ϕ is a k-surface in Ω, we define the boundary $\partial[\phi]$ of $[\phi]$ by

$$\partial[\phi] = \sum_{j=0}^{k}(-1)^j[\phi\circ\sigma_j^{(\ell)}]$$

Thus $\partial[\phi]$ is a $(k-1)$-chain in Ω.

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

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

4.4.5 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.

4.4.6 $\partial^2 = 0$. For any k-chain ψ we have $\partial^2\psi = \partial(\partial\psi) = 0$.

Sketch of proof: It is enough to prove this claim when ψ is a k-surface. Then

$$\partial^2[\psi] = \sum_{j=0}^{k}(-1)^j\partial[\psi\circ\sigma_j^{(k)}]$$

where, as before, $\sigma_j^{(k)} = \langle e_0^{(k)}, ..., \cancel{e_j^{(k)}}, ..., e_k^{(k)}\rangle$. Hence

$$\partial^2[\psi] = \sum_{j=0}^{k}(-1)^j\Big(\sum_{\ell=1}^{j-1}(-1)^\ell[\psi\circ\sigma_{\ell,j}^{(k)}] + \sum_{\ell=j+1}^{k}(-1)^{\ell-1}[\psi\circ\sigma_{j,\ell}^{(k)}]\Big)$$

where, for $p<q$,

$$\sigma_{p,q}^{(k)} = \langle e_0^{(k)}, ..., \cancel{e_p^{(k)}}, ..., \cancel{e_q^{(k)}}, ..., e_k^{(k)}\rangle.$$

Hence

$$\partial^2[\psi] = \sum_{1\le \ell<j\le k} (-1)^{j+\ell}[\psi\circ\sigma^{(k)}_{\ell,j}] + \sum_{1\le j<\ell\le k} (-1)^{j+\ell-1}[\psi\circ\sigma^{(k)}_{j,\ell}] = 0. \qquad \square$$

Rudin's proof is not quite complete and it is unclear how to complete it.

4.4.7 The fundamental theorem of calculus. Suppose $n = 1$, $\Omega = \mathbb{R}$ and $a, b \in \mathbb{R}$ with $a < b$. Let ϕ 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 $f'\mathbf{e}_1$ 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 ϕ 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 $\partial\phi = [\phi(1)] - [\phi(0)]$

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

$\square$

4.4.8 Stokes' theorem. Let Ω be an open subset of $\mathbb{R}^n$ and $k \in \mathbb{N}$. If ϕ is a k-chain in Ω and ω is a $(k-1)$-form of class C^1 in Ω, then

$$\int_\phi d\omega = \int_{\partial\phi} \omega. \tag{8}$$

Sketch of proof: Note that $d\omega$ is a k-form in Ω and that $\partial\phi$ is a chain of $(k-1)$-surfaces.

We begin by showing the claim when $n = k$, $Q^k \subset \Omega$, $\phi = \mathbb{1}$ on Q^k, and $\omega = f\mathbf{e}^{(n)}_\alpha$ where $f \in C^1(\Omega, \mathbb{R})$ and $\alpha = (1, ..., \cancel{r}, ..., k)$.

Recall that with $\sigma_j = \langle e_0, ..., \cancel{e_j}, ..., e_k\rangle$ the boundary of $\phi = \mathbb{1}$ is given by $\sum_{j=0}^k (-1)^j[\sigma_j]$. Here we have

$$\sigma_0 = (1 - t_1 - ... - t_{k-1}, t_1, t_2, ..., t_{k-1})^\top$$

and, if $j \ne 0$,

$$\sigma_j(t) = (t_1, ..., t_{j-1}, 0, t_j, ..., t_{k-1})^\top.$$

Therefore

$$J(\sigma_j, \alpha) = \begin{cases} (-1)^{r-1} & \text{if } j = 0 \\ 1 & \text{if } j = r, \\ 0 & \text{if } 0 \ne j \ne r. \end{cases}$$

For $j = 0$ delete row r in $(\sigma^{(0)})'$. For $r > 1$ flip row 1 with row 2, then 2 with 3 until it becomes row r. This creates a matrix for which it is easy to see that the determinant is -1. Since $r - 2$ flips are necessary we get the claim for $j = 0$. For $j = r$ one has an identity

matrix. Otherwise the matrix has a row of zeros.
Hence we get for the right-hand side of (8)

$$\int_{\partial \mathbb{1}} f\mathbf{e}_\alpha^{(n)} = \sum_{j=0}^{k}(-1)^j \int_{\sigma_j} f\mathbf{e}_\alpha^{(n)} = (-1)^{r-1}\int_{Q^{k-1}} f\circ\sigma^{(0)} + (-1)^r \int_{Q^{k-1}} f\circ\sigma^{(r)}. \tag{9}$$

To compute the left-hand side note first that $d(f\mathbf{e}_\alpha) = (D_r f)\mathbf{e}_{r,\alpha}$ and that $J(\mathbb{1},(r,\alpha)) = (-1)^{r-1}$. Hence $\int_{\mathbb{1}}(D_r f)\mathbf{e}_{r,\alpha} = (-1)^{r-1}\int_{Q^k} D_r f$. To evaluate this we use iterated integrals and integrate first over the r-th variable in Q^k. Fix a point $t = (t_1, ..., t_{k-1}) \in Q^{k-1}$. Then $(t_1, ..., t_{r-1}, s, t_r, ..., t_{k-1})$ is in Q^k precisely if s is in $[0, s_0]$ where $s_0 = 1 - t_1 - ... - t_{k-1}$. Thus, using the fundamental theorem of calculus,

$$\begin{aligned}\int_{\mathbb{1}}(D_r f)\mathbf{e}_{r,\alpha} &= (-1)^{r-1}\int_{Q^{k-1}}\int_{[0,s_0]}(D_r f)(t_1, ..., t_{r-1}, s, t_r, ..., t_{k-1})\,ds\,d(t_1, ..., t_{k-1})\\ &= (-1)^{r-1}\int_{Q^{k-1}}\big(f(t_1, ..., t_{r-1}, s_0, t_r, ..., t_{k-1}) - f(t_1, ..., t_{r-1}, 0, t_r, ..., t_{k-1})\big)\\ &= (-1)^{r-1}\int_{Q^{k-1}}\big(f((\sigma_0\circ T)(t))) - f(\sigma_r(t)\big)\end{aligned}$$

when $T : \mathbb{R}^{k-1} \to \mathbb{R}^{k-1}$ is the linear transformation defined by the requirement $(\sigma_0\circ T)(t) = (t_1, ..., t_{r-1}, s_0, t_r, ..., t_{k-1})$. $T(t) = (t_2, ..., t_{r-1}, 1 - t_1 - ... - t_{k-1}, t_r, ..., t_{k-1})$. Since $T' = T$ and $|\det T| = 1$ we have, according to 3.0.14, $\int_{Q^{k-1}} f\circ\sigma_0\circ T|\det T| = \int_{Q^{k-1}} f\circ\sigma_0$. Hence

$$\int_{\mathbb{1}}(D_r f)\mathbf{e}_{r,\alpha} = (-1)^{r-1}\int_{Q^{k-1}}(f\circ\sigma_0 - f\circ\sigma_r). \tag{10}$$

Finally note that (9) and (10) are identical to finish the proof under the current special circumstances.

The claim also holds for a general $(k-1)$-form ω by the linearity of the integrals.

Next suppose that ϕ is a general k-surface in Ω without assuming $n = k$. Hence see footnote on p.5 there is an open set $U \subset \mathbb{R}^k$ and including Q^k such that ϕ can be extended to U as a continuously differentiable function. If we call the extension T we have $\phi = T\circ\mathbb{1}$ where $\mathbb{1}$ is the identity function on Q^k. We have now, using 4.3.18, part (3) of 4.3.16, and what we just proved,

$$\int_\phi d\omega = \int_{T\circ\mathbb{1}} d\omega = \int_{\mathbb{1}}(d\omega)_T = \int_{\mathbb{1}} d(\omega_T) = \sum_{j=0}^{k}(-1)^j\int_{\sigma_j}\omega_T,$$

where ω_T is a $(k-1)$-form in Ω. On the other hand we have $\partial[\phi] = \sum_{j=0}^{k}(-1)^j[\phi_j]$ where $\phi_j = \phi\circ\sigma_j = T\circ\sigma_j$ using $\phi = T$ on Q^k in the last equality. Therefore

$$\int_{\partial[\phi]}\omega = \sum_{j=0}^{k}(-1)^j\int_{[\phi_j]}\omega = \sum_{j=0}^{k}(-1)^j\int_{T\circ\sigma_j}\omega = \sum_{j=0}^{k}(-1)^j\int_{\sigma_j}\omega_T.$$

This completes the proof for general k-surfaces.

The last thing to be mentioned is that claim now follows also for general chains by the definition of integration over chains. □

4.4.9 Green's theorem. Let $n = 2$. If ϕ is 2-chain in $\Omega \subset \mathbb{R}^2$ and if $\omega = f\mathbf{e}_1 + g\mathbf{e}_2$ is a 1-form of class C^1, then $d\omega = (D_1 g - D_2 f)\mathbf{e}_{1,2}$. Stokes' theorem is called Green's theorem in this case

$$\int_\phi (D_1 g - D_2 f)\mathbf{e}_{1,2} = \int_{\partial\phi} (f\mathbf{e}_1 + g\mathbf{e}_2).$$

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

$$\int_\phi \mathbf{e}_{1,2} = \frac{1}{2}\int_{\partial\phi} (x_1\mathbf{e}_2 - x_2\mathbf{e}_1).$$

This quantity is called the area of (the range of) ϕ.

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 $a, b, c > 0$.

4.5. Closed and exact forms

4.5.1 Closed and exact forms. A form ω is called *exact* if there is another form λ such that $d\lambda = \omega$. A continuously differentiable form ω is called *closed* if $d\omega = 0$.

Every continuously differentiable exact form is closed but the converse is not true in general. See next topic.

4.5.2 A closed form which is not exact. Let $\Omega = \mathbb{R}^2 \setminus \{0\}$ and consider the 1-form

$$\omega = \frac{y}{x^2+y^2}\mathbf{e}_1 - \frac{x}{x^2+y^2}\mathbf{e}_2.$$

Then ω is closed but not exact. If $f = \arctan(x/y)$, then $f_x = y/(x^2+y^2)$ and $f_y = -x/(x^2+y^2)$. Since f is not defined on the x-axis ω is not exact in $\mathbb{R}^2 \setminus \{0\}$. It is how exact in both the upper and the lower half of the x-y-plane.

4.5.3 Poincaré's lemma. If Ω is star-shaped, then every closed form in Ω is exact.

Sketch of proof: Given a point $p \in \Omega$ and any k-form $\mu = \sum_{\alpha\in I_n^k} \mu_\alpha \mathbf{e}_\alpha$ define the $(k-1)$-form

$$\mathfrak{i}(\mu) = \sum_{\alpha\in I_n^k}\sum_{j=1}^k (-1)^{j-1}(x_{\alpha_j} - p_{\alpha_j}) \int_{[0,1]} t^{k-1}\mu_\alpha(p + t(x-p))dt\ \mathbf{e}_{(\alpha_1,\ldots,\cancel{\alpha_j},\ldots,\alpha_k)}.$$

Then verify that $d\mathfrak{i}(\mu) + \mathfrak{i}(d\mu) = \mu$.

Thus, if $d\mu = 0$, then $\mu = d\mathfrak{i}(\mu)$. □

The example $\omega = \frac{2x}{x^2+y^2}\mathbf{e}_1 + \frac{2y}{x^2+y^2}\mathbf{e}_2 = d\ln(x^2+y^2)$ defined on $\mathbb{R}^2 \setminus \{0\}$ shows that Poincaré's lemma provides only a sufficient condition for exactness.

Sketch of proof: Given a point $p \in \Omega$ and any k-form $\mu = \sum_{\alpha\in I_n^k} \mu_\alpha \mathbf{e}_\alpha$ define the $(k-1)$-form

$$\mathfrak{i}(\mu) = \sum_{\alpha\in I_n^k}\sum_{j=1}^k (-1)^{j-1}(x_{\alpha_j} - p_{\alpha_j}) \int_{[0,1]} t^{k-1}\mu_\alpha(p + t(x-p))dt\ \mathbf{e}_{(\alpha_1,\ldots,\cancel{\alpha_j},\ldots,\alpha_k)}.$$

We compute $\mathfrak{i}(d\mu)$ and, using 3.0.15, $d\mathfrak{i}(\mu)$ in the special case when $k = 1$. With the abbreviation $r(x,t) = p + t(x-p)$ we have now

$$d\mu = \sum_{1\le\alpha<\beta\le n} (D_\alpha\mu_\beta - D_\beta\mu_\alpha)\mathbf{e}_{\alpha,\beta}$$

and

$$\mathfrak{i}(\mu) = \sum_{\alpha=1}^{n} (x_\alpha - p_\alpha) \int_{[0,1]} \mu_\alpha(r(x,t))dt.$$

Thus

$$\mathfrak{i}(d\mu) = \sum_{1\le\alpha<\beta\le n} \int_{[0,1]} t(D_\alpha\mu_\beta - D_\beta\mu_\alpha)(r(x,t))\ dt\ \big((x_\alpha - p_\alpha)\mathbf{e}_\beta - (x_\beta - p_\beta)\mathbf{e}_\alpha\big)$$

and

$$\begin{aligned} d\mathfrak{i}(\mu) &= \sum_{\alpha=1}^{n} \int_{[0,1]} \mu_\alpha(r(x,t))\ dt\ \mathbf{e}_\alpha + \sum_{\alpha=1}^{n}(x_\alpha - p_\alpha)\sum_{\beta=1}^{n}\int_{[0,1]} t(D_\beta\mu_\alpha)(r(x,t))\ dt\ \mathbf{e}_\beta \\ &= \sum_{\alpha=1}^{n}\int_{[0,1]} \big[\mu_\alpha(r(x,t)) + (x_\alpha - p_\alpha)t(D_\alpha\mu_\alpha)(r(x,t))\big] dt\ \mathbf{e}_\alpha \\ &+ \sum_{1\le\alpha<\beta\le n}\int_{[0,1]} [(x_\alpha - p_\alpha)t(D_\beta\mu_\alpha)(r(x,t))\mathbf{e}_\beta + (x_\beta - p_\beta)t(D_\alpha\mu_\beta)(r(x,t))\mathbf{e}_\alpha]dt. \end{aligned}$$

Adding the last two expressions gives

$$\begin{aligned} d\mathfrak{i}(\mu) + \mathfrak{i}(d\mu) &= \sum_{\alpha=1}^{n}\int_{[0,1]} \big[\mu_\alpha(r(x,t)) + (x_\alpha - p_\alpha)t(D_\alpha\mu_\alpha)(r(x,t))\big] dt\mathbf{e}_\alpha \\ &+ \sum_{1\le\alpha<\beta\le n}\int_{[0,1]} t[(x_\alpha - p_\alpha)(D_\alpha\mu_\beta)(r(x,t))\mathbf{e}_\beta + (x_\beta - p_\beta)(D_\beta\mu_\alpha)(r(x,t))\mathbf{e}_\alpha]dt \end{aligned}$$

This should be μ. The general case may be worked similarly, it is, however, even more involved. $\square$

4.6. Vector Analysis

Obviously, the most important case for physics is the case $n = 3$ which we will now investigate. Therefore, in this section, Ω denotes an open subset of $\mathbb{R}^3$. We often use bold symbols to denote vectors in $\mathbb{R}^3$.

4.6.1 Dot and cross product in $\mathbb{R}^3$. Given two vectors $\mathbf{a} = (a_1, a_2, a_3)^\top$ and $\mathbf{b} = (b_1, b_2, b_3)^\top$ in $\mathbb{R}$ recall that their dot product $\mathbf{a}\cdot\mathbf{b}$ and their cross product $\mathbf{a}\times\mathbf{b}$ are respectively defined by

$$\mathbf{a}\cdot\mathbf{b} = a_1b_1 + a_2b_2 + a_3b_3$$

and

$$\mathbf{a}\times\mathbf{b} = (a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1)^\top.$$

4.6.2 Vector fields. A *vector field* on Ω is a continuous function from Ω to $\mathbb{R}^3$.

Given a vector field $\mathbf{F}$ we have the 1-form

$$\omega_{\mathbf{F}}^{(1)} = F_1\mathbf{e}_1 + F_2\mathbf{e}_2 + F_3\mathbf{e}_3$$

and the 2-form

$$\omega_{\mathbf{F}}^{(2)} = F_1\mathbf{e}_{2,3} + F_2\mathbf{e}_{3,1} + F_3\mathbf{e}_{1,2}.$$

Conversely, every 1-form and every 2-form gives rise to a vector field.

4.6.3 Gradient, divergence and curl. Suppose u is a continuously differentiable real-valued function on Ω. Then $\operatorname{grad} u = \nabla u = (D_1 u, D_2 u, D_3 u)^\top$ is a vector field, called the *gradient* of u.

If $\mathbf{F}$ is a continuously differentiable vector field on Ω, then the scalar-valued function

$$\operatorname{div} \mathbf{F} = \nabla \cdot \mathbf{F} = D_1 F_1 + D_2 F_2 + D_3 F_3$$

is called the *divergence* of $\mathbf{F}$ while the vector field

$$\operatorname{curl} \mathbf{F} = \nabla \times \mathbf{F} = (D_2 F_3 - D_3 F_2, D_3 F_1 - D_1 F_3, D_1 F_2 - D_2 F_1)^\top$$

is called the *curl* of $\mathbf{F}$.

Note that

$$du = \omega^{(1)}_{\operatorname{grad} u}, \quad d\omega^{(1)}_{\mathbf{F}} = \omega^{(2)}_{\operatorname{curl} \mathbf{F}}, \quad \text{and} \quad d\omega^{(2)}_{\mathbf{F}} = \operatorname{div} \mathbf{F}\, \mathbf{e}_{1,2,3}.$$

4.6.4 Reparametrization. Suppose ϕ is a k-surface in $\mathbb{R}^n$ and $T : Q^k \to Q^k$ is continuously differentiable and bijective. Remember footnote on p. 5. Then $\psi = \phi \circ T$ is also a k-surface in $\mathbb{R}^n$ and $\psi(Q^k) = \phi(Q^k)$, a subset of $\mathbb{R}^n$ which we denote by S. We call both ϕ and ψ parametrizations of S.

Since

$$\int_\psi \omega = \int_{Q^k} \sum_{\alpha \in I^k_n} \omega_\alpha \circ \phi \circ T J(\phi \circ T, \alpha)$$

we have a closer look at $J(\phi \circ T, \alpha)$. Denote the vector $(\phi_{\alpha_1}, ..., \phi_{\alpha_k})^\top$ by Φ. Then

$$J(\phi \circ T, \alpha) = \det(\Phi \circ T)' = \det(\Phi' \circ T) T' = ((\det \Phi') \circ T)(\det T') = J(\phi, \alpha) \det T'.$$

Hence the change of variables formula 3.0.14 shows that

$$\int_\psi \omega = \int_{Q^k} \sum_{\alpha \in I^k_n} (\omega_\alpha \circ \phi \circ T) J(\phi \circ T, \alpha) = \int_{Q^k} \sum_{\alpha \in I^k_n} (\omega_\alpha \circ \phi) J(\phi, \alpha) = \int_\phi \omega$$

assuming that $\det T' > 0$ everywhere. Otherwise, if $\det T' < 0$ everywhere, we get $\int_\psi \omega = -\int_\phi \omega$.

In vector analysis one thinks of curves and surfaces as parametrizable sets. A curve requires one parameter and a surface two. The parameter domains are finitely many copies of Q^1 or Q^2, respectively. We also require that restrictions to the interior of the parameter domains are injective. Thus a closed curve, for instance, is admissible, as is the figure 8.

This should perhaps be moved to an earlier section.

4.6.5 Potential functions. If u is a twice continuously differentiable real-valued function on Ω, we have

$$0 = d^2 u = d\omega^{(1)}_{\operatorname{grad} u} = \omega^{(2)}_{\operatorname{curl} \operatorname{grad} u}.$$

Hence $\operatorname{curl} \operatorname{grad} u = 0$.

Conversely, assume that $\mathbf{F}$ is a continuously differentiable vector field on a star-shaped set Ω such that $\operatorname{curl} \mathbf{F} = 0$. Then

$$0 = \omega^{(2)}_{\operatorname{curl} \mathbf{F}} = d\omega^{(1)}_{\mathbf{F}},$$

$\omega^{(1)}_{\mathbf{F}}$ is closed. Poincaré's lemma shows that the 1-form $\omega^{(1)}_{\mathbf{F}}$ is exact, i.e., there is a twice continuously differentiable 0-form u such that

$$\omega^{(1)}_{\operatorname{grad} u} = du = \omega^{(1)}_{\mathbf{F}}$$

i.e., $\mathbf{F} = \operatorname{grad} u$.

A function continuously differentiable function u such that $\mathbf{F} = -\operatorname{grad} u$ is called a *potential function* for $\mathbf{F}$ and $\mathbf{F}$ is then called a *conservative vector field*.

4.6.6 Finding the potential of a conservative vector field. Suppose $\mathbf{F}$ is a conservative vector field in Ω and that any two points in Ω can be connected by a smooth path in Ω. Thus there is a continuously differentiable function u so that $\mathbf{F} = -\operatorname{grad} u$. Indeed, define

$$v(\mathbf{r}) = -\int_{[0,1]} \mathbf{F}(\gamma(t)) \cdot \gamma'(t) dt = -\int_\gamma \omega_{\mathbf{F}}^{(1)}$$

where γ is a smooth path in Ω (1-surface) connecting a fixed point $\mathbf{r}_0$ to $\mathbf{r}$. Then

$$v(\mathbf{r}) = -\int_{[0,1]} \mathbf{F}(\gamma(t)) \cdot \gamma'(t) dt = \int_{[0,1]} (u \circ \gamma)'(t) dt = u(\gamma(1)) - u(\gamma(0)).$$

It follows that $v - u$ is constant and hence the v is also a potential function for $\mathbf{F}$.

This is an extension of Poincaré's lemma for $k = 1$.

4.6.7 Vector potentials. Suppose $\mathbf{F}$ is a twice continuously differentiable vector field in Ω. Then

$$0 = d^2\omega_{\mathbf{F}}^{(1)} = d\omega_{\operatorname{curl}\mathbf{F}}^{(2)} = (\operatorname{div}\operatorname{curl}\mathbf{F})\mathbf{e}_{1,2,3}.$$

Hence $\operatorname{div}\operatorname{curl}\mathbf{F} = 0$.

Conversely, if $\mathbf{G}$ is a continuously differentiable vector field on a star-shaped set Ω such that $\operatorname{div}\mathbf{G} = 0$, then $d\omega_{\mathbf{G}}^{(2)} = 0$, i.e., $\omega_{\mathbf{G}}^{(2)}$ is a closed 2-form. Again Poincaré's lemma guarantees the existence of a 1-form λ such that $d\lambda = \omega_{\mathbf{G}}^{(2)}$. Define $\mathbf{F}$ to be the vector field whose components are the coefficients of $\mathbf{e}_j$ of $\lambda = \omega_{\mathbf{F}}^{(1)} = F_1\mathbf{e}_1 + F_2\mathbf{e}_2 + F_3\mathbf{e}_3$.

The vector field $\mathbf{F}$ is then called a *vector potential* for $\mathbf{G}$.

Poincaré's lemma gives $\mathbf{F} = \mathbf{I} \times \mathbf{r}$ where $I_j = \int_{[0,1]} tG_j(\mathbf{r}_0 + t(\mathbf{r} - \mathbf{r}_0)dt,\ j = 1, 2, 3$ as one possibility.

4.6.8 Positively oriented n-surfaces in $\mathbb{R}^n$. An n-surface ϕ in $\mathbb{R}^n$ is called *positively oriented* if $\det\phi' > 0$ everywhere on the parameter domain Q^n of ϕ.

4.6.9 Orientable surfaces. Let $\phi : Q^2 \to \mathbb{R}^3$ be a parametrization of a surface S in $\mathbb{R}^3$ and define $\mathbf{n} = (D_1\phi) \times (D_2\phi)$. One verifies that $\mathbf{n} = (J(\phi, (2, 3)), J(\phi, (3, 1)), J(\phi, (1, 2)))^\top$. Note that $\mathbf{n}$ is perpendicular to the vectors $D_1\phi$ and $D_2\phi$ which are tangent to the surface S. The vector $\mathbf{n}(s, t)$ is called a *normal vector* to S at the point $\phi(s, t)$.

The normal vectors at (s, t) of any parametrization $\phi \circ T$ of S point either into the same direction as $\mathbf{n}(s, t)$ or in the opposite direction depending on whether $\det T'(s, t)$ is positive or negative.

Suppose $\ell : Q^1 \to \mathbb{Q}^2$ is a smooth path such that $\phi(\ell(0)) = \phi(\ell(1))$. A surface S is called *orientable* if $\mathbf{n}(\ell(0)) = \mathbf{n}(\ell(1))$ for all ℓ.

4.6.10 The Möbius strip. Consider the chain

$$\phi(s, t) = \begin{pmatrix} 2\cos(2\pi t) + (2s - 1)\cos(\pi t) \\ 2\sin(2\pi t) + (2s - 1)\cos(\pi t) \\ (2s - 1)\sin(\pi t) \end{pmatrix}$$

for $(s, t) \in [0, 1] \times [0, 1]$. Let $\ell(u) = (1/2, u)^\top$. Then

$$\phi(\ell(u)) = \begin{pmatrix} 2\cos(2\pi u) \\ 2\sin(2\pi u) \\ 0 \end{pmatrix}$$

and, in particular, $\phi(\ell(0)) = \phi(\ell(1))$. Also

$$\mathbf{n}(s,t) = 2\pi \begin{pmatrix} 2\sin(\pi t) - 2\sin(3\pi t) + 2s - 1 \\ 2\cos(3\pi t) - 2\cos(\pi t) - 2s + 1 \\ 4\cos(\pi t)(\cos(2\pi t) + \sin(2\pi t)) \end{pmatrix}$$

so that $\mathbf{n}(\ell(0)) = -\mathbf{n}(\ell(1))$.

We conclude that the Möbius strip is not orientable.

4.6.11 Positively oriented boundaries. The boundary of a positively oriented 3-surface is orientable and is called *positively oriented*. For $\phi = \mathbb{1}^{(3)} : Q^3 \to Q^3$ we have

$$\sigma_0^{(3)}(s,t) = \left(\begin{smallmatrix} 1-s-t \\ s \\ t \end{smallmatrix}\right), \quad \sigma_1^{(3)}(s,t) = \left(\begin{smallmatrix} 0 \\ s \\ t \end{smallmatrix}\right), \quad \sigma_2^{(3)}(s,t) = \left(\begin{smallmatrix} s \\ 0 \\ t \end{smallmatrix}\right), \quad \sigma_3^{(3)}(s,t) = \left(\begin{smallmatrix} s \\ t \\ 0 \end{smallmatrix}\right).$$

The corresponding normal vectors are

$$\mathbf{n}_0 = (1,1,1)^\top, \quad \mathbf{n}_1 = (1,0,0)^\top, \quad \mathbf{n}_2 = (0,-1,0)^\top, \quad \mathbf{n}_3 = (0,0,1)^\top.$$

It follows that the positive orientation of $\partial\mathbb{1}^{(3)}$ points outward. By continuity this is also true for the boundaries of any positively oriented 3-surface.

The boundary of a orientable 2-surface is also called positively oriented. For instance if $\phi(s,t) = (s,t,0)$, i.e., if ϕ is the identity, then the (positive) orientation of $\partial\phi$ is counter-clockwise. $\sigma_0^{(2)}(t) = (1-t,t)^\top$, $\sigma_1^{(2)}(t) = (0,t)^\top$, $\sigma_2^{(2)}(t) = (t,0)^\top$.If $\phi(s,t) = (t,s,0)^\top$, then the (positive) orientation of $\partial\phi$ is clockwise.

4.6.12 Line integrals. Suppose $\phi : Q^1 \to \mathbb{R}^3$ is a parametrization of a curve C in $\mathbb{R}^3$. Then $\int_C \mathbf{F} \cdot d\mathbf{r}$ is called the *line integral* of $\mathbf{F}$ along C and is defined by

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_{Q^1} \mathbf{F} \circ \phi \cdot \phi'.$$

Hence

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_\phi \mathbf{F} \cdot \left(\begin{smallmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \mathbf{e}_3 \end{smallmatrix}\right) = \int_\phi \omega_{\mathbf{F}}^{(1)}.$$

4.6.13 Flux. Suppose $\phi : Q^2 \to \mathbb{R}^3$ is a parametrization of a surface S in $\mathbb{R}^3$. Then $\iint_S \mathbf{F} \cdot d\mathbf{S}$ is called the *flux* of $\mathbf{F}$ through S and is defined by

$$\iint_S \mathbf{F} \cdot d\mathbf{S} = \iint_{Q^2} \mathbf{F} \circ \phi \cdot \mathbf{n}.$$

We now obtain

$$\begin{aligned} \iint_S \mathbf{F} \cdot d\mathbf{S} = \iint_{Q^2} &[(F_1 \circ \phi)J(\phi,(2,3)) + (F_2 \circ \phi)J(\phi,(3,1)) + (F_3 \circ \phi)J(\phi,(1,2)) \\ &= \int_\phi (F_1\mathbf{e}_{2,3} + F_2\mathbf{e}_{3,1} + F_3\mathbf{e}_{1,2}) = \int_\phi \omega_{\mathbf{F}}^{(2)}. \end{aligned}$$

4.6.14 Line elements. The vector $\phi'(t)$ occurring in 4.6.12 is a tangent vector to the curve C at the point $\phi(t)$. Since ϕ' represent velocity and $|\phi'|$ speed we define the length L of the curve as

$$L = \int_{Q^1} |\phi'| dt. \tag{11}$$

The mass (or charge) M of a wire stretched out along C whose mass (or charge) density at the point $\mathbf{r}$ is $\rho(\mathbf{r})$ is given by

$$M = \int_{Q^1} \rho \circ \phi |\phi'| dt. \tag{12}$$

The expression $|\phi'|dt$ is a called the line element for C. The integrals occurring in (11) and (12) are *not* integrals of 1-forms.

4.6.15 Area elements. The normal vector $\mathbf{n}$ occurring in 4.6.9 is perpendicular to the surface S and its magnitude is an approximation to the area of the surface patch $\phi(R)$ when R is a small rectangle in Q^2. $D_1\phi$ and $D_2\phi$ span a parallelogram in $\mathbb{R}^3$ whose area is $|\mathbf{n}| = |(D_1\phi) \times (D_2\phi)|$. The area of the patch is almost the same by linear approximation. Therefore the area A of S is given by

$$A = \int_{Q^2} |\mathbf{n}| d(s,t). \tag{13}$$

One also defines the integrals

$$\int_{Q^2} f \circ \phi |\mathbf{n}| d(s,t) \tag{14}$$

when f is a real-valued function defined on S.

The expression $|\mathbf{n}|d(s,t)$ is called the area element of S. Again neither integral in (13) and (14) is the integral of a 2-form.

4.6.16 Volume elements. The n-form represented by $\mathbf{e}_{1,\dots,n}$ is called the volume element in $\mathbb{R}^n$. If ϕ is a positively oriented n-surface and f is a 0-form, then, using 3.0.14,

$$\int_\phi f\mathbf{e}_{1,\dots,n} = \int_{Q^n} (f \circ \phi) \det \phi' = \int_{\phi(Q^n)} f.$$

In particular, if $f = 1$ then we get the volume $\mathrm{vol}(\phi(Q^n))$ of $\phi(Q^n)$. If $f = \rho$, the mass or charge density of a body occupying $\phi(Q^n)$, then we get the total mass or charge of the body.

This is an appropriate definition since the volume of the 3-simplex is $1/6$ and the unit cube is a chain of 6 congruent tetrahedra.

4.6.17 The classical version of Stokes's theorem. Suppose $\mathbf{F}$ is a continuously differentiable vector field in Ω and S is surface in Ω which can be realized as a chain of 2-surfaces. Then

$$\int_S \operatorname{curl} \mathbf{F} \cdot d\mathbf{S} = \int_C \mathbf{F} \cdot \mathbf{r}$$

where C is the positively oriented boundary of S.

Sketch of proof: If $S = \phi(Q^2)$ the abstract version of Stokes's theorem gives

$$\begin{aligned}\int_S \operatorname{curl} \mathbf{F} \cdot d\mathbf{S} = \int_\phi \omega^{(2)}_{\operatorname{curl} \mathbf{F}} = \int_\phi d\omega^{(1)}_{\mathbf{F}} = \int_{\partial\phi} \omega^{(1)}_{\mathbf{F}} &= \sum_{j=0}^{2} (-1)^j \int_{\phi \circ \partial \sigma_j^{(2)}} \omega^{(1)}_{\mathbf{F}} \\ &= \sum_{j=0}^{2} (-1)^j \int_{C_j} \mathbf{F} \cdot \mathbf{r} = \int_C \mathbf{F} \cdot \mathbf{r}\end{aligned}$$

where the C_j are the curves parametrized by $\phi \circ \partial \sigma_j^{(2)}$. □

4.6.18 Gauss's theorem. Suppose $\mathbf{F}$ is a continuously differentiable vector field in Ω and V a compact subset of Ω which can be realized as a chain of positively oriented 3-surfaces. Then

$$\int_V \operatorname{div} \mathbf{F} = \int_S \mathbf{F} \cdot d\mathbf{S}$$

where S is the positively oriented boundary of V.

Sketch of proof: If $V = \phi(Q^3)$ the abstract version of Stokes's theorem gives

$$\begin{aligned}\int_V \operatorname{div} \mathbf{F} = \int_\phi (\operatorname{div} \mathbf{F})\mathbf{e}_{1,2,3} = \int_\phi d\omega_{\mathbf{F}}^{(2)} = \int_{\partial\phi} \omega_{\mathbf{F}}^{(2)} &= \sum_{j=0}^{3} (-1)^j \int_{\phi\circ\sigma_j^{(3)}} \omega_{\mathbf{F}}^{(2)} \\ &= \sum_{j=0}^{3} (-1)^j \int_{S_j} \mathbf{F} \cdot d\mathbf{S}_j = \int_S \mathbf{F} \cdot d\mathbf{S}\end{aligned}$$

where the S_j are the surfaces parametrized by $\phi \circ \sigma_j^{(3)}$. $\square$

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 α 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 α. 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 α 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 λ such that $Ax = \lambda x$, then λ is called an *eigenvalue* of A and x an *eigenvector* associated with λ.

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

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. $x^\top Q x = q(x) = \overline{q(x)} = x^\top Q^\top x$ and hence $q(x) = \frac{1}{2} x^\top (Q + Q^\top) x$.

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 α_j, $j = 1, ..., n$, such that $\alpha_1 + ... + \alpha_n = k$. Induction over n assuming the binomial theorem as given.

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 τ 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 π has an even number of factors then this is true for all factorizations of π. One defines therefore the *parity* of a permutation π, denoted by $(-1)^\pi$, to be $(-1)^\ell = \pm 1$, if it has a factorization consisting of ℓ transpositions. If ℓ is even π is called an even permutation and otherwise an odd permutation.

List of special symbols

Index