Applied Sciences * Monographs # 4**

Constantin UDRIŞTE **Vladimir BALAN**

Linear Algebra and Analysis

Geometry Balkan Press

Bucharest, Romania

2005

Linear Algebra and Analysis (Romanian)
Monographs # 4

Applied Sciences * Monographs
Editor-in-Chief Prof.Dr. Constantin Udrişte
Managing Editor Prof.Dr. Vladimir Balan
University Politehnica of Bucharest

Linear Algebra and Analysis
Constantin Udrişte and Vladimir Balan.
Bucharest: Applied Sciences * Monographs, 2005

Includes bibliographical references.

Preface

This book is intended for an introductory course in Linear Algebra and Analysis, being organized as follows.
The first chapter studies Tensor Algebra and Analysis, insisting on tensor fields, index calculus, covariant derivative, Riemannian metrics, orthogonal coordinates and differential operators. It is written for students who have prior knowledge of linear algebra.
The second chapter sets out to familiarize the students with the fundamental ideas of Field Lines, Field Hypersurfaces, Integral Manifolds, and their descriptions as solutions of differential equations, partial differential equations of first order and Pfaff equations. It requires basic knowledge of differential calculus.
The third chapter is intended as an introduction to Fourier Series. The topics (Hilbert spaces, orthonormal basis, Fourier series, etc) are developed gently, with preference of clarity of exposition over elegance in stating and proving results. However, the students must have some knowledge of linear algebra and integral calculus.
The fourth chapter is an introduction to Numerical Methods in Linear Algebra, focusing on algorithms regarding triangularization of matrices, approximate solutions of linear systems, numerical computation of eigenvalues and eigenvectors, etc.
This book is designed for a course in the second semester, at Department of Engineering, University Politehnica of Bucharest. It enhances the students' knowledge of linear algebra and differential-integral calculus, and develops basic ideas for advanced mathematics, theoretical physics and applied sciences. That is why, as a rule, each paragraph contains definitions, theorems, remarks, examples and exercises-problems.
The volume involves the didactic experience of the authors as members of the Department of Mathematics at University Politehnica of Bucharest, enhanced further by the lecturing in English since 1990, at Department of Engineering. The goals of the text are:
- to spread mathematical knowledge and to cover the basic requirements in major areas of modelling,
- to acquaint the students with the fundamental concepts of the presented topics.

We owe a considerable debt to former leading textbook authors quoted in References, and to colleagues and students who have influenced our didactic work.

June 30, 2005 The authors

Contents

1 Tensor Algebra and Analysis **5**
1.1 Contravariant and covariant vectors 5
1.2 Tensors 9
1.3 Raising and lowering of the indices of a tensor 15
1.4 Vector fields and covector fields 17
1.5 Tensor fields 29
1.6 Linear connections 32
1.7 Riemannian metrics and orthogonal coordinates 38
1.8 Differential operators 43
1.9 $q-$forms 59
1.10 Differential $q-$forms 61

2 Field Lines and Hypersurfaces **69**
2.1 Field lines and first integrals 69
2.2 Field hypersurfaces and linear PDEs 82
2.3 Nonhomogeneous linear PDEs 89
2.4 Pfaff equations and integral submanifolds 95

3 Hilbert Spaces **107**
3.1 Euclidean and Hilbert spaces 107
3.2 Orthonormal basis for a Hilbert space 114
3.3 Fourier series 120
3.4 Continuous linear functionals 125
3.5 Trigonometric Fourier series 127

4 Numerical Methods in Linear Algebra **131**
4.1 The norm of a matrix 131
4.2 The inverse of a matrix 136
4.3 Triangularization of a matrix 139
4.4 Iterative methods for solving linear systems 143
4.5 Solving linear systems in the sense of least squares 146
4.6 Numerical computation of eigenvectors and eigenvalues 149

Index 153

References 153

Chapter 1

Tensor Algebra and Analysis

1.1 Contravariant and covariant vectors

1.1.1 Definition. Let $\mathbf{V}$ be an $\mathbb{R}$-vector space of dimension n. Its elements are called *(*contravariant*) vectors*.

Let $B = \{e_i |\ i = \overline{1,n}\} \subset \mathbf{V}$ be a *basis* in $\mathbf{V}$. Then for all $v \in \mathbf{V}$, there exist $v^i \in \mathbb{R},\ i = \overline{1,n}$ such that

$$v = v^1 e_1 + \ldots + v^n e_n = \sum_{i=1}^{n} v^i e_i.$$

Using the implicit *Einstein rule of summation*, we can write in brief $v = v^i e_i$.

The scalars $\{v^i |\ i = \overline{1,n}\} = \{v^1, \ldots, v^n\}$ are called the *contravariant components* of the vector v.

Let be another basis $B' = \{e_{i'} |\ i' = \overline{1,n}\} \subset \mathbf{V}$, related to B via the relations

$$e_{i'} = A^i_{i'} e_i, \ \ i' = \overline{1,n}. \tag{1}$$

Then the vector v decomposes relative to B' like $v = v^{i'} e_{i'}$.

The connection between the *components* of v relative to B and B' is given by

$$v^i = A^i_{i'} v^{i'} \tag{2}$$

or in matrix notation, denoting $X = {}^t(v^1, \ldots, v^n),\ \ X' = {}^t(v^{1'}, \ldots, v^{n'})$ and $A = (A^i_{i'})_{i,i'=\overline{1,n}}$, the relations (2) rewrite

$$X = AX'.$$

If we introduce the matrix $A^{-1} = (A^{i'}_i)_{i,i'=\overline{1,n}}$, defined by the relations

$$\left\{ \begin{array}{l} AA^{-1} = I_n \\ A^{-1}A = I_n \end{array} \right. \ \Leftrightarrow \ \left\{ \begin{array}{l} A^i_{i'} A^{i'}_j = \delta^i_j, \ \ i, j = \overline{1,n} \\ A^i_{i'} A^{j'}_i = \delta^{j'}_{i'}, \ \ i', j' = \overline{1,n}, \end{array} \right. \tag{3}$$

where δ^i_j and $\delta^{i'}_{j'}$ are the Kronecker symbols, we infer

$$A^{j'}_i v^i = A^{j'}_i A^i_{i'} v^{i'} = \delta^{j'}_{i'} v^{i'} = v^{j'},$$

and hence an equivalent form of (2) is

$$v^{i'} = A^{i'}_i v^i$$

or in condensed form,

$$X' = A^{-1} X.$$

1.1.2 Definition. Any linear form $\omega : \mathbf{V} \to \mathbb{R}$ is called *1-form*, *covariant vector* or *covector*.

We denote by $L(\mathbf{V}, \mathbb{R})$ the set of all 1-forms on $\mathbf{V}$. This has a canonical structure of vector space of dimension n and is called also *the dual space* of $\mathbf{V}$, denoted briefly by $\mathbf{V}^*$.

For a given basis $B = \{e_i | i = \overline{1,n}\}$ of $\mathbf{V}$, we can associate naturally a basis $B^* = \{e^i | i = \overline{1,n}\}$ of the dual vector space $\mathbf{V}^*$, called *dual basis*, by means of the relations

$$e^i(e_j) = \delta^i_j, \;\; i,j = \overline{1,n}. \tag{4}$$

Then any covector $\omega \in \mathbf{V}^*$ can be decomposed with respect to B^* like

$$\omega = \omega_i e^i, \;\; \omega_i \in \mathbb{R}, i = \overline{1,n}.$$

The scalars $\{\omega_i \mid i = \overline{1,n}\}$ are called *the components of the covector* [1] ω.

If one chooses another basis of $\mathbf{V}^*$, say $B^{*'} = \{e^{i'} | i' = \overline{1,n}\}$ dual to $B' = \{e_{i'} \mid i' = \overline{1,n}\}$, and (1) holds true, then we have

$$e^{i'} = A^{i'}_i e^i.$$

If the covector ω decomposes in $B^{*'}$ like $\omega = \omega_{i'} e^{i'}$, then the relation between the components of ω with respect to B^* and $B^{*'}$ is

$$\omega_i = A^{i'}_i \omega_{i'}$$

or, in equivalent form

$$\omega_{i'} = A^i_{i'} \omega_i,$$

where the coefficients $A^{j'}_k$ and $A^l_{i'}$ are related by (3).

The dual vector space $\mathbf{V}^{**}$ of $\mathbf{V}^*$ is isomorphic to $\mathbf{V}$ and therefore it can be identified to $\mathbf{V}$ via the formula $v(\omega) = \omega(v)$.

[1] In matrix language, the contravariant vector will be represented by a column-matrix, and a covariant vector, by a row-matrix.

1.1.3. Exercises

1. Compute the dual basis $B^* = \{f^i\}_{\overline{1,n}} \subset \mathbf{V}^*$ corresponding to the basis $B = \{f_i\}_{\overline{1,n}} \subset \mathbf{V} = \mathbb{R}^n$, in each of the following cases

a) $f_1 = {}^t(1,0), f_2 = {}^t(1,1), \quad (n=2)$;

b) $f_1 = {}^t(1,0,0), f_2 = {}^t(1,1,0), f_3 = {}^t(1,1,1), \quad (n=3)$.

Solution. a) The duality of $\{f^i\}_{i=\overline{1,2}} \in \mathbf{V}^*$ w.r.t. $\{f_i\}_{i=\overline{1,2}} \in \mathbf{V}$, writes

$$f^j(f_i) = \delta^j_i, \quad \text{for all } i,j = \overline{1,2}.$$

The matrix of change from the natural basis $B = \{e_1 = (1,0), e_2 = (0,1)\} \subset \mathbb{R}^2$ to $\{f_1, f_2\}$ is $A = [f_1, f_2] = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Then, denoting by $\{e^1, e^2\}$ the dual natural basis of $(\mathbb{R}^2)^*$, we remark that the duality relations rewrite

$$[f^j]A = [e^j],\ j = \overline{1,n} \quad \Leftrightarrow \quad [f^j] = A^{-1}[e^j],\ j = \overline{1,n}.$$

Consequently

$$\begin{bmatrix} f^1 \\ f^2 \end{bmatrix} = A^{-1} \begin{bmatrix} e^1 \\ e^2 \end{bmatrix} = A^{-1} I_2 = A^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix},$$

and hence

$$\begin{cases} [f^1] = (1,-1) \\ [f^2] = (0,1) \end{cases} \Rightarrow \begin{cases} f^1 = e^1 - e^2 \\ f^2 = e^2. \end{cases}$$

b) Following the same proofline, we have in this case

$$A = [f_1, f_2, f_3] = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad \Rightarrow \quad \begin{bmatrix} f^1 \\ f^2 \\ f^3 \end{bmatrix} = A^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix},$$

and hence

$$f^1 = e^1 - e^2, \qquad f^2 = e^2 - e^3, \qquad f^3 = e^3.$$

2. Consider the vector space $V = \mathbb{R}^3$ endowed with the canonical basis $B = \{e_i | i = \overline{1,3}\}$, the vector $v = v^i e_i \in \mathbf{V}$ of components $X = {}^t(v^1, v^2, v^3) = {}^t(1, 0, -1)$ and the 1-form $\omega = 5e^1 + e^2 - e^3 \in (\mathbb{R}^3)^*$.

Let $B' = \{e_{i'} | i' = \overline{1,3}\}$, be a new basis, with $e_{i'} = A^i_{i'} e_i$, where

$$A = (A^i_{i'})_{i,i'=\overline{1,3}} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix}.$$

Compute the components of X and ω w.r.t. the new bases of V and V^*, respectively.

Solution. The contravariant components $v^{i'}$ of v $(v = v^{i'} e_{i'})$ obey the rule

$$v^{i'} = A^{i'}_i v^i, \; i' = \overline{1,3}. \tag{5}$$

We obtain $A^{-1} = (A^{i'}_i)_{i,i'=\overline{1,3}} = \begin{pmatrix} -2 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & -1 \end{pmatrix}$ and using (5), it follows

$$X' = {}^t(v^{1'}, v^{2'}, v^{3'}) = {}^t(-3, 1, 2),$$

and hence the expressions of v with respect to the two bases are

$$v = e_1 - e_3 = -3e_{1'} + e_{2'} + 2e_{3'}.$$

Also, the 1-form $\omega = 5e^1 + e^2 - e^3 \in (\mathbb{R}^3)^*$, has relative to $B^{*'}$ the components $\omega_{i'}$ $(\omega = \omega_{i'} e^{i'})$, given by

$$\omega_{i'} = A^i_{i'} \omega_i, \; i' = \overline{1,3}. \tag{6}$$

Using (6), we obtain $\omega_{1'} = 1, \omega_{2'} = 5, \omega_{3'} = 2$, so that $\omega = e^{1'} + 5e^{2'} + 2e^{3'}$.

3. Compute the components of the following vectors with respect to the new basis $B' = \{f_1, f_2\} \subset \mathbb{R}^2 = \mathbf{V}$, where $f_1 = {}^t(1,0), f_2 = {}^t(1,1)$, or to its dual.

a) $v = 3e_1 + 2e_2$;

b) $\eta = e^1 - e^2$.

Solution. a) The old components are $\{v^i\} = \{v^1 = 3, v^2 = 2\}$ and form the matrix $[v]_B = {}^t(3,2)$. They are related to the new components via the relations $v^{i'} = A^{i'}_i v^i, \; i' = \overline{1,2}$. The change of coordinates in $\mathbf{V}$ is given by $X = AX'$, i.e., in our notations, $v^i = A^i_{i'} v^{i'}, \; i = \overline{1,2}$, where

$$A = [f_1, f_2] = (A^i_{j'}) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \;\Rightarrow\; (A^{i'}_j) = A^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Hence the new components of v are

$$\begin{cases} v^{1'} = A^{1'}_1 v^1 + A^{1'}_2 v^2 = 1 \cdot 3 + (-1) \cdot 2 = 1 \\ v^{2'} = A^{2'}_j v^j = 0 \cdot 3 + 1 \cdot 2 = 2 \end{cases} \quad\Rightarrow\quad v = 1f'_1 + 2f'_2 = f'_1 + 2f'_2,$$

and the new matrix of components of v is $[v]_{B'} = {}^t(1,2)$.

↟ Hw.[2] Check that $[v]_{B'} = A^{-1}[v]_B$.

b) The components change via $\eta_{i'} = A^i_{i'} \eta_i, \; i' = \overline{1,2} \;\Leftrightarrow\; [\eta]_{B'} = [\eta]_B A$.

↟ Hw. Check that $[\eta]_{B'} = (1, 0)$.

[2] Homework.

1.2 Tensors

We shall generalize the notions of contravariant vector, covariant vector (1-form), and bilinear forms. Let $\mathbf{V}$ be an n-dimensional $\mathbb{R}$ - vector space, and $\mathbf{V}^*$ its dual. We shall denote hereafter the vectors in $\mathbf{V}$ by $u, v, w, \ldots$ and the covectors in $\mathbf{V}^*$ by $\omega, \eta, \theta, \ldots$, etc.

Let us denote in the following

$$\mathbf{V}^{*p} = \underbrace{\mathbf{V}^* \times \ldots \times \mathbf{V}^*}_{p \text{ times}}, \text{ and } \mathbf{V}^q = \underbrace{\mathbf{V} \times \ldots \times \mathbf{V}}_{q \text{ times}}.$$

The previously introduced notions of vectors and covectors can be generalized in the following manner.

1.2.1 Definition. A function $T : \mathbf{V}^{*p} \times \mathbf{V}^q \to \mathbb{R}$ which is linear in each argument (i.e., multilinear) is called a *tensor of type* (p, q) on $\mathbf{V}$.

The numbers p and q are called *orders of contravariance*, and *covariance*, respectively. The number $p + q$ is called *order of the tensor*.

Let $T^p_q(\mathbf{V})$ be the set of all tensors of type (p, q) on $\mathbf{V}$. This can be organized canonically as a real vector space of dimension n^{p+q}. Remark that the definition imposes the following identifications

- $T^0_0(\mathbf{V}) = \mathbb{R}$ (the space of scalars),
- $T^1_0(\mathbf{V}) = \mathbf{V}$ (the space of contravariant vectors),
- $T^0_1(\mathbf{V}) = \mathbf{V}^*$ (the space of covectors),
- $T^0_2(\mathbf{V}) =$ (the space of bilinear forms on $\mathbf{V}$, denoted by $B(\mathbf{V}, \mathbb{R})$).

1.2.2 Definition. We call *tensor product* the mapping

$$\otimes : (S, T) \in T^p_q(\mathbf{V}) \times T^r_s(\mathbf{V}) \to S \otimes T \in T^{p+r}_{q+s}(\mathbf{V})$$

given by

$$\begin{aligned} S \otimes T(\omega^1, \ldots, \omega^{p+r}, v_1, \ldots, v_{q+s}) &= S(\omega^1, \ldots, \omega^p, v_1, \ldots, v_q) \cdot \\ &\cdot T(\omega^{p+1}, \ldots, \omega^{p+r}, v_{q+1}, \ldots, v_{q+s}), \end{aligned} \tag{7}$$

for all $\omega^i \in \mathbf{V}^*, i = \overline{1, p+r}$, and $v_k \in \mathbf{V} \equiv (\mathbf{V}^*)^*, k = \overline{1, q+s}$, where $p, q, r, s \in \mathbb{N}$ are arbitrary and fixed. It can be proved that $\otimes$ is an $\mathbb{R}-$bilinear, associative mapping and that $T^p_q = \mathbf{V}^p \otimes \mathbf{V}^{*q}$.

1.2.3 Theorem. *Let* $B = \{e_i | i = \overline{1, n}\} \subset \mathbf{V}$ *be a basis in* $\mathbf{V}$*, and*

$$B^* = \{e^i | i = \overline{1, n}\} \subset \mathbf{V}^*$$

its dual basis. Then the set $B^p_q \subset T^p_q(V)$*,*

$$B^p_q = \{E^{j_1\dots j_q}_{i_1\dots i_p} \equiv e_{i_1} \otimes \dots \otimes e_{i_p} \otimes e^{j_1} \otimes \dots \otimes e^{j_q} \mid i_1, \dots, i_p, j_1, \dots, j_q = \overline{1,n}\} \quad (8)$$

represents a basis in $T^p_q(\mathbf{V})$*, and it has* n^{p+q} *elements.*

Proof. The proof for the general case can be performed by analogy with the proof for $p = q = 1$, which we give below. So we prove that $B^1_1 = \{e_{i_1} \otimes e^{j_1} | i_1, j_1 = \overline{1,n}\}$ is a basis in $T^1_1(\mathbf{V}) = \mathbf{V}^* \otimes \mathbf{V}$.

Using the *Einstein summation convention rule*, let $t^{i_1}_{j_1} e_{i_1} \otimes e^{j_1} = 0$ be a vanishing linear combination of the n^2 vectors of B^1_1. But as

$$\mathbf{V}^* \otimes \mathbf{V} \equiv \mathbf{V}^* \otimes \mathbf{V}^{**} \equiv (\mathbf{V} \otimes \mathbf{V}^*)^* = L(\mathbf{V} \times \mathbf{V}^*, \mathbb{R}),$$

we have

$$(t^{i_1}_{j_1} e_{i_1} \otimes e^{j_1})(e^{k_1}, e_{n_1}) = 0, \quad \text{for all } e^{k_1} \in \mathbf{V}^*, e_{n_1} \in \mathbf{V}.$$

Therefore, using (4) for $p = q = 1$, and the multilinearity of $\otimes$, we infer

$$0 = t^{i_1}_{j_1} e_{i_1}(e^{k_1}) e^{j_1}(e_{n_1}) = t^{i_1}_{j_1} \delta^{k_1}_{i_1} \delta^{j_1}_{n_1} = t^{k_1}_{n_1}.$$

So that $t^{k_1}_{n_1} = 0$, for all $k_1, n_1 = \overline{1,n}$, and thus the set B^1_1 is linearly independent.

It can be also proved that the set B^1_1 provides a system of generators of $T^1_1(\mathbf{V})$, and hence it is a basis of $T^1_1(\mathbf{V})$. □

Any tensor $T \in T^p_q(\mathbf{V})$ can be decomposed with respect to B^p_q (8) like

$$T = T^{i_1\dots i_p}_{j_1\dots j_q} E^{j_1\dots j_q}_{i_1\dots i_p}, \quad (9)$$

and the set of real numbers

$$\{T^{i_1\dots i_p}_{j_1\dots j_q} \mid i_1, \dots, i_p, j_1, \dots, j_q = \overline{1,n}\}$$

is called the *set of components of* T with respect to B^p_q.

Examples. 1. A (1,0)-tensor (a vector) $v \in T^1_0(\mathbf{V}) \equiv \mathbf{V}$ decomposes like $v = v^i e_i$.

2. A (0,1)-tensor (a covector) $\omega \in T^0_1(\mathbf{V}) \equiv \mathbf{V}^*$ decomposes like $\omega = \omega_i e^i$.

3. A (1,1)-tensor (assimilated to a linear operator) $T \in T^1_1(\mathbf{V}) \equiv End(\mathbf{V})$ decomposes like $T = T^i_j e_i \otimes e^j$.

4. A (0,2)-tensor (assimilated to a bilinear form) $Q \in T^0_2(\mathbf{V}) \equiv B(\mathbf{V}, \mathbb{R})$ decomposes like $Q = Q_{ij} e^i \otimes e^j$.

1.2.4 Remarks. 1°. The components of the tensor product of two tensors S and T in (7) are given by

$$(S \otimes T)^{i_1\dots i_{p+r}}_{j_1\dots j_{q+s}} = S^{i_1\dots i_p}_{j_1\dots j_q} \, T^{i_{p+1}\dots i_{p+r}}_{j_{q+1}\dots j_{q+s}},$$

for $i_1, \dots, i_{p+r}, j_1, \dots, j_{q+s} = \overline{1,n}$.

2^o. Let B^p_q be the basis (8) of $T^p_q(\mathbf{V})$ induced by the given basis $B = \{e_i \mid i = \overline{1,n}\}$ of $\mathbf{V}$, and let B'^p_q be the basis induced similarly by another basis $B' = \{e_{i'} \mid i' = \overline{1,n}\}$ of V, which is connected to B via (1). The basis B^p_q is changed into the basis B'^p_q by the formulas

$$E^{j'_1\dots j'_q}_{i'_1\dots i'_p} = A^{j'_1}_{j_1}\dots A^{j'_q}_{j_q}\cdot A^{i_1}_{i'_1}\dots A^{i_p}_{i'_p}\cdot E^{j_1\dots j_q}_{i_1\dots i_p},$$

where

$$E^{j'_1\dots j'_q}_{i'_1\dots i'_p} = e_{i'_1}\otimes\dots\otimes e_{i'_p}\otimes e^{j'_1}\otimes\dots\otimes e^{j'_q}.$$

Let $T \in T^p_q(\mathbf{V})$ be decomposed like (9) with respect to B^p_q, and also like

$$T = T^{i'_1\dots i'_p}_{j'_1\dots j'_q}\cdot E^{j'_1\dots j'_q}_{i'_1\dots i'_p},$$

with respect to B'^p_q. Then the relation between the two sets of components of T is given by

$$T^{i'_1\dots i_{p'}}_{j'_1\dots j_{q'}} = A^{i'_1}_{i_1}\dots A^{i'_p}_{i_p}\cdot A^{j_1}_{j_{1'}}\dots A^{j_q}_{j_{q'}}\cdot T^{i_1\dots i_p}_{j_1\dots j_q},$$

with $A^i_{i'}$, $A^{i'}_i$ given by (1) and (3).

1.2.6 Definition. We call *transvection of tensors on the indices* of positions (r, s), the mapping $tr^r_s : T^p_q(\mathbf{V}) \to T^{p-1}_{q-1}(\mathbf{V})$ given by

$$[(tr^r_s)\,(T)]^{i_1\dots i_{r-1}i_{r+1}\dots i_p}_{j_1\dots j_{s-1}j_{s+1}\dots j_q} = \sum_{k=1}^n T^{i_1\dots i_{r-1}ki_{r+1}\dots ip}_{j_1\dots j_{s-1}kj_{s+1}\dots jq}, \quad \text{for all } T\in T^p_q(\mathbf{V}).$$

Remark. Using a vector $v \in T^1_0(\mathbf{V}) = \mathbf{V}$, one can define the transvection with v of each tensor $T \in T^p_q(\mathbf{V})$, $q \geq 1$. Say, for $v = v^i e_i$ and $T = T_{jk}e^j \otimes e^k \in T^0_2(\mathbf{V})$, the transvected tensor $tr_v(T) = (tr^1_1)(v\otimes T) \in T^0_1(\mathbf{V})$ has the components given by $[tr_v(T)]_i = v^s T_{si}$.

1.2.7. Exercises

1. Compute the components of the following tensors with respect to the corresponding tensorial basis associated to the new basis $B' = \{f_1, f_2\} \subset \mathbb{R}^2 = \mathbf{V}$, where $f_1 = {}^t(1, 0), f_2 = {}^t(1, 1)$.

a) $A = e_1 \otimes e^2 - 3e_2 \otimes e^2 \in T^1_1(\mathbb{R}^2)$;

b) $Q = e^1 \otimes e^2 - e^2 \otimes e^1 \in T^0_2(\mathbb{R}^2)$;

c) $T = e_1 \otimes e^2 \otimes e^1 - 2e_2 \otimes e^1 \otimes e^2 \in T^1_2(\mathbb{R}^2)$.

Solution. a) The formulas of change of components are

$$A^{i'}_{j'} = C^{i'}_i C^j_{j'} A^i_j,\ i', j' = \overline{1,2} \Leftrightarrow [A]_{B'} = C^{-1}[A]_B C,$$

where $[A]_B = (A^i_j)_{i,j=\overline{1,n}} = \begin{pmatrix} 0 & 1 \\ 0 & -3 \end{pmatrix}$, and the matrices C and C^{-1} are computed above. ♣ Hw. Check that

$$A^{1'}_{2'} = C^{1'}_i C^j_{2'} A^i_j = 4, \quad A^{1'}_{1'} = A^{2'}_{1'} = 0, \quad A^{2'}_{2'} = -3.$$

Then we have

$$A = 4e^{2'} \otimes e_{1'} - 3e^{2'} \otimes e_{2'}, \ [A]_{B'} = \begin{pmatrix} 0 & 4 \\ 0 & -3 \end{pmatrix},$$

and $[A]_{B'} = C^{-1}[A]_B C$.

b) The change of component rules are $Q_{i'j'} = C^i_{i'} C^j_{j'} Q_{ij}, \ i', j' = \overline{1,2}$. ♣ Hw. Check that

$$[Q]_{B'} = (Q_{i'j'})_{i',j'=\overline{1,n}} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

and that $[Q]_{B'} =^t C[Q]_B C$.

c) The change of components obey the rule

$$T^{i'}_{j'k'} = C^j_{j'} C^k_{k'} C^{i'}_i T^i_{jk}, \quad i, j, k = \overline{1,2}.$$

2. Let be given the following tensors, expressed in the appropriate basis, associated to the natural basis $B = \{e_1, \ldots, e_n\}$ of $\mathbf{V} = \mathbb{R}^n$, where $n = 2$ for $a), b), c)$, and $n = 3$ for $d)$ and $e)$.

a) $a = 5$;

b) $A = 2e_1 \otimes e^2 + 3e_2 \otimes e^2$;

c) $\eta = 3e^2 - e^1$;

d) $Q = e^2 \otimes e^3 - e^1 \otimes e^2 + 5e^2 \otimes e^2$;

e) $v = 3e_1 + 4e_2 - e_3$.

1°. Determine the type of these tensors and the vector space to which they belong.

2°. Indicate the general form of a basis and the dimension d for each space.

3°. Determine the tensor components in the appropriate basis.

Solution. We display in the following the space of tensors, the types, the basis, the dimension of the corresponding module of the given tensors:

a)	$a = 5 \in \mathbb{R} \equiv T^0_0(\mathbf{V})$,	type (0, 0),	$B^0_0 = \{1\}$,	$d = 2^{0+0} = 1$,
b)	$A \in T^1_1(\mathbf{V}) = \mathbf{V} \otimes \mathbf{V}^*$,	type (1, 1),	$B^1_1 = \{e_i \otimes e^j\}$,	$d = 2^{1+1} = 4$,
c)	$\eta \in T^0_1(\mathbf{V}) = \mathbf{V}^*$,	type (0, 1),	$B^0_1 = \{e^j\}$,	$d = 2^{0+1} = 2$,
d)	$Q \in T^0_2(\mathbf{V}) = \mathbf{V}^* \otimes \mathbf{V}^*$,	type (0, 2),	$B^0_2 = \{e^i \otimes e^j\}$,	$d = 3^{0+2} = 9$,
e)	$v \in T^1_0(\mathbf{V}) = \mathbf{V}$,	type (1, 0),	$B^1_0 = \{e_i\}$,	$d = 3^{1+0} = 3$,

The corresponding components are:

$$a = 5 \cdot 1,\ a \equiv \{5\}; \qquad \{A^1_2 = 2, A^2_2 = 3,\ \text{oth. null}\ \}$$
$$\{\eta_1 = -1, \eta_2 = 3\}; \qquad \{Q_{23} = 1, Q_{12} = -1, Q_{22} = 5,\ \text{oth. null}\ \}$$
$$\{v^1 = 3,\ v^2 = 4,\ v^3 = -1\}.$$

Remarks. a) The tensor a is a *scalar*. The field of scalars $\mathbf{K} = \mathbb{R}$ is an 1-dimensional vector space over $\mathbb{R}$.

b) We associate to the tensor A the matrix

$$[A] = (a^i_j)_{i,j=\overline{1,n}} = \begin{pmatrix} 0 & 2 \\ 0 & 3 \end{pmatrix};$$

this defines a *linear transformation* $\tilde{A} \in End(\mathbb{R}^2)$, given by

$$\tilde{A}\begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = \begin{pmatrix} 2x^2 \\ 3x^2 \end{pmatrix},\ \text{for all } x = \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} \in \mathbb{R}^2;$$

c) The tensor $\eta \in \mathbf{V}^*$ defines a *linear form* $\tilde{\eta}$ of matrix $[\eta] = (-1, 3)$, given by

$$\tilde{\eta} : \mathbf{V} \to \mathbb{R},\ \ \tilde{\eta}(x^1, x^2) = (-1, 3)\begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = -x^1 + 3x^2,\ \text{for all } \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} \in \mathbb{R}^2.$$

d) The tensor $Q \in T^0_2(\mathbb{R}^3)$ defines a *bilinear non-symmetrical form* of matrix

$$[B] = (b_{ij})_{i,j=\overline{1,n}} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

given by

$$\tilde{B} : \mathbf{V}\times\mathbf{V} \to \mathbb{R},\ \ \tilde{B}(u, v) = (u^1, u^2, u^3)[B]\begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} = b_{ij}u^iv^j = -u^1v^2 + 5u^2v^2 + u^2v^3,$$

for all $u = u^ie_i, v = v^je_j \in \mathbf{V}$.

e) The tensor $v = v^ie_i \in \mathbf{V}$ is a *vector*, of associated matrix (column-vector)

$$[v] = {}^t(v^1, v^2, v^3) = {}^t(3, 4, -1).$$

3. Check that the types of tensors in exercise 2 define canonically $\mathbb{R}$ - valued multilinear applications. Specify the domain and the correspondence laws in each case.

Solution.

a) $a \in T^0_0(V) = \mathbb{R} \quad \Rightarrow \quad \hat{a} : \mathbb{R} \to \mathbb{R},\ \hat{a}(k) = ak \in \mathbb{R},\ \ \forall k \in \mathbb{R}.$

b) $A \in T^1_1(V) \quad \Rightarrow \quad \hat{A} : \mathbf{V}^* \times \mathbf{V} \to \mathbb{R},\ \hat{A}(\eta, v) = A^i_j \eta_i v^j \in \mathbb{R},\ \ \forall \eta \in \mathbf{V}^*, v \in \mathbf{V}.$

c) $\eta \in T^0_1(V) = \mathbf{V}^* \quad \Rightarrow \quad \hat{\eta} : \mathbf{V} \to \mathbb{R},\ \hat{\eta}(v) = \eta_i v^i \in \mathbb{R},\ \ \forall v \in \mathbf{V}.$

d) $B \in T^0_2(V) = \mathbf{V}^* \otimes \mathbf{V}^* \quad \Rightarrow \quad \hat{B} : \mathbf{V} \otimes \mathbf{V} \to \mathbb{R},\ \hat{B}(v, w) = B_{ij} v^i w^j \in \mathbb{R},\ \ \forall v, w \in \mathbf{V},$

e) $v \in T^1_0(V) = \mathbf{V} \quad \Rightarrow \quad \hat{v} : \mathbf{V}^* \to \mathbb{R},\ \hat{v}(\eta) = v^i \eta_i \in \mathbb{R},\ \ \forall \eta \in \mathbf{V}^*.$

4. Compute the following tensors expressed in the corresponding basis associated to the natural basis $B^1_0 \subset \mathbf{V} = \mathbb{R}^n$:

a) $w = 3v + u$, where $u = e_1 - e_2$, and $v \equiv \{v^1 = 5, v^2 = 0, v^3 = 7\}, (n = 3)$;

b) $R = P + 5Q$, where $P = e_1 \otimes e_3 \otimes e^2$ and

$$Q = e_2 \otimes e_1 \otimes e^2 - 5e_1 \otimes e_3 \otimes e^2, (n = 3);$$

c) $R = tr^2_1(Q)$, where

$$Q = 5e_1 \otimes e_2 \otimes e^1 \otimes e^3 - 4e_2 \otimes e_2 \otimes e^2 \otimes e^3 - e_1 \otimes e_2 \otimes e^2 \otimes e^3, (n = 3);$$

d) $k = tr^1_1(A)$, where $A = 5e_1 \otimes e^1 + 6e_1 \otimes e^2 - e_2 \otimes e^2, (n = 2)$;

e) $w = tr^2_1(T)$, where

$$T = A \otimes v, A = 5e_1 \otimes e^2 - 3e_2 \otimes e^3, v = 2e_2 - e_1, (n = 3);$$

f) $k = tr^1_1(\eta \otimes v)$, where

$$\eta = e^1 + 2e^2 \text{ and } v = 2e_2, (n = 2);$$

g) $a = tr^1_1 tr^2_2(B \otimes u \otimes v)$, where

$$B = e^1 \otimes e^2 - 2e^2 \otimes e^2 \text{ and } u = e_1, v = e_2 - 3e_1, (n = 2).$$

Solution. a) $w = 3(5e_1 + 7e_3) + (e_1 - e_2) = 16e_1 - e_2 + 21e_3$;

b) $R = 5e_2 \otimes e_1 \otimes e^2 - 24e_1 \otimes e_3 \otimes e^2$;

c) We remark that $R = tr^2_1(Q)$ and $Q \in T^2_2(\mathbf{V})$; hence $R \in T^1_1(\mathbf{V})$. The components of the transvected tensor are $R^i_j = Q^{is}_{sj}$. We use that

$$Q^{12}_{13} = 5, Q^{22}_{23} = -4, Q^{12}_{23} = -1$$

and the other components are null, whence, e.g.,

$$R^2_3 = Q^{21}_{13} + Q^{22}_{23} = -4.$$

♣ Hw. Compute the other components of R, and show that $tr_1^2(Q) \neq tr_1^1(Q)$, though the two tensors are of the same type. Hence we remark that the positions of the transvection indices are essential.

d) $A \in T_1^1(\mathbf{V}) \Rightarrow tr_1^1(A) \in \mathbb{R}, k = A_1^1 + A_2^2 = 5 - 1 = 4 \in \mathbb{R}$;

♣ Hw. Check that $k = Tr[A]$.

e) We obtain

$$T = (5e_1 \otimes e^2 - 3e_2 \otimes e^3) \otimes (2e_2 - e_1) \quad = 10e_1 \otimes e^2 \otimes e_2 - 5e_1 \otimes e^2 \otimes e_1 - \\ -6e_2 \otimes e^3 \otimes e_2 + 3e_2 \otimes e^3 \otimes e_1.$$

The components of the tensor T are

$$\{T_2^{12} = 10, T_2^{11} = -5, T_3^{22} = -6, T_3^{21} = 3, \text{ oth. null }\}.$$

The transvected components are $w = tr_1^2(T) \in T_0^1(\mathbf{V}), w^i = T_s^{is}$,

$$\begin{cases} w^1 = T_1^{11} + T_2^{12} + T_3^{13} = 10 \\ w^2 = T_1^{21} + T_2^{22} + T_3^{23} = 0 \\ w^3 = T_1^{31} + T_2^{32} + T_3^{33} = 0, \end{cases}$$

and thus $w = 10e_1$. ♣ Hw. Check that $[A][v] = [w]$.

f) $\eta \otimes v \in T_1^1(\mathbf{V})$, whence $k \in \mathbb{R}$. We get

$$k = tr_1^1[(e^1 + 2e^2) \otimes 2e_2] = tr_1^1(2e^1 \otimes e_2 + 4e^2 \otimes e_2) = 4.$$

♣ Hw. Show that $k = \tilde{\eta}(\tilde{v}) = [\eta][v]$;

g) Let $R = B \otimes u \otimes v$. Hence its components are

$$\{R_{12}^{12} = 1, R_{12}^{11} = -3, R_{22}^{11} = 6, R_{22}^{12} = -2, \text{ oth. null }\},$$

so that $a = tr_1^1 tr_2^2(R) = 1$. ♣ Hw. Check that $a = {}^t[u][B][v] = \tilde{B}(u, v)$, and that $tr_2^2 tr_1^1(R) = tr_1^1 tr_2^2(R)$.

1.3 Raising and lowering of the indices of a tensor

Let $\mathbf{V}$ be an n-dimensional real Euclidean vector space. Its scalar product is also called *Riemannian metric on* $\mathbf{V}$. This is defined as a symmetric positively defined bilinear form $\langle \cdot, \cdot \rangle : \mathbf{V} \times \mathbf{V} \to \mathbb{R}$. Its components g_{ij} with respect to some fixed basis $B = \{e_i \mid i = \overline{1, n}\} \subset \mathbf{V}$ of $\mathbf{V}$ are given by $g_{ij} = \langle e_i, e_j \rangle$. Generally, we can write

$$\langle u, v \rangle = g_{ij} u^i v^j, \text{ for all } u = u^i e_i, v = v^j e_j \in \mathbf{V}. \tag{10}$$

Any arbitrary fixed vector $u \in \mathbf{V}$ defines a covector

$$\langle u, \cdot \rangle \in L(\mathbf{V}, \mathbb{R}) = \mathbf{V}^*,$$

of components $g_{ij}u^i$, via the linear mapping given by

$$G : \mathbf{V} \to \mathbf{V}^*, \ (G(u))(v) = \langle u, v\rangle, \quad \text{for all } u, v \in \mathbf{V}.$$

Properties: 1°. The mapping G is linear and bijective, hence an isomorphism.

2°. Using (10), one can see that G is characterized by the matrix (denoted also by G), $G = (g_{ij})_{i,j=\overline{1,n}}$, of inverse $G^{-1} = (g^{kl})_{k,l=\overline{1,n}}$, where $g^{kl}g_{ls} = \delta^k_s$.

3°. If B is an orthonormal basis with respect to the scalar product $\langle\cdot,\cdot\rangle$, we have $G(e_i) = e^i \in \mathbf{V}^*$, $i = \overline{1,n}$, and we notice that the dual basis B^* is also orthonormal with respect to the scalar product on $\mathbf{V}^*$ given by

$$\langle\omega, \eta\rangle = \omega_i\eta_j g^{ij}, \quad \text{for all } \omega = \omega_i e^i, \ \eta = \eta_j e^j \in \mathbf{V}^*.$$

Using G and G^{-1} one can *lower*, respectively *raise* the indices of a given tensor.

1.3.1 Definition. Let $T = \{T^{i_1\ldots i_p}_{j_1\ldots j_q}\} \in T^p_q(V)$ and $s \in \overline{1,p}$ and $t \in \overline{1,q}$. The function defined by

$$(G_{s,t}T)^{i_1\ldots i_{s-1}i_{s+1}\ldots i_p}_{j_1\ldots j_{t-1}j_tj_{t+1}\ldots j_{q+1}} = g_{j_ti_s}T^{i_1\ldots i_{s-1}i_si_{s+1}\ldots i_p}_{j_1\ldots j_{t-1}j_{t+1}\ldots j_{q+1}}$$

is called *lowering of the indices*. Analogously, using G^{-1}, we define the *raising of the indices*. The lowering and raising produce new tensors, since they are in fact the tensor products $g \otimes T$, $g^{-1} \otimes T$ followed by suitable transvections.

The real vector spaces of tensors of order $p+q$ are isomorphic via raising and lowering of indices. For instance one can lower the index of a vector $v = v^ie_i \in \mathbf{V}$, by $v_k = g_{ks}v^s$, obtaining the covector $\omega = v_ke^k \in \mathbf{V}^*$, or raise the index of a covector $\omega = \omega_ke^k \in \mathbf{V}^*$, by $v^k = g^{ks}\omega_s$, obtaining the vector $v = v^ke_k \in \mathbf{V}$.

Remark. The scalar product g_{ij} on V induces the scalar product g^{kl} on V^*, and the scalar product on $T^p_q(V)$ given by the mapping

$$G^{k_1l_1\ldots k_ql_q}_{i_1j_1\ldots i_pj_p} = g_{i_1j_1}\ldots g_{i_pj_p}\cdot g^{k_1l_1}\ldots g^{k_ql_q}.$$

1.3.2. Exercises

1. Find $a, b, c \in \mathbb{R}$ such that the mapping $\langle\ \cdot\ ,\ \cdot\ \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$,

$$\langle u, v\rangle = u_1v_1 - 2au_1v_2 - 2u_2v_1 + cu_2v_2 + b, \quad \text{for all } u = (u_1, u_2), \ v = (v_1, v_2) \in \mathbb{R}^2,$$

defines a scalar product on $\mathbb{R}^2$. Find its components relative to the canonic basis.

Solution. The bilinearity implies $b = 0$, the symmetry implies $a = 1$ and the positive definiteness implies $c > 4$. Hence $\langle\ \cdot\ ,\ \cdot\ \rangle$ defines a scalar product iff $a = 1, b = 0, c > 4$. In this case we have

$$\langle u, v\rangle = u_1v_1 - 2u_1v_2 - 2u_2v_1 + cu_2v_2, \quad \text{for all } u = (u_1, u_2), \ v = (v_1, v_2) \in \mathbb{R}^2,$$

and its components are $(g_{ij})_{i,j=\overline{1,2}} = \begin{pmatrix} 1 & -2 \\ -2 & c \end{pmatrix}$.

2. Lower the index of the vector $u = 3e_1 + 2e_2 - e_3 \in V = \mathbb{R}^3$; raise the index of the covector $\omega = 2e^1 - 3e^3 \in V^*$, using the metric

$$\langle u, v\rangle = u^1v^1 + 2u^2v^2 + 3u^3v^3, \quad \text{for all } u = (u^1, u^2, u^3), v = (v^1, v^2, v^3) \in \mathbb{R}^3.$$

Solution. The 1-form $\eta = G(u)$ has the components

$$\begin{cases} \eta_1 = g_{1s}u^s = g_{11}u^1 = 1 \cdot 3 = 3 \\ \eta_2 = g_{2s}u^s = g_{22}u^2 = 2 \cdot 2 = 4 \\ \eta_3 = g_{3s}u^s = g_{33}u^3 = 3 \cdot (-1) = -3, \end{cases}$$

and consequently $\eta = 3e^1 + 4e^2 - 3e^3$. Also, for $v = G^{-1}(\omega)$, we obtain

$$\begin{cases} v^1 = g^{1j}\omega_j = g^{11}\omega_1 = 1 \cdot 2 = 2 \\ v^2 = g^{2j}\omega_j = g^{22}\omega_2 = \frac{1}{2} \cdot 0 = 0 \\ v^3 = g^{3j}\omega_j = g^{33}\omega_3 = \frac{1}{3} \cdot (-3) = -1, \end{cases}$$

and hence $v = 2e_1 - e_3 \in \mathbf{V}$.

3. Lower the third index on first position and raise the second index on position three, for a tensor $S \in T^3_3(V)$, where V is endowed with the metric g_{ij}. Raise the second index of the metric g on the first position.

Solution. We have

$$(G_{31}S)^{ir}_{ujkl} = S^{irs}_{jkl}g_{us} \in T^2_4(V), \qquad (G^{23}S)^{irts}_{jl} = S^{irs}_{jkl}g^{kt} \in T^4_2(V).$$

We get also $(G^{21}g)^i_j = g^{is}g_{js} = \delta^i_j$.

1.4 Vector fields and covector fields

Classically, a vector field in $\mathbb{R}^3$ is given by

$$\vec{v}(x, y, z) = v_1(x, y, z)\vec{i} + v_2(x, y, z)\vec{j} + v_3(x, y, z)\vec{k}.$$

We shall consider the more general case, the space $\mathbb{R}^n$, and the point $x = (x^1, \ldots, x^n) \in D \subset \mathbb{R}^n$, where D is an open subset of $\mathbb{R}^n$. Also, to avoid repetitions, we assume the differentiability of class $\mathcal{C}^\infty$ in definitions.

1.4.1 Definition. A differentiable function $f : D \to \mathbb{R}$ is called a *scalar field*. We denote by $\mathcal{F}(D)$ the set of all scalar fields on D.

It can be shown that $\mathcal{F}(D)$ can be organized as a real commutative algebra [3] in a natural way, considering the addition and multiplication of functions in $\mathcal{F}(D)$, and their multiplication with real scalars.

1.4.2 Definition. Let x be a fixed point of $D \subset \mathbb{R}^n$. If a function $X_x : \mathcal{F}(D) \to \mathbb{R}$ satisfies the conditions:

1) X_x is $\mathbb{R}-$linear,

2) X_x is a *derivation*, i.e.,

$$X_x(fg) = X_x(f) \cdot g(x) + f(x) \cdot X_x(g), \quad \forall f, g \in \mathcal{F}(D), \tag{11}$$

then it is called a *tangent vector* to D at the point x. The set of all tangent vectors at x is denoted by T_xD.

Example. The mapping $\frac{\partial}{\partial x^i}\big|_x : \mathcal{F}(D) \to \mathbb{R}$, given by

$$\left(\frac{\partial}{\partial x^i} \bigg|_x \right)(f) = \frac{\partial f}{\partial x^i}(x), \quad \text{for all } f \in \mathcal{F}(D),$$

is a tangent vector at x.

Remarks. 1°. For any vector X_x, we have $X_x(c) = 0$, for any $c \in \mathcal{F}(\mathbb{R})$, that is, for the constant functions $f(x) = c$. Indeed, from (11) and $f = g = 1$, we find

$$X_x(1 \cdot 1) = 1 \cdot X_x(1) + 1 \cdot X_x(1),$$

whence $X_x(1) = 0$; then $X_x(c) = X_x(c \cdot 1) = cX_x(1) = c \cdot 0 = 0$, for all $c \in \mathbb{R}$.

2°. We define the null operator $O_x : \mathcal{F}(D) \to \mathbb{R}$, by $O_x(f) = 0$. Then O_x is a tangent vector at x.

3°. If $a, b \in \mathbb{R}$ and X_x, Y_x are tangent vectors, then $aX_x + bY_x$ is a tangent vector at x too.

4°. By the addition and the multiplication with scalars, the set T_xD has a structure of a real vector space.

1.4.3 Theorem. *The set*

$$\left\{ \frac{\partial}{\partial x^i} \bigg|_{x_0}, \quad i = \overline{1, n} \right\}$$

is a basis for $T_{x_0}D$. This is called *the natural frame at x_0*.

Proof. First we check the linear independence. Let $a^i \in \mathbb{R}$, $i = \overline{1, n}$ such that $a^i \frac{\partial}{\partial x^i}\, |_{x_0} = 0$. Applying the tangent vector $a^i \frac{\partial}{\partial x^i}\, |_{x_0} \in T_{x_0}D$, to the *coordinate function* x^j, we obtain

$$0 = O_{x_0}(x^j) = \left(a^i \frac{\partial}{\partial x^i} \bigg|_{x_0} \right)(x^j) = a^i \frac{\partial x^j}{\partial x^i}(x_0) = a^i \delta_i^j = a^j.$$

[3] An *algebra* is a vector space which is endowed with a third (internal multiplicative) operation which is associative, distributive with respect to the addition of vectors, and associative with respect to multiplication with scalars.

Thus we have $a^j = 0, j = \overline{1,n}$, whence the set is linearly independent.

Now we check that the set $\left\{ \frac{\partial}{\partial x^i}\Big|_{x_0},\ \ i = \overline{1,n} \right\}$ generates (spans) the space $T_{x_0}D$. Let $f \in \mathcal{F}(D)$. Then, applying the rule of derivation of composed functions, we have

$$\begin{aligned} f(x) &= f(x_0) + \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} f\left(x_0 + t(x - x_0)\right) dt = \\ &= f(x_0) + \int_0^1 \sum_{i=1}^n \frac{\partial f}{\partial x^i}\Bigg|_{x_0 + t(x - x_0)} (x^i - x_0^i) dt. \end{aligned}$$

Denoting $g_i(x) = \displaystyle\int_0^1 \frac{\partial f}{\partial x^i}\,|_{x_0+t(x-x_0)}\, dt$, we notice that

$$g_i(x_0) = \frac{\partial f}{\partial x^i}(x_0),$$

and $f(x) = f(x_0) + \sum\limits_{i=1}^n g_i(x)(x^i - x_0^i)$. Then applying an arbitrary tangent vector $X_x \in T_xD$ to this function, we obtain

$$X_x(f) = 0 + \sum_{i=1}^n \left[X_x(g_i)(x)(x^i - x_0^i) + g_i(x) X_x(x^i - x_0^i) \right],$$

which becomes, for $x = x_0$,

$$X_{x_0}(f) = \sum_{i=1}^n X_{x_0}(x^i) \frac{\partial f}{\partial x^i}(x_0).$$

Denoting $a^i = X_{x_0}(x^i)$ we have

$$X_{x_0}(f) = a^i \frac{\partial}{\partial x^i}\Bigg|_{x_0} (f).$$

Since f is arbitrary, we infer $X_{x_0} = a^i \dfrac{\partial}{\partial x^i}\,|_{x_0}$, whence X_{x_0} is generated by the set

$$\left\{ \frac{\partial}{\partial x^i}\Big|_{x_0},\ \ i = \overline{1,n} \right\}. \qquad \square$$

Example. The object $X_P = 2\frac{\partial}{\partial x}|_P + 3\frac{\partial}{\partial y}|_P \in T_PD$, is a tangent vector at the point $P(x_0, y_0) \in D \subset \mathbb{R}^2$, which is decomposed with respect to the basis

$$\left\{ \frac{\partial}{\partial x}\Bigg|_P, \frac{\partial}{\partial y}\Bigg|_P \right\} \subset T_PD.$$

1.4.4 Definition. Let $D \subset \mathbb{R}^n$. A differentiable function $X : D \to \bigcup\limits_{x \in D} T_xD$, with $X(x) \in T_xD$, for each $x \in D$, is called *a vector field on D*. We denote by $\mathcal{X}(D)$ the set of all vector fields on D.

The operations

$$(X+Y)(x) = X(x)+Y(x)$$
$$(\lambda X)(x) = \lambda X(x), \qquad \text{for all } \lambda \in \mathbb{R},\ X, Y \in \mathcal{X}(D),\ x \in D,$$

determine on $\mathcal{X}(D)$ a structure of a real vector space.

A basis of the $\mathcal{F}(D)-$module [4] $\mathcal{X}(D)$ is provided by the set of vector fields $\{\frac{\partial}{\partial x^i},\ i=\overline{1,n}\}$, where

$$\frac{\partial}{\partial x^i} : D \to \bigcup_{x\in D} T_x D,\ \ \frac{\partial}{\partial x^i}(x) = \left.\frac{\partial}{\partial x^i}\right|_x,\ \ \text{for all } x \in D,\ i=\overline{1,n}. \tag{12}$$

They determine *a natural field of frames* for $\mathcal{X}(D)$, and are called *fundamental vector fields.*

1.4.5 Theorem. *Let* $X \in \mathcal{X}(D)$. *There exist the real functions*

$$X^i \in \mathcal{F}(D),\ i=\overline{1,n},$$

such that $X = X^i \frac{\partial}{\partial x^i}$.

The differentiable functions X^i are called *the components* of the vector field X with respect to the natural frame field.

Proof. For $x \in D,\ X(x) = X^i(x)\frac{\partial}{\partial x^i}\,|_x,\ X^i(x) \in \mathbb{R},\ i=\overline{1,n}$. Thus

$$X^i : x \in D \to X^i(x) \in \mathbb{R},\ i=\overline{1,n}$$

are the required components. □

Example. $X = \frac{x}{x^2+y^2}\frac{\partial}{\partial x} + e^{xy}\frac{\partial}{\partial y} \in \mathcal{X}(D)$, where $D = \mathbb{R}^2\backslash\{(0,0)\} \subset \mathbb{R}^2$, is a vector field on D.

1.4.6 Definition. Let $X, Y \in \mathcal{X}(D)$ be two vector fields (having their components X^i, Y^j of class $\mathcal{C}^\infty$). We call *the Lie bracket of* X *and* Y, the field $[X,Y] \in \mathcal{X}(D)$ given by

$$[X,Y](f) = X(Y(f)) - Y(X(f)),\ \text{for all } f \in \mathcal{F}(D), \tag{13}$$

where we denoted $X(f) = X^i\frac{\partial f}{\partial x^i}$, for all $f \in \mathcal{F}(D)$.

The following properties hold true:

a) $[X,Y] = -[Y,X]$,
b) $[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$, (*the Jacobi identity*);
c) $[X,X] = 0$, for all $X, Y, Z \in \mathcal{X}(D)$,

and we have *a*) ⇔ *c*). Also, the Lie bracket is $\mathbb{R}$-bilinear with respect to X and Y.

[4] We call an $R-$module a set M endowed with two operations (one internal - *addition*, and the second external - *multiplication with scalars from R)*, which obey the same properties as the ones of a vector space, with the essential difference that R is not a field, but *a ring.*

The real vector space $\mathcal{X}(D)$ together with the product given by the bracket

$$[\ .\ ,\ .\ ] : \mathcal{X}(D) \times \mathcal{X}(D) \to \mathcal{X}(D)$$

defined in (13) determine a real *Lie algebra*.

For $D \subset \mathbb{R}^3$, any vector field $v \in \mathcal{X}(D)$ can be rewritten in the classical sense

$$v(x, y, z) = v_1(x, y, z)\vec{i} + v_2(x, y, z)\vec{j} + v_3(x, y, z)\vec{k},$$

with $v_1, v_2, v_3 \in \mathcal{F}(D)$, replacing $\left\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right\}$ with $\{\vec{i}, \vec{j}, \vec{k}\}$.

For any $x_0 \in D \subset \mathbb{R}^n$, consider the dual vector space $(T_{x_0}D)^*$ and denote it by $T^*_{x_0}D$. Its elements are the linear mappings $\omega_{x_0} \in T^*_{x_0}D$, $\omega_{x_0} : T_{x_0}D \to \mathbb{R}$ called *covectors* (*covariant vectors*, $1-$*forms*) at x_0. The space $T^*_{x_0}D$ has a canonical structure of a vector space and is called *the cotangent space*.

Example. For a given mapping $f \in \mathcal{F}(D)$, for each $x_0 \in D$, its differential $df\ |_{x_0} \in T^*_{x_0}D$ at x_0 is an $\mathbb{R}$ - linear mapping on $T_{x_0}D$, hence an example of covector at x_0, since

$$(df\ |_{x_0})(X_{x_0}) = X_{x_0}(f),\ \text{for all}\ X_{x_0} \in T_{x_0}D.$$

Theorem. *Let* $\left\{\left.\frac{\partial}{\partial x^i}\right|_{x_0},\ i = \overline{1, n}\right\}$ *be the natural frame in* $T_{x_0}D$*, and* $x^j : D \to \mathbb{R}$ *the coordinate functions. Then the set*

$$\{dx^i\ |_{x_0}, i = \overline{1, n}\}$$

is a basis in $T^*_{x_0}D$*, called also (natural)* coframe *at* x_0*.*

Proof. We remark that $dx^j\ |_{x_0}\left(\frac{\partial}{\partial x^i}\ |_{x_0}\right) = \frac{\partial x^j}{\partial x^i}(x_0) = \delta^j_i$, i.e., we have a dual basis. Now we want to analyse directly if the given set is a basis. Consider the vanishing linear combination $a_j dx^j\ |_{x_0} = 0, a_j \in \mathbb{R}, j = \overline{1, n}$. Applying this covector to $\left.\frac{\partial}{\partial x^i}\right|_{x_0}$, we obtain

$$0 = (a_j dx^j\ |_{x_0})\left(\left.\frac{\partial}{\partial x^i}\right|_{x_0}\right) = a_j\delta^j_i = a_i \Leftrightarrow a_i = 0, i = \overline{1, n},$$

hence the set of covectors is linearly independent.

Consider now a covector $\omega_{x_0} \in T^*_{x_0}D$. Since $\omega_{x_0} : T_{x_0}D \to \mathbb{R}$ is linear, for any tangent vector $X_{x_0} = X^i_{x_0}\frac{\partial}{\partial x^i}\ |_{x_0} \in T_{x_0}D$, we have

$$\omega_{x_0}(X_{x_0}) = \omega_{x_0}\left(X^i_{x_0}\left.\frac{\partial}{\partial x^i}\right|_{x}\right) = X^i_{x_0}\omega_{x_0}\left(\left.\frac{\partial}{\partial x^i}\right|_{x_0}\right).$$

Similarly, for any $i = \overline{1,n}$, we find

$$dx^i\ |_{x_0}\ (X_{x_o}) = dx^i\ |_{x_0}\ \left(X^k_{x_0} \frac{\partial}{\partial x^k}\bigg|_{x_0} \right) = X^k_{x_0} dx^i\ |_{x_0}\ \left(\frac{\partial}{\partial x^k}\bigg|_{x_0} \right) = X^k_{x_0}\delta^i_k = X^i_{x_0},$$

whence, denoting $\omega_i(x_0) = \omega_{x_0}\left(\frac{\partial}{\partial x^i}\big|_{x_0}\right)$, we infer

$$\omega_{x_0}(X_{x_0}) = \omega_i(x_0) dx^i\ |_{x_0}\ (X_{x_0}) = (\omega_i(x_0) dx^i\ |_{x_0})(X_{x_0}),\ \text{for all } X_{x_0} \in T_{x_0}D.$$

Thus, the covector decomposes $\omega_{x_0} = \omega_i(x_0) dx^i\ |_{x_0}$, and hence is generated by the set $\{dx^i\ |_{x_0}, i = \overline{1,n}\}$. □

Definition. A differentiable mapping $\omega : D \to \bigcup\limits_{x\in D} T^*_x D$, with $\omega(x) \in T^*_x D$ is called *differential 1-form* (or *covariant vector field*, *covector field* on D). The set of differential 1-forms on D will be denoted by $\mathcal{X}^*(D)$.

The addition of forms and their multiplication with real scalar functions endows $\mathcal{X}^*(D)$ with a structure of a real vector space. The set of 1-forms

$$dx^i : D \to \bigcup_{x\in D} T^*_x D,\ \ dx^i(x) = dx^i\ |_x,\ \ \text{for all } x \in D,\ i = \overline{1,n} \tag{14}$$

determines a basis in the $\mathcal{F}(D)$-module $\mathcal{X}^*(D)$. Any differential 1−form can be written $\omega = \omega_i dx^i$. The *components* ω_i are differentiable functions.

Examples. 1. $\omega_x = 2dx^1\ |_x + 3dx^2\ |_x \in T^*_x D$, $x \in D \subset \mathbb{R}^n$ is a covector.

2. $X_x = 5\frac{\partial}{\partial x^1}\ |_x - \frac{\partial}{\partial x^2}\ |_x \in T_x D$ is a vector.

3. $\omega = x^1 x^2 dx^1 - (\sin x^2) dx^2 \in \mathcal{X}^*(D)$ is a covector field (1-form).

4. $X = e^{-x_2}\frac{\partial}{\partial x^1} + x_2\frac{\partial}{\partial x^3} \in \mathcal{X}(D)$, where $D \subset \mathbb{R}^3$, is a vector field.

1.4.7 Let $(x^1, \ldots, x^n)$ be the coordinates of an arbitrary point $x \in D \subset \mathbb{R}^n$. If the point x has the new coordinates $(x^{1'}, \ldots, x^{n'})$, then these are related to the old ones by a change of coordinates

$$x^{i'} = x^{i'}(x^i), i' = \overline{1,n}, \quad \det\left(\frac{\partial x^{i'}}{\partial x^i}(x^j)\right) \neq 0 \tag{15}$$

which are reverted locally to

$$x^i = x^i(x^{i'}), i = \overline{1,n}, \quad \det\left(\frac{\partial x^i}{\partial x^{i'}}(x^{j'})\right) \neq 0. \tag{16}$$

Then the transformation (16) induces a change of basis in $T_x D$,

$$B_x = \left\{ \frac{\partial}{\partial x^i}\bigg|_x,\ \ i = \overline{1,n} \right\} \ \ \to\ \ B'_x = \left\{ \frac{\partial}{\partial x^{i'}}\bigg|_x,\ \ i' = \overline{1,n} \right\},$$

and also of the corresponding dual basis in T_x^*D,

$$B_x^* = \{dx^i \mid_x, \ \ i = \overline{1,n}\} \quad \to \quad B_x^{*'}\{dx^{i'} \mid_x, \ \ i = \overline{1,n}\}.$$

Proposition. The bases above are pairwise related by the formulas

$$\left.\frac{\partial}{\partial x^i}\right|_x = \frac{\partial x^{i'}}{\partial x^i}(x^j) \left.\frac{\partial}{\partial x^{i'}}\right|_x, \ i = \overline{1,n}$$

$$dx^i \mid_x = \frac{\partial x^i}{\partial x^{i'}}(x^{j'})dx^{i'} \mid_x \ , \ \ i = \overline{1,n}.$$

The corresponding bases of *fields of frames* of the $\mathcal{F}(D)$-modules $\mathcal{X}(D)$ and $\mathcal{X}^*(D)$, are respectively

$$B = \left\{\frac{\partial}{\partial x^i}, \ \ i = \overline{1,n}\right\}, \ \ B' = \left\{\frac{\partial}{\partial x^{i'}}, \ \ i' = \overline{1,n}\right\},$$

and

$$B^* = \{dx^i, \ \ i = \overline{1,n}\}, \ \ B^{*'} = \{dx^{i'}, \ \ i = \overline{1,n}\}.$$

They are related by

$$\frac{\partial}{\partial x^i} = \frac{\partial x^{i'}}{\partial x^i}(x^j)\frac{\partial}{\partial x^{i'}}; \quad dx^i = \frac{\partial x^i}{\partial x^{i'}}(x^{j'})dx^{i'}. \tag{17}$$

Let $X \in \mathcal{X}(D), \ X = X^i \frac{\partial}{\partial x^i} = X^{i'} \frac{\partial}{\partial x^{i'}}$ be a vector field, and let

$$\omega \in \mathcal{X}^*(D), \quad \omega = \omega_i dx^i = \omega_{i'} dx^{i'}$$

be an 1-form on D.

Then their old/new components of X and ω are respectively related by

$$X^{i'} = \frac{\partial x^{i'}}{\partial x^i} X^i, \qquad \omega_{i'} = \frac{\partial x^i}{\partial x^{i'}} \omega_i.$$

Remark. Consider a point $P = (x, y) \in D \subset \mathbb{R}^2$, and a vector field

$$X = f\frac{\partial}{\partial x} + g\frac{\partial}{\partial y} \in \mathcal{X}(D), \ \ f, g \in \mathcal{F}(D).$$

We may identify $\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\} \equiv \{\vec{i}, \vec{j}\}$, since the vectors $\frac{\partial}{\partial x} \mid_P$ and $\frac{\partial}{\partial x} \mid_Q$ are equipolent for all $P, Q \in D$; $\vec{i}$ is exactly the class of equipolence $\frac{\partial}{\partial x}$; the similar happens to $\frac{\partial}{\partial y}$. Therefore we can write $X = f\vec{i} + g\vec{j}$.

When the metric on $D \subset \mathbb{R}^2$ is the *canonic metric*, $g_{ij} = \delta_{ij}, \ i, j = \overline{1,2}$, we might also identify $\{dx, dy\} \equiv \{\vec{i}, \vec{j}\}$, respectively.

Similar considerations hold true for $D \subset \mathbb{R}^3$, where we identify $\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\} \equiv \{\vec{i}, \vec{j}, \vec{k}\}$, and, in the case of canonic metric, $\{dx, dy, dz\} \equiv \{\vec{i}, \vec{j}, \vec{k}\}$.

Example. We consider in $D = \mathbb{R}^2 \backslash \{(0,0)\}$ the change of coordinates

$$(x^1, x^2) \equiv (x, y) \in \mathbb{R}^2 \backslash \{(0,0)\} \quad \rightarrow \quad (x^{1'}, x^{2'}) \equiv (\rho, \theta) \in (0, \infty) \times [0, 2\pi),$$

given by $x^i = x^i(x^{i'}), i = \overline{1,2}$, with

$$\begin{cases} x^1 = x^{1'} \cos x^{2'} \\ x^2 = x^{1'} \sin x^{2'} \end{cases} \text{or} \begin{cases} x = \rho \cos\theta \\ y = \rho \sin\theta. \end{cases}$$

Then the change of basis $B = \{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\} \rightarrow B' = \{\frac{\partial}{\partial \rho}, \frac{\partial}{\partial \theta}\}$ is described by the relations

$$\frac{\partial}{\partial x} = \frac{\partial \rho}{\partial x}\frac{\partial}{\partial \rho} + \frac{\partial \theta}{\partial x}\frac{\partial}{\partial \theta}, \qquad \frac{\partial}{\partial y} = \frac{\partial \rho}{\partial y}\frac{\partial}{\partial \rho} + \frac{\partial \theta}{\partial y}\frac{\partial}{\partial \theta}.$$

and the change of dual bases $B^* \rightarrow B^{*'}$ is performed according to the rules

$$dx = \frac{\partial x}{\partial \rho}\,\mathrm{d}\rho + \frac{\partial x}{\partial \theta}\,\mathrm{d}\theta, \qquad dy = \frac{\partial y}{\partial \rho}\,\mathrm{d}\rho + \frac{\partial y}{\partial \theta}\,\mathrm{d}\theta.$$

1.4.8. Exercises

1. Let be the vector field

$$X = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z} \in \mathcal{X}(D), \;\; D \subset \mathbb{R}^3.$$

Find the components of X in cylindrical and spherical coordinates.

Solution. Let be the change from Cartesian to cylindrical coordinates

$$(x^1, x^2, x^3) \equiv (x, y, z) \rightarrow (x^{1'}, x^{2'}, x^{3'}) = (\rho, \theta, z),$$

given by the formulas

$$\begin{cases} x = \rho\cos\theta \\ y = \rho\sin\theta \\ z = z \end{cases} \Leftrightarrow \begin{cases} x^1 = x^{1'}\cos x^{2'} \\ x^2 = x^{1'}\sin x^{2'} \\ x^3 = x^{3'} \end{cases} \Rightarrow \begin{cases} \rho = \sqrt{x^2+y^2} \\ \theta = \arctan\frac{y}{x} \\ z = z. \end{cases} \tag{18}$$

Since the vector field rewrites

$$X = x^1\frac{\partial}{\partial x^1} + x^2\frac{\partial}{\partial x^2} + x^3\frac{\partial}{\partial x^3},$$

we get the components $X^1 = x^1, X^2 = x^2, X^3 = x^3$. We have to determine the new components $X^{1'}, X^{2'}, X^{3'}$, which satisfy $X = X^{i'}\frac{\partial}{\partial x^{i'}}$, by means of the relations which link the two sets of components of X:

$$X^i = \frac{\partial x^i}{\partial x^{i'}} X^{i'}, \quad X^{i'} = \frac{\partial x^{i'}}{\partial x^i} X^i, \;\; i, i' = \overline{1,3}. \tag{19}$$

Method 1. Compute the matrix

$$C^{-1} = \left(\frac{\partial x^i}{\partial x^{i'}} \right)^{-1} = \left(\frac{\partial x^{j'}}{\partial x^j} \right)$$

and then replace its elements in formula (19).

Method 2. We have the relations

$$\begin{aligned} \frac{\partial}{\partial x} &= \frac{\partial \rho}{\partial x}\frac{\partial}{\partial \rho} + \frac{\partial \theta}{\partial x}\frac{\partial}{\partial \theta} + \frac{\partial z}{\partial x}\frac{\partial}{\partial z} \\ \frac{\partial}{\partial y} &= \frac{\partial \rho}{\partial y}\frac{\partial}{\partial \rho} + \frac{\partial \theta}{\partial y}\frac{\partial}{\partial \theta} + \frac{\partial z}{\partial y}\frac{\partial}{\partial z} \\ \frac{\partial}{\partial z} &= \frac{\partial}{\partial z}. \end{aligned}$$

derived from the formulas (18). Then we replace $\frac{\partial}{\partial x^1} = \frac{\partial}{\partial x}$, $\frac{\partial}{\partial x^2} = \frac{\partial}{\partial y}$, and $\frac{\partial}{\partial x^3} = \frac{\partial}{\partial z}$ in

$$X = X^1 \frac{\partial}{\partial x} + X^2 \frac{\partial}{\partial y} + X^3 \frac{\partial}{\partial z},$$

and express $X^1 = x^1, X^2 = x^2, X^3 = x^3$ via (18) with respect to the new coordinates (ρ, θ, z); as result, we yield X in cylindrical coordinates.

2. Consider in $\mathbb{R}^2$ the change of coordinates

$$\begin{cases} x' = x + y \\ y' = x - y. \end{cases}$$

Determine the components of X and df in the new coordinates, and compute $X(f)$, for

$$X = x\frac{\partial}{\partial y} - \frac{\partial}{\partial x} \in \mathcal{X}(\mathbb{R}^2) \quad \text{and} \quad f(x,y) = x^2 - y, \ f \in \mathcal{F}(\mathbb{R}^2) = \mathcal{T}_0^0(\mathbb{R}^2).$$

Solution. We identify $(x, y) = (x^1, x^2)$ and $(x', y') = (x^{1'}, x^{2'})$. The components of X have the rules of change

$$\begin{cases} X^i = C^i_{i'} X^{i'}, \ i = \overline{1,2} \\ X^{i'} = C^{i'}_i X^i, \ i' = \overline{1,2}, \end{cases} \tag{20}$$

where

$$(C^{i'}_i)_{i',i=\overline{1,2}} = \left(\frac{\partial x^{i'}}{\partial x^i} \right)_{i',i=\overline{1,2}} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

and

$$(C^i_{i'})_{i,i'=\overline{1,2}} = \left(\frac{\partial x^i}{\partial x^{i'}} \right)_{i,i'=\overline{1,2}} = \left(\frac{\partial x^{i'}}{\partial x^i} \right)^{-1}_{i',i=\overline{1,2}} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix}. \tag{21}$$

The relations (20) rewrite explicitly

$$\begin{cases} X^{1'} = C_1^{1'} X^1 + C_2^{1'} X^2 = 1(-1) + 1(x) = x - 1 \\ X^{2'} = C_1^{2'} X^1 + C_2^{2'} X^2 = 1(-1) + (-1)(x) = -x - 1. \end{cases}$$

Since

$$\begin{cases} x = (x' + y')/2 \\ y = (x' - y')/2, \end{cases} \tag{22}$$

we get the components of the vector field X in the new coordinates, i.e.,

$$X = \frac{x' + y' - 2}{2} \cdot \frac{\partial}{\partial x^{1'}} - \frac{x' + y' + 2}{2} \cdot \frac{\partial}{\partial x^{2'}}.$$

Note that the matrix of change of coordinates results from the two relations in the statement

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1} = \left(\frac{\partial x^i}{\partial x^{i'}} \right)_{i,i'=\overline{1,2}}.$$

Also, using (22), we have the function $X(f)$ computed in new coordinates straightforward

$$\begin{aligned} X(f) &= x \frac{\partial}{\partial y}(x^2 - y) - \frac{\partial}{\partial x}(x^2 - y) = -x - 2x = -3x = \\ &= -3(x' + y')/2 = -3(x^{1'} + x^{2'})/2. \end{aligned}$$

Regrading the differential form $df = 2xdx - dy$, one can use two methods.

Method 1. The rules of change $dx^i = \frac{\partial x^i}{\partial x^{i'}} dx^{i'}$ which are produced via differentiating the relations (22), we get

$$\begin{cases} dx^1 = dx = (dx' + dy')/2 = (dx^{1'} + dx^{2'})/2 \\ dx^2 = dy = (dx' - dy')/2 = (dx^{1'} - dx^{2'})/2, \end{cases}$$

which together with (22) provides by direct replacement in df

$$\begin{aligned} df &= (x^{1'} + x^{2'}) \cdot \tfrac{1}{2}(dx^{1'} + dx^{2'}) - \tfrac{1}{2}(dx^{1'} - dx^{2'}) = \\ &= \frac{x^{1'} + x^{2'} - 1}{2} \cdot dx^{1'} + \frac{x^{1'} + x^{2'} + 1}{2} \cdot dx^{2'}. \end{aligned}$$

Method 2. The differential of f writes $df = \omega_i dx^i = \omega_{i'} dx^{i'}$ and its old/new components are related by $\omega_{i'} = \frac{\partial x^i}{\partial x^{i'}} \omega_i$, whence, replacing the entries of the matrix $C = \left(\frac{\partial x^i}{\partial x^{i'}} \right)_{i,i'=\overline{1,2}}$ from (21) and the old components $\omega_1 = 2x^1$, $\omega_2 = -1$ of df, we obtain

$$\begin{cases} \omega_{1'} = \frac{\partial x^1}{\partial x^{1'}} \omega_1 + \frac{\partial x^2}{\partial x^{1'}} \omega_2 = \frac{1}{2}(\omega_1 + \omega_2) \\ \omega_{2'} = \frac{\partial x^1}{\partial x^{2'}} \omega_1 + \frac{\partial x^2}{\partial x^{2'}} \omega_2 = \frac{1}{2}(\omega_1 - \omega_2). \end{cases}$$

Hence we find

$$\begin{aligned} df &= \omega_{1'} dx^{1'} + \omega_{2'} dx^{2'} = \\ &= \frac{x^{1'} + x^{2'} - 1}{2} \cdot dx^{1'} + \frac{x^{1'} + x^{2'} + 1}{2} \cdot dx^{2'}. \end{aligned}$$

3. For $D = \mathbb{R}^3$, compute the components of the vector field $X = x\frac{\partial}{\partial x} \in \mathcal{X}(D)$ and of the differential form $\omega = xy\, dz \in \mathcal{X}^*(D)$ in cylindrical and spherical coordinates.

Solution. a) The change from Cartesian to cylindrical coordinates

$$(x^1, x^2, x^3) \to (\rho, \theta, z) = (x^{1'}, x^{2'}, x^{3'})$$

is given by

$$\begin{cases} x = \rho \cos\theta \\ y = \rho \sin\theta \\ z = z. \end{cases} \tag{23}$$

The associated Jacobian matrix

$$C = \left(\frac{\partial x^i}{\partial x^{i'}} \right)_{i,i'=\overline{1,3}} = \begin{pmatrix} \cos\theta & -\rho\sin\theta & 0 \\ \sin\theta & \rho\cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{24}$$

has the inverse

$$C^{-1} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\frac{\sin\theta}{\rho} & \frac{\cos\theta}{\rho} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \left(\frac{\partial x^{i'}}{\partial x^i} \right)_{i,i'=\overline{1,3}}$$

and finally, the new components of X and ω can be computed considering their relations with the old ones.

Method 1. Consider the relations between the new/old components

$$X^{i'} = \frac{\partial x^{i'}}{\partial x^i} X^i \quad \text{and} \quad \omega_{i'} = \frac{\partial x^i}{\partial x^{i'}} \omega_i,$$

where $X^1 = x,\ X^2 = X^3 = 0$, and respectively $\omega_1 = \omega_2 = 0,\ \ \omega_3 = xy$.

Method 2. Use the compound partial derivative rule

$$\begin{cases} \dfrac{\partial}{\partial x} = \dfrac{\partial}{\partial \rho}\dfrac{\partial \rho}{\partial x} + \dfrac{\partial}{\partial \theta}\dfrac{\partial \theta}{\partial x} + \dfrac{\partial}{\partial z}\dfrac{\partial z}{\partial x} \\ \dfrac{\partial}{\partial y} = \dfrac{\partial}{\partial \rho}\dfrac{\partial \rho}{\partial y} + \dfrac{\partial}{\partial \theta}\dfrac{\partial \theta}{\partial y} + \dfrac{\partial}{\partial z}\dfrac{\partial z}{\partial y} \\ \dfrac{\partial}{\partial z} = \dfrac{\partial}{\partial z}. \end{cases}$$

consider the reverse of the relation between coordinates (23),

$$\begin{cases} \rho = \sqrt{x^2+y^2} \\ \theta = \arctan \dfrac{y}{x} \\ z = z, \end{cases} \tag{25}$$

and replace (24) and (25) in $X = X^1\frac{\partial}{\partial x} + X^2\frac{\partial}{\partial y} + X^3\frac{\partial}{\partial z}$.

As for ω, use the straightforward differentiation compound rule,

$$\begin{cases} dx = \dfrac{\partial x}{\partial \rho}\mathrm{d}\rho + \dfrac{\partial x}{\partial \theta}\mathrm{d}\theta + \dfrac{\partial x}{\partial z}dz = \cos\theta \mathrm{d}\rho - \rho\sin\theta \mathrm{d}\theta \\ dy = \dfrac{\partial y}{\partial \rho}\mathrm{d}\rho + \dfrac{\partial y}{\partial \theta}\mathrm{d}\theta + \dfrac{\partial y}{\partial z}dz = \sin\theta \mathrm{d}\rho + \rho\cos\theta \mathrm{d}\theta \\ dz = dz, \end{cases}$$

and replace them and (23) in $\omega = \omega_1 dx + \omega_2 dy + \omega_3 dz$, using (18).

The answers are

$$X = \rho\cos^2\theta\frac{\partial}{\partial\rho} - \sin\theta\cos\theta\frac{\partial}{\partial\theta} + 0\frac{\partial}{\partial z}, \quad \omega = \rho^2\sin\theta\cos\theta dz.$$

Remark that for finding the new components of X, method 1 works faster, while for ω, method 2 is straightforward.

b) The change between Cartesian to spherical coordinates

$$(x,y,z) = (x^1,x^2,x^3) \to (r,\varphi,\theta) = (x^{1'},x^{2'},x^{3'})$$

is given by

$$\begin{cases} x = r\sin\varphi\cos\theta \\ y = r\sin\varphi\sin\theta \\ z = r\cos\varphi. \end{cases} \tag{26}$$

♣ Hw. Find the new (spherical) components for X and ω.

4. Determine the components of the vector field $X = \frac{\partial}{\partial x} - \frac{y}{x}\frac{\partial}{\partial y} \in \mathcal{X}(\mathbb{R}^2\backslash Oy)$ in polar coordinates.

Solution. The old components of X are $X^1 = 1,\ X^2 = -y/x$ and the change of coordinates

$$(x^1,x^2) = (x,y) \to (x^{1'},x^{2'}) = (\rho,\theta) \tag{27}$$

is given by

$$\begin{cases} x = \rho\cos\theta \\ y = \rho\sin\theta \end{cases} \Leftrightarrow \begin{cases} \rho = \sqrt{x^2+y^2} \\ \theta = \arctan\,(y/x). \end{cases} \tag{28}$$

The formulas of change of components write

$$X^{i'} = \frac{\partial x^{i'}}{\partial x^i} X^i$$

and involves the Jacobian matrix

$$\left(\frac{\partial x^{i'}}{\partial x^i} \right)_{i',i=\overline{1,2}} = \left(\frac{\partial x^i}{\partial x^{i'}} \right)^{-1}_{i,i'=\overline{1,2}} = \begin{pmatrix} \cos\theta & -\rho\sin\theta \\ \sin\theta & \rho\cos\theta \end{pmatrix}^{-1} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\dfrac{\sin\theta}{\rho} & \dfrac{\cos\theta}{\rho} \end{pmatrix}.$$

The old components of X rewrite in polar coordinates

$$\begin{cases} X^1 = 1 \\ X^2 = -\tan\theta, \end{cases}$$

hence its new ones will be

$$\begin{cases} X^{1'} = \frac{\partial x^{1'}}{\partial x^i} X^i = \frac{\partial x^{1'}}{\partial x^1} X^1 + \frac{\partial x^{1'}}{\partial x^2} X^2 = \cos\theta \cdot 1 + \sin\theta \cdot (-\tan\theta) \\ X^{2'} = \frac{\partial x^{2'}}{\partial x^i} X^i = \frac{\partial x^{2'}}{\partial x^1} X^1 + \frac{\partial x^{2'}}{\partial x^2} X^2 = -\frac{\sin\theta}{\rho} \cdot 1 + \frac{\cos\theta}{\rho} \cdot (-\tan\theta), \end{cases}$$

so we obtained the decomposition of X w.r.t. the new (polar) coordinates

$$X \quad = X^{1'} \frac{\partial}{\partial x^{1'}} + X^{2'} \frac{\partial}{\partial x^{2'}} = \frac{\cos 2\theta}{\cos\theta} \frac{\partial}{\partial\rho} - \frac{2\sin\theta}{\rho} \frac{\partial}{\partial\theta}.$$

1.5 Tensor fields

Let $D \subset \mathbb{R}^n$ a domain. Using the notion of tensor of type (p,q) on the vector space T_xD, we can define the *tensor fields.*

1.5.1 Definition. A tensor field of type (p,q) on D is a differentiable mapping

$$T : D \to \bigcup_{x\in D} T^p_q(T_xD), \text{ such that } T(x) \in T^p_q(T_xD), \ \forall x \in D.$$

Note. The set of all tensor fields of type (p,q) on D is denoted by $\mathcal{T}^p_q(D)$; this set admits a canonical structure of real vector space and of $\mathcal{F}(D)-$module.

Using the bases B and B^* of $\mathcal{X}(D)$ and $\mathcal{X}^*(D)$ considered in (12) and (14), it can be proved that the following set of tensor fields

$$B^p_q = \{\mathcal{E}^{j_1\ldots j_q}_{i_1\ldots i_p} \equiv \frac{\partial}{\partial x^{i_1}} \otimes \ldots \otimes \frac{\partial}{\partial x^{i_p}} \otimes dx^{j_1} \otimes \ldots \otimes dx^{j_q} \ , \ i_1,\ldots,i_p, j_1,\ldots,j_q = \overline{1,n}, \}$$

of type (p,q), represents a basis in the $\mathcal{F}(D)-$module $\mathcal{T}^p_q(D)$. This basis contains n^{p+q} elements, and hence $\dim \mathcal{T}^p_q(D) = n^{p+q}$.

Any tensor field $T \in \mathcal{T}^p_q(D)$ can be expressed in this basis,

$$T = T^{i_1 \ldots i_p}_{j_1 \ldots j_q}\, \mathcal{E}^{j_1 \ldots j_q}_{i_1 \ldots i_p},$$

and the differentiable functions $T^{i_1 \ldots i_p}_{j_1 \ldots j_q} \in \mathcal{F}(D)$, are called *the components of the tensor field* T with respect to the natural basis B^p_q.

Examples. $T \in \mathcal{T}^2_0(D), \quad T = T^{ij} \frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^j}$.
$g \in \mathcal{T}^0_2(D), \quad g = g_{ij} dx^i \otimes dx^j$.
$\omega \in \mathcal{T}^0_1(D) \equiv \mathcal{X}^*(D), \quad \omega = \omega_i dx^i$.
$X \in \mathcal{T}^1_0(D) \equiv \mathcal{X}(D), \quad X = X^i \frac{\partial}{\partial x^i}$.
$U \in \mathcal{T}^1_1(D), \quad U = U^i_j \frac{\partial}{\partial x^i} \otimes dx^j$.
The tensor field

$$g = x dx \otimes dx + (\sin y) dx \otimes dy \equiv x^1 dx^1 \otimes dx^1 + (\sin x^2) dx^1 \otimes dx^2 \in \mathcal{T}^0_2(D), D \subset \mathbb{R}^2,$$

where we denoted $(x, y) \equiv (x^1, x^2)$, has the components

$$\{g_{11} = x, g_{12} = \sin y, g_{22} = g_{21} = 0\}.$$

Then we can associate to g the matrix $[g]$ of its components

$$[g] = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} = \begin{pmatrix} x & \sin y \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} x^1 & \sin x^2 \\ 0 & 0 \end{pmatrix}.$$

If a change of coordinates (15) occurs then, considering the change of bases (17) in $\mathcal{X}(D)$ and $\mathcal{X}^*(D)$, we remark that the *basis* B^p_q,

$$\left\{ \mathcal{E}^{j_1 \ldots j_q}_{i_1 \ldots i_p} = \frac{\partial}{\partial x^{i_1}} \otimes \ldots \otimes \frac{\partial}{\partial x^{i_p}} \cdot dx^{j_1} \otimes \ldots \otimes dx^{j_q}, \quad i_1, \ldots, i_p, j_1, \ldots, j_q = \overline{1, n} \right\}$$

changes to the corresponding basis B'^p_q by the formulas

$$\mathcal{E}^{j'_1 \ldots j'_q}_{i'_1 \ldots i'_p} = \frac{\partial x^{i_1}}{\partial x^{j'_1}} \ldots \frac{\partial x^{i_p}}{\partial x^{j'_p}} \cdot \frac{\partial x^{i'_1}}{\partial x^{j_1}} \ldots \frac{\partial x^{i'_q}}{\partial x^{j_q}} \cdot \mathcal{E}^{j_1 \ldots j_q}_{i_1 \ldots i_p}, \;\; i'_1, \ldots, i'_p, j'_1, \ldots, j'_q = \overline{1, n}.$$

Then the components of a tensor field $T \in \mathcal{T}^p_q(D)$ expressed w.r.t. these bases, change by the rule

$$T^{i'_1 \ldots i'_p}_{j'_1 \ldots j'_q} = \frac{\partial x^{j_1}}{\partial x^{j'_1}} \ldots \frac{\partial x^{j_q}}{\partial x^{j'_q}} \cdot \frac{\partial x^{i'_1}}{\partial x^{i_1}} \ldots \frac{\partial x^{i'_p}}{\partial x^{i_p}}\, T^{i_1 \ldots i_p}_{j_1 \ldots j_q}.$$

1.5.2. Exercises

1. In $D = \mathbb{R}^2 \backslash \{(0,0)\}$, compute the components of the tensor field

$$T = \sqrt{x^2 + y^2} dx \otimes dx + xy\, dy \otimes dy \in \mathcal{T}_2^0(D),$$

in polar coordinates.

Solution. Denoting $(x, y) = (x^1, x^2)$, $T \in \mathcal{T}_2^0(D)$ decomposes as

$$T = T_{ij} dx^i \otimes dx^j, \tag{29}$$

whence its components are

$$T_{11} = \sqrt{x^2 + y^2}, \quad T_{22} = xy, \quad T_{12} = T_{21} = 0. \tag{30}$$

The change from Cartesian to polar coordinates

$$(x, y) = (x^1, x^2) \to (\rho, \theta) = (x^{1'}, x^{2'})$$

is described by (28). We look for the components $T_{i'j'}$ of T in its decomposition

$$T = T_{i'j'} dx^{i'} \otimes dx^{j'}.$$

Method 1. Use the relation between the new/old components of T,

$$T_{i'j'} = \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^j}{\partial x^{j'}} T_{ij}, \quad i', j' = \overline{1, 2},$$

taking into account that the Jacobian matrix entry factors come from (28)

$$C = \begin{pmatrix} \cos\theta & -\rho\sin\theta \\ \sin\theta & \rho\cos\theta \end{pmatrix} = \left(\frac{\partial x^i}{\partial x^{i'}}\right)_{i,i'=\overline{1,2}}.$$

Method 2. Replace (28) in the Cartesian expression of T, compute then $dx \otimes dx$, and $dy \otimes dy$ using the relations

$$\begin{cases} dx = \mathrm{d}\,(\rho\cos\theta) = \cos\theta \cdot \mathrm{d}\rho - \rho\sin\theta \cdot \mathrm{d}\theta \\ dy = \mathrm{d}\,(\rho\sin\theta) = \sin\theta \cdot \mathrm{d}\rho + \rho\cos\theta \cdot \mathrm{d}\theta, \end{cases}$$

and fill in (29).

↟ Hw. Perform the computations by both methods, and check that the result is the same.

2. Find the components of the canonic metric

$$g = \delta_{ij}\, dx^i \otimes dx^j \in \mathcal{T}_2^0(\mathbb{R}^2)$$

in polar coordinates.

Answer. Proceeding as in the previous exercise, one gets

$$g = \mathrm{d}\rho \otimes \mathrm{d}\rho + \rho^2 \mathrm{d}\theta \otimes \mathrm{d}\theta.$$

1.6 Linear connections

Let D be an *open subset of* $\mathbb{R}^n$, and $Y \in \mathcal{X}(D)$. Let also $x \in D$ fixed, and $X_x \in T_xD$.

Definition. The vector

$$\nabla_{X_x} Y \equiv \frac{d}{dt} Y(x + tX_x) \Big|_{t=0} \in T_xM \tag{31}$$

is called *the (flat) covariant derivative of* Y *with respect to* X_x.

If $Y = Y^i \frac{\partial}{\partial x^i}$, it can be proved that

$$\nabla_{X_x} Y = \nabla_{X_x}(Y^i) \frac{\partial}{\partial x^i} \Big|_x = X_x(Y^i) \frac{\partial}{\partial x^i} \Big|_x .$$

We can extend the definition (31). The (flat) covariant derivative of Y with respect to the vector field $X \in \mathcal{X}(D)$ will be given by

$$\nabla_X Y = X(Y^i) \frac{\partial}{\partial x^i}, \quad \text{for all } X, Y \in \mathcal{X}(D). \tag{32}$$

The operator (32) has the following *properties*

a) $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$,

b) $\nabla_X(aY + bZ) = a\nabla_X Y + b\nabla_X Z$,

c) $\nabla_X(fY) = (\nabla_X f)Y + f\nabla_X Y$,

d) $\nabla_X f = X(f)$,

for all $f, g \in \mathcal{F}(D)$, $a, b \in \mathbb{R}$, $X, Y, Z \in \mathcal{X}(D)$. More general than (32), we can introduce the following extension

1.6.1 Definition. A mapping $\nabla : \mathcal{X}(D)^2 \to \mathcal{X}(D)$, described by the correspondence

$$(X, Y) \in \mathcal{X}(D) \times \mathcal{X}(D) \to \nabla_X Y \in \mathcal{X}(D),$$

which satisfies the properties above is called *linear connection* or *covariant derivative* on D.

Since $\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} \in \mathcal{X}(D)$, we can write

$$\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \Gamma^h_{ij} \frac{\partial}{\partial x^h}, \quad \Gamma^h_{ij} \in \mathcal{F}(D),\ i, j, h = \overline{1, n}.$$

The n^3 functions $\{\Gamma^h_{ij}\}_{i,j,h=\overline{1,n}}$ of this decomposition are called *the components of the linear connection* ∇.

For two arbitrary given vector fields $X = X^i \frac{\partial}{\partial x^i}$, $Y = Y^j \frac{\partial}{\partial x^j} \in \mathcal{X}(D)$, we obtain

$$\begin{aligned}
\nabla_X Y &= \nabla_{X^i \frac{\partial}{\partial x^i}} \left(Y^j \tfrac{\partial}{\partial x^j}\right) = X^i \nabla_{\frac{\partial}{\partial x^i}} \left(Y^j \tfrac{\partial}{\partial x^j}\right) = \\
&= X^i (\nabla_{\frac{\partial}{\partial x^i}} Y^j) \tfrac{\partial}{\partial x^j} + X^i Y^j \nabla_{\frac{\partial}{\partial x^i}} \tfrac{\partial}{\partial x^j} = \\
&= X^i \tfrac{\partial}{\partial x^i}(Y^s) \tfrac{\partial}{\partial x^s} + X^i Y^j \Gamma^s_{ij} \tfrac{\partial}{\partial x^s} = \\
&= X^i \left(\tfrac{\partial Y^s}{\partial x^i} + Y^j \Gamma^s_{ij}\right) \tfrac{\partial}{\partial x^s} \overset{not}{=} X^i Y^s_{,i} \tfrac{\partial}{\partial x^s},
\end{aligned}$$

where the functions $Y^s_{,i} = \dfrac{\partial Y^s}{\partial x^i} + Y^j \Gamma^s_{ij}$ are called *the components of the covariant derivative* of the field Y. Note that these are basically defined by the relation

$$\nabla_{\frac{\partial}{\partial x^i}} Y = Y^k_{,i} \frac{\partial}{\partial x^k}.$$

1.6.2 The linear connection ∇ determines two tensor fields:

1°. The *torsion field of* ∇ is a tensor field $T \in \mathcal{T}^1_2(D)$, provided by the $\mathcal{F}(M)-$linear mapping $T : \mathcal{X}(D)^2 \to \mathcal{X}(D)$,

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y] \in \mathcal{X}(D), \quad \text{for all } X, Y \in \mathcal{X}(D).$$

2. The *curvature tensor field of* ∇ is a tensor field $R \in \mathcal{T}^1_3(D)$, given by the $\mathcal{F}(M)-$linear mapping $R : \mathcal{X}(D)^3 \to \mathcal{X}(D)$,

$$R(X,Y)Z = \nabla_X(\nabla_Y Z) - \nabla_Y(\nabla_X Z) - \nabla_{[X,Y]} Z, \quad \text{for all } X, Y, Z \in \mathcal{X}(D).$$

The components of these tensor fields, computed relative to the components $\{\Gamma^i_{jk}\}_{i,j,k=\overline{1,n}}$ of ∇ are, respectively,

$$\begin{aligned} T^i_{jk} &= \Gamma^i_{jk} - \Gamma^i_{kj}, \\ R^h_{ijk} &= \frac{\partial \Gamma^h_{ki}}{\partial x^j} - \frac{\partial \Gamma^h_{ji}}{\partial x^k} + \Gamma^h_{js}\Gamma^s_{ki} - \Gamma^h_{ks}\Gamma^s_{ji}. \end{aligned} \tag{33}$$

If $T^i_{jk} = 0,\ i,j,k = \overline{1,n}$, then the connection ∇ is called *symmetrical connection.*

If $R^i_{jkl} = 0,\ i,j,k,l = \overline{1,n}$, then we say that the connection ∇ is *flat.*

It can be proved that for a change of coordinates (15), the components $\{\Gamma^i_{jk}\}$ and $\{\Gamma^{i'}_{j'k'}\}$ of ∇ are connected by the relations

$$\Gamma^{i'}_{j'k'} \frac{\partial x^{j'}}{\partial x^j} \frac{\partial x^{k'}}{\partial x^k} = \Gamma^h_{jk} \frac{\partial x^{i'}}{\partial x^h} - \frac{\partial^2 x^{i'}}{\partial x^j \partial x^k}, \quad i', j, k = \overline{1,n}.$$

We remark that for $X \in \mathcal{X}(D)$, the mapping $\nabla_X : \mathcal{X}(D) \to \mathcal{X}(D)$, given by

$$(\nabla_X)(Y) = \nabla_X Y, \quad \text{for all } Y \in \mathcal{X}(D), \tag{34}$$

defines an $\mathbb{R}$-linear endomorphism of the real vector space $\mathcal{X}(D)$. We can extend it to $\mathcal{T}^p_q(D),\ p, q \in \mathbb{N}$, by postulating the following conditions for this extended map:

a) $\nabla_X f = X(f),$ for all $f \in \mathcal{F}(D)$,

b) $\nabla_X Y,$ for all $Y \in \mathcal{X}(D)$ as defined in (34),

c) $\nabla_X \omega \in \mathcal{X}^*(D),$ for all $\omega \in \mathcal{X}^*(D)$, is defined by

$$(\nabla_X \omega)(Y) = X(\omega(Y)) - \omega(\nabla_X Y), \quad \text{for all } Y \in \mathcal{X}(D),$$

d) $\nabla_X T \in \mathcal{T}_q^p(D)$, for all $T \in \mathcal{T}_q^p(D)$ is defined by

$$(\nabla_X T)(\omega^1, \dots, \omega^p; Y_1, \dots, Y_q) = X(T(\omega^1, \dots, \omega^p; Y_1, \dots, Y_q)) -$$
$$-\sum_{k=1}^{p} T(\omega^1, \dots, \nabla_X \omega^k, \dots, \omega^p, Y_1, \dots, Y_q) -$$
$$-\sum_{k=1}^{q} T(\omega^1, \dots, \omega^p, Y_1, \dots, \nabla_X Y_k, \dots, Y_q).$$

Then the extended mapping ∇_X described by a)–d) is subject of the following

1.6.3 Theorem. *The operator ∇_X has the properties*

a) ∇_X preserves the type of the tensor fields;

b) ∇_X commutes with transvections;

c) $\nabla_X(S \otimes T) = (\nabla_X S) \otimes T + S \otimes \nabla_X T$, for all $S \in \mathcal{T}_q^p(D)$, $T \in \mathcal{T}_s^r(D)$.

1.6.4 Definition. If a tensor field $T \in \mathcal{T}_s^r(D)$ satisfies the condition

$$\nabla_X T = 0, \quad \text{for all } X \in \mathcal{X}(D),$$

then T is called a *parallel tensor field.*

Examples. 1. For $Y = Y^i \frac{\partial}{\partial x^i} \in \mathcal{X}(D) = \mathcal{T}_0^1(D)$, we have

$$\nabla_{\frac{\partial}{\partial x^k}} Y = Y^s_{,k} \frac{\partial}{\partial x^s}, \text{ with } Y^s_{,k} = \frac{\partial Y^s}{\partial x^k} + \Gamma^s_{ki} Y^i.$$

2. For $\omega = \omega_i dx^i \in X^*(D) = \mathcal{T}_1^0(D)$, we have

$$\nabla_{\frac{\partial}{\partial x^k}} \omega = \omega_{s,k}\, dx^s, \text{ with } \omega_{s,k} = \frac{\partial \omega_s}{\partial x^k} - \Gamma^i_{ks} \omega_i.$$

3. For $T = T^i_j \frac{\partial}{\partial x^i} \otimes dx^j \in \mathcal{T}_1^1(D)$, we have

$$\nabla_{\frac{\partial}{\partial x^k}} T = T^i_{j,k} \frac{\partial}{\partial x^i} \otimes dx^j,$$

with

$$T^i_{j,k} = \frac{\partial T^i_j}{\partial x^k} + \Gamma^i_{kh} T^h_j - \Gamma^h_{kj} T^i_h.$$

1.6.4. Exercises

1. Let $\{\Gamma^k_{ij}\}_{i,j,k=\overline{1,n}}$ be the components of a linear connection in $\mathbb{R}^2$, where $\Gamma^1_{12} = xy$, and the other components are zero. Let also $f, g \in \mathcal{F}(\mathbb{R}^2)$ and $X, Y, Z \in \mathcal{X}(\mathbb{R}^2)$ be given by

$$f(x,y) = x + y, \quad g(x,y) = xy, \qquad X = y\frac{\partial}{\partial x},\ Y = x\frac{\partial}{\partial y}, Z = x\frac{\partial}{\partial x} + y^2\frac{\partial}{\partial y}.$$

Verify the relations

a) $\nabla_{fX+gY}Z = f\nabla_X Z + g\nabla_Y Z$;

b) $\nabla_X(\alpha Y + \beta Z) = \alpha\nabla_X Y + \beta\nabla_X Z$;

c) $\nabla_X fY = f\nabla_X Y + (\nabla_X f)Y$;

for all $X, Y, Z \in \mathcal{X}(D),\ \alpha, \beta \in \mathbb{R},\ f, g \in \mathcal{F}(D)$.

Solution. Given a linear connection ∇ on $D \subset \mathbb{R}^n$ and considering the natural basis

$$\left\{\frac{\partial}{\partial x^i} \overset{not}{=} \partial_i\ ,\ \ i = \overline{1, n}\right\} \subset \mathcal{X}(D),$$

the components $\{\Gamma^k_{ij}\}_{i,j,k=\overline{1,n}}$ of the linear connection ∇ are determined by the relations

$$\nabla_{\partial i}\partial j = \Gamma^k_{ij}\partial_k,$$

where we denoted $\partial_1 = \frac{\partial}{\partial x^1} = \frac{\partial}{\partial x}, \partial_2 = \frac{\partial}{\partial x^2} = \frac{\partial}{\partial y}$. In our case, we have $\nabla_{\frac{\partial}{\partial x}}\frac{\partial}{\partial y} = xy\frac{\partial}{\partial x}$ and all the other components of ∇ in the local basis of the $\mathcal{F}(\mathbb{R}^2)$-module $\mathcal{X}(\mathbb{R}^2)$ are null.

a) We have $V = fX + gY = (xy + y^2)\frac{\partial}{\partial x} + x^2y\frac{\partial}{\partial y}$, whence $V^1 = (xy + y^2), V^2 = x^2y$, and $Y^1 = 0, Y^2 = x,\ Z^1 = x, Z^2 = y^2$. Then, using the relation

$$\nabla_V Z = V^s(\partial_s Z^i + \Gamma^i_{st}Z^t)\partial_i,$$

one finds that

$$\begin{aligned}
\nabla_V Z &= V^s(\partial_s Z^i + \Gamma^i_{st}Z^t)\partial_i = V^s(\partial_s Z^i)\partial_i + V^s\Gamma^i_{st}Z^t\partial_i = \\
&= V^1\tfrac{\partial Z^i}{\partial x}\partial_i + V^2\tfrac{\partial Z^i}{\partial y}\partial_i + V^1\Gamma^1_{12}Z^2\partial_1 = \\
&= (xy + y^2)\cdot 1\tfrac{\partial}{\partial x} + x^2y\cdot 2y\cdot\tfrac{\partial}{\partial y} + (xy + y^2)xy\cdot y^2\tfrac{\partial}{\partial x} = \\
&= (xy + y^2)(1 + xy^3)\tfrac{\partial}{\partial x} + 2x^2y^2\tfrac{\partial}{\partial y}.
\end{aligned}$$

As well, we have $X^1 = y, X^2 = 0$ and hence

$$\begin{aligned}
f\nabla_X Z &= fX^s(\partial_s Z^i + \Gamma^i_{st}Z^t)\tfrac{\partial}{\partial x^i} = fX^1\left(\tfrac{\partial Z^i}{\partial x} + \Gamma^i_{12}Z^2\right)\tfrac{\partial}{\partial x^i} = \\
&= fX^1\left[\left(\tfrac{\partial x}{\partial x} + \Gamma^1_{12}\cdot y^2\right)\tfrac{\partial}{\partial x} + \left(\tfrac{\partial y^2}{\partial x} + \Gamma^2_{12}\cdot y^2\right)\tfrac{\partial}{\partial y}\right] = \\
&= (x + y)y\cdot(1 + xy^3)\tfrac{\partial}{\partial x}, \\
g\nabla_Y Z &= gY^s(\partial_s Z^i + \Gamma^i_{st}Z^t)\tfrac{\partial}{\partial x^i} = gY^2\left(\tfrac{\partial Z^i}{\partial y} + \Gamma^i_{2t}Z^t\right)\tfrac{\partial}{\partial x^i} = \\
&= gY^2\tfrac{\partial Z^2}{\partial y}\tfrac{\partial}{\partial x^2} = xy\cdot x\cdot 2y\tfrac{\partial}{\partial y} = 2x^2y^2\tfrac{\partial}{\partial y},
\end{aligned}$$

whence the stated equality. ⇞ Hw. Solve the items b) and c).

2. Let $\nabla \equiv \{\Gamma^k_{ij}\}_{i,j,k=\overline{1,n}}$ be a linear connection in $\mathbb{R}^2$, where $\Gamma^1_{12} = xy$, and the other components are zero. Let also $f \in \mathcal{F}(\mathbb{R}^2)$ and $X, Y \in \mathcal{X}(\mathbb{R}^2)$, $\omega \in \mathcal{X}^*(\mathbb{R}^2)$ and $g \in \mathcal{T}^0_2(\mathbb{R}^2)$ be respectively given by $f(x,y) = x + y, X = y\frac{\partial}{\partial x}, Y = x\frac{\partial}{\partial y}, \omega = xdx + xydy, \frac{1}{y^2}dx \otimes dx + dy \otimes dy$. Compute:

a) $\nabla_X f = X(f)$;

b) $\nabla_X Y$;

c) $\nabla_X \omega$;

d) $\nabla_X g$.

Solution. a) Using the definition, we find $X(f) \in \mathcal{F}(\mathbb{R}^2)$,

$$X(f) = X^i \frac{\partial f}{\partial x^i} = y\frac{\partial f}{\partial x} = y.$$

b) We notice that

$$\nabla_X Y = X^k Y^i_{;k} \frac{\partial}{\partial x^i} = yY^2_{;1}\frac{\partial}{\partial y} = y\left(\frac{\partial Y^2}{\partial x^1} + \Gamma^2_{1j}Y^j\right)\frac{\partial}{\partial y},$$

and hence $\nabla_X Y = y\frac{\partial}{\partial x}(x)\frac{\partial}{\partial y} = y\frac{\partial}{\partial y} \in \mathcal{X}(\mathbb{R}^2)$.

c) We have

$$\begin{aligned}(\nabla_X\omega)(Z) &= X(\omega(Z)) - \omega(\nabla_X Z) = \\ &= X(\omega_i Z^i) - \omega\left(X^i(\partial_i Z^j)\frac{\partial}{\partial x^j} + X^i Z^j \Gamma^k_{ij}\frac{\partial}{\partial x^k}\right),\end{aligned}$$

for all $Z \in \mathcal{X}(\mathbb{R}^2)$, whence

$$(\nabla_X\omega)_j = (\nabla_X\omega)(\frac{\partial}{\partial x^j}) = X^i\omega_{j;i}, \quad \text{with} \quad \omega_{j;i} = \frac{\partial\omega_j}{\partial x^i} - \Gamma^s_{ij}\omega_s.$$

Hw. Compute $\omega_{j;i}$ for $i, j = \overline{1,2}$ and then determine $\nabla_X\omega$.

d) We have

$$(\nabla_X g) = X^k g_{ij;k}dx^i \otimes dx^j = X^k(\partial_k g_{ij} - g_{sj}\Gamma^s_{ki} - g_{si}\Gamma^s_{kj})dx^i \otimes dx^j.$$

3. Compute the torsion components and the curvature components (33) produced by the connection from the previous exercise. Is ∇ a *symmetrical connection* ?

4. Compute the following tensor fields

a) $\nabla_X(Y \otimes g)$;

b) $\nabla_X[tr^1_2(Y \otimes g)] \equiv \theta$;

c) Check that $tr^1_2(\nabla_X(Y \otimes g)) = \theta$.

Hint. a) The tensor field $U = Y \otimes g \in \mathcal{T}^1_2(M)$ has the covariantly derived components

$$U^i_{jk;l} = \partial_l U^i_{jk} + U^s_{jk}\Gamma^i_{ls} - U^i_{sk}\Gamma^s_{lj} - U^i_{js}\Gamma^s_{lk}.$$

5. Consider on $\mathbb{R}^2$ the linear connection of components

$$\Gamma^i_{jk} = \begin{cases} 1, & \text{for } i = j = 1, k = 2 \\ 0, & \text{for all the other indices} \end{cases}$$

i.e., the only nonzero component is $\Gamma^1_{12} = 1$. Compute $\nabla_X Y$, $\nabla_X \omega$, $\nabla_X (Y \otimes \omega)$ for the following fields

$$X = x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}, \quad Y = x^2\frac{\partial}{\partial y}, \quad \omega = xdy + dx.$$

Solution. a) We apply

$$\nabla_X Y = X^i Y^j_{,i}\frac{\partial}{\partial x^j} = X^i\left(Y^1_{,i}\frac{\partial}{\partial x} + Y^2_{,i}\frac{\partial}{\partial y}\right)$$

Since $Y^1 = 0,\ Y^2 = x^2$, we find

$$Y^1_{,i} = \tfrac{\partial Y^1}{\partial x^i} + \Gamma^1_{is}Y^s = \begin{cases} 1\cdot x^2, & \text{for } i = 1 \\ 0, & \text{for } i = 2 \end{cases}$$

$$Y^2_{,i} = \tfrac{\partial Y^2}{\partial x^i} + \Gamma^2_{is}Y^s = \begin{cases} 2x, & \text{for } i = 1 \\ 0, & \text{for } i = 2, \end{cases}$$

so the only nonzero components are $Y^1_{,1} = x^2,\ Y^2_{,1} = 2x$. Then we have

$$\nabla_X Y = -y\cdot x^2\frac{\partial}{\partial x} + (-y)\cdot 2x\frac{\partial}{\partial y} = -x^2y\frac{\partial}{\partial x} - 2xy\frac{\partial}{\partial y}.$$

b) We apply the formulas

$$\nabla_X\omega = X^i\omega_{j,i}dx^j = X^i(\omega_{1,i}dx + \omega_{2,i}dy), \quad \omega_{k,i} = \frac{\partial\omega_k}{\partial x^i} - \Gamma^s_{ik}\omega_s.$$

c) We apply for $T^i_k = Y^i\omega_k$ the formula

$$\begin{aligned} \nabla_X(Y\otimes\omega) &= X^i T^j_{k,i}\tfrac{\partial}{\partial x^j}\otimes dx^k = \\ &= X^i\left(T^1_{1,i}\tfrac{\partial}{\partial x}\otimes dx + T^1_{2,i}\tfrac{\partial}{\partial x}\otimes dy + T^2_{1,i}\tfrac{\partial}{\partial y}\otimes dx + T^2_{2,i}\tfrac{\partial}{\partial y}\otimes dy\right) \end{aligned}$$

and

$$T^j_{k,i} = \frac{\partial T^j_k}{\partial x^i} + \Gamma^j_{is}T^s_k - \Gamma^t_{ik}T^j_t.$$

↟ Hw. Compute $\nabla_x\omega$ and $\nabla_X(Y\otimes\omega)$ effectively.

1.7 Riemannian metrics and orthogonal coordinates

In the Euclidean space $\mathbb{R}^n$ related to an orthonormal basis, the *length* of a vector $v = (v^1, \ldots, v^n) \in \mathbb{R}^n$ is given by

$$||v|| = \sqrt{(v^1)^2 + \ldots + (v^n)^2} = \sqrt{\delta_{ij} v^i v^j} = \sqrt{\langle v, v\rangle}.$$

The tensor $G = \delta_{ij} dx^i \otimes dx^j \in \mathcal{T}_2^0(\mathbb{R}^n)$, which provides the scalar product at each point of $D = \mathbb{R}^n$, will be called *Riemannian metric on* $\mathbb{R}^n$.

More general, for a real Euclidean space V_n, the scalar product provides a tensor of type $(0,2)$ which is symmetrical and positive definite, called *Riemannian metric on* V_n.

Let now D be an open subset of $\mathbb{R}^n$.

1.7.1 Definition. A tensor field $g \in \mathcal{T}_2^0(D)$ such that for all $x \in D$, $g(x)$ is a Riemannian metric on $T_x D$, is called *Riemannian metric on the set* D and the pair (D, g) is called *Riemannian manifold.*

Examples. 1. On $D = \mathbb{R}^n$, the tensor field

$$g_{ij}(x) = \delta_{ij}, \quad \text{for all } i, j \in \overline{1,n}, x \in \mathbb{R}^n,$$

is a Riemannian metric, the so-called *flat metric.*

2. On $D = \{(x,y) \in \mathbb{R}^2 | y > 0\}$, the tensor field

$$g_{ij}(x,y) = \frac{1}{y^2}\,\delta_{ij}, \text{ for all } i, j \in \overline{1,n},\ (x,y) \in D,$$

is a Riemannian metric, called *the Poincaré metric.*

For a change of coordinates (15), the components of the metric g change via

$$g_{i'j'} = g_{ij} \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^j}{\partial x^{j'}}, \quad i', j' = \overline{1,n}. \tag{35}$$

Using the Riemannian metric $g \in \mathcal{T}_2^0(D)$, we compute the following geometric objects

a) the *scalar product* of $X, Y \in \mathcal{X}(D)$, given by $\langle X, Y\rangle = g(X,Y)$;

b) the *norm of a vector field* $X \in \mathcal{X}(D)$, given by $||X|| = \sqrt{g(X,X)}$;

c) the *angle* $\theta \in \mathcal{F}(D)$ between two vector fields $X, Y \in \mathcal{X}(D)$, given by the relation

$$\cos\theta = \frac{g(X,Y)}{||X|| \cdot ||Y||},\ \theta \in [0, \pi];$$

the two vector fields are called *orthogonal*, and we write $X \perp_g Y$ iff $g(X,Y) = 0$;

d) the *length of a curve* $\gamma : I = [a,b] \to D$, given by

$$l(\gamma) = \int_I \sqrt{g(\gamma'(t), \gamma'(t))}dt.$$

For a metric $g = g_{ij}(x)dx^i \otimes dx^j$, we can associate the matrix $[g] = (g_{ij})_{i,j=\overline{1,n}}$. The entries of the inverse matrix $[g]^{-1} = (g^{ij})_{i,j=\overline{1,n}}$ define *the reciprocal tensor field*

$$g^{-1} = g^{ij}\frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^j} \in \mathcal{T}^2_0(D)$$

of the metric g. Then, denoting $[g^{-1}] = (g^{ij})_{\ i,j=\overline{1,n}} = [g]^{-1}$, we have the obvious relations

$$[g] \cdot [g^{-1}] = I_n \quad \Leftrightarrow \quad g_{ij}g^{jk} = \delta^k_i,\ i,k = \overline{1,n}.$$

Remarks. 1°. For $X = X^i\frac{\partial}{\partial x^i}, Y = Y^j\frac{\partial}{\partial x^j} \in \mathcal{X}(D)$, we have $g(X,Y) = g_{ij}X^iY^j$.

2°. The Riemannian metric g and its reciprocal tensor field g^{-1} are used for "*raising*" and "*lowering*" the indices of a tensor field, like shown below:

a) The correspondence

$$X \in \mathcal{T}^1_0(D) \to \omega \in \mathcal{T}^0_1(D)$$

is produced by $\omega_i = X^s g_{si}$;

b) The correspondence

$$\omega \in \mathcal{T}^0_1(D) \to X \in \mathcal{T}^1_0(D)$$

is produced by $X^i = \omega_s g^{is}$;

c) The tensor field $S \in \mathcal{T}^1_1(D)$ produces the tensor field

$$U \in \mathcal{T}^0_2(D),\ U_{ij} = g_{ir}S^r_j,$$

or the tensor field $V \in \mathcal{T}^2_0(D),\ V^{ij} = g^{ir}S^j_r$.

1.7.2 Theorem. *The Riemannian metric* $g \in \mathcal{T}^0_2(D)$ *determines a symmetric connection* ∇ *which obeys* $\nabla_X g = 0$, *for all* $X \in \mathcal{X}(D)$.

Proof. The covariant derivative $\nabla_X g \in \mathcal{T}^0_2(D)$ is given by

$$(\nabla_X g)(Y,Z) \equiv \nabla_X(g(Y,Z)) - g(\nabla_X Y, Z) - g(Y, \nabla_X Z), \quad \text{for all } X,Y,Z \in \mathcal{X}(D).$$

From $(\nabla_X g)(Y,Z) = 0$, for all $X,Y,Z \in \mathcal{X}(D)$, i. e., $g_{ij,k} = 0,\ i,j,k = \overline{1,n}$ and ∇ symmetrical, i. e.,

$$\Gamma^i_{jk} = \Gamma^i_{kj},$$

we obtain the n^3 linear equations

$$0 = g_{ij;k} \equiv \frac{\partial g_{ij}}{\partial x^k} - \Gamma^h_{ki}g_{jh} - \Gamma^h_{kj}g_{ih} \tag{36}$$

with n^3 unknowns $\{\Gamma^i_{jk}; i,j,k = \overline{1,n}\}$. To produce a linear connection, we solve the system in a suitable way; we use the consequence

$$g_{jk;i} + g_{ki;j} - g_{ij;k} = 0$$

of (36), and we obtain the solution

$$\Gamma_{ij}^{h} = \frac{1}{2} g^{hs} \left(\frac{\partial g_{js}}{\partial x^{i}} + \frac{\partial g_{is}}{\partial x^{j}} - \frac{\partial g_{ij}}{\partial x^{s}} \right), \tag{37}$$

which determines a connection. □

The connection ∇ having these components is called *the Levi-Civita connection* associated to g, and its components $\{\Gamma_{ij}^{h}\}_{i,j,h=\overline{1,n}}$ are called *the Christoffel symbols of second kind.*

Remarks. 1. Since $g_{is} = g_{si}, \forall s, i \in \overline{1,n}$, the coefficients above are symmetric in lower indices.

2. Denoting $\frac{\partial g_{js}}{\partial x^{i}} = \partial_i g_{js}$, the formula (37) rewrites

$$\Gamma_{jk}^{i} = \frac{1}{2} g^{is} (\partial_j g_{sk} + \partial_k g_{sj} - \partial_s g_{jk}).$$

3. In the Euclidean case (for $g_{ij} = \delta_{ij}, \forall i, j \in \mathbb{R}^n$), all the components Γ_{jk}^{i} of the Levi-Civita connection ∇ vanish. In this case, the covariant derivation reduces to the usual derivation with respect to a given vector field.

Orthogonal systems of coordinates. Lamé coefficients.

Let $(x^1, \ldots, x^n)$ be a coordinate system, and $(x^{1'}, \ldots, x^{n'})$ another one, connected to the former one by (15), and let $g \in \mathcal{T}_2^0(D)$ a Riemannian metric on D. The components of g relative to the two systems of coordinates are related by

$$g_{i'j'} = \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^j}{\partial x^{j'}} g_{ij}. \tag{38}$$

Definition. The system of coordinates $(x^{1'}, \ldots, x^{n'})$ is called *orthogonal* iff $g_{i'j'} = 0$ for all $i' \neq j'$. The positive definiteness of g implies $g_{i'i'} > 0,\ i' = \overline{1,n}$. The functions $H_{i'} = \sqrt{g_{i'i'}}$ are called *the coefficients of Lamé* .

Let $(x^{1'}, \ldots, x^{n'})$ be an orthogonal system. Starting with a vector field

$$X = X^{i'} \frac{\partial}{\partial x^{i'}}, \in \mathcal{X}(D)$$

and lowering the upper index of X via $g_{i'j'}$, we obtain the covariant field $\omega = X_{i'} dx^{i'} \in \mathcal{X}^*(D)$, with

$$X_{i'} = g_{i'j'} X^{j'}.$$

If we denote by

$$t_{i'} = \frac{\partial}{\partial x^{i'}} = \frac{\partial x^j}{\partial x^{i'}} \frac{\partial}{\partial x^j},\ i' = \overline{1,n},$$

the elements of the basis $B' \subset \mathcal{X}(D)$ associated to the orthogonal system $(x^{1'}, \ldots, x^{n'})$, we can build the orthogonal unit vectors

$$e_{i'} = \frac{1}{H_{i'}} t_{i'}\ \ (!),\ i' = \overline{1,n},$$

where the sign "(!)" indicates no summation (i.e., *exception from the Einstein rule*) in indices i'. Then we have

$$X = X^{i'}\frac{\partial}{\partial x^{i'}} = X^{i'}\frac{\partial x^i}{\partial x^{i'}}\frac{\partial}{\partial x^i} \equiv X^{i'}t_{i'} = \sum_{i'=1}^{n}(H_{i'}X^{i'})e_{i'}.$$

1.7.3 Definition. The components of the vector field $X \in \mathcal{X}(D)$ introduced above have the following names

$$\begin{array}{ll} X^{i'} & = \textit{contravariant components} \\ X_{j'} & = \textit{covariant components} \\ H_{i'}X^{i'}(!) & = \textit{physical components.} \end{array}$$

Generally, the contravariant components of a covariant field $\omega = \omega_{i'}dx^{i'} \in \mathcal{X}^*(D)$ are

$$\omega^{i'} = g^{i's'}\omega_{s'}, \text{ and } \omega^{i'}\frac{\partial}{\partial x^{i'}} \in \mathcal{X}(D).$$

Consequently,

$$\begin{array}{ll} \omega_{i'} & = \textit{covariant components} \\ \omega^{i'} = g^{i'i'}\omega_{i'} = \dfrac{1}{H_{i'}^2}\omega_{i'} \text{ no summation} & = \textit{contravariant components} \\ H_{i'}\omega^{i'} = \frac{1}{H_{i'}}\omega_{i'} \text{ no summation} & = \textit{physical components.} \end{array}$$

1.7.3. Exercises

1. Compute the components of the metric, the Lamé coefficients and the Christoffel symbols in spherical and cylindrical coordinates, in $M = \mathbb{R}^3\backslash Oz$.

Solution. We change the Cartesian coordinates to the cylindrical ones,

$$(x^1, x^2, x^3) = (x, y, z) \to (x^{1'}, x^{2'}, x^{3'}) = (\rho, \theta, z),$$

with the formulas of change

$$\begin{cases} x = \rho\cos\theta \\ y = \rho\sin\theta \\ z = z. \end{cases}$$

The metric g writes in old/new coordinates

$$g = g_{ij}dx^i \otimes dx^j = \delta_{ij}dx^i \otimes dx^j, \qquad g = g_{i'j'}dx^{i'} \otimes dx^{j'}$$

and the new *components* are related to the old ones via the formulas

$$g_{i'j'} = g_{ij}\cdot\frac{\partial x^i}{\partial x^{i'}}\frac{\partial x^j}{\partial x^{j'}} = \delta_{ij}\cdot\frac{\partial x^i}{\partial x^{i'}}\frac{\partial x^j}{\partial x^{j'}} = \sum_{i=1}^{3}\frac{\partial x^i}{\partial x^{i'}}\frac{\partial x^i}{\partial x^{j'}}, \quad i', j' = \overline{1,3}. \tag{39}$$

We remark that the right term is the scalar product between the column-vectors i' and j' of the Jacobian matrix

$$\left(\frac{\partial x^i}{\partial x^{i'}}\right)_{i,i'=\overline{1,3}} = \begin{pmatrix} \cos\theta & -\rho\sin\theta & 0 \\ \sin\theta & \rho\cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Replacing the corresponding entries in (39), one gets the matrix of the new components

$$(g_{i'j'})_{i',j'\in\overline{1,3}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \rho^2 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and g decomposes in cylindrical coordinates as

$$g = 1\mathrm{d}\rho\otimes\mathrm{d}\rho + \rho^2\mathrm{d}\theta\otimes\mathrm{d}\theta + 1dz\otimes dz.$$

Variant. One can compute the square of the arc element $ds^2 = dx^2 + dy^2 + dz^2$, i.e.,

$$g = dx\otimes dx + dy\otimes dy + dz\otimes dz$$

by direct replacement of the differentials

$$\begin{cases} dx = \cos\theta\ \mathrm{d}\rho - \rho\sin\theta\ \mathrm{d}\theta \\ dy = \sin\theta\ \mathrm{d}\rho + \rho\cos\theta\ \mathrm{d}\theta \\ dz = dz, \end{cases}$$

and infer the same result. The components of g provide the *Lamé coefficients for cylindrical coordinates*

$$\begin{cases} H_\rho = \sqrt{g_{1'1'}} = 1 \\ H_\theta = \sqrt{g_{2'2'}} = \rho \\ H_z = \sqrt{g_{3'3'}} = 1. \end{cases}$$

Using non-primed indices now in cylindrical coordinates, the coefficients of the Levi-Civita connection are given by the formulas

$$\Gamma^i_{jk} = \frac{1}{2}g^{is}(\partial_j g_{sk} + \partial_k g_{sj} - \partial_s g_{jk}),\ \ i,j,k = \overline{1,3},$$

where we denoted $\partial_i = \partial/\partial x^i$, $i = \overline{1,3}$. For computing them one determines first the matrix of the reciprocal tensor field g^{-1} of g,

$$(g^{ij})_{i,j=\overline{1,3}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/\rho^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then one obtains, e.g., from the total amount of $3^3 = 27$ components,

$$\begin{aligned} \Gamma^1_{11} &= \tfrac{1}{2}g^{1s}(\partial_1 g_{s1} + \partial_1 g_{s1} - \partial_s g_{11}) = 0 \\ \Gamma^1_{12} &= \Gamma^1_{21} = \tfrac{1}{2}g^{1s}(\partial_2 g_{s1} + \partial_1 g_{s2} - \partial_s g_{12}) = 0. \end{aligned}$$

1.8 Differential operators

Let $D \subset \mathbb{R}^3$ be an open and connected set and $\delta_{ij},\ i, j = 1, 2, 3$, be the Riemannian metric on D. We shall implicitly use the Cartesian coordinates on D.

1.8.1 Consider the following operators in $D \subset \mathbb{R}^3$:

1) The *gradient* $\nabla : \mathcal{F}(D) \to \mathcal{X}(D)$,

$$\operatorname{grad} f = \nabla f = \frac{\partial f}{\partial x}\vec{i} + \frac{\partial f}{\partial y}\vec{j} + \frac{\partial f}{\partial z}\vec{k} \in \mathcal{X}(D).$$

2) The *Hessian* Hess $: \mathcal{F}(D) \to \mathcal{T}_2^0(D)$,

$$\operatorname{Hess} f \equiv \left(\frac{\partial^2 f}{\partial x^i \partial x^j}\right)_{i,j=\overline{1,3}}.$$

3) The *divergence*, div $: \mathcal{X}(D) \to \mathcal{F}(D)$, associates to a vector field

$$v = v_1(x, y, z)\vec{i} + v_2(x, y, z)\vec{j} + v_3(x, y, z)\vec{k} \in \mathcal{X}(D),$$

the scalar field (function)

$$\operatorname{div}\ v = \frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} + \frac{\partial v_3}{\partial z} \in \mathcal{F}(D).$$

Remark that, formally, we can write $\operatorname{div} v = \langle \operatorname{grad}, v \rangle$.

4) The *curl* (*rotor*) curl $: \mathcal{X}(D) \to \mathcal{X}(D)$, associates to any vector field $v \in \mathcal{X}(D)$ as above, the vector field

$$\operatorname{curl} v = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ v_1 & v_2 & v_3 \end{vmatrix} \in \mathcal{X}(D).$$

In the following, we generalize these operators for the case when the coordinates on $D \subset \mathbb{R}^n$ are arbitrary, and consequently D is endowed with a Riemannian metric $g \in \mathcal{T}_2^0(D)$, having the covariant components g_{ij} and the contravariant ones g^{ij}, related by (38). The Riemannian metric g induces the Riemannian connection ∇.

1.8.2 The *gradient*, grad $: f \in \mathcal{F}(D) \to \operatorname{grad} f \in \mathcal{X}(D)$ is defined by the relation

$$g(X,\ \operatorname{grad} f) = \mathrm{d}f(X) \equiv X(f), \quad \text{for all } X \in \mathcal{X}(D). \tag{40}$$

Since $\mathrm{d}f = \omega_i dx^i \in \mathcal{X}^*(D)$ has the components $\omega_i = \frac{\partial f}{\partial x^i},\ i = \overline{1, n}$, we easily obtain the (contravariant !) components of $\operatorname{grad} f$, namely

$$X^i = g^{ij}\frac{\partial f}{\partial x^j} \equiv g^{ij}\omega_j, \quad \operatorname{grad} f = X^i \frac{\partial}{\partial x^i} \in \mathcal{X}(D),$$

so that

$$\operatorname{grad} f = \left(g^{ij}\frac{\partial f}{\partial x^j}\right)\frac{\partial}{\partial x^i} \in \mathcal{X}(D).$$

This operator has the following properties

a) $\operatorname{grad}(a_1 f_1 + a_2 f_2) = a_1 \operatorname{grad} f_1 + a_2 \operatorname{grad} f_2$;

b) $\operatorname{grad}\left(\frac{f_1}{f_2}\right) = \frac{1}{f_2^2}\left[f_2 \operatorname{grad} f_1 - f_1 \operatorname{grad} f_2\right]$;

c) $\operatorname{grad}\varphi(f) = (\varphi' \circ f)\operatorname{grad} f$,

for all $a_{1,2} \in \mathbb{R},\ f, f_1, f_2 \in \mathcal{F}(D), \varphi \in \mathcal{C}^\infty(\mathbb{R})$.

Remarks. 1°. Sometimes, instead of the symbol "grad" we shall use ∇ (nabla).

2°.The solutions of the system

$$\frac{\partial f}{\partial x^i} = 0, \quad i \in \overline{1,n},$$

give the *critical points* of $f \in \mathcal{F}(D)$. The system can be briefly rewritten in the form

$$\operatorname{grad} f = 0.$$

Theorem. *If f has no critical points, then*

a) grad f is orthogonal to the hypersurfaces described by $f(x) = k, (k \in \mathbb{R})$, which are called constant level sets*;*

b) grad f shows at each point the sense and direction of the steepest *(highest) increase of the function f.*

Proof. We prove a). Let $\Sigma : f(x) = c,\ \ c \in \mathbb{R}$ arbitrary and fixed. Consider the coordinates $x = (x^1, \ldots, x^n) \in \mathbb{R}^n$, and a curve on Σ,

$$\alpha : I \subset \mathbb{R} \to \Sigma \subset \mathbb{R}^n,$$

given by $\alpha(t) = (x^1(t), \ldots, x^n(t)), t \in I$. Then

$$f(\alpha(t)) = f(x^1(t), \ldots, x^n(t)) = c, \quad \text{for all } t \in I. \tag{41}$$

Differentiating (41) by t and considering the relations

$$\operatorname{grad} f = g^{ij}\frac{\partial f}{\partial x^j}\frac{\partial}{\partial x^i} \quad \text{and} \quad \alpha' = \left(\frac{dx^1}{dt}, \ldots, \frac{dx^n}{dt}\right),$$

we infer that

$$\frac{\partial f}{\partial x^i}\frac{dx^i}{dt} = 0 \quad \Leftrightarrow \quad g(\operatorname{grad} f, \alpha') = 0.$$

Indeed

$$g(\text{ grad } f, \alpha') = g_{rs}\left(g^{ri}\frac{\partial f}{\partial x^i}\right)\frac{dx^s}{dt} = \delta^i_s \frac{\partial f}{\partial x^i}\frac{dx^s}{dt} = \frac{\partial f}{\partial x^i}\frac{dx^i}{dt} = 0,$$

and hence $\text{grad} f \perp_g \alpha'$. □

1.8.3 Let ∇ be the Riemannian connection on D. The *Hessian* is the operator $\text{Hess} : \mathcal{F}(D) \to \mathcal{T}^0_2(D)$, which associates to any function $f \in \mathcal{F}(D)$, a $(0, 2)$-type tensor, given by $\text{Hess} f = \nabla(\mathrm{d}f)$, i.e.,

$$\text{Hess} f(X, Y) = (\nabla_X \, \mathrm{d}f)(Y), \ \ , \text{ for all } X, Y \in \mathcal{X}(D).$$

In coordinates,

$$\text{Hess} f = (\text{ Hess} f)_{jk} \, dx^j \otimes dx^k \tag{42}$$

with the components

$$(\text{ Hess} f)_{jk} = \frac{\partial^2 f}{\partial x^j \partial x^k} - \Gamma^h_{jk}\frac{\partial f}{\partial x^h}, \ \ j, k = \overline{1, n}, \tag{43}$$

where Γ^i_{jk} are the Christoffel symbols of g. The tensor field $\text{Hess} f$ is called *the Hessian of* f. The operator Hess is intensively used in the theory of extrema.

1.8.4 The *divergence* is the operator $\text{div} : \mathcal{X}(D) \to \mathcal{F}(D)$, given by

$$\text{div} X = X^i_{,i} = \frac{\partial X^i}{\partial x^i} + \Gamma^i_{ik}X^k, \ \text{ for all } X = X^i\frac{\partial}{\partial x^i} \in \mathcal{X}(D). \tag{44}$$

We remind that $X^i_{,j} = \frac{\partial X^i}{\partial x^j} + \Gamma^i_{jk}X^k$, and that the sum about i above means the transvection tr^1_1 of the $(1, 1)$-type field $T = X^i_{,j}\frac{\partial}{\partial x^i} \otimes dx^j \in \mathcal{T}^1_1(D)$.

Proposition. Denoting $G = \det(g_{ij})$, an equivalent expression for $\text{div} X$ is

$$\text{div} X = \frac{1}{\sqrt{G}}\frac{\partial}{\partial x^i}(\sqrt{G}X^i), \ \text{ for all } X \in \mathcal{X}(D).$$

Proof. By direct computation, we obtain

$$\Gamma^i_{ij} = \frac{1}{2}g^{ik}\left(\frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k}\right) = \frac{1}{2}g^{ik}\frac{\partial g_{ik}}{\partial x^j}, \tag{45}$$

and the partial derivative of the determinant G can be written

$$\frac{\partial G}{\partial x^j} = \frac{\partial G}{\partial g_{ik}}\frac{\partial g_{ik}}{\partial x^j} = Gg^{ik} \cdot \frac{\partial g_{ik}}{\partial x^j}. \tag{46}$$

Using (46), the relation (45) becomes

$$\Gamma^i_{ij} = \frac{1}{2G}\frac{\partial G}{\partial x^j} = \frac{1}{\sqrt{G}}\frac{\partial \sqrt{G}}{\partial x^j},$$

and, replacing in the formula (44), we infer

$$\operatorname{div} X = X^i_{,i} = \frac{\partial X^i}{\partial x^i} + \frac{1}{\sqrt{G}}\left(\frac{\partial\sqrt{G}}{\partial x^k}\right)X^k = \frac{1}{\sqrt{G}}\left(\sqrt{G}\frac{\partial X^i}{\partial x^i} + X^i\frac{\partial\sqrt{G}}{\partial x^i}\right),$$

whence the statement holds true. □

Remarks. The divergence has the following properties

1) $\operatorname{div}(fX) = X(f) + f\operatorname{div} X$;

2) $\operatorname{div}(aX + bY) = a\operatorname{div} X + b\operatorname{div} Y$, for all $a, b \in \mathbb{R}, X, Y \in \mathcal{X}(D), f \in \mathcal{F}(D)$.

1.8.5 Definitions. a) The *Laplacian* is the operator $\Delta : \mathcal{F}(D) \to \mathcal{F}(D)$, given by

$$\Delta f = \operatorname{div}(\operatorname{grad} f), \quad \text{for all } f \in \mathcal{F}(D).$$

In coordinates, it has the expression

$$\Delta f = \frac{1}{\sqrt{G}}\frac{\partial}{\partial x^l}\left(\sqrt{G}g^{kl}\frac{\partial f}{\partial x^k}\right) = g^{kl}f_{kl},$$

where $\{f_{kl}\}_{k,l=\overline{1,n}}$ are the components of the Hessian (43).

b) A function $f \in \mathcal{F}(D)$ which obeys $\Delta f = 0$ is called *harmonic*.

1.8.6 Definition. The *curl* (*rotor*) is the operator $\operatorname{curl} : \mathcal{X}(D) \to \mathcal{T}^0_2(D)$, given by

$$\operatorname{curl} X = (\omega_{j;i} - \omega_{i;j})\, dx^i \otimes dx^j \in \mathcal{T}^0_2(D), \quad \text{for all } X \in \mathcal{X}(D),$$

with $\omega_j = g_{js}X^s$, i.e., the 1-form $\omega = \omega_j dx^j \in \mathcal{X}^*(D)$ is obtained from X by lowering the index and $\omega_{j;i} = \frac{\partial\omega_j}{\partial x^i} - \Gamma^k_{ij}\omega_k$, $i, j = \overline{1,n}$, where Γ^k_{ij} are the Christoffel symbols of g. In the Euclidean case, when $g_{ij} = \delta_{ij}$, we have

$$\operatorname{curl} X = \left(\frac{\partial}{\partial x^i}X^j - \frac{\partial}{\partial x^j}X^i\right)dx^i \otimes dx^j, \quad \text{for all } X \in \mathcal{X}(D).$$

Remarks. 1°. The tensor $\operatorname{curl} X \in \mathcal{T}^0_2(D)$ is antisymmetrical in lower indices (i.e., denoting the components by a_{ij}, we have $a_{ij} = -a_{ji}$, for all $i, j = \overline{1,n}$).

2°. In the case $D \subset \mathbb{R}^3$, we can attach to $\operatorname{curl} X$ a contravariant vector field, namely

$$(\operatorname{curl} X)^i = \frac{1}{\sqrt{G}}\left(\frac{\partial}{\partial x^j}(g_{ks}X^s) - \frac{\partial}{\partial x^k}(g_{js}X^s)\right),$$

where the triple $\{i, j, k\}$ is a cyclic permutation of the numbers $\{1, 2, 3\}$.

3°. For $D \subset \mathbb{R}^3$, to $\operatorname{curl} X$ we can associate the (antisymmetrical) matrix of the form

$$[\operatorname{curl} X] = \begin{pmatrix} 0 & a_{12} & a_{13} \\ -a_{12} & 0 & a_{23} \\ -a_{13} & -a_{23} & 0 \end{pmatrix}.$$

1.8.7 Let us change the Cartesian coordinates (x, y, z) to the orthogonal coordinates (u, v, w), via the formulas

$$\begin{cases} x = x(u, v, w) \\ y = y(u, v, w) \\ z = z(u, v, w). \end{cases}$$

Fix a point $P(x_0, y_0, z_0) \in \mathbb{R}^3$,

$$\begin{cases} x_0 = x(u_0, v_0, w_0) \\ y_0 = y(u_0, v_0, w_0) \\ z_0 = z(u_0, v_0, w_0). \end{cases}$$

Consider the three coordinate lines and the three coordinate surfaces through this point; for example, the (u)-line is given parametrically by

$$\begin{cases} x = x(u, v_0, w_0) \\ y = y(u, v_0, w_0) \\ z = z(u, v_0, w_0) \end{cases}$$

and the $u = u_0$ surface is given parametrically by

$$\begin{cases} x = x(u_0, v, w) \\ y = y(u_0, v, w) \\ z = z(u_0, v, w). \end{cases}$$

We suppose that the new system is orthogonal (i.e., the tangents to the coordinate lines are mutually orthogonal).

Let $\vec{r} = x(u, v, w)\vec{i} + y(u, v, w)\vec{j} + z(u, v, w)\vec{k}$ be the position vector and the partial derivatives

$$\vec{r}_u = \frac{\partial \vec{r}}{\partial u},\ \vec{r}_v = \frac{\partial \vec{r}}{\partial v},\ \vec{r}_w = \frac{\partial \vec{r}}{\partial w}.$$

These vectors are mutually orthogonal, and the Lamé coefficients of the new coordinate system are the norms

$$H_u = ||\vec{r}_u||,\ H_v = ||\vec{r}_v||,\ H_w = ||\vec{r}_w||,$$

and one can normalize the family of these vectors, obtaining the orthonormal basis $\{\vec{e}_u, \vec{e}_v, \vec{e}_w\} \subset \mathcal{X}(\mathbb{R}^3)$ given by

$$\vec{e}_u = \frac{1}{H_u}\vec{r}_u,\ \vec{e}_v = \frac{1}{H_v}\vec{r}_v,\ \vec{e}_w = \frac{1}{H_w}\vec{r}_w.$$

The area element in the plane $x = x_0$, expressed in Cartesian coordinates is

$$\mathrm{d}\sigma = dy \wedge dz \overset{not}{=} dy\, dz,$$

and the area element on the curvilinear surface $u = u_0$ is given by

$$d\sigma = H_v\ H_w dv \wedge dw \overset{not}{=} H_v\ H_w\ dv\ dw.$$

Also, *the volume element* in the Cartesian system is

$$dV = dx \wedge dy \wedge dz \overset{not}{=} dx\ dy\ dz$$

and in curvilinear coordinates,

$$dV = H_u H_v H_w\ du \wedge dv \wedge dw \overset{not}{=} H_u H_v H_w\ du\ dv\ dw.$$

1.8.8 Let $D \subset \mathbb{R}^3$. We pass from the Cartesian coordinates $(x^1, x^2, x^3) = (x, y, z)$. We pass to a general orthogonal system of coordinates $(x^{1'}, x^{2'}, x^{3'}) = (u, v, w)$.

The *gradient*

$$\operatorname{grad} f = \frac{\partial f}{\partial x}\vec{i} + \frac{\partial f}{\partial y}\vec{j} + \frac{\partial f}{\partial z}\vec{k} \in \mathcal{X}(D)$$

rewrites

$$\operatorname{grad} f = \frac{1}{H_u}\frac{\partial f}{\partial u}\vec{e}_u + \frac{1}{H_v}\frac{\partial f}{\partial v}\vec{e}_v + \frac{1}{H_w}\frac{\partial f}{\partial w}\vec{e}_w = \underset{uvw}{S}\frac{1}{H_u}\frac{\partial f}{\partial u}\vec{e}_u, \tag{47}$$

where the coefficients of the linear combination above are the *physical components of* grad *f*.

Also, *in physical components*, we have

$$\vec{V} = V_u\vec{e}_u + V_v\vec{e}_v + V_w\vec{e}_w \in \mathcal{X}(D),\ D \subset \mathbb{R}^3;$$

the *rotor* and the *divergence* of the vector field $\vec{V}$ rewrite

$$\begin{aligned} \operatorname{curl}\vec{V} &= \frac{1}{H_u H_v H_w}\begin{vmatrix} H_u\vec{e}_u & H_v\vec{e}_v & H_w\vec{e}_w \\ \frac{\partial}{\partial u} & \frac{\partial}{\partial v} & \frac{\partial}{\partial w} \\ H_u V_u & H_v V_v & H_w V_w \end{vmatrix} \\ \operatorname{div}\vec{V} &= \frac{1}{H_u H_v H_w}\left[\frac{\partial}{\partial u}(V_u H_v H_w) + \frac{\partial}{\partial v}(V_v H_w H_u) + \frac{\partial}{\partial w}(V_w H_u H_v)\right] = \\ &= \frac{1}{H}\underset{uvw}{\mathrm{S}}\frac{\partial}{\partial u}(V_u H_v H_w), \end{aligned} \tag{48}$$

where $H = H_u H_v H_w$. Moreover, since $\Delta f = \operatorname{div}(\operatorname{grad} f)$, for all $f \in \mathcal{F}(D)$, replacing $\vec{V}$ with $\operatorname{grad} f$, we obtain the expression of the *Laplacian* of f,

$$\begin{aligned} \Delta f &= \frac{1}{H_u H_v H_w}\left[\frac{\partial}{\partial u}\left(\frac{\partial f}{\partial u}\frac{H_v H_w}{H_u}\right) + \frac{\partial}{\partial v}\left(\frac{\partial f}{\partial v}\frac{H_w H_u}{H_v}\right) + \frac{\partial}{\partial w}\left(\frac{\partial f}{\partial w}\frac{H_u H_v}{H_w}\right)\right] = \\ &= \frac{1}{H}\underset{uvw}{S}\frac{\partial}{\partial u}\left(\frac{\partial f}{\partial u}\frac{H_v H_w}{H_u}\right). \end{aligned} \tag{49}$$

Since $\vec{V}(f) = g(\vec{V}, \text{grad}\, f)$, we find also

$$\vec{V}(f) = \frac{V_u}{H_u}\frac{\partial f}{\partial u} + \frac{V_v}{H_v}\frac{\partial f}{\partial v} + \frac{V_w}{H_w}\frac{\partial f}{\partial w}, \quad \text{for all } f \in \mathcal{F}(D).$$

1.8.9. Exercises

1. Find the Lamé coefficients for cylindrical coordinates.

Solution. Considering the change from Cartesian coordinates to cylindrical coordinates

$$(x^1, x^2, x^3) = (x, y, z) \to (x^{1'}, x^{2'}, x^{3'}) = (\rho, \theta, z),$$

we obtain the vector field $\vec{r} = \rho\cos\theta\vec{i} + \rho\sin\theta\vec{j} + z\vec{k}$, and the partial derivatives

$$\begin{cases} \vec{r}_u = \vec{r}_\rho = \cos\theta\vec{i} + \sin\theta\vec{j} \\ \vec{r}_v = \vec{r}_\theta = -\rho\sin\theta\vec{i} + \rho\cos\theta\vec{j} \\ \vec{r}_w = \vec{r}_z = \vec{k}. \end{cases}$$

The Lamé coefficients are

$$\begin{cases} H_u = ||\vec{r}_u|| = 1 \\ H_v = ||\vec{r}_v|| = \rho \\ H_w = ||\vec{r}_w|| = 1, \end{cases}$$

and the orthonormal basis is

$$\begin{cases} \vec{e}_u = \cos\theta\vec{i} + \sin\theta\vec{j} \\ \vec{e}_v = -\sin\theta\vec{i} + \cos\theta\vec{j} \\ \vec{e}_w = \vec{k}. \end{cases}$$

Since in the new orthogonal coordinates (u, v, w), the *natural metric*

$$g = \delta_{ij}dx^i dx^j = dx^2 + dy^2 + dz^2$$

of $\mathbb{R}^3$ rewrites

$$g = H_u^2 \mathrm{d}u^2 + H_v^2 \mathrm{d}v^2 + H_w^2 \mathrm{d}w^2,$$

we obtain in cylindrical coordinates

$$g = \mathrm{d}s^2 = \mathrm{d}\rho^2 + \rho^2 \mathrm{d}\theta^2 + dz^2.$$

♣ Hw. Show that for spherical coordinates (r, φ, θ) provided by the relations between the Cartesian/spherical coordinates (26), we obtain the *Lamé coefficients*

$$H_r = 1,\ H_\varphi = r,\ H_\theta = r\sin\varphi.$$

♣ Hw. Determine in cylindrical and spherical coordinates the covariant, contravariant components, divergence and curl of the vector field

$$\vec{V} = x\vec{i} + 2xy\vec{j} + z\vec{k} = X \in \mathcal{X}(D).$$

Hint. Compute X, $\operatorname{div} X$, $\operatorname{curl} X$ in Cartesian coordinates, and then determine the components in the new orthogonal systems of coordinates.

2. In the Poincaré plane $H = \{(x, y) \in \mathbb{R}^2 \mid y > 0\}$, with the *metric* $g \in \mathcal{T}_2^0(H)$ of components

$$g_{11} = g_{22} = \frac{1}{y^2}, \; g_{12} = g_{21} = 0,$$

compute

a) the *angle* $\theta = (\widehat{X, Y})$ between the vector fields X, Y, where

$$X = xy\frac{\partial}{\partial x} + \frac{\partial}{\partial y}, \quad Y = y\frac{\partial}{\partial y} \in \mathcal{X}(H);$$

b) the *length* (*norm*) $||X|| \in \mathcal{F}(H)$ of the vector field $X \in \mathcal{X}(H)$;

c) the *curve-length* $l(\gamma|_{[0,1]})$ of $\gamma(t) = (t, t^2 + 1)$, $t \in [0, 1]$.

Solution. a) We have

$$\cos\theta = \frac{\langle X, Y\rangle_g}{||X|| \cdot ||Y||} = \frac{1}{y^2} \cdot y \cdot \frac{1}{\frac{1}{y}\sqrt{x^2y^2 + 1}} = \frac{1}{\sqrt{1 + x^2y^2}} \in (0, 1],$$

whence $\theta = \arccos \frac{1}{\sqrt{1+x^2y^2}} \in \left[0, \frac{\pi}{2}\right]$.

b) We obtain $||X|| = \frac{\sqrt{1+x^2y^2}}{y}$.

c) The *velocity* of the parametrized curve and its *speed* are respectively

$$\gamma'(t) = (1, 2t)$$

and

$$\begin{aligned} ||\gamma'(t)|| &= \sqrt{\langle\gamma'(t), \gamma'(t)\rangle_g} = \sqrt{(1, 2t)\begin{pmatrix} 1/y^2 & 0 \\ 0 & 1/y^2 \end{pmatrix}\Bigg|_{x=t, y=t^2+1}\begin{pmatrix} 1 \\ 2t \end{pmatrix}} = \\ &= \sqrt{\tfrac{1}{y^2} + \tfrac{1}{y^2} \cdot 4t^2}\,\Bigg|_{x=t,y=t^2+1} = \frac{\sqrt{1 + 4t^2}}{t^2 + 1}. \end{aligned}$$

Then its *length* is

$$l_{[0,1],\gamma} = \int_0^1 \frac{\sqrt{1 + 4t^2}}{t^2 + 1} dt = -\sqrt{3} \cdot \operatorname{arctanh}\sqrt{3/5} - 2\log(\sqrt{5} - 2).$$

3. Compute the *Christoffel symbols* for the metric from the previous exercise.

Solution. The *reciprocal tensor field* of g has the matrix

$$(g^{ij}) = \begin{pmatrix} y^2 & 0 \\ 0 & y^2 \end{pmatrix}.$$

We replace its components in the expressions of the Levi-Civita connection components,

$$\Gamma^i_{jk} = \frac{1}{2} g^{is}(\partial_{\{j} g_{sk\}} - \partial_s g_{jk}),$$

where $\tau_{\{i\ldots j\}} \equiv \tau_{i\ldots j} + \tau_{j\ldots i}$, and obtain the nontrivial components

$$\Gamma^1_{12} = \Gamma^1_{21} = \Gamma^2_{22} = -\frac{1}{y}, \Gamma^2_{11} = \frac{1}{y},$$

all the other components being null.

4. a) Compute the Lamé coefficients for polar coordinates.

b) Find the covariant and the physical components of the vector field

$$X = \frac{\theta}{\rho}\frac{\partial}{\partial\rho} - \theta\frac{\partial}{\partial\theta}.$$

Solution. a) The coordinate change from Cartesian to polar coordinates (27) is provided by (28). The relation between the new/old components of the metric tensor are

$$g_{i'j'} = \frac{\partial x^i}{\partial x^{i'}}\frac{\partial x^j}{\partial x^{j'}} g_{ij} = \frac{\partial x^i}{\partial x^{i'}}\frac{\partial x^j}{\partial x^{j'}}\delta_{ij},$$

where $\left(\frac{\partial x^i}{\partial x^{i'}}\right)_{i,i'=\overline{1,2}} = \begin{pmatrix} \cos\theta & -\rho\sin\theta \\ \sin\theta & \rho\cos\theta \end{pmatrix}$. We obtain the new components

$$(g_{i'j'})_{i',j'=\overline{1,2}} = \begin{pmatrix} 1 & 0 \\ 0 & \rho^2 \end{pmatrix},$$

whence the Lamé coefficients follow straightforward

$$\begin{cases} H_{1'} = \sqrt{g_{1'1'}} = 1 \\ H_{2'} = \sqrt{g_{2'2'}} = \rho. \end{cases}$$

b) We have $X = X^{i'}\frac{\partial}{\partial x^{i'}} = (X^{i'}H_{i'})e_{i'}$. Then the physical components are given by $\tilde{X}^{i'} = X^{i'}H_{i'}$ and the covariant (lowered) ones, by $X_{j'} = g_{i'j'}X^{i'}$.

♣ Hw. Compute these components.

5. Let $D = \mathbb{R}^3\backslash Oz$ be endowed with the metric $g = \delta_{ij}dx^i \otimes dx^j \in \mathcal{T}^0_2(D)$. Determine the components of the metric g in cylindrical and spherical coordinates, and the Christoffel symbols.

Hint. The cylindrical coordinates were described in (18); the relation between the Cartesian coordinates $(x, y, z) = (x^1, x^2, x^3)$ and the spherical coordinates $(x^{1'}, x^{2'}, x^{3'}) = (r, \varphi, \theta)$ is described by

$$\begin{cases} x = r\sin\varphi\cos\theta \\ y = r\sin\varphi\sin\theta \\ z = r\cos\varphi \end{cases}, \quad r \in (0,\infty),\ \varphi \in [0,\pi],\ \theta \in [0,2\pi). \tag{50}$$

Method 1. Since $g = \delta_{ij} dx^i \otimes dx^j = g_{i'j'} dx^{i'} \otimes dx^{j'}$ and the relation between the new/old components is

$$g_{i'j'} = \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^j}{\partial x^{j'}} \, g_{ij}, \tag{51}$$

using (50), one determines the matrix $C = \left(\frac{\partial x^i}{\partial x^{i'}}\right)_{i,i'=\overline{1,3}}$ and replaces in (51).

Method 2. By straightforward differentiation in (50), one gets

$$dx = \mathrm{d}(r \sin\varphi \cos\theta) = \sin\varphi\cos\theta \;\mathrm{d}r + r\cos\varphi\cos\theta \;\mathrm{d}\varphi - r\sin\varphi\sin\theta \;\mathrm{d}\theta,$$

and dy, dz similarly expressed in terms of $\mathrm{d}r, \mathrm{d}\varphi, \mathrm{d}\theta$. Replacing in

$$g = \delta_{ij} dx^i \otimes dx^j \equiv dx^2 + dy^2 + dz^2,$$

we find the new components of g.

6. Write the formulas (47), (48), (49) for $\operatorname{grad} f$, $\operatorname{div} X$, Δf, $\operatorname{curl} X$ in cylindrical and spherical coordinates, on $D \subset \mathbb{R}^3 \backslash Oz$.

7. Let be $f \in \mathcal{F}(\mathbb{R}^2)$, $f(x,y) = xy^2$. Compute:

a) $\operatorname{grad} f \in \mathcal{X}(\mathbb{R}^2)$;

b) $\operatorname{Hess} f \in \mathcal{T}^0_2(\mathbb{R}^2)$.

Solution. a) The gradient is $\operatorname{grad} f = \frac{\partial f}{\partial x}\vec{i} + \frac{\partial f}{\partial y}\vec{j} = y^2\vec{i} + 2xy\vec{j}$.

b) The Hessian writes

$$\operatorname{Hess} f = \frac{\partial^2 f}{\partial x^i \partial x^j} dx^i \otimes dx^j; \;\; [\operatorname{Hess} f] = \begin{pmatrix} 0 & 2y \\ 2y & 2x \end{pmatrix}.$$

8. Let be the vector field $X = yx\frac{\partial}{\partial y} - y^2\frac{\partial}{\partial x} + z\frac{\partial}{\partial z} \in \mathcal{X}(\mathbb{R}^3)$. Compute

a) $\operatorname{div} X \in \mathcal{F}(\mathbb{R}^3)$;

b) $\operatorname{curl} X \in \mathcal{X}(\mathbb{R}^3)$.

Solution. a) The divergence of X is $\operatorname{div} X = \displaystyle\sum_{i=1}^{3} \frac{\partial X^i}{\partial x^i} = x + 1$.

b) The curl of X is

$$\operatorname{curl} X = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ X^1 & X^2 & X^3 \end{vmatrix} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ -y^2 & xy & z \end{vmatrix} = 0\vec{i} + 0\vec{j} + (y + 2y)\vec{k} = 3y\vec{k}.$$

9. Let be given $X = xy\frac{\partial}{\partial x}$, $f(x,y) = x - y^2$, $Y = xy\frac{\partial}{\partial y}$. Check the following relations:

a) $\mathrm{div}\,(fX) = X(f) + f\ \mathrm{div}\,X$;

b) $\mathrm{div}\,(aX + bY) = a\ \mathrm{div}\,X + b\ \mathrm{div}\,Y$, (for $a = 5$ and $b = -1$).

10. Let $D = \mathbb{R}^2$ endowed with the metric g of matrix $[g] = \begin{pmatrix} 1+x^2 & 0 \\ 0 & 1 \end{pmatrix}$.

a) Compute the gradient (40) of the function $f(x,y) = xy^2, f \in \mathcal{F}(\mathbb{R}^2)$.

b) For $a = 5, b = 7$ and $g(x,y) = x - y$, check that

$$\begin{array}{rcl} \mathrm{grad}\,(af + bg) & = & a\ \mathrm{grad}\,f + b\ \mathrm{grad}\,g \\ \mathrm{grad}\left(\dfrac{f}{g}\right) & = & \dfrac{1}{g^2}(g\ \mathrm{grad}\,f - f\ \mathrm{grad}\,g). \end{array}$$

c) For $f(x,y) = x^2 + y^2, f \in \mathcal{F}(D), D = \mathbb{R}^2\backslash\{(0,0)\}$, show that:

- Crit $f = \emptyset$;
- $\mathrm{grad}\,f \perp \Gamma$, where Γ is the curve described by $f(x,y) = c,\ c \in \mathbb{R}^*_+$;
- $\mathrm{grad}\,f$ indicates the sense of the *steepest increase* of f.

Solution. a) The reciprocal tensor field g^{-1} of g has the associated matrix

$$[g^{-1}] = (g^{ij})_{i,j\in\overline{1,2}} = \begin{pmatrix} \dfrac{1}{1+x^2} & 0 \\ 0 & 1 \end{pmatrix}.$$

Hence applying the formula (40), we get $\mathrm{grad}\,f = \frac{y^2}{1+x^2}\frac{\partial}{\partial x} + 2xy\frac{\partial}{\partial y}$.

♣ Hw. Solve b). For c), we remark that

$$\mathrm{grad}\,f = 0 \Leftrightarrow \frac{2x}{1+x^2}\frac{\partial}{\partial x} + 2y\frac{\partial}{\partial y} = 0 \Rightarrow (x,y) = (0,0) \notin \ \mathrm{Dom}\ f,$$

so that we obtain $Crit(f) = \emptyset$. The curves $f(x,y) = c$ are described by the equations of the form $x^2 + y^2 = c,\ c > 0$, hence are circles parametrized by

$$\gamma(t) = (\sqrt{c}\cos t, \sqrt{c}\sin t) \ \Rightarrow\ \gamma'(t) = (-\sqrt{c}\sin t, \sqrt{c}\cos t),$$

whence

$$g(\gamma', \nabla f) = \frac{2\sqrt{c}\cos t}{1 + c\cos^2 t}(-\sqrt{c}\sin t)(1 + c^2\cos^2 t) + 2\sqrt{c}\sin t \cdot \sqrt{c}\cos t = 0,$$

for all $t \in I$, so $\mathrm{grad}\,f \perp \Gamma$.

The *variation* $D_u(f)$ of f in the direction provided by a unit vector $u = (a, b)$ (hence satisfying $\sqrt{a^2+b^2} = 1$) is given by $D_u(f) = \langle u, \ \mathrm{grad}\, f\rangle$. Its absolute value is subject to the *Cauchy-Schwartz inequality*

$$|\langle u, \ \mathrm{grad}\, f\rangle| \leq ||u|| \cdot ||\ \mathrm{grad}\, f\ || = ||\ \mathrm{grad}\, f\ || = \sqrt{\left(\frac{y^2}{1+x^2}\right)^2 + (2xy)^2}$$

and is maximized only when the relation above becomes equality, which takes place iff u is collinear to $\mathrm{grad}\, f$, i.e., when the variation is considered in the direction of $\mathrm{grad}\, f$. Hence $\mathrm{grad}\, f$ shows the direction of the *steepest* increase of f at any point of $\mathbb{R}^2$.

11. Compute the Hessian of the function $f(x, y) = yx + y^2$.

Hint. Apply the formulas (42), (43) and identify $(x, y) \equiv (x^1, x^2)$.

12. Compute the divergence of the vector field

$$X = xy\frac{\partial}{\partial y} - z\frac{\partial}{\partial x} \in \mathcal{X}(\mathbb{R}^3),$$

for $\mathbb{R}^3$ endowed with the metric g of matrix

$$[g] = \begin{pmatrix} 1+x^2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and then in the Poincaré plane $H, g)$ with

$$H = \{(x, y) \in \mathbb{R}^2 | y > 0\}, \qquad [g] = \begin{pmatrix} 1/y^2 & 0 \\ 0 & 1/y^2 \end{pmatrix},$$

using the following equivalent formulas:

a) $\mathrm{div}\, X = X^i_{;i}$, where $X^i_{;j} = \dfrac{\partial X^i}{\partial x^j} + \Gamma^i_{jk} X^k$;

b) $\mathrm{div}\, X = \dfrac{1}{\sqrt{G}} \dfrac{\partial}{\partial x^i}(\sqrt{G} X^i)$, where $G = \det\, [g]$.

Hint. The only nonzero components of the Levi-Civita connection on the Poincaré plane are

$$\Gamma^1_{12} = \Gamma^1_{21} = \Gamma^2_{22} = -\frac{1}{y}, \Gamma^2_{11} = \frac{1}{y}.$$

Also, since $g_{11} = g_{22} = \dfrac{1}{y^2}$, $g_{12} = g_{21} = 0$, we have $G = \det\, (g_{ij})_{i,j=\overline{1,2}} = \dfrac{1}{y^4}$.

13. Compute *the curl (rotor)* of the vector field $X = x\frac{\partial}{\partial y} - z\frac{\partial}{\partial z} \in \mathcal{X}(\mathbb{R}^3)$.

Hint. Use the formula

$$\omega = \text{curl}\, X = \left[\frac{\partial}{\partial x^i}(g_{hj}X^h) - \frac{\partial}{\partial x^j}(g_{hi}X^h)\right] dx^i \otimes dx^j.$$

14. Let $D \subset \mathbb{R}^3$ be endowed with the metric $g_{ij} = \delta_{ij}$. Consider the function $f(x, y, z) = x^2 z$, $f \in \mathcal{F}(D)$, and the vector field $X = xz\frac{\partial}{\partial y}$, $X \in \mathcal{X}(D)$. Compute $\text{grad}\, f$, $\text{Hess}\, f, \Delta f$, $\text{curl}\, X$ and $\text{div}\, X$ in cylindrical and in spherical coordinates.

Solution. We present two equivalent approaches.

Method 1. Denoting $(x, y, z) = (x^1, x^2, x^3)$, we use the formulas

$$\begin{aligned}
\text{grad}\, f &= g^{ij}\frac{\partial f}{\partial x^i}\frac{\partial}{\partial x^j} \\
\text{Hess}\, f &= \left(\frac{\partial^2 f}{\partial x^j \partial x^k} - \Gamma^i_{jk}\frac{\partial f}{\partial x^i}\right) dx^j \otimes dx^k \\
\Delta f &= \text{div}\,(\,\text{grad}\, f) = \frac{1}{\sqrt{G}}\frac{\partial}{\partial x^s}\left(\sqrt{G}\, g^{rs}\frac{\partial f}{\partial x^r}\right) \\
\text{curl}\, X &= \frac{1}{\sqrt{G}} \underset{ijk}{\mathrm{S}} \left[\frac{\partial}{\partial x^j}(g_{ks}X^s) - \frac{\partial}{\partial x^k}(g_{js}X^s)\right]\frac{\partial}{\partial x^i} \\
\text{div}\, X &= \frac{1}{\sqrt{G}}\frac{\partial}{\partial x^i}(\sqrt{G}X^i),
\end{aligned} \tag{52}$$

where by $\underset{ijk}{\mathrm{S}}$ we denoted the cyclic sum over the indices $\{1, 2, 3\} = \{i, j, k\}$. In our case, we have

$$g_{ij} = \delta_{ij}; \ \ g^{ij} = \delta^{ij}; \ \ G = \det\,(g_{ij})_{i,j=\overline{1,3}} = 1,$$

and $\Gamma^i_{jk} = \frac{1}{2}g^{is}\left(\frac{\partial g_{js}}{\partial x^k} + \frac{\partial g_{ks}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^s}\right) \equiv 0$, whence the formulas become

$$\begin{aligned}
\text{grad}\, f &= \frac{\partial f}{\partial x}\frac{\partial}{\partial x} + \frac{\partial f}{\partial y}\frac{\partial}{\partial y} + \frac{\partial f}{\partial z}\frac{\partial}{\partial z} \\
\text{Hess}\, f &= \frac{\partial^2 f}{\partial x^j \partial x^k} dx^j \otimes dx^k \\
\Delta f &= \sum_{i=1}^{3} \frac{\partial^2 f}{\partial {x^i}^2} \\
\text{curl}\, X &= \underset{ijk}{\mathrm{S}}\left(\frac{\partial X^k}{\partial x^j} - \frac{\partial X^j}{\partial x^k}\right)\frac{\partial}{\partial x^i}, \ \text{where } \{i, j, k\} = \{1, 2, 3\} \\
\text{div}\, X &= \sum_{i=1}^{3} \frac{\partial X^i}{\partial x^i}.
\end{aligned}$$

Since $X^2 = xz,\ X^1 = X^3 = 0$, by direct computation we obtain

$$\begin{aligned} \operatorname{grad} f &= 2zx\tfrac{\partial}{\partial x} + x^2\tfrac{\partial}{\partial z}, \\ \operatorname{Hess} f &= 2z dx \otimes dx + 2x dx \otimes dz + 2x dz \otimes dx, \\ \Delta f &= 2z, \\ \operatorname{div} X &= \tfrac{\partial(xz)}{\partial y} = 0 \end{aligned}$$

and

$$\operatorname{curl} X = -\frac{\partial X^2}{\partial x^3}\frac{\partial}{\partial x^1} + \frac{\partial X^2}{\partial x^1}\frac{\partial}{\partial x^3}2,$$

whence $\operatorname{curl} X = -x\frac{\partial}{\partial x} + z\frac{\partial}{\partial z}$. Then, after changing the coordinates, one can compute the new components of the tensor fields

$$\begin{aligned} &U = \operatorname{grad} f \in \mathcal{X}(D), \quad H = \operatorname{Hess} f \in \mathcal{T}^0_2(D),\ Y = \operatorname{curl} X \in \mathcal{X}(D), \\ &\varphi = \operatorname{div} X \in \mathcal{F}(D), \qquad \psi = \Delta f \in \mathcal{F}(D), \end{aligned}$$

with the corresponding formulas of change of components, and express these tensor fields with respect to the new coordinates $(x^{1'}, x^{2'}, x^{3'}) \equiv (\rho, \theta, z)$ or, respectively (r, φ, θ) for spherical coordinates.

Method 2.

- Determine the new components $(X^{1'}, X^{2'}, X^{3'})$ of X and rewrite f in new coordinates.
- We have $G = \delta_{ij} dx^i \otimes dx^j = g_{i'j'} dx^{i'} \otimes dx^{j'}$. For computing the components $g_{i'j'}$, we either use the new/old components relations $g_{i'j'} = \frac{\partial x^i}{\partial x^{i'}}\frac{\partial x^j}{\partial x^{j'}} g_{ij}$, or compute directly $dx^i = \frac{\partial x^i}{\partial x^{i'}} dx^{i'}$ and replace in g. As consequence, one gets for cylindrical coordinates

  $$g = \mathrm{d}\rho \otimes \mathrm{d}\rho + \rho^2\ \mathrm{d}\theta \otimes \mathrm{d}\theta + dz \otimes dz,$$

  and

  $$g = dr \otimes dr + r^2 d\varphi \otimes \mathrm{d}\varphi + r^2 \sin^2 \varphi\ \mathrm{d}\theta \otimes \mathrm{d}\theta,$$

  in spherical coordinates.
- Apply the general formulas (35) and (18), whence

  $$g_{1'1'} = 1,\ g_{2'2'} = \rho^2,\ g_{3'3'} = 1;\ \ G = \rho^2$$

  for cylindrical coordinates, and respectively (35) and (50) for spherical coordinates, obtaining

  $$g_{1'1'} = 1,\ g_{2'2'} = r^2,\ g_{3'3'} = r^2 \sin^2 \varphi;\ \ G = r^4 \sin^2\varphi.$$

- Apply formulas (47),(48),(49) to the fields expressed in new coordinates.

15. Consider the cylindrical coordinates $(u, v, w) = (\rho, \theta, z)$ in $D \subset \mathbb{R}^3$, the mapping

$$f \in \mathcal{F}(D),\ f(\rho, \theta, z) = \rho \cos\theta + z,$$

and the vector field $V \in \mathcal{X}(D),\ V = \sin\theta \frac{\partial}{\partial\rho} - z\frac{\partial}{\partial\theta}$.

Find the physical components of V, and compute $\operatorname{grad} f, \Delta f,\ \operatorname{curl} V$ and $\operatorname{div} V$.

Solution. Remark that $(x^{1'}, x^{2'}, x^{3'}) = (u, v, w) = (\rho, \theta, z)$ are orthogonal coordinates. We compute first the Lamé coefficients in one of the following two ways

- $H_u = \sqrt{g_{1'1'}}, H_v = \sqrt{g_{2'2'}}, H_w = \sqrt{g_{3'3'}}$, or
- $H_u = ||\vec{r}_u||, H_v = ||\vec{r}_v||, H_w = ||\vec{r}_w||$, i.e., the lengths of the partial derivatives of the position vector

$$\vec{r}(u, v, w) = \vec{r}(\rho, \theta, z) = \rho\cos\theta\vec{i} + \rho\sin\theta\vec{j} + z\vec{k}$$

of a point $(x, y, z) = (\rho\cos\theta, \rho\sin\theta, z)$, where we have denoted

$$\vec{r}_u = \frac{\partial\vec{r}}{\partial u},\ \ \vec{r}_v = \frac{\partial\vec{r}}{\partial v},\ \ \vec{r}_w = \frac{\partial\vec{r}}{\partial w}.$$

We obtain the Lamé coefficients

$$H_u = 1, H_v = \rho, H_w = 1,$$

and consider the attached orthonormal basis $\{\vec{e}_u, \vec{e}_v, \vec{e}_w\}$ of the $\mathcal{F}(D)$-module $\mathcal{X}(D)$ given by the relations

$$\vec{e}_u = \frac{1}{H_u}\frac{\partial}{\partial u},\ \vec{e}_v = \frac{1}{H_v}\frac{\partial}{\partial v},\ \vec{e}_w = \frac{1}{H_w}\frac{\partial}{\partial w}.$$

Then we have

$$V = V^{1'}\frac{\partial}{\partial u} + V^{2'}\frac{\partial}{\partial v} + V^{3'}\frac{\partial}{\partial w} = V_u\vec{e}_u + V_v\vec{e}_v + V_w\vec{e}_w,$$

whence we infer *the physical components* of V

$$\begin{cases} V_u = H_u X^{1'} \\ V_v = H_v X^{2'} \\ V_w = H_w X^{3'} \end{cases} \Rightarrow \begin{cases} V_u = \sin v \cdot H_u = \sin v \\ V_v = (-w)\cdot H_v = -wu \\ V_w = 0. \end{cases}$$

It follows that $V = \sin v\ \vec{e}_u - uw\ \vec{e}_v = \sin\theta\ \vec{e}_\rho - \rho z\ \vec{e}_\theta$.

Then we apply the formulas

$$\begin{aligned}
\operatorname{grad} f &= \underset{uvw}{\mathrm{S}} \tfrac{1}{H_u} \tfrac{\partial f}{\partial u} \vec{e}_u \\
\Delta f &= \tfrac{1}{H} \underset{uvw}{\mathrm{S}} \tfrac{\partial}{\partial u} \left(\tfrac{\partial f}{\partial u} \tfrac{H_v H_w}{H_u} \right), \\
\operatorname{curl} V &= \tfrac{1}{H} \begin{vmatrix} H_u \vec{e}_u & H_v \vec{e}_v & H_w \vec{e}_w \\ \tfrac{\partial}{\partial u} & \tfrac{\partial}{\partial v} & \tfrac{\partial}{\partial w} \\ H_u V_u & H_v V_v & H_w V_w \end{vmatrix} \\
\operatorname{div} V &= \tfrac{1}{H} \underset{uvw}{\mathrm{S}} \tfrac{\partial}{\partial u} \left(V_u H_v H_w \right),
\end{aligned}$$

with $H \equiv H_u H_v H_w = \rho$, where by $\underset{uvw}{\mathrm{S}}$ we denoted the cyclic sum about u, v, w.

16. ♣ Hw. The same questions as in the previous exercise, for

$$f(r, \varphi, \theta) = r \cos \varphi - \sin \theta \quad \text{and} \quad V = r \frac{\partial}{\partial \varphi} - \sin \varphi \frac{\partial}{\partial \theta},$$

expressed in spherical coordinates $(u, v, w) = (r, \varphi, \theta)$.

Hint. Compute first the spherical Lamé coefficients

$$H_u = 1, \ H_v = u = r, \ H_w = u \sin v = r \sin \varphi.$$

17. Consider the *Poincaré 2-plane* $H = \{(x, y) \in \mathbb{R}^2 \mid y > 0\}$, endowed with the *metric*

$$g_{ij}(x, y) = \frac{1}{y^2} \, \delta_{ij}, \quad i, j = \overline{1, 2}.$$

Let be given the function $f(x, y) = \sqrt{x^2 + y^2}$ and the vector field

$$X = y \frac{\partial}{\partial x} + x^2 y \frac{\partial}{\partial y} \in \mathcal{X}(M).$$

Determine $\operatorname{grad} f$, $\operatorname{Hess} f, \operatorname{div} X, \Delta f$ and $\operatorname{curl} X$.

Hint. We use the formulas (52).

1°. The metric g and its reciprocal tensor g^{-1} have respectively the components

$$(g_{ij})_{i,j=\overline{1,2}} = \begin{pmatrix} \frac{1}{y^2} & 0 \\ 0 & \frac{1}{y^2} \end{pmatrix}, \quad (g^{ij})_{i,j=\overline{1,2}} = \begin{pmatrix} y^2 & 0 \\ 0 & y^2 \end{pmatrix}.$$

Then, denoting $Y = \operatorname{grad} f$, we compute its components

$$\begin{cases} Y^1 = g^{11} \frac{\partial f}{\partial x} = y^2 \cdot \dfrac{x}{f} \\ Y^2 = g^{22} \frac{\partial f}{\partial y} = y^2 \cdot \dfrac{y}{f} \end{cases} \Rightarrow \operatorname{grad} f = \frac{y^2}{\sqrt{x^2 + y^2}} \left(x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} \right).$$

2°. As for the components of the Levi-Civita connection $\{\Gamma^i_{jk}\}_{i,j,k=\overline{1,3}}$, we obtain

$$\Gamma^1_{12} = \Gamma^1_{21} = \Gamma^2_{22} = -\frac{1}{y}, \Gamma^2_{11} = \frac{1}{y},$$

and all the other components are null.

3°. The determinant of $[g]$ is $G = \frac{1}{y^4}$, so $\frac{1}{\sqrt{G}} = y^2$, and applying the formula, we find

$$\begin{aligned} \operatorname{div} X &= y^2 \cdot \frac{\partial(\frac{1}{y^2} X^i)}{\partial x^i} = y^2 \left(\frac{\partial}{\partial x}\left(\frac{1}{y^2} \cdot y\right) + \frac{\partial}{\partial y}\left(\frac{1}{y^2} \cdot x^2 y\right)\right) = \\ &= y^2 \cdot x^2 \cdot \frac{-1}{y^2} = -x^2. \end{aligned}$$

4°. ♣ Hw. Compute Δf.

5°. Lowering the indices of X and denoting $\omega_k = g_{sk} X^s$ we rewrite

$$\operatorname{curl} X = \left(\frac{\partial \omega_j}{\partial x^i} - \frac{\partial \omega_i}{\partial x^j}\right) dx^i \otimes dx^j.$$

By direct computation, we find

$$\begin{cases} \omega_1 = g_{11} X^1 = \frac{1}{y^2} \cdot y = \frac{1}{y} \\ \omega_2 = g_{22} X^2 = \frac{1}{y^2} \cdot x^2 y = \frac{x^2}{y}. \end{cases}$$

Then, denoting $\Omega = \operatorname{curl} X$, we obtain the components

$$\begin{cases} \Omega_{11} = \Omega_{22} = 0 \\ \Omega_{12} = -\Omega_{21} = \dfrac{\partial \omega_2}{\partial x} - \dfrac{\partial \omega_1}{\partial y} = \dfrac{2x}{y} + \dfrac{1}{y^2} = \dfrac{2xy+1}{y^2}. \end{cases}$$

1.9 $q-$forms

Let V be an n-dimensional real vector space.

1.9.1 Definition. A *q-form* (or a *form of order* q) over the space V is a tensor $\omega \in T^0_q(V)$ which is antisymmetric.

Let $\omega \in T^0_q(V)$ be a q-form. The antisymmetry of ω can be expressed by the relation

$$\omega(v_{\sigma(1)}, \ldots, v_{\sigma(q)}) = (\text{ sign } \sigma)\, \omega(v_1, \ldots, v_q),$$

for all the permutations σ of $\{1, \ldots, q\}$ and for all $v_1, \ldots, v_q \in V$.

Denote by $F_q(V)$ the set of all q-forms on V. This has a natural structure of real vector space. Its dimension is

$$\dim F_q(V) = \begin{cases} C^q_n, & \text{for } q \in \{0, 1, 2, \ldots, n\} \\ 0, & \text{for } q > n. \end{cases}$$

As particular cases, we have $F_0(V) = \mathbb{R}, \ F_1(V) = V^*$.

1.9.2 Definition. Let $p, q > 0$. We call *exterior product* the operator

$$\wedge : \ (\omega^1, \omega^2) \in F_p(V) \times F_q(V) \to \omega^1 \wedge \omega^2 \in F_{p+q}(V)$$

given by

$$\begin{aligned}(\omega^1 \wedge \omega^2)(v_1, \dots v_p, v_{p+1}, \dots, v_{p+q}) &= \frac{1}{(p+q)!} \sum (\text{ sign } \sigma \ ) \ \omega^1(v_{\sigma(1)}, \dots, v_{\sigma(p)}) \cdot \\ &\cdot \omega^2(v_{\sigma(p+1)}, \dots, v_{\sigma(p+q)}),\end{aligned}$$

where $\omega^1 \in F_p(V), \ \omega^2 \in F_q(V); \ v_1, \dots, v_{p+q} \in V$ and σ is a permutation of the numbers $\{1, \dots, p+q\}$. We call $\omega^1 \wedge \omega^2$ *the exterior product of the forms* ω^1 *and* ω^2.

Remarks. 1°. The exterior product is bilinear and associative.

2°. $\omega^1 \wedge \omega^2 = (-1)^{pq} \omega^2 \wedge \omega^1$, for all $\omega^1 \in F_p(V), \ \omega^2 \in F_q(V)$.

1.9.3 Theorem. *Let* $\{e_i \mid i = \overline{1, n}\} \subset V$ *be a basis in* V *and* $\{e^j \mid j = \overline{1, n}\} \subset V^*$, *the corresponding dual basis of* V^*. *Then*

a) the set $\{e^{i_1} \wedge \dots \wedge e^{i_q} \mid i_1, \dots, i_q = \overline{1, n}\}$ *generates* $F_q(V)$;

b) the set

$$\{e^{i_1} \wedge \dots \wedge e^{i_q} \mid 1 \le i_1 < \dots < i_q \le n\}$$

is a basis (called exterior product basis*) of* $F_q(V)$.

Consequently, any form $\omega \in F_q(V)$ can be expressed as

$$\omega = \omega_{i_1 \dots i_q} e^{i_1} \wedge \dots \wedge e^{i_q} = q! \cdot \sum_{1 \le j_1 < j_2 < \dots < j_q \le n} \omega_{j_1 \dots j_q} e^{j_1} \wedge \dots \wedge e^{j_q}.$$

The numbers $\omega_{i_1 \dots i_q} \in \mathbb{R}$ are called *the components of the* q*−form.*

1.9.4. Exercises

1. Compute the form $\eta = \omega \wedge \theta$, for $\omega = e^1 \wedge e^2 - 2e^2 \wedge e^3$ and $\theta = e^1 - e^3$, and find the components of these forms.

Solution. We obtain

$$\begin{aligned}\omega \wedge \theta \ &= (e^1 \wedge e^2 - 2e^2 \wedge e^3) \wedge (\theta = e^1 - e^3) = -2e^2 \wedge e^3 \wedge e^1 - e^1 \wedge e^2 \wedge e^3 = \\ &= -3e^1 \wedge e^2 \wedge e^3.\end{aligned}$$

The components are:

$$\omega_{12} = -\omega_{21} = \tfrac{1}{2}, \omega_{23} = \omega_{32} - 1, \omega_{13} = -\omega_{31} = \omega_{11} = \omega_{22} = \omega_{33} = 0;$$

$$\theta_1 = 1, \theta_2 = 0, \theta = -1;$$

$$\eta_{123} = \tfrac{1}{3!}(-3) = -\tfrac{1}{2} = \eta_{231} = \eta_{312} = -\eta_{213} = -\eta_{132} = \eta_{321}.$$

2. Is $\omega = 4e^1 \wedge e^2 \wedge e^3$ a basis in $F_3(\mathbb{R}^3)$? But in $F_3(\mathbb{R}^4)$?

Solution. The 3-form ω is a basis in the vector space $F_3(\mathbb{R}^3)$, since it is linearly independent (being nonzero form) and $dim\ F_3(\mathbb{R}^3) = 1$.

On the other hand, ω is **not** a basis of $F_3(\mathbb{R}^4)$, since it generates just an 1-dimensional vector subspace of the C_4^3=4-dimensional vector space $F_3(\mathbb{R}^4)$. However, it can be completed to a basis of $F_3(\mathbb{R}^4)$, e.g.,

$$\{\omega, e^1 \wedge e^3 \wedge e^4, e^2 \wedge e^3 \wedge e^4, e^1 \wedge e^2 \wedge e^4\}.$$

1.10 Differential q−forms

Let us consider the path integral

$$\int_\gamma [P(x,y,z)dx + Q(x,y,z)dy + R(x,y,z)dz]$$

and the surface integral

$$\int_\sigma [P(x,y,z)dy \wedge dz + Q(x,y,z)dz \wedge dx + R(x,y,z)dx \wedge dy]\,.$$

If P, Q, R are differentiable functions, then the mathematical object

$$P(x,y,z)dx + Q(x,y,z)dy + R(x,y,z)dz$$

is called *differential 1-form*, and the object

$$P(x,y,z)\ dy \wedge dz + Q(x,y,z)\ dz \wedge dx + R(x,y,z)\ dx \wedge dy$$

is called *differential 2-form.*

We introduce the general notion of differential q-form as follows. Let $D \subset \mathbb{R}^n$ be an open subset of $\mathbb{R}^n$, let $x = (x^1, \ldots, x^n) \in D,\ \ V = T_xD$, and let $F_q(T_xD)$ be the set of all q-forms on T_xD.

1.10.1 Definition. A differentiable function $\omega : D \to \bigcup\limits_{x \in D} F_q(T_xD)$ such that

$$\omega(x) \in F_q(T_xD),\ \text{for all } x \in D,$$

is called *differential q−form* (or *a field of q-forms*).

Let $\mathcal{F}_q(D)$ be the set of all differential q-forms on D. It has a canonical structure of real vector space and of an $\mathcal{F}(D)$−module. Regarding the last structure, if $B = \left\{\frac{\partial}{\partial x^i}, i = \overline{1,n}\right\}$ is a basis in $\mathcal{X}(D)$ and $B^* = \left\{dx^i, i = \overline{1,n}\right\} \subset \mathcal{X}^*(D)$ is the corresponding dual basis, then the induced set of generators is

$$\{dx^{i_1} \wedge \ldots \wedge dx^{i_q} \mid i_1, \ldots, i_q \in \overline{1,n}\},$$

and *the induced basis in* $\mathcal{F}_q(D)$ is

$$\left\{dx^{i_1} \wedge \ldots \wedge dx^{i_q} \mid 1 \leq i_1 < i_2 < \ldots < i_q \leq n\right\} \subset \mathcal{F}_q(D).$$

Consequently, any differential form $\omega \in \mathcal{F}_q(D)$ decomposes like

$$\begin{aligned}\omega &= \omega_{i_1 \ldots i_q} dx^{i_1} \wedge \ldots \wedge dx^{i_q} = \\ &= q! \sum_{1 \leq j_1 < \ldots < j_q \leq n} \omega_{j_1 \ldots j_q} e^{j_1} \wedge \ldots \wedge e^{j_q},\end{aligned}$$

where $\omega_{i_1,\ldots,i_q} \in \mathcal{F}(D)$ are called *the components of the differential* $q-$*form.*

1.10.2 Operations with forms

1.10.2.1 Definition. The *interior product* is the application

$$\lrcorner : \mathcal{X}(D) \times \mathcal{F}_q(D) \to \mathcal{F}_{q-1}(D),$$

given by

$$(X \lrcorner \omega)(X_1, \ldots, X_{q-1}) = \omega(X, X_1, \ldots, X_{q-1}),$$

for all $X_i \in \mathcal{X}(D), i = \overline{1, q-1}$. Then the differential $(p-1)-$form $X \lrcorner \omega \in \mathcal{F}_{q-1}(D)$ is called also *the interior product between* X *and* ω.

The components of $X \lrcorner \omega$ are given by

$$(X \lrcorner \omega)_{i_2 \ldots i_q} = X^{i_1} \omega_{i_1 i_2 \ldots i_q}, \ i_2, \ldots, i_q = \overline{1, n}.$$

Consequently, $X \lrcorner \omega$ is the tensor product $X \otimes \omega$ followed by a suitable transvection.

1.10.2.2 Definition. The function $d : \mathcal{F}_q(D) \to \mathcal{F}_{q+1}(D),$ for all $q \geq 0$, which satisfies

a) $d^2 = 0$;

b) $d(\omega^1 \wedge \omega^2) = d\omega^1 \wedge \omega^2 + (-1)^{pq} \omega^1 \wedge d\omega^2,$ for all $\omega^1 \in \mathcal{F}_p(D), \omega^2 \in \mathcal{F}_q(D)$;

c) $X(d\omega) = d(X(\omega))$, where

$$X(\omega) = X \lrcorner d\omega + \mathrm{d}(X \lrcorner \omega), \quad \text{for all } X \in \mathcal{X}(D), \omega \in \mathcal{F}_p(D);$$

d) $d(f) = df,$ for all $f \in \mathcal{F}_0(D) \equiv \mathcal{F}(D)$,

is called *exterior differentiation.*

Generally, for a differential $q-$form $\omega = \omega_{i_1 \ldots i_p} dx^{i_1} \wedge \ldots \wedge dx^{i_q}$, we have

$$\mathrm{d}\omega = \mathrm{d}\omega_{i_1 \ldots i_q} \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_q}.$$

In particular:

- for $\omega = \omega_i dx^i \in \mathcal{F}_1(D), D \subset \mathbb{R}^3$, we obtain

$$\begin{aligned} \mathrm{d}\omega \quad &= \mathrm{d}\omega_i \wedge dx^i = \frac{\partial \omega_i}{\partial x^j} dx^j \wedge dx^i = \left(\frac{\partial \omega_3}{\partial x^2} - \frac{\partial \omega_2}{\partial x^3} \right) dx^2 \wedge dx^3 + \\ &+ \left(\frac{\partial \omega_1}{\partial x^3} - \frac{\partial \omega_3}{\partial x^1} \right) dx^3 \wedge dx^1 + \left(\frac{\partial \omega_2}{\partial x^1} - \frac{\partial \omega_1}{\partial x^2} \right) dx^1 \wedge dx^2, \end{aligned}$$

- for $\eta = \eta_{ij} dx^i \wedge dx^j \in \mathcal{F}_2(D),\ D \subset \mathbb{R}^3$, we obtain

$$\begin{aligned} \mathrm{d}\eta \quad &= \tfrac{\partial \eta_{ij}}{\partial x^k} dx^i \wedge dx^j \wedge dx^k = \\ &= 2 \left(\tfrac{\partial \eta_{12}}{\partial x^3} + \tfrac{\partial \eta_{31}}{\partial x^2} + \tfrac{\partial \eta_{23}}{\partial x^1} \right) dx^1 \wedge dx^2 \wedge dx^3, \end{aligned}$$

where we used the relations

$$\mathrm{d}\omega_i = \frac{\partial \omega_i}{\partial x^j} dx^j, \quad dx^i \wedge dx^i = 0, \quad dx^i \wedge dx^j = -dx^j \wedge dx^i, \quad \text{for all } i, j = \overline{1, n}.$$

1.10.2.3 Definitions. a) A differential form $\omega \in \mathcal{F}_q(D)$ is called *closed* iff $\mathrm{d}\omega = 0$.
b) If there exist $\eta \in \mathcal{F}_{q-1}(D)$ such that $\omega = \mathrm{d}\eta$, then ω is called *an exact form*.

1.10.2.4 Remarks. For $D \subset \mathbb{R}^3$ endowed with the metric $g = \delta_{ij} dx^i \otimes dx^j$, the 1-forms and the 2-forms can be converted into vector fields using the rules

$$\omega = \omega_i\ dx^i \in \mathcal{F}_1(D) \underset{(1)}{\leftrightarrow} \sum_{i=1}^{3} \omega_i \frac{\partial}{\partial x^i} \in \mathcal{X}(D)$$

$$\omega = \omega_i\ dx^i \in \mathcal{F}_1(D) \underset{(2)}{\leftrightarrow} \omega_1\ dx^2 \wedge dx^3 + \omega_2\ dx^3 \wedge dx^1 + \omega_3\ dx^1 \wedge dx^2 \in \mathcal{F}_2(D).$$

The correspondence $\mathrm{d}f \in \mathcal{F}_1(D) \underset{(1)}{\leftrightarrow}$ grad $f \in \mathcal{X}(D)$, can be extended to

$$\omega = \omega_i\ dx^i \underset{(1)}{\leftrightarrow} X = X^i \frac{\partial}{\partial x^i} \in \mathcal{X}(D), \quad \omega \in \mathcal{F}_1(D)$$

by means of the relations

$$\omega_i = \delta_{is} X^s; \quad X^i = \delta^{is} \omega_s. \tag{53}$$

Particularly, we have the correspondence

$$\eta = \mathrm{d}\omega \in \mathcal{F}_2(D) \underset{(2)}{\leftrightarrow} Y = \operatorname{curl} X \in \mathcal{X}(D),$$

which is given in general by

$$\eta = \sum_{i<j} \eta_{ij} dx^i \wedge dx^j \in \mathcal{F}_2(D) \underset{(2)}{\leftrightarrow} Y^i \frac{\partial}{\partial x^i} \in \mathcal{X}(D),$$

with $\eta_{12} = Y^3,\ \eta_{23} = Y^1,\ \eta_{31} = Y^2$.

Also the correspondence $\eta \underset{(2)}{\leftrightarrow} Y$ yields the relation

$$\mathrm{d}\eta = (\operatorname{div} Y)\, dx^1 \wedge dx^2 \wedge dx^3.$$

1.10.3. Exercises

1. Let $X = x\frac{\partial}{\partial z} \in \mathcal{X}(D),\ D \subset \mathbb{R}^3$ and

$$\omega = xzdx \wedge dy - ydy \wedge dz.$$

Compute the following tensor fields:

a) $X \lrcorner \omega$, where $X = x\dfrac{\partial}{\partial z},\ \omega = xz\ dx \wedge dy - y\ dy \wedge dz$;

b) $\mathrm{d}\omega$, where $\omega = xz\ dx \wedge dy - y\ dy \wedge dz$;

c) $\omega \wedge \eta,\ \eta \wedge \omega$, where $\eta = xdx + 2ydz$;

d) $\mathrm{d}f, \mathrm{d}^2 f$, where $f(x, y, z) = x + 2yz$.

Consider as well the case $\omega = xdy + ydz$.

Solution. a) If $X = X^i \frac{\partial}{\partial x^i}$ and $\omega = \omega_{i_1 \dots i_p} dx^{i_1} \wedge \dots \wedge dx^{i_p}$, then

$$(X \lrcorner \omega) = X^{i_1} \omega_{i_1 i_2 \dots i_p} dx^{i_2} \wedge \dots \wedge dx^{i_p}.$$

For $\omega = xzdx \wedge dy - ydy \wedge dz \in \mathcal{F}_2(D)$ and $X = x\frac{\partial}{\partial z} \in \mathcal{X}(D)$, denoting $(x^1, x^2, x^3) \equiv (x, y, z)$ and identifying then with the decompositions

$$\omega = \omega_{ij} dx^i \wedge dx^j,\quad X = X^i \frac{\partial}{\partial x^i},$$

we obtain the components

$$\omega_{12} = xz = -\omega_{21},\ \ \omega_{23} = -y = -\omega_{32},\ \ \omega_{31} = \omega_{13} = 0 \Rightarrow [\omega] = \begin{pmatrix} 0 & xz & 0 \\ -xz & 0 & -y \\ 0 & y & 0 \end{pmatrix}$$

and respectively $X^1 = X^2 = 0, X^3 = x$. Then, for $\theta = (X \lrcorner \omega) = \theta_i dx^i \in \mathcal{F}_1(D)$, with $\theta_i = X^k \omega_{ki},\ i = \overline{1,3}$, we obtain

$$\begin{cases} \theta_1 = X^1\omega_{11} + X^2\omega_{21} + X^3\omega_{31} = 0 \\ \theta_2 = X^1\omega_{12} + X^2\omega_{22} + X^3\omega_{32} = xy \\ \theta_3 = X^1\omega_{13} + X^2\omega_{23} + X^3\omega_{33} = 0 \end{cases} \Rightarrow \quad X \lrcorner \omega = xydy.$$

For $\omega = xdy + ydz = \omega_i dx^i \in \mathcal{F}_1(D)$, and $X = x\frac{\partial}{\partial z} \equiv X^i \frac{\partial}{\partial x^i} \in \mathcal{X}(D)$, the components of ω are $\omega_1 = 0, \omega_2 = x, \omega_3 = y$ and then for $\theta = (X \lrcorner \omega) \in \mathcal{F}_0(D) = \mathcal{F}(D)$, we find

$$\theta = X^k \omega_k = X^1\omega_1 + X^2\omega_2 + X^3\omega_3 \Rightarrow \omega \lrcorner X = xy.$$

b) Generally, for $\omega = \omega_{i_1 \ldots i_p} dx^{i_1} \wedge \ldots \wedge dx^{i_p}$, we have

$$\mathrm{d}\omega = \mathrm{d}\omega_{i_1 \ldots i_p} \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_p}.$$

For $\omega = xzdx \wedge dy - ydy \wedge dz \in \mathcal{F}_2(D)$, using

$$dx^i \wedge dx^i = 0, \;\; dx^i \wedge dx^j = -dx^j \wedge dx^i, \text{ for all } i, j = \overline{1,3},$$

we infer

$$\mathrm{d}\omega = (xdz + zdx) \wedge dx \wedge dy - \underbrace{dy \wedge dy}_{=0} \wedge dz = xdz \wedge dx \wedge dy = xdx \wedge dy \wedge dz \in \mathcal{F}_3(D).$$

For $\omega = xdy + ydz \in \mathcal{F}_1(D)$, we infer $\mathrm{d}\omega = dx \wedge dy + dy \wedge dz \in \mathcal{F}_2(D)$.

c) For $\omega = xzdx \wedge dy - ydy \wedge dz \in \mathcal{F}_2(D)$, $\eta = xdx + 2ydz \in \mathcal{F}_1(D)$, we find

$$\begin{aligned} \omega \wedge \eta &= (xzdx \wedge dy - ydy \wedge dz) \wedge (xdx + 2ydz) = \\ &= x^2 z \underbrace{dx \wedge dy \wedge dx}_{=0} - xydy \wedge dz \wedge dx + 2xyz\, dx \wedge dy \wedge dz - 2y^2 \underbrace{dy \wedge dz \wedge dz}_{=0} = \\ &= (2xyz - xy) dx \wedge dy \wedge dz = xy(2z-1) dx \wedge dy \wedge dz \in \mathcal{F}_3(D). \end{aligned}$$

♣ Hw. For $\omega = xdy + ydz$, determine $\omega \wedge \eta$. Also, in both cases, compute

$$\omega \wedge \omega, \;\; \eta \wedge \eta, \;\; \omega \wedge \omega \wedge \eta.$$

d) By straightforward computation, we find $\mathrm{d}f = dx + 2zdy + 2ydz \in \mathcal{F}_1(D)$, and

$$\mathrm{d}^2 f = \mathrm{d}(\mathrm{d}f) = 0 \in \mathcal{F}_2(D).$$

2. In $D \subset \mathbb{R}^3$ endowed with the metric $g_{ij} = \delta_{ij}$, establish the relations

$$\begin{cases} \omega \in \mathcal{F}_1(D) \overset{(1)}{\leftrightarrow} X \in \mathcal{X}(D) \\ \eta \in \mathcal{F}_2(D) \overset{(2)}{\leftrightarrow} Y \in \mathcal{X}(D) \end{cases},$$

provided by

$$\begin{cases} \alpha_i dx^i \overset{(1)}{\leftrightarrow} \alpha_i \frac{\partial}{\partial x^i} \\ a_{ij} dx^i \wedge dx^j \overset{(2)}{\leftrightarrow} a_k \frac{\partial}{\partial x^k}, \text{ with } a_k = a_{ij}, \text{ for } \{i,j,k\} = \{1,2,3\}, \end{cases}$$

for the given data

a) $\omega = xdz - ydx \in \mathcal{F}_1(D)$ and $\eta = xydz \wedge dx - y^2 dy \wedge dx \in \mathcal{F}_2(D)$;

b) $X = x\dfrac{\partial}{\partial y} - z\dfrac{\partial}{\partial x}$ and $Y = 2z\dfrac{\partial}{\partial y} - x\dfrac{\partial}{\partial z} \in \mathcal{X}(D)$.

Solution. a) The direct associations provide

$$\begin{array}{ll} \omega = xdz - ydx & \overset{(1)}{\leftrightarrow} \; X = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial z} \\ \eta = y^2 dx \wedge dy + xydz \wedge dx & \overset{(2)}{\leftrightarrow} \; Y = y^2\frac{\partial}{\partial z} + xy\frac{\partial}{\partial y}. \end{array}$$

b) The reverse associations yield

$$\begin{array}{ll} X = x\frac{\partial}{\partial y} - z\frac{\partial}{\partial x} & \overset{(1)}{\leftrightarrow} \; \omega = -zdx + xdy, \\ Y = 2z\frac{\partial}{\partial y} - x\frac{\partial}{\partial z} & \overset{(2)}{\leftrightarrow} \; \eta = 2zdz \wedge dx - xdx \wedge dy. \end{array}$$

3. Check that in $D \subset \mathbb{R}^3$ endowed with the metric $g_{ij} = \delta_{ij}$, the following relations hold true (i.e., the 1-forms and 2-forms can be converted to vector fields)

$$\begin{array}{lcl} \mathrm{d}f & \overset{(1)}{\leftrightarrow} & \operatorname{grad} f \\ \mathrm{d}\omega & \overset{(2)}{\leftrightarrow} & \operatorname{curl} X, \text{ in case that } \omega \overset{(1)}{\leftrightarrow} X \\ \mathrm{d}\eta & = & (\operatorname{div} Y)\, dx^1 \wedge dx^2 \wedge dx^3, \text{ in case that } \eta \overset{(2)}{\leftrightarrow} Y. \end{array}$$

Apply these correspondences for the fields $f(x,y,z) = xy$, $f \in \mathcal{F}_0(D) = \mathcal{F}(D)$, for

$$\omega = xydz \in \mathcal{F}_1(D) \quad \text{and for} \quad \eta = xzdy \wedge dz \in \mathcal{F}_2(D).$$

Solution. By straightforward calculation, we get

- $\mathrm{d}f = ydx + xdy$ and $\operatorname{grad} f = g^{ij}\frac{\partial f}{\partial x^j}\frac{\partial}{\partial x^i} = y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}$. Then one can easily see that $\mathrm{d}f \overset{(1)}{\leftrightarrow} \operatorname{grad} f$.

- The relation $\omega = xydz \overset{(1)}{\leftrightarrow} X = xy\frac{\partial}{\partial z}$ implies $X^1 = X^2 = 0, X^3 = xy$. Then one gets

$$\mathrm{d}\omega = -ydz \wedge dx + xdy \wedge dz$$

and, based on the formula

$$\operatorname{curl} X = \sum_{i<j} \left(\frac{\partial X^j}{\partial x^i} - \frac{\partial X^i}{\partial x^j} \right) dx^i \wedge dx^j \overset{(2)}{\leftrightarrow} Z^k \frac{\partial}{\partial x^k}, \;\; \{i,j,k\} = \{1,2,3\},$$

we compute

$$Z^1 = \frac{\partial X^3}{\partial x^2} - \frac{\partial X^2}{\partial x^3} = x, Z^2 = -y, Z^3 = 0 \Rightarrow Z = x\frac{\partial}{\partial x} - y\frac{\partial}{\partial y},$$

which shows that $\mathrm{d}\omega \overset{(2)}{\leftrightarrow} \operatorname{curl} X$.

- The relation $\eta \overset{(2)}{\leftrightarrow} Y$ provides $Y = xz\frac{\partial}{\partial x}$, whence $\operatorname{div} Y = \frac{\partial(xz)}{\partial x} = z$. On the other hand,
$$\mathrm{d}\eta = z dx \wedge dy \wedge dz,$$
whence the equality
$$\mathrm{d}\eta = (\operatorname{div} Y) dx^1 \wedge dx^2 \wedge dx^3$$
holds true.

4. Compute the exterior derivative $\mathrm{d}\omega$ and the exterior product $\omega \wedge \theta$ for
$$\theta = dx + y dz \in \mathcal{F}_1(\mathbb{R}^3)$$
and for ω given below:
a) $\omega = z^2\ dx \wedge dy - y^2 x\ dy \wedge dz \in \mathcal{F}_2(\mathbb{R}^3)$,
b) $\omega = (x - yz)\ dx \wedge dy \wedge dz \in \mathcal{F}_3(\mathbb{R}^3)$.

Solution. For $\omega = \omega_{i_1 \dots i_p} dx^{i_1} \wedge \dots \wedge dx^{i_p} \in \mathcal{F}_p(\mathbb{R}^3)$ we have
$$\mathrm{d}\omega = (\mathrm{d}\omega_{i_1 \dots i_p}) \wedge dx^{i_1} \wedge \dots \wedge dx^{i_p} \in \mathcal{F}_{p+1}(\mathbb{R}^3),$$
with $\mathrm{d}\omega_{i_1 \dots i_p} = \dfrac{\partial \omega_{i_1 \dots i_p}}{\partial x^i} dx^i$. We know that
$$dx^i \wedge dx^i = 0, \quad dx^j \wedge dx^i = -dx^i \wedge dx^j, \text{ for all } i, j = \overline{1, n}. \tag{54}$$
a) The exterior differential of ω writes
$$\begin{aligned} \mathrm{d}\omega &= \left(\frac{\partial z^2}{\partial x} dx + \frac{\partial z^2}{\partial y} dy + \frac{\partial z^2}{\partial z} dz \right) \wedge dx \wedge dy - \\ &- \left(\frac{\partial y^2 x}{\partial x} dx + \frac{\partial y^2 x}{\partial y} dy + \frac{\partial y^2 x}{\partial z} dz \right) \wedge dy \wedge dz = \\ &= (0 + 0 + 2z dz) \wedge dx \wedge dy - (y^2 dx + 2xy dy + 0) \wedge dy \wedge dz = \\ &= 2z dx \wedge dy \wedge dz - y^2 dx \wedge dy \wedge dz - 2xy dy \wedge dy \wedge dz. \end{aligned}$$
Hence, considering (54), we find $\mathrm{d}\omega = (2z - y^2)\ dx \wedge dy \wedge dz \in \mathcal{F}_3(\mathbb{R}^3)$.

Also the exterior product of ω with $\theta = dx + y dz$ is
$$\begin{aligned} \omega \wedge \theta &= (z^2\ dx \wedge dy - y^2 x\ dy \wedge dz) \wedge (dx + y dz) = \\ &= z^2\ dx \wedge dy \wedge dx + z^2 y\ dx \wedge dy \wedge dz - y^2 x\ dy \wedge dz \wedge dx - y^3 x\ dy \wedge dz \wedge dz = \\ &= (z^2 y - y^2 x) dx \wedge dy \wedge dz. \end{aligned}$$
b) The exterior differential is
$$\begin{aligned} \mathrm{d}\omega &= (dx - z dy - y dz) \wedge dx \wedge dy \wedge dz = dx \wedge dx \wedge dy \wedge dz - \\ &\quad -z\ dy \wedge dx \wedge dy \wedge dz - y dz \wedge dx \wedge dy \wedge dz =) + 0 + 0 = 0. \end{aligned}$$

The exterior product writes

$$\omega \wedge \theta = [(x - yz)dx \wedge dy \wedge dz] \wedge (dx + ydz).$$

This is a 4-form. Taking into account (54) we have that all the 4-forms in $\mathbb{R}^3$ are null, hence we have $\omega \wedge \theta = 0$.

↟ Hw. Check this by direct computation, applying (54).

5. Find dω in the following cases

a) $\omega = x^2dx + y^2dy + z^2dz$;

b) $\omega = e^{xy}dx \wedge dz$;

c) $\omega = \dfrac{1}{x+y+z}(dx + dy + dz)$.

Solution. a) Based on the formula

$$\mathrm{d}\omega = \mathrm{d}\omega_{i1,\ldots,ir} \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_r} = \frac{\partial \omega_{i_1\ldots i_r}}{\partial x^i}\, dx^i \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_r},$$

for all $\omega \in \mathcal{F}_r(D)$, we find

a) d$\omega = 0$;

b) $\mathrm{d}\omega = xe^{xy}dy \wedge dx \wedge dz = -xe^{xy}\, dx \wedge dy \wedge dz$;

c) $\mathrm{d}\omega = \dfrac{-1}{(x+y+z)^2}(dx\wedge dy + dx\wedge dz + dy\wedge dx + dy\wedge dz + dz\wedge dx + dz\wedge dy) = 0.$

In the last case, we remark that ω is an exact form,

$$\omega = \mathrm{d}(\ln|x + y + z|),$$

and hence it is closed, i.e., d$\omega = 0$.

6. Compute $\omega \lrcorner X$, for $\omega = x^2dx \wedge dy \in \mathcal{F}_2(\mathbb{R}^2)$ and $X = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} \in \mathcal{X}(\mathbb{R}^2)$.

Solution. By straightforward computation, we obtain

$$\begin{aligned} \omega \lrcorner X &= tr_1^1\left[\left(\frac{x^2}{2}(dx \otimes dy - dy \otimes dx)\right) \otimes X\right] = \\ &= \frac{x^2}{2}\,[dx(X)dy - dy(X)dx] = \frac{x^2}{2}(xdy - ydx) \in \mathcal{F}_1(\mathbb{R}^2). \end{aligned}$$

Chapter 2

Field Lines and Hypersurfaces

2.1 Field lines and first integrals

2.1.1 Let $D \subset \mathbb{R}^n$ be an open connected set, and X a vector field of class $\mathcal{C}^1$ on D. A curve $\alpha : I \to D$ of class $\mathcal{C}^1$ which satisfies $\alpha' = X \circ \alpha$ is called a *field line* [1] of X. The image $\alpha(I)$ is called *orbit of* X (see the figure).

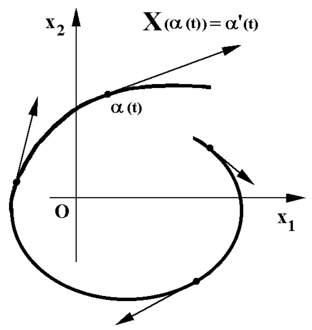

Fig. 1

The differential equation

$$\alpha'(t) = X(\alpha(t)) \tag{55}$$

is equivalent to the integral equation

$$\alpha(t) = \alpha(t_0) + \int_{t_0}^{t} X(\alpha(s))ds,$$

and hence the curve α which satisfies (55) is of class $\mathcal{C}^2$.

[1]a field line is also called: *vector line*, *force line* or *flow line of* X.

Explicitly, for $X = (X_1, \ldots, X_n)$ and

$$\alpha = (x_1, \ldots, x_n), \quad x_1 = x_1(t), \ldots, x_n = x_n(t),\ t \in I,$$

the vector differential equation (55) is written under the form of an *autonomous differential system*

$$\begin{cases} \dfrac{dx_1}{dt} = X_1(x_1, \ldots, x_n) \\ \ldots \\ \dfrac{dx_n}{dt} = X_n(x_1, \ldots, x_n). \end{cases} \tag{56}$$

The existence and uniqueness theorem for a *Cauchy problem* associated to this system can be reformulated as

2.1.2 Theorem. *If X is a vector field of class $\mathcal{C}^1$, then for any $x_0 \in D,\ t_0 \in \mathbb{R}$, there exists an open interval I and a field line $\alpha : I \to D$ with the properties:*

a) $\alpha(t_0) = x_0$;

b) if $\beta : J \to D$ is any other field line of X such that $\beta(t_1) = x_0$, then $J \subseteq I$ and

$$\beta(t) = \alpha(t - t_0 + t_1), \quad \textit{for all } t \in J.$$

The curve α is called *the maximal field line of X which passes through x_0.*

Definition. The solutions $a \in D$ of the algebraic autonomous system

$$X_1(x_1, \ldots, x_n) = 0, \ldots, X_n(x_1, \ldots, x_n) = 0, \tag{57}$$

are called *zeros* of X. The constant solutions $x(t) = a,\ t \in \mathbb{R}$ of the differential system (56) are called *equilibrium points of X*.

Example. Find the field lines and the equilibrium points of the field

$$X(x_1, x_2) = (-x_2, x_1).$$

Solution. By integration, one obtains *the general solution*

$$\begin{cases} x_1(t) = C_1 \cos t + C_2 \sin t \\ x_2(t) = C_1 \sin t - C_2 \cos t, \end{cases}$$

which for the initial conditions

$$\begin{cases} x_1(0) = a \\ x_2(0) = b \end{cases}$$

provides the particular solution which passes through the point (a, b) at the moment $t_0 = 0$, given by

$$\begin{cases} x_1(t) = a \cos t - b \sin t \\ x_2(t) = a \sin t + b \cos t. \end{cases}$$

The unique equilibrium point of X is the origin $O(0,0)$.

Properties of field lines.

1) The reparametrizations by translations are admitted.

2) Two distinct maximal field lines of X have in common no point.

3) Any maximal field line is either injective, either simple and closed, or constant. Assume that $\varphi : D \to D_*,\ y = \varphi(x)$ is a diffeomorphism of class $\mathcal{C}^1$. The system

$$\frac{dx_i}{dt} = X_i(x)$$

is equivalent to

$$\frac{dy_j}{dt} = \sum_{i=1}^{n} \left(\frac{\partial y_j}{\partial x_i} X_i \right) (\varphi^{-1}(y)), \quad j = \overline{1,n},$$

or, equivalently,

$$\frac{dy_j}{dt} = (D_X y_j)(\varphi^{-1}(y)).$$

The following theorem gives conditions under which $D_X y_j = \begin{cases} 0, \text{ for } j = \overline{1, n-1} \\ 1, \text{ for } j = n. \end{cases}$

2.1.3 Theorem (Rectification Theorem). *If $X = (X_1, \ldots, X_n)$ is a vector field of class $\mathcal{C}^1$ on D and x_0 satisfies $X(x_0) \neq 0$, then there exists a neighborhood U of x_0 and a diffeomorphism $\varphi : U \to V \subset \mathbb{R}^n,\ y = \varphi(x)$ of class $\mathcal{C}^1$, such that the system $\frac{dx_i}{dt} = X_i(x)$ on U_{x_0} is reduced to*

$$\frac{dy_1}{dt} = 0, \ldots, \frac{dy_{n-1}}{dt} = 0,\ \frac{dy_n}{dt} = 1$$

on the domain $V = \varphi(U)$.

Definition. A function $f : D \to \mathbb{R}$ is called *first integral for the differential system (56)* iff $D_X f = 0$.

This definition is equivalent to any of the two following assertions:

a) f is constant along any field line of X;

b) any orbit of X is contained in one single set of constant level of f.

♣ Hw. Check that the *Hamiltonian function* $H : \mathbb{R}^{2n} \to \mathbb{R}$ is a (global) first integral of *the Hamiltonian system*

$$\frac{dx_i}{dt} = -\frac{\partial H}{\partial y_i}; \quad \frac{dy_i}{dt} = \frac{\partial H}{\partial x_i},$$

i.e., it satisfies $D_X H = 0$, where X is the associated *Hamiltonian field*

$$X = (X_1, \ldots, X_n, X_{n+1}, \ldots, X_{2n}),\ \ X_i = -\frac{\partial H}{\partial y_i},\ X_{n+i} = \frac{\partial H}{\partial x_i},\ i = \overline{1,n}.$$

2.1.4 Theorem. *Let X be a field of class $\mathcal{C}^1$ on D, and $x_0 \in D$, such that $X(x_0) \neq 0$. Then there exists a neighborhood U_{x_0} of x_0 and $n-1$ first integrals $f_1,\ f_2, \dots, f_{n-1}$, which are functional independent on U_{x_0} and any other first integral is a function of class $\mathcal{C}^1$ of $f_1,\ f_2, \dots, f_{n-1}$. The orbits of X are described by the system of Cartesian equations $f_1(x) = C_1, \dots, f_{n-1}(x) = C_{n-1}$, where $C_1, \dots, C_{n-1} \in \mathbb{R}$.*

Proof. By Theorem 1.3, there exists a diffeomorphism

$$y_1 = f_1(x), \dots, y_{n-1} = f_{n-1}(x), y_n = f_n(x), \quad x \in U,$$

such that $D_X f_\alpha = 0,\ \alpha = \overline{1, n-1}$ and hence f_α are first integrals.

Let $f_1, \dots, f_{n-1}$ be first integrals; if f is another first integral, then $D_X f_\alpha = 0$ and $D_X f = 0$. Hence $\frac{D(f, f_1, \dots, f_{n-1})}{D(x_1, \dots, x_n)} = 0$, whence the functions $f, f_1, \dots, f_{n-1}$ are functionally dependent, i.e., $f = \phi(f_1, \dots, f_{n-1})$. □

We note that the differential system of n equations (56) rewrites

$$dt = \frac{dx_1}{X_1(x_1, \dots, x_n)} = \dots = \frac{dx_n}{X_n(x_1, \dots, x_n)},$$

or, omitting dt, as

$$\frac{dx_1}{X_1(x_1, \dots, x_n)} = \dots = \frac{dx_n}{X_n(x_1, \dots, x_n)}, \tag{58}$$

which is called *the symmetric differential system associated to the vector field X*. Then the first integrals of X can be determined from (58) using *the method of integrable combinations*: if there exist the functions $\lambda_j : D \to \mathbb{R},\ j \in \overline{1, n}$ and $f : D \to \mathbb{R}$ such that

$$\frac{dx_1}{X_1} = \dots = \frac{dx_n}{X_n} = \frac{\sum_{j=1}^n \lambda_j(x) dx_j}{\sum_{j=1}^n \lambda_j X_j} = \frac{df}{0}, \tag{59}$$

then it follows that $df = 0$, whence f is constant along the orbits of X, hence f is a first integral for X. In this case, the differential

$$df = \sum_{j=1}^n \lambda_j(x) dx_j$$

which satisfies the condition

$$\sum_{i=1}^n \lambda_i(x) X_i(x) = 0,$$

is called *integrable combination* of the symmetric system (58).

Example. The following symmetric differential system admits the integrable combinations and the outgoing first integrals pointed out below:

$$\frac{dx}{bz-cy}=\frac{dy}{cx-az}=\frac{dz}{ay-bx} = \frac{adx+bdy+cdz}{0}=\frac{d(ax+by+cz)}{0}$$

provides $f_1(x,y,z)=ax+by+cz$, and

$$\frac{dx}{bz-cy}=\frac{dy}{cx-az}=\frac{dz}{ay-bx} = \frac{xdx+ydy+zdz}{0}=\frac{d\left(\frac{x^2+y^2+z^2}{2}\right)}{0}$$

provides $f_2(x,y,z)=\frac{x^2+y^2+z^2}{2}$.

In general, one may use previously found first $n-2$ integrals $f_1,\ldots,f_{n-2}$ for finding a new (functional independent) first integral f_{n-1}, as follows:

- express, e.g., $x_1,\ldots,x_{n-2}$ from the equations $f_1(x)=C_1,\ldots,f_{n-2}(x)=C_{n-2}$ in terms of $C_1,\ldots,C_{n-2}$;
- replace them into the system (58); the original symmetric differential system reduces to a differential equation in x_{n-1} and x_n;
- solve this equation, and find its general solution via

$$g_{n-1}(x_{n-1},x_n;C_1,\ldots,C_{n-2})=C_{n-1};$$

- replace in g the constants $C_1,\ldots C_{n-2}$ respectively by $f_1(x),\ldots f_{n-2}(x)$, and obtain the first integral

$$f_{n-1}(x)=g_{n-1}(x_{n-1},x_n;f_1(x),\ldots,f_{n-2}(x)).$$

We should note that the procedure applies for any subset of $n-2$ variables expressed and substituted in the original symmetric differential system.

Example. The symmetric differential system

$$\frac{dx}{x^2}=\frac{dy}{xy}=\frac{dz}{y^2}$$

admits a first integral which can be derived by integrable combinations:

$$\frac{dx}{x^2}=\frac{dy}{xy}=\frac{dz}{y^2} = \frac{ydx-xdy+0dz}{0}=\frac{d(y/x)}{0} \Rightarrow \frac{y}{x}=C \Rightarrow f_1(x,y,z)=y/x.$$

We obtain the second first integral by using the first one: we use the relation

$$f_1(x,y,z)=C_1 \Leftrightarrow y/x=C_1 \Rightarrow y=C_1x,$$

and replacing y in the symmetric system, we obtain from the equality of the extreme ratios

$$\frac{dx}{x^2} = \frac{dz}{y^2} \quad \Leftrightarrow \quad \frac{dx}{x^2} = \frac{dz}{C_1^2 x^2} \quad \Rightarrow \quad dx = \frac{1}{C_1^2} dz,$$

whence, by integration one gets $x = \frac{z}{C_1^2} + C_2$; then, substituting $C_1 = y/x$, we infer the second *first integral* of the symmetric differential system,

$$f_2(x, y, z) = x - \frac{z}{(y/x)^2} = \frac{x(y^2 - zx)}{y^2}.$$

♠ Hw. Find the first integrals for the following vector fields:

a) $\vec{V} = xz\vec{i} + z(2x - y)\vec{j} - x^2\vec{k}$;

b) $\vec{V} = x^2(y + z)\vec{i} - y^2(z + x)\vec{j} + z^2(y - x)\vec{k}$;

c) $\vec{V} = y^2z^2\vec{i} + xyz^2\vec{j} + xy^2\vec{k}$.

Let X be a $\mathcal{C}^\infty$ vector field on $\mathbb{R}^n$. It defines the corresponding differential equation

$$\frac{d\alpha}{dt} = X(\alpha(t)),$$

with an arbitrary initial condition $\alpha(0) = x$. The solution $\alpha_x(t),\ t \in I = (-\varepsilon, \varepsilon)$ $(x \in U \subset \mathbb{R}^n)$, generates a local diffeomorphism $T^t : U \to \mathbb{R}^n,\ \ T^t(x) = \alpha_x(t),\ t \in I$, solution of the operator differential equation

$$\frac{dT^t}{dt} = X \circ T^t, \quad T^0 = id,$$

which we call *local flow* on $\mathbb{R}^n$ generated by X (*local group with one parameter of diffeomorphisms*). The approximation $y = x + tX(x)$ of $T^t(x)$ is called *infinitesimal transformation* generated by X.

If the solution $\alpha_x(t)$ exists for arbitrary $x \in \mathbb{R}^n,\ t \in \mathbb{R}$, the field X is called *complete*, and the flow $T^t(x) = \alpha_x(t)$ is called *global.*

2.1.5 For determining numerically some points of a field line of a given vector field X and which passes through the initial point x_0 at the moment $t = 0$, one can apply, e.g.,

The Runge-Kutta method. *Let X be a vector field of class $\mathcal{C}^2$ on $D \subset \mathbb{R}^n$ and the Cauchy problem*

$$\frac{dx}{dt} = X(x), \quad x(t_0) = x_0.$$

The iterative method of Runge-Kutta for determining the approximate solution of the Cauchy problem is based on formulas

$$x^*_{k+1} = x^*_k + \frac{1}{6}[k_1 + 2(k_2 + k_3) + k_4], \quad k = 0, 1, \ldots, n,$$

where

$$k_1 = hX(x_k^*), k_2 = hX\left(x_k^* + \frac{k_1}{2}\right), k_3 = hX\left(x_k^* + \frac{k_2}{2}\right), k_4 = hX(x_k^* + k_3)$$

and $h = \Delta t > 0$ *is the* $t-$*advance ("the step").*

2.1.6. Exercises

1. Determine the field lines of the vector field

$$X = (x - y + z,\ 2y - z, z),\quad X \in \mathcal{X}(\mathbb{R}^3),$$

and two first integrals which fix the orbits.

Solution. The field lines are solutions of the homogeneous linear differential system with constant coefficients

$$\frac{dx}{dt} = x - y + z,\quad \frac{dy}{dt} = 2y - z,\quad \frac{dz}{dt} = z.$$

The third equation is with separable variables; we obtain $z = ce^t,\ \ t \in \mathbb{R}$. Replacing in the second equation, we obtain the linear differential equation $\dfrac{dy}{dt} = 2y - ce^t$ with the general solution $y = be^{2t} + ce^t,\ \ t \in \mathbb{R}$. Introducing z and y in the first equation, we find a linear differential equation $\dfrac{dx}{dt} = x - be^{2t}$ with the general solution $x = ae^t - be^{2t},\ \ t \in \mathbb{R}$. Obviously, $(0, 0, 0)$ is the only equilibrium point and it is not stable.

The orbits are fixed by the implicit Cartesian equations $x + y = c_1 z$ (family of planes), $y - z = c_2 z^2$ (family of quadric surfaces). Then

$$f : \mathbb{R}^3 \setminus xOy \to \mathbb{R},\qquad f(x, y, z) = \frac{x + y}{z}$$

$$g : \mathbb{R}^3 \setminus xOy \to \mathbb{R},\qquad g(x, y, z) = \frac{y - z}{z^2}$$

are first integrals. The first integrals f and g are functionally independent since

$$rank\begin{pmatrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} & \dfrac{\partial f}{\partial z} \\ \dfrac{\partial g}{\partial x} & \dfrac{\partial g}{\partial y} & \dfrac{\partial g}{\partial z} \end{pmatrix} = rank\begin{pmatrix} \dfrac{1}{z} & \dfrac{1}{z} & -\dfrac{x+y}{z^2} \\ 0 & \dfrac{1}{z^2} & \dfrac{z - 2y}{z^3} \end{pmatrix} = 2.$$

Alternative. The differential system can be written in the symmetric form

$$\frac{dx}{x - y + z} = \frac{dy}{2y - z} = \frac{dz}{z},$$

which provides the linear equation $\frac{dy}{dz} = \frac{2}{z}y - 1$ with the general solution $y = c_2 z^2 + z$. Then $\frac{d(x+y)}{x+y} = \frac{dz}{z}$ has the general solution $x + y = c_1 z$.

Note. Since $divX = 4 > 0$, the global flow $T^t : \mathbb{R}^3 \to \mathbb{R}^3$,

$$T^t(x, y, z) = (xe^t - ye^{2t},\ ye^{2t} + ze^t,\ ze^t), \quad t \in \mathbb{R}$$

associated to X enlarges the volume.

2. Determine a local diffeomorphism that rectifies the vector field from the previous problem and check the result by direct computations.

Solution. Since

$$f(x, y, z) = \frac{x+y}{z}, \quad g(x, y, z) = \frac{y-z}{z^2}, \quad z \neq 0,$$

are functionally independent first integrals and $\frac{d \ln z}{dt} = 1$, the desired diffeomorphism is defined by

$$x' = \frac{x+y}{z},\ y' = \frac{y-z}{z^2},\ z' = \ln|z|,\ z \neq 0.$$

In order to check this, we denote $X_1 = x - y + z$, $X_2 = 2y - z$, $X_3 = z$ and calculate

$$X_{1'} = \tfrac{\partial x'}{\partial x} X_1 + \tfrac{\partial x'}{\partial y} X_2 + \tfrac{\partial x'}{\partial z} X_3 = 0,$$
$$X_{2'} = \tfrac{\partial y'}{\partial x} X_1 + \tfrac{\partial y'}{\partial y} X_2 + \tfrac{\partial y'}{\partial z} X_3 = 0,$$
$$X_{3'} = \tfrac{\partial z'}{\partial x} X_1 + \tfrac{\partial z'}{\partial y} X_2 + \tfrac{\partial z'}{\partial z} X_3 = \tfrac{1}{z} z = 1.$$

Note. Consider the vector field $(X_1, \ldots, X_n)$ and the diffeomorphism $x_{i'} = x_{i'}(x_i)$, $i' = \overline{1, n}$. The new components of the vector field are $(X_{1'}, \ldots, X_{n'})$, where

$$X_{i'} = \sum_{i=1}^{n} \frac{\partial x_{i'}}{\partial x_i} X_i.$$

Therefore, the rectifying diffeomorphism is fixed by the system of equations with partial derivatives

$$\sum_{i=1}^{n} \frac{\partial x_{i'}}{\partial x_i} X_i = 0,\ i' = \overline{1, n-1} \qquad \text{(homogeneous)}$$

$$\sum_{i=1}^{n} \frac{\partial x_{n'}}{\partial x_i} X_i = 1 \qquad \text{(non-homogeneous).}$$

The symmetric differential system

$$\frac{dx_1}{X_1} = \ldots = \frac{dx_n}{X_n} = \frac{dx_{n'}}{1},$$

assigned to the non-homogeneous equation, shows that $x_{1'}, \ldots, x_{(n-1)'}$ are the first integrals that determine the field lines, and $x_{n'}$ is determined from

$$x_{1'}(x_1, \ldots, x_n) = c_1, \quad \ldots, \quad x_{(n-1)'}(x_1, \ldots, x_n) = c_{n-1}, \quad dx_{n'} = \frac{dx_n}{X_n(x_1, \ldots, x_n)}.$$

3. For each of the following vector fields, determine the field lines by the method of the integrable combinations and fix a local rectifying diffeomorphism.

a) $\vec{V} = xz\vec{i} + z(2x - y)\vec{j} - x^2\vec{k}$;

b) $\vec{V} = x^2(y + z)\vec{i} - y^2(z + x)\vec{j} + z^2(y - x)\vec{k}$;

c) $\vec{V} = y^2z^2\vec{i} + xyz^2\vec{j} + xy^2z\vec{k}$;

d) $\vec{V} = xz\vec{i} + yz\vec{j} - (x^2 + y^2)\vec{k}$;

e) $\vec{V} = (y - z)\vec{i} + (z - x)\vec{j} + (x - y)\vec{k}$;

f) $\vec{V} = x(y - z)\vec{i} + y(z - x)\vec{j} + z(x - y)\vec{k}$.

Solution. a) The symmetric differential system that determines the field lines is

$$\frac{dx}{xz} = \frac{dy}{z(2x - y)} = \frac{dz}{-x^2}.$$

The solutions of the algebraic system $xz = 0$, $z(2x - y) = 0$, $x^2 = 0$, i.e., the points $(0, y, 0)$, $y \in \mathbb{R}$ and $(0, 0, z)$, $z \in \mathbb{R}$ are equilibrium points of $\vec{V}$.

The differential equation $\dfrac{dx}{xz} = \dfrac{dz}{-x^2}$ is equivalent to $xdx + zdz = 0$ and has the general solution defined by $x^2+z^2 = c_1$ (family of cylindrical surfaces with generatrices parallel to Oy).

We notice that we have

$$\frac{dx}{x} = \frac{dy}{2x - y} = \frac{d(x - y)}{y - x}.$$

Using either the linear differential equation $\dfrac{dy}{dx} = 2 - \dfrac{1}{x}y$, or the equation with separated variables $\dfrac{dx}{x} = \dfrac{d(x - y)}{y - x}$, we find $x(x - y) = c_2$ (family of cylindrical surfaces with the generatrices parallel to Oz).

We obtain the field lines (orbits)

$$x^2 + z^2 = c_1, \quad x(x - y) = c_2,$$

and specify that the *Jacobi matrix* assigned to the first integrals $x^2 + z^2$ and $x^2 - xy$ is

$$\begin{pmatrix} 2x & 0 & 2z \\ 2x - y & -x & 0 \end{pmatrix}.$$

Hence the first integrals are functionally independent on $\mathbb{R}^3 \setminus yOz$.

Since $v_3 = -x^2 = z^2 - c_1$, the rectifying diffeomorphism is defined by

$$x' = x^2 + z^2, \quad y' = x(x-y), \quad z' = \left(\int \frac{dz}{z^2 - c_1} \right)_{c_1 = x^2+z^2}.$$

The function z' rewrites

$$z' = \frac{1}{2\sqrt{c_1}} \ln \left| \frac{z - \sqrt{c_1}}{z + \sqrt{c_1}} \right|_{c_1 = x^2+z^2} = \frac{1}{2\sqrt{x^2+z^2}} \ln \left| \frac{z - \sqrt{x^2+z^2}}{z + \sqrt{x^2+z^2}} \right|,$$

where we neglect the constant of integration that yields just a translation !).

Note. Since $div\vec{V} = 0$ (i.e., $\vec{V}$ is a solenoidal field), the associated flow of $\vec{V}$ perserves the volume.

b) The symmetric differential system

$$\frac{dx}{x^2(y+z)} = \frac{dy}{-y^2(z+x)} = \frac{dz}{z^2(y-x)}$$

may be written in the following forms

$$\frac{x^{-1}dx}{x(y+z)} = \frac{y^{-1}dy}{-y(z+x)} = \frac{z^{-1}dz}{z(y-x)} = \frac{x^{-1}dx + y^{-1}dy + z^{-1}dz}{0},$$
$$\frac{x^{-2}dx}{y+z} = \frac{y^{-2}dy}{-z-x} = \frac{-z^{-2}dz}{-y+x} = \frac{x^{-2}dx + y^{-2}dy - z^{-2}dz}{0}.$$

From $x^{-1}dx + y^{-1}dy + z^{-1}dz = 0$ we find $xyz = c_1$ (family of Ţiţeica surfaces), while from $x^{-2}dx + y^{-2}dy - z^{-2}dz = 0$ we obtain $\frac{1}{x} + \frac{1}{y} - \frac{1}{z} = c_2$ (the domain of definition of the function (first integral) defined by the left side is not connected). Thus, the orbits are depicted by the implicit Cartesian equations

$$\begin{cases} xyz = c_1 \\ \frac{1}{x} + \frac{1}{y} - \frac{1}{z} = c_2. \end{cases}$$

To these, we add the equilibrium points which are generated by the solutions of the algebraic system

$$x^2(y+z) = 0, \quad y^2(z+x) = 0, \quad z^2(y-x) = 0.$$

We notice that

$$v_3 = z^2(y-x) = z^2\sqrt{(x+y)^2 - 4xy} = z^2\sqrt{\left(c_2 + \frac{1}{z}\right)^2 \frac{c_1^2}{z^2} - 4\frac{c_1}{z}}.$$

Therefore the rectifying diffeomorphism is defined by

$$x' = xyz, \quad y' = \frac{1}{x} + \frac{1}{y} - \frac{1}{z}, \quad z' = \left(\int \frac{dz}{\sqrt{(zc_2+1)^2 c_1^2 - 4c_1 z^3}} \right) \Bigg|_{\substack{c_1 = xyz \\ c_2 = \frac{1}{x}+\frac{1}{y}-\frac{1}{z}}}$$

(an *elliptic antiderivative*).

c) From the differential system

$$\frac{dx}{dt} = y^2 z^2, \quad \frac{dy}{dt} = xyz^2, \quad \frac{dz}{dt} = xy^2 z,$$

we obtain

$$y\frac{dy}{dt} - z\frac{dz}{dt} = 0 \Rightarrow \frac{d}{dt}(y^2 - z^2) = 0 \Rightarrow y^2 - z^2 = c_1,$$

(a family of hyperbolic cylinders with the generatrices parallel to Ox). Moreover, we find

$$x\frac{dx}{dt} - y\frac{dy}{dt} = 0 \Rightarrow \frac{d}{dt}(x^2 - y^2) = 0 \Rightarrow x^2 - y^2 = c_2,$$

(a family of hyperbolic cylinders with the generatrices parallel to Oz).

It follows that the family of orbits is given by

$$\begin{cases} y^2 - z^2 = c_1 \\ x^2 - y^2 = c_2. \end{cases}$$

The already determined first integrals are functionally independent on $\mathbb{R}^3 \setminus (Ox \cup Oy \cup Oz)$. The equilibrium points are generated by the solutions of the algebraic system

$$y^2 z^2 = 0, \quad xyz^2 = 0, \quad xy^2 z = 0.$$

We can write

$$v_3 = xy^2 z = \sqrt{c_1 + c_2 + z^2}(c_1 + z^2)z,$$

function that determines the rectifying diffeomorphism

$$x' = y^2 - z^2, \; y' = x^2 - y^2, \; z' = \left(\int \frac{dz}{\sqrt{c_1 + c_2 + z^2}(c_1 + z^2)z} \right) \Bigg|_{\substack{c_1 = y^2 - z^2 \\ c_2 = x^2 - y^2}}.$$

Note. We can write $\vec{V} = yz \; \text{grad}\,(xyz)$, that is, $\vec{V}$ is a *bi-scalar vector field.*

d) Consider the first-order differential system

$$\frac{dx}{dt} = xz, \quad \frac{dy}{dt} = yz, \quad \frac{dz}{dt} = -x^2 - y^2.$$

We notice that the axis Oz consists of equilibrium points. Then,

$$\frac{d}{dt}\ln|x| = \frac{d}{dt}\ln|y|, \quad x \neq 0, \quad y \neq 0$$

implies $x = c_1 y, \quad x \neq 0,\ y \neq 0$ (family of parts of planes). By analogy,

$$x\frac{dx}{dt} + y\frac{dy}{dt} + z\frac{dz}{dt} = 0$$

gives

$$\frac{d}{dt}(x^2 + y^2 + z^2) = 0 \quad \text{or} \quad x^2 + y^2 + z^2 = c_2,$$

a family of spheres of centre at the origin. All in all, the orbits are arcs of circles (which may degenerate in particular to points).

From

$$x = c_1 y, \quad x^2 + y^2 + z^2 = c_2, \quad v_3 = -x^2 - y^2,$$

we find $v_3 = z^2 - c_2$. Thus,

$$\begin{aligned} x' &= \frac{x}{y}, \qquad y' = x^2 + y^2 + z^2, \\ z' &= \int \frac{dz}{z^2 - c_2}\bigg|_{c_2 = x^2+y^2+z^2} = \frac{1}{2\sqrt{c_2}}\ln\left|\frac{z - \sqrt{c_2}}{z + \sqrt{c_2}}\right|_{c_2 = x^2+y^2+z^2} = \\ &= \frac{1}{2\sqrt{x^2+y^2+z^2}}\ln\left|\frac{z - \sqrt{x^2+y^2+z^2}}{z + \sqrt{x^2+y^2+z^2}}\right|, \end{aligned}$$

where we neglected the constant of integration, define a rectifying diffeomorphism.

Note. The vector $\vec{V} = \dfrac{z^3}{2}\,\text{grad}\,\dfrac{x^2+y^2}{z^2},\ z \neq 0$ is a *bi-scalar vector field.*

e) The homogeneous differential linear system

$$\frac{dx}{dt} = y - z, \quad \frac{dy}{dt} = z - x, \quad \frac{dz}{dt} = x - y$$

admits a straight line of equilibrium points of equations $x = y = z$. We also infer

$$\begin{aligned} &\frac{dx}{dt} + \frac{dy}{dt} + \frac{dz}{dt} = 0 &&\Rightarrow x + y + z = c_1 \\ &x\frac{dx}{dt} + y\frac{dy}{dt} + z\frac{dz}{dt} = 0 &&\Rightarrow x^2 + y^2 + z^2 = c_2, \end{aligned}$$

which represent respectively a family of planes and a family of spheres.

Hence the orbits are circles laying in planes which are perpendicular to the straight line $x = y = z$. f) ♣ Hw.

4. For each of the following vector fields, determine the field lines and fix a local rectifying diffeomorphism:

a) $\vec{V} = (x^2 - y^2 - z^2)\vec{i} + 2xy\vec{j} + 2xz\vec{k}$;

b) $\vec{V} = x\vec{i} + y\vec{j} + (z + \sqrt{x^2 + y^2 + z^2})\vec{k}$;

c) $\vec{V} = (x - y)\vec{i} + (x + y)\vec{j} + z\vec{k}$;

d) $\vec{V} = (x + y)\vec{i} - x(x + y)\vec{j} + (x - y)(2x + 2y + z)\vec{k}$;

e) $\vec{V} = x\vec{i} + y\vec{j} + (z - x^2 - y^2)\vec{k}$;

f) $\vec{V} = y\vec{i} + x\vec{j} + 2xy\sqrt{a^2 - z^2}\vec{k}$;

g) $\vec{V} = (x^2 - yz)\vec{i} + (y^2 - zx)\vec{j} + (z^2 - xy)\vec{k}$;

h) $\vec{V} = (xy - 2z^2)\vec{i} + (4xz - y^2)\vec{j} + (yz - 2x^2)\vec{k}$.

In each case, find the field line that passes through the point (1,1,1).

Hint. The field lines are: a) $y = c_1 z,\ x^2 + y^2 + z^2 = 2c_2 z$,

b) $y = c_1 x,\ z = c_2 + \sqrt{x^2 + y^2 + z^2}$;

c) $\dfrac{dz}{z} = \dfrac{1}{2}\dfrac{d(x^2 + y^2)}{x^2 + y^2} = \dfrac{d(x^{-1}y)}{1 + (x^{-1}y)^2}$;

d) $x^2 + y^2 = c_1,\ (x + y)(x + y + z) = c_2$;

e) $y = c_1 x,\ z + x^2 + y^2 = c_2 x$;

f) $x^2 - y^2 = c_1,\ x^2 + y^2 = 2\arcsin\dfrac{z}{a} + c_2$;

g) $x - y = c_1(y - z),\ xy + yz + zx = c_2$;

h) $z^2 + xy = c_1,\ x^2 + yz = c_2$.

5. Determine the vector fields whose orbits are

a) $\begin{cases} x + \arcsin\dfrac{y}{z} = c_1 \\ y + \arcsin\dfrac{x}{z} = c_2, \end{cases}$

b) $\begin{cases} \cosh xy + z = c_1 \\ x + \sinh yz = c_2. \end{cases}$

6. Approximate the solution of the Cauchy problem

$$\frac{dx}{dt} = x(1 - 2y^2), \quad \frac{dy}{dt} = -y(1 + 2x^2), \quad \frac{dz}{dt} = 2z(x^2 + y^2),$$
$$x(0) = y(0) = z(0) = 1, \quad t \in [0, 10], \quad h = 0.2,$$

by the Runge-Kutta method. Verify the result using first integrals.

7. Study the completeness of the vector fields from the problem 4.

8. Consider the Hamiltonian

$$H(x,y) = \frac{1}{2}\sum_{k=1}^{3} y_k^2 + V(x).$$

In each of the following cases determine the flow generated by the Hamiltonian vector field associated to H and specify if it is a global flow.

a) $V(x) = x_1^2 - x_2^2 + x_3$,

b) $V(x) = x_1^2 + x_2 + x_3$,

Solution. a) We obtain

$$\begin{cases} x_1 = a_1 \cos\sqrt{2}t + b_1 \sin\sqrt{2}t \\ x_2 = a_2 e^{\sqrt{2}t} + b_2 e^{-\sqrt{2}t} \\ x_3 = -\dfrac{1}{2}t^2 + a_3 t + b_3 \end{cases} ; \quad \begin{cases} y_1 = \sqrt{2}a_1 \sin\sqrt{2}t - 2b_1 \cos\sqrt{2}t \\ y_2 = -\sqrt{2}a_2 e^{\sqrt{2}t} + 2b_2 e^{-\sqrt{2}t} \\ y_3 = t - a_3, \end{cases}$$

$t \in \mathbb{R}$ (global flow).

9. In each of the following cases, determine the infinitesimal transformation and the flow generated by the vector field X. Study if the corresponding flow reduces, enlarges or preserves the volume and specify the fixed points.

a) $X = (-x(x^2 + 3y^2), -2y^3, -2y^2 z)$;

b) $X = (3xy^2 - x^3, -2y^3, -2y^2 z)$;

c) $X = (xy^2, x^2 y, z(x^2 + y^2))$;

d) $X = (x(y^2 - z^2), -y(x^2 + z^2), z(x^2 + y^2))$;

e) $X = (x(y - z), y(z - x), z(x - y))$;

f) $X = ((z - y)^2, z, y)$.

Hint. a) and b) reduce the volume; c) and d) enlarge the volume; e) and f) preserve the volume.

2.2 Field hypersurfaces and linear PDEs

2.2.1 Let $D \subset \mathbb{R}^n$ be an open and connected subset, and let X be a vector field of class $\mathcal{C}^1$ on D of components $(X_1, \ldots, X_n)$, $X_i : D \to \mathbb{R}$, $i = \overline{1,n}$. For a given function $f : D \to \mathbb{R}$ of class $\mathcal{C}^1$ and for $C \in \mathbb{R}$, the set of constant level

$$M : \ f(x_1, \ldots, x_n) = C$$

determines a *hypersurface* of $\mathbb{R}^n$ iff

$$\operatorname{grad} f \,|_M \neq 0.$$

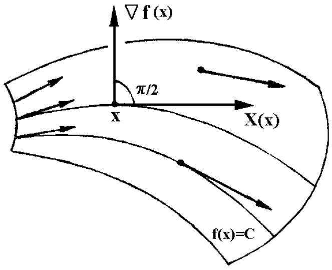

Fig. 2

If the vector field X is tangent to M (see the figure), then one has

$$\langle X_x,\ \operatorname{grad} f(x)\rangle = 0, \quad \text{for all } x \in M, \tag{60}$$

which rewrites in brief $D_X f|_M = 0$. In this case, we say that M is a *field hypersurface of* X. The relation (60) rewrites explicitly as a *homogeneous linear equation with partial derivatives (PDE) of the first order*

$$X_1(x)\frac{\partial f}{\partial x_1} + \ldots + X_n(x)\frac{\partial f}{\partial x_n} = 0, \tag{61}$$

with the unknown function f. To this we attach the symmetric differential system

$$\frac{dx_1}{X_1(x)} = \ldots = \frac{dx_n}{X_n(x)}, \tag{62}$$

that determines the field lines, often called the *characteristic differential system* of the equation (61).

A field hypersurface of X is generated by the field lines produced by the symmetric differential system (62). Also, paraphrasing the definition of first integrals, we note that any first integral of (62) is a solution of (61). Moreover, for any $n-1$ first integrals $f_1, \ldots, f_n$ of (62), which are functionally independent in a neighborhood of a point $x_0 \in D$ at which $X(x_0) \neq 0$, a function f is a solution of the PDE (61) iff it is a functional combination of $f_1, \ldots, f_n$, i.e., there exists a mapping $\Phi : \mathbb{R}^n \to \mathbb{R}$ of class $\mathcal{C}^1$ such that

$$f(x) = \Phi(f_1(x), \ldots, f_n(x)). \tag{63}$$

Then, *the general solution of the symmetric system* (62), given by the field lines

$$f_1(x) = C_1, \ldots, f_{n-1}(x) = C_{n-1}$$

provides the *general solution* (63) of (61).

2.2.2 Example. Consider the homogeneous linear PDE:

$$(x-a)\frac{\partial f}{\partial x}+(y-b)\frac{\partial f}{\partial y}+(z-c)\frac{\partial f}{\partial z}=0.$$

The associated symmetric differential system

$$\frac{dx}{x-a}=\frac{dy}{y-b}=\frac{dz}{z-c}$$

admits the first integrals (obtained from the ODE provided by both the equalities)

$$f_1(x,y,z)=\frac{x-a}{y-b}, \quad f_2(x,y,z)=\frac{y-b}{z-c}.$$

Hence the field lines of the field $X(x,y,z)=(x-a,y-b,z-c)$ are given by the two-parameter family of curves

$$\left\{\begin{array}{l} f_1=C_1 \\ f_2=C_2 \end{array}\right. \Leftrightarrow \left\{\begin{array}{l} \dfrac{x-a}{y-b}=C_1 \\ \dfrac{y-b}{z-c}=C_2 \end{array}\right. \Leftrightarrow \left\{\begin{array}{l} x-a-C_1(y-b)=0 \\ y-b-C_2(z-c)=0, \end{array}\right.$$

i.e., a family of straight lines in $\mathbb{R}^3$. Moreover, the first integrals f_1 and f_2 provide the general solution f of the form

$$f(x,y,z)=\Phi\left(\frac{x-a}{y-b},\frac{y-b}{z-c}\right),$$

with arbitrary function $\Phi : R^2 \to \mathbb{R}$ of class $\mathcal{C}^1$.

2.2.3 Let $f=(f_1,\ldots,f_{n-1})$ be the general solution of the system (62), and

$$M_C: \quad \Phi(f_1(x),\ldots,f_{n-1}(x))=0$$

the associated family of field hypersurfaces. For determining a *unique* solution of (61), (i.e., a determined hypersurface $M_0: \ \Phi(f_1(x),\ldots,f_{n-1}(x))=0$ attached to X) one needs to impose certain restrictive conditions besides (61). This leads to

The Cauchy problem for the linear homogeneous PDE. Find the field hypersurface which contains the submanifold with $n-2$ dimensions $\Gamma: \left\{\begin{array}{l} g(x)=0 \\ h(x)=0 \end{array}\right.$ (see the figure).

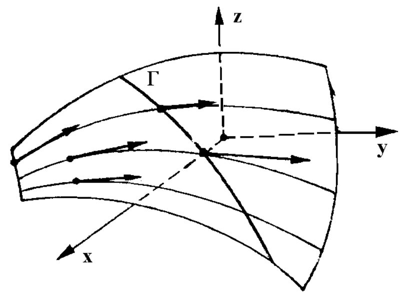

Fig. 3

For solving this problem, we consider the system of $n+1$ equations

$$f_1(x) = C_1, \ \ldots, \ f_{n-1}(x) = C_{n-1}, \quad g(x) = 0, \quad h(x) = 0$$

in the n unknowns $x_1, \ldots, x_n$, and derive the compatibility condition of the form

$$\Phi(C_1, \ldots, C_{n-1}) = 0.$$

Then, under the hypothesis that the field lines are not included in Γ, the desired geometric locus is

$$M: \ \Phi(f_1(x), \ldots, f_{n-1}(x)) = 0.$$

Example. Find the solution of the Cauchy problem

$$\begin{cases} (x^2+y^2)\dfrac{\partial f}{\partial x} + 2xy\dfrac{\partial f}{\partial y} + xz\dfrac{\partial f}{\partial z} = 0 \\ \Gamma: \ \ x = 1, \ \ y^2 + z^2 = 1. \end{cases}$$

From the geometric point of view, find the field surface $M: \ \ f(x,y,z) = 0$ of the field

$$X(x,y,z) = (x^2+y^2, 2xy, xz),$$

containing the circle Γ located in the plane $z = 1$, which has radius 1 and the center at $C(0,0,1)$.

2.2.4. Exercises

1. Show that the functions f and g defined by

$$f(x) = x_i + x_j + x_k, \qquad g(x) = x_i x_j x_k, \ \ x = (x_1, \ldots, x_n) \in \mathbb{R}^n,$$

with $i, j, k \in \overline{1,n}$ fixed, are solutions of the PDE

$$x_i(x_j - x_k)\frac{\partial f}{\partial x_i} + x_j(x_k - x_i)\frac{\partial f}{\partial x_j} + x_k(x_i - x_j)\frac{\partial f}{\partial x_k} = 0.$$

Solution. We notice that

$$\frac{\partial f}{\partial x_i} = 1, \quad \frac{\partial f}{\partial x_j} = 1, \quad \frac{\partial f}{\partial x_k} = 1, \quad \frac{\partial f}{\partial x_l} = 0, \qquad l \neq i, j, k,$$

and

$$\frac{\partial g}{\partial x_i} = x_j x_k, \quad \frac{\partial g}{\partial x_j} = x_i x_k, \quad \frac{\partial g}{\partial x_k} = x_i x_j, \quad \frac{\partial g}{\partial x_l} = 0, \ l \neq i, j, k.$$

Introducing these relations in the equations, we obtain identities.

2. Let the vector fields

$$X_1 = (1, 0, 0), \qquad X_2 = (0, 1, 0), \qquad X_3 = (0, 0, 1),$$

$$X_4 = (-x_2, x_1, 0), \quad X_5 = (0, -x_3, x_2), \quad X_6 = (x_3, 0, -x_1), \quad (x_1, x_2, x_3) \in \mathbb{R}^3$$

be linearly independent Killing vector fields in $\mathcal{X}(\mathbb{R}^\ni)$. Determine the field surfaces of these fields.

Solution. A vector field $X = (X_1, \ldots, X_n)$ on $\mathbb{R}^n$ that satisfies the system with partial derivatives $\dfrac{\partial X_i}{\partial x_j} + \dfrac{\partial X_j}{\partial x_i} = 0$, $i, j = 1, \ldots, n$ is called a *Killing field.*

Taking account of the definition of the field hypersurfaces, we write the homogeneous linear PDEs of the first order and find their solutions.

We get $\dfrac{\partial f}{\partial x_1} = 0 \Rightarrow f_1(x_2, x_3) = c_1$ (families of cylindrical surfaces with generatrices parallel to Ox_2). By cyclic permutations, we get the field surfaces of X_2 and X_3 as being $f_2(x_3, x_1) = c_2$ (families of cylindrical surfaces with generatrices parallel to Ox_2) respectively $f_3(x_1, x_2) = c_3$ (families of cylindrical surfaces with generatrices parallel to Ox_1).

For X_4 we obtain $-x_2 \dfrac{\partial f}{\partial x_1} + x_1 \dfrac{\partial f}{\partial x_2} = 0$. To this equation, we assign the characteristic differential system $\dfrac{dx_1}{-x_2} = \dfrac{dx_2}{x_1} = \dfrac{dx_3}{0}$ and the general solution of the equation with partial derivatives is $f_4(x_1^2 + x_2^2, x_3) = c_4$ (families of surfaces of revolution). By circular permutations, we find other two families of surfaces of revolution $f_5(x_2^2 + x_3^2, x_1) = c_5$ and $f_6(x_3^2 + x_1^2, x_2) = c_6$ corresponding to X_5, respectively, X_6.

3. For each of the following vector fields, determine the family of their field surfaces.

a) $\vec{V} = x(x+z)\vec{i} + y(y+z)\vec{j} + (z^2 - xy)\vec{k}$,

b) $\vec{V} = (xy^3 - 2x^4)\vec{i} + (2y^4 - x^3 y)\vec{j} + 9z(x^3 - y^3)\vec{k}$,

c) $\vec{V} = (x^2 + y^2)\vec{i} + 2xy\vec{j} + xz\vec{k}$.

Solution. a) To the equation

$$x(x+z)\frac{\partial f}{\partial x} + y(y+z)\frac{\partial f}{\partial y} + (z^2 - xy)\frac{\partial f}{\partial z} = 0,$$

we assign the symmetric differential system

$$\frac{dx}{x(x+z)} = \frac{dy}{y(y+z)} = \frac{dz}{z^2 - xy}.$$

From $\dfrac{dx}{x(x+z)} = \dfrac{ydz - zdy}{-y^2(x+z)}$ it follows $\dfrac{dx}{x} = -\dfrac{ydz - zdy}{y^2}$, and then $\ln|x| + \dfrac{z}{y} = c_1$ (the domain of definition of the first integral $(x, y, z) \to \ln|x| + \dfrac{z}{y}$ is not connected).

From $\dfrac{dy}{y(y+z)} = \dfrac{xdz - zdx}{-x^2(y+z)}$, we find $\dfrac{dy}{y} = -\dfrac{xdz - zdx}{x^2}$ and hence $\ln|y| + \dfrac{z}{x} = c_2$.

The family of orbits

$$\ln|x| + \frac{z}{y} = c_1, \;\; \ln|y| + \frac{z}{x} = c_2, \;\; x \neq 0, \; y \neq 0$$

produces the family of the field surfaces

$$\Phi\left(\ln|x| + \frac{z}{y}, \ln|y| + \frac{z}{x}\right) = 0, \;\; x \neq 0, \; y \neq 0,$$

where Φ is an arbitrary function of class $\mathcal{C}^1$, and $f = \Phi(c_1, c_2)$.

Note. It remains to find the field lines on $x = 0, \; y \neq 0$. The restriction of $\vec{V}$ to $x = 0, \; y \neq 0$ is

$$\vec{V}(0, y, z) = y(y+z)\vec{j} + z^2\vec{k}.$$

From $\dfrac{dy}{y(y+z)} = \dfrac{dz}{z^2}$, we obtain $\dfrac{dz}{dy} = \dfrac{z^2}{y(y+z)}$ (homogeneous differential equation) with the general solution $z = c_3 e^{-z/y}$. Then $x = 0, \; z = c_3 e^{-z/y}$ are field lines of $\vec{V}$.

By analogy, we discuss the cases $x \neq 0, \; y = 0$, respectively $x = 0, \; y = 0$.

The equation

$$(xy^3 - 2x^4)\frac{\partial f}{\partial x} + (2y^4 - x^3 y)\frac{\partial f}{\partial y} + 9z(x^3 - y^3)\frac{\partial f}{\partial z} = 0$$

corresponds to the characteristic system

$$\frac{dx}{xy^3 - 2x^4} = \frac{dy}{2y^4 - x^3 y} = \frac{dz}{9z(x^3 - y^3)}.$$

We notice that $3x^2y^3zdx + 3x^3y^2zdy + x^3y^3dz = 0$ and thus $x^3y^3z = c_1$. Then, we see that

$$(y^{-2} - 2x^{-3}y)dx + (x^{-2} - 2y^{-3})dy = 0$$

and so $xy^{-2} + yx^{-2} = c_2$ (the domain of definition of this first integral, characterized by $x \neq 0, \; y \neq 0$, is not connected).

The family of surfaces that contains the family of orbits

$$x^3y^3z = c_1, \;\; xy^{-2} + x^{-2}y = c_2, \;\; x \neq 0, \;\; y \neq 0$$

has the equation

$$\Phi(x^3y^3z, \; xy^{-2} + x^{-2}y) = c, \;\; x \neq 0, \;\; y \neq 0,$$

where Φ is an arbitrary function of class $\mathcal{C}^1$.

Note. a) For $x = 0, \; y \neq 0$ we find $\vec{V}(0, y, z) = 2y^4\vec{j} - 9zy^3\vec{k}$. The differential equation $\dfrac{dy}{2y^4} = \dfrac{dz}{-9zy^3}$ has the general solution $y^9z^2 = c_3$. Hence the field lines are

$$x = 0, \;\; y^9z^2 = c_3.$$

By analogy, we treat the case $x \neq 0, \; y = 0$.

b) The relation $\Phi(x^3y^3z, xy^{-2} + x^{-2}y) = c$ rewrites $x^3y^3z = \varphi(xy^{-2} + x^{-2}y)$, where φ is an arbitrary function of class $\mathcal{C}^1$.

c) To the equation

$$(x^2 + y^2)\frac{\partial f}{\partial x} + 2xy\frac{\partial f}{\partial y} + xz\frac{\partial f}{\partial z} = 0$$

we associate the symmetric differential system

$$\frac{dx}{x^2 + y^2} = \frac{dy}{2xy} = \frac{dz}{zx}.$$

We infer

$$\frac{dy}{y} = 2\frac{dz}{z} \qquad \Rightarrow z^2 = c_1 y$$

$$\frac{d(x-y)}{(x-y)^2} = \frac{d(x+y)}{(x+y)^2} \quad \Rightarrow x^2 - y^2 = c_2 y.$$

Hence the family of the field surfaces has the equation

$$\Phi\left(\frac{z^2}{y}, \frac{x^2 - y^2}{y}\right) = c,$$

$y \neq 0$, where Φ is an arbitrary function of class $\mathcal{C}^1$.

For $y = 0$, we obtain $\vec{V}(x, 0, z) = x^2\vec{i} + xz\vec{k}$, and $\dfrac{dx}{x^2} = \dfrac{dz}{xz}$ yields $x = c_3z, \; z \neq 0$. Thus, $y = 0, \; x = c_3z; \; z \neq 0$ are field lines.

4. Consider the vector field

$$\vec{V} = x^2(y + z)\vec{i} - y^2(x + z)\vec{j} + z^2(y - x)\vec{k}.$$

Determine the field surface that passes through the curve of equations $\begin{cases} xy = 1 \\ x + y = 1. \end{cases}$

Solution. We notice that

$$\begin{aligned}\frac{dx}{x^2(z+y)} &= \frac{dy}{-y^2(z+x)} = \frac{dz}{z^2(y-x)} = \\ &= \frac{x^{-1}dx + y^{-1}dy + z^{-1}dz}{0} = \frac{x^{-2}dx + y^{-2}dy - z^{-2}dz}{0}.\end{aligned}$$

Then the family of orbits of $\vec{V}$ is

$$\begin{aligned}&xyz = c_1 \\ &\frac{1}{x} + \frac{1}{y} - \frac{1}{z} = c_2, \quad x \neq 0, y \neq 0, z \neq 0.\end{aligned} \tag{64}$$

The condition of compatibility of the system

$$\begin{aligned}&xyz = c_1, \quad xy = 1, \quad y + x = 1, \\ &\frac{1}{x} + \frac{1}{y} - \frac{1}{z} = c_2, \quad x \neq 0, y \neq 0, z \neq 0,\end{aligned}$$

selects the orbits which lean upon the given curve. We find $1 - c_1^{-1} = c_2$. We reconsider the desired surface as being the locus of the orbits (64) which satisfy the condition $1 - c_1^{-1} = c_2$. Eliminating the parameters c_1, c_2, we obtain the surface

$$1 - \frac{1}{xyz} = \frac{1}{x} + \frac{1}{y} - \frac{1}{z}, \quad x \neq 0, y \neq 0, z \neq 0.$$

Note. *div* $\vec{V} = 0$, and hence $\vec{V}$ is a solenoidal field.

2.3 Nonhomogeneous linear PDEs

2.3.1 Definition. We call *nonhomogeneous linear equation with partial derivatives of order one*, a PDE with unknown function $f : D \subset \mathbb{R}^n \to \mathbb{R}$ of the form

$$X_1(x, f(x))\frac{\partial f}{\partial x_1} + \ldots + X_n(x, f(x))\frac{\partial f}{\partial x_n} = F(x, f(x)),$$

where $X_1, \ldots, X_n, F : \mathbb{R}^{n+1} \to \mathbb{R}$ are differentiable functions of class $\mathcal{C}^1$.

For solving this type of PDE, we look for the solution f given in implicit form $\Phi(x, f) = 0$, provided by a PDE of form (61) with the associated symmetric differential system

$$\frac{dx_1}{X_1(x, f)} = \ldots \frac{dx_n}{X_n(x, f)} = \frac{df}{F(x, f)}.$$

2.3.2 Example. Find the general solution of the PDE

$$xz\frac{\partial z}{\partial x} - yz\frac{\partial z}{\partial y} = x^2 + y^2.$$

Solution. The characteristic symmetric differential system is

$$\frac{dx}{xz} = \frac{dy}{-yz} = \frac{dz}{x^2 + y^2}.$$

Its first integrals are

$$f_1(x, y, z) = xy, f_2(x, y, z) = (x^2 - y^2 + z^2)/2,$$

hence the general solution is given implicitly by equation of the form

$$\Phi(xy, (x^2 - y^2 + z^2)/2) = 0.$$

2.3.3. Exercises

1. Solve the following Cauchy problems

a) $z(x + z)\frac{\partial z}{\partial x} - y(y + z)\frac{\partial z}{\partial y} = 0,\ z(1, y) = \sqrt{y};$

b) $\frac{\partial z}{\partial y} = xz\frac{\partial z}{\partial x},\ z(x, 0) = x.$

Solution. a) This is a non-homogeneous linear equation with partial derivatives of the first order. The associate symmetric differential system is

$$\frac{dx}{z(x + z)} = \frac{dy}{-y(y + z)} = \frac{dz}{0},$$

and $dz = 0$ implies $z = c_1$. Then the system rewrites

$$\frac{dy}{y(y + c_1)} = \frac{-dx}{-c_1(x + c_1)} \quad \Leftrightarrow \quad \frac{1}{y(y + c_1)} = \frac{1}{c_1 y} - \frac{1}{c_1(y + c_1)},$$

whence $\frac{y(x + c_1)}{y + c_1} = c_2$ and after plugging in $c_1 = z$, we find $\frac{y(x + z)}{y + z} = c_2$.

The general solution of the equation with partial derivatives is $\frac{y(x + z)}{y + z} = \varphi(z)$, where φ is an arbitrary function of class $\mathcal{C}^1$. Putting the condition $z(1, y) = \sqrt{y}$, it follows $\frac{y(1 + \sqrt{y})}{y + \sqrt{y}} = \varphi(\sqrt{y})$ or, otherwise, $\varphi(u) = \frac{u^2(1 + u)}{u^2 + u} = u$. Coming back to the general solution, we find $\frac{y(x + z)}{y + z} = z$ or $z = \sqrt{xy}$.

b) Non-homogeneous linear equation with partial derivatives of the first order. The characteristic differential system $\frac{dx}{-xz} = \frac{dy}{1} = \frac{dz}{0}$ becomes $z = c_1$, $\frac{dx}{x} = -c_1 dy$ or $z = c_1$, $x = c_2 e^{-yz}$.

The general solution of the equation with partial derivatives is $z = \Phi(xe^{yz})$, where Φ is an arbitrary function of class $\mathcal{C}^1$. Putting the condition $z(x, 0) = x$, we find $\Phi(x) = x$, or otherwise, $\Phi(u) = u$. Coming back with this function Φ in the general solution, we obtain $z = xe^{yz}$.

2. Determine the general solution of the Euler equation

$$\sum_{i=1}^{n} x_i \frac{\partial f}{\partial x_i}(x) = pf(x).$$

Solution. Assuming $p \neq 0$, the characteristic system

$$\frac{dx_1}{x_1} = \ldots = \frac{dx_n}{x_n} = \frac{df}{pf}$$

admits the general solution

$$x_2 = c_1 x_1, \quad x_3 = c_2 x_1, \ldots, x_n = c_{n-1} x_1, \quad f = c_n x_1^p.$$

The general solution of the Euler equation is the homogeneous function

$$f(x) = x_1^p \varphi\left(\frac{x_2}{x_1}, \ldots, \frac{x_n}{x_1}\right),$$

where φ is a function of class $\mathcal{C}^1$; we note that f has the degree of homogeneity p.

3. Let there be given the vector fields

$$X = (y, -x, 0),\ Y = (x, y, -z),\ Z = (x^2, yx, xz),\ U = (xy, y^2, yz),\ V = (xz, yz, z^2).$$

Show that the family of cones

$$x^2 + y^2 = c^2 z^2,$$

where c is a real parameter, is invariant with respect to the flow generated by each of the given vector fields.

Solution. Consider the function $f : \mathbb{R}^3 \setminus xOy \to \mathbb{R}$, $f(x, y, z) = \frac{x^2+y^2}{z^2}$.

We find

$$D_X f = y\frac{\partial f}{\partial x} - x\frac{\partial f}{\partial y} = y\frac{2x}{z^2} - x\frac{2y}{z^2} = 0,$$

$$D_Y f = x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} - z\frac{\partial f}{\partial z} = x\frac{2x}{z^2} + y\frac{2y}{z^2} - z\frac{2(x^2+y^2)}{z^3} = 0,$$

$$D_Z f = 0, \quad D_U f = 0, \quad D_V f = 0.$$

Note. f is a first integral that is common to all the systems that determine the field lines of X, Y, U, Z and respectively V.

4. Determine the implicit Cartesian equation of the cylinder with generatrices parallel to $\vec{V} = \vec{i}+\vec{j}-\vec{k}$ and having as a director curve the circle $C : x^2+y^2 = 1,\ z = 1$.

Solution. We have to solve the following Cauchy problem: determine the field surface of $\vec{V}$ that leans upon the circle C. The equation

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} - \frac{\partial f}{\partial z} = 0$$

has the characteristic system

$$\frac{dx}{1} = \frac{dy}{1} = \frac{dz}{-1},$$

with the general solution $y-x = c_1$, $z+x = c_2$. Among these straight lines, we select those which lean upon the circle C, that is, we find the condition of compatibility of the system

$$y - x = c_1,\ \ z + x = c_2,\ \ x^2 + y^2 = 1,\ \ z = 1.$$

We reconsider the desired surface as being the locus of the straight lines $y - x = c_1$, $z + x = c_2$ which satisfy

$$(c_2 - 1)^2 + (c_2 + c_1 - 1)^2 = 1.$$

After replacing c_1, c_2 with the corresponding first integrals, we get

$$(z + x - 1)^2 + (z + y - 1)^2 = 1.$$

5. Determine the implicit Cartesian equation of the cone that has the vertex at the origin of the axes and the director curve $C : x^2 + y^2 = x,\ z = 1$.

Solution. The concurrent vector field determined by $O(0, 0, 0)$ and $M(x, y, z)$ is $X = (x, y, z)$. By this notice, we get a Cauchy problem: determine the field surface of X that leans upon the circle C. The field surfaces of X are the sets of constant level assigned to the solutions of the PDE

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} + z\frac{\partial f}{\partial z} = 0.$$

To this, it corresponds the characteristic system

$$\frac{dx}{x} = \frac{dy}{y} = \frac{dz}{z}$$

with the general solution $y = c_1x$, $z = c_2x$.

The condition of compatibility of the algebraic system

$$y = c_1x,\ \ z = c_2x,\ \ x^2 + y^2 = x,\ z = 1$$

is $1+c_1^2=c_2$. Plugging in $c_1=y/x,\ c_2=z/x$ in this relation, we find

$$1+\frac{y^2}{x^2}=\frac{z}{x}\quad\Leftrightarrow\quad x^2+y^2=zx.$$

6. Determine the implicit Cartesian equation of the surface obtained by the rotation of the parabola $P: y^2=2z+1,\ x=0$ about the axis Oz.

Solution. Consider the vector field $\vec{X}=x\vec{i}+y\vec{j}+z\vec{k}$ and $y=\vec{k}$ the *parallel vector field* that gives the direction of the axis. It follows the Killing vector field $\vec{Z}=\vec{X}\times\vec{Y}=y\vec{i}-x\vec{j}$. By this remark, it remains to find the field surface of $\vec{Z}$ that leans upon the parabola P. We find the orbits of $\vec{Z}$, that is, the solutions of the system

$$\frac{dx}{y}=\frac{dy}{-x}=\frac{dz}{0}.$$

These are $z=c_1$, $x^2+y^2=c_2$ (family of circles). We select the circles that lean upon P; the condition of compatibility of the algebraic system

$$z=c_1,\quad x^2+y^2=c_2,\quad y^2=2z+1,\quad x=0$$

is $c_2=2c_1+1$. We consider the surface as being the locus of the circles $z=c_1$, $x^2+y^2=c_2$ for which $c_2=2c_1+1$. Eliminating c_1,c_2 we find the paraboloid

$$x^2+y^2-2z=1.$$

7. There are given the following vector fields:

$$X=(1,2x,3y),\quad Y=(x,2y,3z),\quad Z=(0,1,3x)$$

and the function $f:\mathbb{R}^3\to\mathbb{R}$, $f(x,y,z)=3xy-z-2x^3$. Show that:

a) the function f is invariant with respect to the groups of one parameter generated by X, respectively Z;

b) the function f is a eigenvector of Y with respect to the eigenvalue 3.

Solution. a) By direct computation, we yield

$$D_Xf=\frac{\partial f}{\partial x}+2x\frac{\partial f}{\partial y}+3y\frac{\partial f}{\partial z}=3y-6x^2+2x\cdot 3x+3y(-1)=0$$
$$D_Zf=\frac{\partial f}{\partial y}+3x\frac{\partial f}{\partial z}=0.$$

b) $D_Yf=x\dfrac{\partial f}{\partial x}+2y\dfrac{\partial f}{\partial y}+3z\dfrac{\partial f}{\partial z}=x(3y-6x^2)+6xy-3z=3f.$

8. Let $x=(x_1,x_2,x_3,x_4)\in\mathbb{R}^4$. Verify that the function defined by

$$f(x)=\frac{x_1-x_3}{x_1-x_4}\,\frac{x_2-x_4}{x_2-x_3}$$

is a solution of the PDEs

$$\sum_{i=1}^{4}\frac{\partial f}{\partial x_i}=0,\quad \sum_{i=1}^{4}x_i\frac{\partial f}{\partial x_i}=0,\quad \sum_{i=1}^{4}x_i^2\frac{\partial f}{\partial x_i}=0.$$

9. For each of the following vector fields, determine the family of the field surfaces.

a) $\vec{V}=2(x^2-y^2)\vec{i}+2xy\vec{j}+xyz\vec{k}$;
b) $\vec{V}=(z+e^x)\vec{i}+(z+e^y)\vec{j}+(z^2-e^{x+y})\vec{k}$;
c) $\vec{V}=xy\vec{i}-y\sqrt{1-y^2}\vec{j}+(z\sqrt{1-y^2}-2axy)\vec{k}$.

Hint. We obtain: a) $z^2=e^y\varphi(y^2e^{x^2/y^2})$, b) $y+ze^{-x}=\varphi(x+ze^{-y})$, c) $yz+ax(y+\sqrt{1-y^2})=\varphi(xe^{\arcsin y})$.

10. Consider the vector field $\vec{V}=xz\vec{i}+z(2x-y)\vec{j}-x^2\vec{k}$. Determine the field surface containing the curve $xy=1$, $xz=1$.

11. Solve the following Cauchy problems:

a) $x\dfrac{\partial z}{\partial y}-y\dfrac{\partial z}{\partial x}=0,\ z(x,0)=\varphi(x)$;
b) $\dfrac{\partial z}{\partial x}+\dfrac{\partial z}{\partial y}=z^2,\ z(x,0)=\psi(x)$;
c) $x\dfrac{\partial f}{\partial x}+y\dfrac{\partial f}{\partial y}+\dfrac{\partial f}{\partial z}=f,\ f(x,y,0)=\chi(x)$.

Hint. a) $z=\varphi(\sqrt{x^2+y^2})$, b) $z=\dfrac{\psi(x-y)}{1-y\psi(x-y)}$, c) $f=\chi(xe^{-z},ye^{-z})e^z$.

12. Show that the harmonic polynomials of two variables are homogeneous polynomials.

Hint. $(x+iy)^n=P_n(x,y)+iQ_n(x,y)$ implies

$$\frac{\partial P_n}{\partial x}=nP_{n-1},\ \frac{\partial P_n}{\partial y}=-nQ_{n-1},\ \frac{\partial Q_n}{\partial x}=nQ_{n-1},\ \frac{\partial Q_n}{\partial y}=nP_{n-1},$$

$$P_{n+1}=xP_n-yQ_n,\quad Q_{n+1}=yP_n+xQ_n,$$

and hence

$$x\frac{\partial P_n}{\partial x}+y\frac{\partial P_n}{\partial y}=nP_n,\quad x\frac{\partial Q_n}{\partial x}+y\frac{\partial Q_n}{\partial y}=nQ_n.$$

13. Let $f:\mathbb{R}^{2n}\to\mathbb{R},\quad (x,y)\to f(x,y)$ be a function of class $\mathcal{C}^\infty$ on $\mathbb{R}^{2n}\backslash\mathbb{R}^n$, where $\mathbb{R}^n=\{(x,y)\in\mathbb{R}^{2n}|y=0\}$. Show that if $f(x,0)=0$ and if the partial function $y \to f(x, y)$ is homogeneous, having the degree of homogeneity $p > 0$, then f is continuous on $\mathbb{R}^n$.

Hint. It has to be shown that $\lim_{(x,y)\to(x,0)} f(x, y) = f(x, 0) = 0$.

14. Let there be given the vector fields $(1, 0, y), (x, 0, z)$, $(x^2, xy - z, xz)$, $(0, 1, x)$, $(0, y, z)$, $(xy - z, y^2, yz)$. Show that the hyperbolic paraboloid $z = xy$ is invariant with respect to the flow generated by each of the given vector fields.

15. The same problem for the vector fields $(1, x)$, $(x, 2y)$, $(x^2 - y, xy)$ and the parabola $x^2 - 2y = 0$.

2.4 Pfaff equations and integral submanifolds

2.4.1 Let $D \subset \mathbb{R}^n$ be an open connected subset, and $X = (X_1, \ldots, X_n)$ be a vector field on D of class $\mathcal{C}^1$ without zeros on D. We call *Pfaff equation* determined by the field X an equation of the form

$$X_1(x)dx_1 + \ldots + X_n(x)dx_n = 0. \tag{65}$$

If M is a submanifold of $\mathbb{R}^n$ and T_xM is the tangent space to M at its point $x \in M$, then we impose the condition that the displacement vector $(dx_1, \ldots, dx_n)$ of the point x belongs to T_xM. Then the orthogonality condition (65) shows that the field X and hence its field lines, are orthogonal to the manifold M. This is why a solution of (65) characterizes a submanifold of $\mathbb{R}^n$, of dimension $k \in \{1, \ldots, n-1\}$.

The submanifolds of $\mathbb{R}^n$ orthogonal to the field lines of X are hence characterized by the Pfaff equation (65) and are also called *integral manifolds of the Pfaff equation* (see the figure).

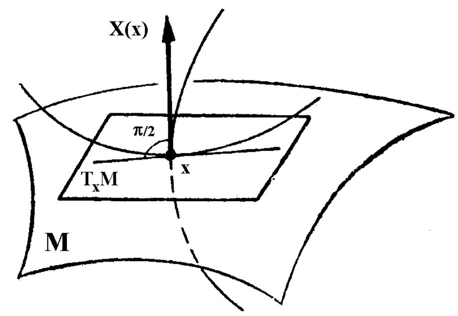

Fig. 4

Remark that for $k = 1$, M is a curve, for $k = 2$, M is a surface and for $k = n - 1$, the solution manifold M is a hypersurface.

The integral manifold M can be described as follows:

- by an immersion $f = (f_1, \ldots, f_n)$, $f(u) = (f_1(u), \ldots, f_n(u))$, $u = (u_1, \ldots, u_k)$, subject to the condition of orthogonality

$$X_1(f(u))\frac{\partial f_1}{\partial u_j} + \ldots + X_n(f(u))\frac{\partial f_n}{\partial u_j} = 0, \quad j = \overline{1,k};$$

- by a submersion $F = (F_1, \ldots, F_{n-k})$,

$$M : \begin{cases} F_1(x) = 0 \\ \ldots \\ F_{n-k}(x) = 0, \end{cases}$$

subject to the condition that (65) is a consequence of the relations

$$\begin{cases} F_1(x) = 0, & \ldots, & F_{n-k}(x) = 0 \\ dF_1 = 0, & \ldots, & dF_{n-k} = 0. \end{cases}$$

2.4.2 The integral curves of the Pfaff equation are characterized by the differential equation

$$X_1(\alpha(t))\frac{dx_1}{dt} + \ldots + X_n(\alpha(t))\frac{dx_n}{dt} = 0.$$

2.4.3 Theorem. *Let $X = (X_1, \ldots, X_n)$ be a vector field which satisfies the condition*

$$\frac{\partial X_i}{\partial x_j} = \frac{\partial X_j}{\partial x_i}, \quad \text{for all } i, j = \overline{1,n}$$

(i.e., is an irrotational vector field*), and $x_0 = (x_{10}, \ldots, x_{n0}) \in \mathbb{R}^n$.*

*a) If D is an $n-$*dimensional interval *(i.e., a Cartesian product of n intervals), then the level hypersurfaces of the function*

$$f : D \to \mathbb{R}, \quad f(x) = \sum_{i=1}^{n} \int_{x_{i0}}^{x_i} X_i(x_{10}, \ldots, x_{i-10}, x_i, \ldots, x_n)dx_i$$

are orthogonal to the field X, and hence the integral hypersurfaces of the Pfaff equation (65).

b) If D is a convex set, then the level hypersurfaces of

$$f(x) = \int_0^1 \langle X_{(x_0 + t(x - x_0))}, x - x_0 \rangle dt$$

are integral hypersurfaces of the Pfaff equation (65).

Remark. Since X is irrotational, it is locally a *potential vector field*, i.e., there exists a function f such that $X = \text{grad}\, f$.

The Pfaff equation (65) is said to be an *exact equation* if there exists a function $f : D \to \mathbb{R}$ of class $\mathcal{C}^2$ such that

$$\frac{\partial f}{\partial x_i}(x) = X_i(x), \quad i = \overline{1, n},$$

or, equivalently,

$$df(x) = \sum_{i=1}^{n} X_i(x) dx_i.$$

If the Pfaff equation (65) is not exact, sometimes there exists a non-constant function $\mu : D \to \mathbb{R} \setminus \{0\}$ of class $\mathcal{C}^1$ such that

$$\mu(x)X_1(x)dx_1 + \ldots + \mu(x)X_n(x)dx_n = 0$$

is an exact equation; in this case the function μ is called an *integrant factor* and it satisfies the system with partial derivatives

$$\frac{\partial(\mu X_i)}{\partial x_j}(x) = \frac{\partial(\mu X_j)}{\partial x_i}(x), \quad i, j = \overline{1, n};\ i \neq j.$$

The existence of an integrant factor implies the existence of a function λ such that $X = \lambda\ \text{grad}\, f$.

The locally exact Pfaff equations and those Pfaff equations which admit a local integrant factor are called *completely integrable equations.* The Pfaff equation (65) is completely integrable if and only if through each point $x_0 \in D$, it passes an integral hypersurface of the equation.

The set of all the integral hypersurfaces of a completely integrable Pfaff equation is called the *general solution* of the equation.

2.4.4 For $n = 2$, the Pfaff equation (65) is completely integrable. For $n \geq 3$, the Pfaff equation (65) is completely integrable if and only if

$$X_i \left(\frac{\partial X_k}{\partial x_j} - \frac{\partial X_j}{\partial x_k} \right) + X_j \left(\frac{\partial X_i}{\partial x_k} - \frac{\partial X_k}{\partial x_i} \right) + X_k \left(\frac{\partial X_j}{\partial x_i} - \frac{\partial X_i}{\partial x_j} \right) = 0, \quad i, j, k = \overline{1, n}.$$

2.4.5 Let X be a vector field of class $\mathcal{C}^1$ on a domain $D \subset \mathbb{R}^n$. We assume that there exist two scalar fields λ of class $\mathcal{C}^1$ and f of class $\mathcal{C}^2$ such that $X = \lambda\ \text{grad}\ f$. If λ and f are functionally independent, then X is called a *bi-scalar* field. If λ and f are functionally dependent, then X is a *potential field.*

X is *locally bi-scalar* or *locally potential vector field* if and only if it admits a family of hypersurfaces of constant level orthogonal to the field lines (see the figure).

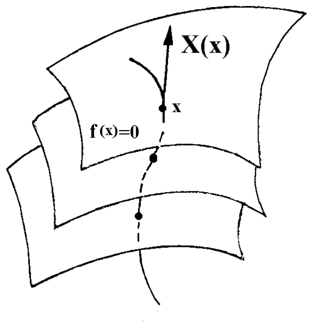

Fig. 5

For $n = 3$, a vector field X is locally bi-scalar or locally potential if and only if

$$\langle X, \text{ curl } X \rangle = 0.$$

2.4.6 If the Pfaff equation (65) is not completely integrable (and hence $n \geq 3$), then the set of all integral manifolds of this equation is called a *non-holonomic hypersurface* defined by the set D and by the vector field X on D.

The previous theory is also extended for the case when X has isolated zeros on D. In this case, the zeros of X on D are called *singular points* of the non-holonomic hypersurface.

Example. The Pfaff equation

$$zdx - ydy = 0$$

describes a non-holonomic surface (family of curves).

2.4.7. Exercises

1. Determine the integral manifolds of the Pfaff equations

a) $(e^{xy} + 1)dx + \dfrac{xy - 1}{y^2} e^{xy} dy = 0, \quad (x, y) \in \mathbb{R}^2 \setminus Ox;$

b) $2xz\ dx + 2yz\ dy + (z^2 - y^2 - x^2)dz = 0, \quad z > 0.$

Solution. We notice that

$$\frac{\partial}{\partial y}(e^{xy} + 1) = \frac{\partial}{\partial x}\left(\frac{xy - 1}{y^2} e^{xy}\right).$$

Therefore, the family of the integral curves is

$$\int_{x_0}^{x} (e^{xy} + 1)dx + \int_{y_0}^{y} \frac{x_0 y - 1}{y^2} e^{x_0 y} dy = c.$$

Alternative. The semiplane $y > 0$ is convex and

$$X(x,y) = \left(e^{xy} + 1, \frac{xy-1}{y^2}e^{xy}\right).$$

Fixing $x_0 = 0,\ y_0 = 1$, we find the family of the integral curves

$$\int_0^1 \left[\left(e^{tx[1+t(y-1)]} + 1\right) + \frac{tx[1+t(y-1)]}{[1+t(y-1)]^2}e^{tx[1+t(y-1)]}(y-1)\right] dt = c_3.$$

↟ Hw. Compute the integrals.

b) Let $\vec{X} = (2xz, 2yz, z^2 - y^2 - x^2)$. It follows

$$\text{curl } \vec{X} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ 2xz & 2yz & z^2-y^2-x^2 \end{vmatrix} = -4y\vec{i} + 4x\vec{j}, \quad \langle \vec{X},\ \text{curl } \vec{X}\rangle = 0.$$

Therefore, the Pfaff equation is completely integrable. By checking, we get the integrant factor $\mu(x,y,z) = \dfrac{1}{z^2}$. Multiplying by this, we find the locally exact Pfaff equation

$$\frac{2x}{z}dx + \frac{2y}{z}dy + \left(1 - \frac{y^2}{z^2} - \frac{x^2}{z^2}\right) dz = 0 \text{ or } d[(x^2+y^2+z^2)z^{-1}] = 0.$$

It follows the family of the integral surfaces

$$x^2 + y^2 + z^2 = c_1 z$$

(family of spheres). The field lines of $\vec{X}$ are circles orthogonal to these spheres.

2. Solve the following systems

a) $\begin{cases} xzdx + z(2x-y)dy - x^2dz = 0 \\ x = y \end{cases}$

b) $\begin{cases} (x^2-yz)dx + (y^2-zx)dy + (z^2-xy)dz = 0 \\ z = xy. \end{cases}$

Solution. a) Let $\vec{X} = xz\vec{i} + z(2x-y)\vec{j} - x^2\vec{k}$. We find

$$\text{curl } \vec{X} = (2x-y)\vec{i} + 3x\vec{j} + 2z\vec{k}, \quad \langle \vec{X},\ \text{curl } \vec{X}\rangle = 6x^2z - 4xyz.$$

Thus, the Pfaff equation defines a non-holonomic surface that contains the plane $xOy : z = 0$.

From $x = y$, we obtain $dx = dy$ and hence $2xzdx - x^2dz = 0$. The solutions are $x = 0,\ y = 0$ (a point) and $x^2z = c_1,\ x = y$ (a family of curves).

b) We denote

$$\vec{X} = (x^2 - yz)\vec{i} + (y^2 - zx)\vec{j} + (z^2 - xy)\vec{k}$$

and we verify curl $\vec{X} = 0$. Therefore, the Pfaff equation is exact. It follows

$$d\left(\frac{x^3 + y^3 + z^3 - 3xyz}{3}\right) = 0 \quad \text{or} \quad x^3 + y^3 + z^3 - 3xyz = c_1.$$

Thus, the solutions of the initial system are the curves

$$x^3 + y^2 + z^3 - 3xyz = c_1, \quad z = xy.$$

3. Show that

$$\vec{X} = z(1 - e^y)\vec{i} + xze^y\vec{j} + x(1 - e^y)\vec{k}$$

is a locally bi-scalar vector field on $\mathbb{R}^3$ and determine the scalar fields λ and f for which $\vec{X} = \lambda\, grad\, f$.

Solution. We obtain curl $\vec{X} = -2xe^y\vec{i} + 2ze^y\vec{k}$ and $\langle \vec{X},\ \text{curl}\ \vec{X}\rangle = 0$. Therefore $\vec{X}$ is a *locally bi-scalar vector field.*

We determine the field lines (orbits) of the curl (rotor), that is the solutions of the system

$$\frac{dx}{-x} = \frac{dy}{0} = \frac{dz}{z}.$$

We get the family of conics $y = c_1,\ xz = c_2$; from $\vec{X} = \alpha \text{grad}\, y + \beta \text{grad}\, xz$, we obtain $\alpha = xze^y$, $\beta = 1 - e^y$ and the completely integrable Pfaff equation

$$z(1 - e^y)dx + zxe^y dy + x(1 - e^y)dz = 0$$

becomes $xze^y dy + (1 - e^y)d(xz) = 0$. We find $\dfrac{e^y - 1}{xz} = c$ (the domain of definition of the function defined by the expression of the left side is not connected).

From the identity $\vec{X} = \lambda$ grad $\frac{e^y-1}{xz}$ it follows $\lambda = x^2z^2$. Obviously, λ and f are functionally independent.

4. Determine the scalar field $\varphi : \mathbb{R}^2 \setminus Ox \to \mathbb{R}$ of class $\mathcal{C}^1$ such that

$$\vec{V} = yz\vec{i} - xz\vec{j} + z\varphi(x, y)\vec{k}$$

is locally bi-scalar.

Solution. Since

$$\text{curl}\ V = \left(z\frac{\partial\varphi}{\partial y} + x\right)\vec{i} - \left(z\frac{\partial\varphi}{\partial x} - y\right)\vec{j} - 2z\vec{k},$$

the problem reduces to finding the general solution of the equation with partial derivatives $\langle \vec{V},\ \text{curl}\ \vec{V}\rangle = 0$ or $y\frac{\partial\varphi}{\partial y} + x\frac{\partial\varphi}{\partial y} - 2\varphi = 0$. The symmetric differential system $\dfrac{dx}{x} = \dfrac{dy}{y} = \dfrac{d\varphi}{2\varphi}$ yields $\dfrac{x}{y} = c_1,\ \dfrac{\varphi}{y^2} = c_2$. Hence $\varphi(x,y) = y^2\psi\left(\dfrac{x}{y}\right)$, where ψ is an arbitrary function of class $\mathcal{C}^1$.

Note. In the particular case $\varphi(x,y) = y^2$, we find $\vec{V} = y^2 z\, grad\, \left(\dfrac{x}{y} + z\right)$.

5. Consider the vector fields:

a) $\vec{V} = xz\vec{i} + z(2x-y)\vec{j} - x^2\vec{k}$;

b) $\vec{V} = x^2(y+z)\vec{i} - y^2(z+x)\vec{j} + z^2(y-x)\vec{k}$;

c) $\vec{V} = y^2z^2\vec{i} + xyz^2\vec{j} + xy^2z\vec{k}$.

Establish if there are level surfaces orthogonal to the field lines. If so, establish the implicit Cartesian equations of these surfaces.

For the non-holonomic surfaces, determine the singular points, the intersections with the axes and with the planes of coordinates.

Solution. a) We compute

$$\text{curl}\ \vec{V} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ xz & z(2x-y) & -x^2 \end{vmatrix} = (y-2x)\vec{i} + 3x\vec{j} + 2z\vec{k}$$

and $\langle \vec{V},\ \text{curl}\ \vec{V}\rangle = 2xz(3x-y)$. Therefore, the Pfaff equation

$$xzdx + z(2x-y)dy - x^2dz = 0$$

depicts a non-holonomic surface (it is not completely integrable). Despite this, it can be noticed that for $z = 0$ both equation and condition $\langle \vec{V},\ \text{curl}\ \vec{V}\rangle = 0$ hold. Therefore, the plane $xOy : z = 0$ is orthogonal to the field lines.

The singular points of the non-holonomic surface are the solutions of the algebraic system $xz = 0,\ z(2x-y) = 0,\ x = 0$. Thus, the axes Oy and Oz consist only of singular points.

The intersection with the axis $Ox :\ y = 0,\ z = 0$, provide $dy = 0,\ dz = 0$ and hence the Pfaff equation

$$xzdx + z(2x-y)dy - x^2dz = 0$$

is fulfilled; so, the axis Ox is an integral curve.

The intersection with the plane yOz :

$$x = 0,\ xzdx + z(2x-y)dy - x^2dz = 0 \Rightarrow x = 0,\ zydy = 0;$$

hence, the intersection consists of the axis Oy, the axis Oz and the straight lines $x = 0, y = c_1$.

b) We find

$$\text{curl } \vec{V} = (y^2 + z^2)\vec{i} + (z^2 + x^2)\vec{j} - (x^2 + y^2)\vec{k}$$

and

$$\langle \vec{V},\ \text{curl } \vec{V} \rangle = x^2y^3 + x^2z^3 + x^3z^2 - x^3y^2 - y^2z^3 - z^2y^3 \neq 0.$$

The Pfaff equation

$$x^2(y + z)dx - y^2(z + x)dy + z^2(y - x)dz = 0$$

defines a non-holonomic surface.

The singular points are the solutions of the algebraic system

$$x^2(y + z) = 0,\ y^2(z + x) = 0,\ z^2(y - x) = 0.$$

The intersection with the axis $Ox:\ y = 0,\ z = 0 \Rightarrow dy = 0,\ dz = 0$ and the Pfaff equation is identically fulfilled; hence, the non-holonomic surface contains the axis Ox.

The intersection with the plane $xOy : z = 0 \Rightarrow dz = 0$ and the Pfaff equation reduces to $xy(xdx - ydy) = 0$; hence, the intersection consists of the axis Ox, the axis Oy and the family of conics of hyperbolic type $z = 0,\ x^2 - y^2 = c_1$.

c) It follows $\text{curl } \vec{V} = y^2z\vec{j} - yz^2\vec{k}$ and $\langle \vec{V},\ \text{curl } \vec{V} \rangle = 0$. In other words, the Pfaff equation $y^2z^2dx + xyz^2dy + xy^2zdz = 0$ is completely integrable (the vector field $\vec{V}$ is *bi-scalar*).

From $yzd(xyz) = 0$, we find the family of the surfaces orthogonal to the lines, $xyz = c$ (family of Ţiţeica surfaces). The planes $xOy : z = 0,\ xOz : y = 0$ consists of singular points (zeros of $\vec{V}$).

We finally notice that $\vec{V} = yz\ \text{grad}\ (xyz)$.

6. Which of the following Pfaff equations define non-holonomic quadrics ?

a) $(x + y)dx + (z - x)dy - zdz = 0$;

b) $(y + z)dx + (x + z)dy + (x + y)dz = 0$.

Determine their intersection with the plane $x = y$.

Solution. If they are not completely integrable, then the Pfaff equations assigned to the linear vector fields on $\mathbb{R}^n$, $n \geq 2$, define non-holonomic hyperquadrics.

a) Let $\vec{X} = (x+y)\vec{i} + (z-x)\vec{j} - z\vec{k}$. We find $\text{curl } X = -\vec{i} - 2\vec{k}$ and $(\vec{X},\ \text{curl } \vec{X}) = -x - y + 2z \neq 0$. Therefore, the Pfaff equation defines a non-holonomic quadric.

The intersection with the plane $x = y$: replacing $dx = dy$ in the Pfaff equation, we find the homogeneous differential equation $\dfrac{dx}{dz} = \dfrac{z}{z + x}$ whose solutions $\varphi(x, z) = c$ can be determined; the intersection consists of the curves $x = y,\ \varphi(x, z) = c$.

b) $\vec{X} = (y+z)\vec{i} + (x+z)\vec{j} + (x+y)\vec{k}$, curl $\vec{X} = 0$. The Pfaff equation

$$(y+z)dx + (x+z)dy + (x+y)dz = 0$$

is completely integrable and has the general solution $xy + yz + zx = c$.

7. Determine the integrant factor that is necessary for the Pfaff equation below to become locally exact and determine its general solution

$$\sum_{i=1}^{n} \left[2x_i + (x_1^2 + \ldots + x_n^2)x_1 \ldots x_{i-1}x_{i+1} \ldots x_n\right] dx_i = 0.$$

Solution. Denoting

$$X_i(x) = 2x_i + (x_1^2 + \ldots + x_n^2)x_1 \ldots x_{i-1}x_{i+1} \ldots x_n,$$

we find

$$\begin{aligned}\frac{\partial X_i}{\partial x_j} &= 2\delta_{ij} + 2x_j x_1 \ldots x_{i-1}x_{i+1} \ldots x_n + \\ &+ (x_1^2 + \ldots x_n^2)x_1 \ldots x_{i-1}x_{i+1} \ldots x_{j-1}x_{j+1} \ldots x_n.\end{aligned}$$

Obviously $\dfrac{\partial X_i}{\partial x_j} \neq \dfrac{\partial X_j}{\partial x_i}$, $i \neq j$, but $\displaystyle\sum_{i=1}^{n} X_i \left(\frac{\partial X_k}{\partial x_j} - \frac{\partial X_j}{\partial x_k}\right) = 0$, that is the Pfaff equation is completely integrable.

We look for a function $\mu : \mathbb{R}^n \to \mathbb{R}$ such that

$$\frac{\partial(\mu X_i)}{\partial x_j} = \frac{\partial(\mu X_j)}{\partial x_i}, \quad i \neq j.$$

This system admits the particular solution $\mu(x) = \exp\,(x_1 \ldots x_n)$. Therefore, the equation

$$e^{x_1 \ldots x_n} \sum_{i=1}^{n} [2x_i + (x_1^2 + \ldots + x_n^2)x_1 \ldots x_{i-1}x_{i+1} \ldots x_n] dx_i = 0$$

is exact and has the general solution

$$(x_1^2 + \ldots + x_n^2)e^{x_1 \ldots x_n} = c.$$

8. We consider the Pfaff equation $xdy - zdz = 0$ which is not completely integrable. Determine the integral curves passing through the point (1,1,1).

Solution. Let $\alpha : (x, y, z) : I \to \mathbb{R}^3$, $x = x(t)$, $y = y(t)$, $z = z(t)$ be an arbitrary curve of class $\mathcal{C}^1$. This satisfies the Pfaff equation if and only if

$$x(t)y'(t) - z(t)z'(t) = 0.$$

From here, we notice that it is sufficient to consider $y(t)$ and $z(t)$ as given functions with $y'(t) \neq 0$ and $x(t) = \dfrac{z(t)z'(t)}{y'(t)}$. Thus,

$$x = \frac{z(t)z'(t)}{y'(t)}, \quad y = y(t), \; z = z(t), \; t \in I$$

depicts the family of the integral curves of the Pfaff equation. Through the point (1,1,1) are passing only the curves for which $y(t_0) = 1, \; z(t_0) = 1, \; y'(t_0) = z'(t_0)$.

Alternative. Let the implicit Cartesian equations

$$f(x, y, z) = 0, \quad g(x, y, z) = 0$$

depict an arbitrary curve of class $\mathcal{C}^1$. This satisfies the Pfaff equation if and only if the homogeneous algebraic system

$$\begin{cases} xdy - zdz = 0 \\ \dfrac{\partial f}{\partial x}dx + \dfrac{\partial f}{\partial y}dy + \dfrac{\partial f}{\partial z}dz = 0 \\ \dfrac{\partial g}{\partial x}dz + \dfrac{\partial g}{\partial y}dy + \dfrac{\partial g}{\partial z}dz = 0 \end{cases}$$

in the unknowns (dx, dy, dz) has nonzero solutions, that is,

$$\begin{vmatrix} 0 & x & -z \\ \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} & \dfrac{\partial f}{\partial z} \\ \dfrac{\partial g}{\partial x} & \dfrac{\partial g}{\partial y} & \dfrac{\partial g}{\partial z} \end{vmatrix} = 0.$$

From this, we notice (for example) that f may be arbitrarily given, while g is determined as a solution of some homogeneous linear equation with partial derivatives of the first order.

9. Determine the general solutions of the following Pfaff equations

$$dx_3 = \frac{x_3 + a}{x_1}dx_1 + \frac{x_3 + a}{x_2}dx_2, \quad dx + \frac{x}{x + y + z}(dy + dz) = 0.$$

Hint. $x_1x_2 = c(x_3 + a)$, family of hyperbolic paraboloids; for the second Pfaff equation, we find $x(x + 2y + 2z) = c$.

10. Solve the following systems

a) $$\begin{cases} (x^2 - y^2 - z^2)dx + 2xydy + 2xzdz = 0 \\ x^2 + y^2 + z^2 = 1 \end{cases}$$

b) $$\begin{cases} (x-y)dx + (x+y)dy + zdz = 0 \\ z = x^2 + y^2. \end{cases}$$

11. Consider the vector fields:

a) $\vec{X}(x,y,z) = x^2yz\vec{i} + xy^2z\vec{j} + xyz^2\vec{k}$,

b) $\vec{X}(\vec{r}) = 2\langle\vec{a},\vec{r}\rangle\vec{a} + (\vec{b}\times\vec{r})\,\vec{b} - \dfrac{\langle\vec{a},\vec{r}\rangle^2 + \langle\vec{b},\vec{r}\rangle^2}{\langle\vec{a},\vec{b}\times\vec{r}\rangle}\vec{a}\times\vec{b}$.

c) $\vec{X}(x,y,z) = ayz\vec{i} + bzx\vec{j} + cxy\vec{k}$ (*Euler field*).

Show that in each case, $\vec{X}$ is a locally bi-scalar field and determine the functions λ and f such that $\vec{X} = \lambda$ grad f.

Hint. a) $\vec{X}(x,y,z) = \dfrac{xyz}{2}$ grad $(x^2+y^2+z^2)$,

b) $\vec{X}(\vec{r}) = \langle\vec{a}\times\vec{b},\vec{r}\rangle$ grad $\dfrac{\langle\vec{a},\vec{r}\rangle^2 + \langle\vec{b},\vec{r}\rangle^2}{\langle\vec{a}\times\vec{b},\vec{r}\rangle}$.

12. Consider the vector field

$$\vec{V} = \text{ grad } \varphi(r) + \varphi(r) \text{ grad } \psi(r),$$

where $r = \sqrt{x^2+y^2+z^2}$, and φ and ψ are functions of class $\mathcal{C}^\infty$.

a) Determine the field lines of $\vec{V}$.

b) Show that $\vec{V}$ is a locally potential field and determine the family of the surfaces orthogonal to the field lines.

Hint. a) The solutions form a family of straight lines

$$y = c_1x,\ z = c_2x,\ (x,y,z) \in \mathbb{R}^3 \setminus \{(0,0,0)\}.$$

b) $\vec{V} = e^{-\psi}$ grad φe^{ψ} and $e^{-\psi}, \varphi e^{\psi}$ are functionally dependent; the spheres $r^2 = c$ are orthogonal to the field lines.

13. We consider the following vector fields:

a) $\vec{V} = xz\vec{i} + yz\vec{j} - (x^2+y^2)\vec{k}$;

b) $\vec{V} = (y-z)\vec{i} + (z-x)\vec{j} + (x-y)\vec{k}$;

c) $\vec{V} = x(y-z)\vec{i} + y(z-x)\vec{j} + z(x-y)\vec{k}$.

Establish if there exist any families of level surfaces orthogonal to the field lines. In case of existence, find the implicit Cartesian equations of these families of surfaces. For the non-holonomic surfaces, determine the singular points, the intersections with the axes and the planes of coordinates.

Hint. a) $\vec{V} = \dfrac{z^3}{2}$ grad $\dfrac{x^2+y^2}{z^2}$;

b) $\vec{V} = (z-x)^2 \operatorname{grad} \dfrac{x-y}{x-z}$, a *Killing vector field*;

c) $(\vec{V},\ \operatorname{curl}\ \vec{V}) = x(z^2-y^2) + y(x^2-z^2) + z(y^2-x^2) \neq 0$.

14. Which of the following Pfaff equations define *non-holonomic quadrics* ?

a) $ydx + zdy - (6x + 11y + 6z)dz = 0$;

b) $(5x - 2y + 4z)dx + (4x - y)dy + 2xdz = 0$;

c) $ydx + (x+z)dy + (y+z)dz = 0$.

In the affirmative cases, determine the integral curves and the intersections of the non-holonomic quadrics with planes passing through the axes of coordinates.

Hint. a)-b) non-holonomic quadrics, c) $z^2 + 2xy + 2yz = c$.

15. Study if the following Pfaff equations admit or not integrant factors:

a) $\dfrac{x_1}{x_2}dx_1 + \dfrac{x_2}{x_3}dx_2 + \ldots + \dfrac{x_{n-1}}{x_n}dx_{n-1} + \dfrac{x_n}{x_1}dx_n = 0$;

b) $x_1 \sin x_n dx_1 + x_2 \sin x_{n-1} dx_2 + \ldots + x_n \sin x_1 dx_n = 0$.

Chapter 3

Hilbert Spaces

3.1 Euclidean and Hilbert spaces

Let $\mathbf{V}$ be a complex vector space. A *scalar product* on $\mathbf{V}$ is a mapping that associates to each ordered pair of vectors x, y a scalar (complex number) denoted $\langle x, y\rangle$ that satisfies

$$\begin{array}{ll} symmetry: & \langle x,y\rangle = \overline{\langle y,x\rangle}, \\ homogeneity: & \langle \alpha x,y\rangle = \alpha\langle x,y\rangle,\ \alpha\in\mathbb{C}, \\ additivity: & \langle x+y,z\rangle = \langle x,z\rangle + \langle y,z\rangle, \\ positive\quad definiteness: & \langle x,x\rangle > 0,\ when\ x\neq 0, \end{array}$$

where the bar stands for the complex conjugate. From these items, we obtain

$$\begin{array}{l} \langle x, y+z\rangle = \langle x,y\rangle + \langle x,z\rangle \\ \langle x,\alpha y\rangle = \bar{\alpha}\langle x,y\rangle \\ \langle 0,x\rangle = 0. \end{array}$$

Usually, the existence of a scalar product is assumed. From a scalar product we can build infinitely many other ones.

A complex vector space together with a scalar product on it is called a *Euclidean space*. The scalar product induces a *Euclidean norm* $||x||$, and a *Euclidean distance*

$$d(x,y) = ||y-x||.$$

The most important relation on a Euclidean space is the *Cauchy-Schwartz inequality*

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

The scalar product is used to define the orthogonality. The norm or the distance are used for topological reasons (to introduce open sets, convergence, continuity, etc). Particularly, the scalar product is a continuous function.

3.1.1 Definition. A Euclidean space is called *complete* if any Cauchy sequence is convergent. A complete Euclidean space is called *Hilbert space* and is denoted by H.

Examples. 1. Let $\mathbf{V}_3$ be the real vector space of free vectors in $\mathbf{E}_3$. The function

$$\langle x, y\rangle = \begin{cases} ||x||\;||y||\cos\theta, & \text{for } x \neq 0 \text{ and } y \neq 0 \\ 0, & \text{for } x = 0 \text{ or } y = 0, \end{cases}$$

where $\theta = (\widehat{x, y})$, is a *scalar product on* $\mathbf{V}_3$. Of course, $\mathbf{V}_3$ is a Hilbert space.

2. Let $\mathbb{R}^n$ be the real vector space of $n-$uples and $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ be two elements of $\mathbb{R}^n$. The function

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i\, y_i$$

is a *scalar product on* $\mathbb{R}^n$. Analogously,

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i \overline{y_i}$$

is a *scalar product on* $\mathbb{C}^n$. Note that $\mathbb{R}^n$ and $\mathbb{C}^n$ are Hilbert spaces.

3. Let l_2 be the complex vector space of all complex sequences $x = \{x_1, \ldots, x_n, \ldots\}$ with the property $\sum\limits_{i=1}^{\infty} |x_i|^2 < \infty$. The function defined by

$$\langle x, y\rangle = \sum_{i=1}^{\infty} x_i \overline{y_i}$$

is a *scalar product on* l_2. Note that l_2 is a Hilbert space.

4. Let $\mathbf{V} = \mathcal{C}^0[a, b]$ be real or the complex vector space of all continuous functions on $[a, b]$ with complex values. Then

$$\langle x, y\rangle = \int_a^b x(t)\bar{y}(t)dt$$

is a *scalar product on* $\mathbf{V}$. It can be proved that $\mathbf{V} = \mathcal{C}^0[a, b]$ *is not a Hilbert space.*

5. Let $\mathbf{V} = L_2[a, b]$ be the vector space of all functions x on $[a, b]$ with real or complex values such that $|x|^2$ is integrable. Then

$$\langle x, y\rangle = \int_a^b x(t)\bar{y}(t)dt$$

is a *scalar product on* $\mathbf{V}$. It can be proved that $\mathbf{V} = L_2[a, b]$ is a Hilbert space, and $L_2[a, b] = \overline{\mathcal{C}^0[a, b]}$, where the bar means closedness.

6. Let $D \subset \mathbb{R}^n$ be a compact set and $\mathbf{V} = \mathcal{C}^1(D)$ the real vector space of $\mathcal{C}^1$ real functions on D. If $f \in \mathbf{V}$, we denote $\nabla f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right)$. Then V can be endowed with the scalar product

$$\langle f, g\rangle = \int_D (fg + \langle \nabla f, \nabla g\rangle)dv.$$

Let $\mathbf{V}$ be a Euclidean space. Two vectors $x, y \in \mathbf{V}$ are said to be orthogonal if $\langle x, y\rangle = 0$. The orthogonality is denoted by $x \perp y$. If $A, B \subset \mathbf{V}$, then $A \perp B$ means $x \perp y$ for all $x \in A, y \in B$.

3.1.2 Pythagorean theorem. *If $x \perp y$, then*

$$||x+y||^2 = ||x||^2 + ||y||^2.$$

Proof. Since $\langle x, y\rangle = 0$, we find

$$||x+y||^2 = \langle x+y, x+y\rangle = \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle = ||x||^2 + ||y||^2. \qquad \square$$

Let M be a proper subset of $\mathbf{V}$. The *orthogonal complement of M* is defined by

$$M^\perp = \{x \in \mathbf{V} \mid \langle x, y\rangle = 0, \text{ for all } y \in M\}.$$

Of course, $x \perp M$ is equivalent to $x \in M^\perp$.

3.1.3 Theorem. *Let $\mathbf{V}$ be a Euclidean vector space and M be a proper subset of $\mathbf{V}$. Then $M^\perp$ is a closed vector subspace of $\mathbf{V}$.*

Proof. Obviously $M^\perp$ is a vector subspace. Assume $x_n \in M^\perp$ with $x_n \to x_0$. Let us show that $x_0 \in M^\perp$. Since the scalar product is a continuous function, we have

$$\langle x_0, y\rangle = \lim_{n\to\infty} \langle x_n, y\rangle = 0, \quad \text{for all } y \in M. \qquad \square$$

3.1.4 Corollary. *If $\mathbf{V}$ is complete, then $M^\perp$ is also complete.*

Now we recall the algebraic concept of *projection*: an endomorphism $\mathcal{P} : \mathbf{V} \to \mathbf{V}$ which satisfies $\mathcal{P}^2 = \mathcal{P}$. A projection is said to be *orthogonal* if its range and null space are orthogonal.

3.1.5 Theorem. *An orthogonal projection is continuous.*

Proof. Always $x = r + n,\ r \in Im\mathcal{P}, n \in Ker\mathcal{P}$. Of course, $r \perp n$ because $\mathcal{P}$ is orthogonal. It follows

$$||x||^2 = ||r||^2 + ||n||^2.$$

Consequently

$$||\mathcal{P}x||^2 = ||r||^2 \leq ||x||^2,$$

and hence $\mathcal{P}$ is continuous (note that a linear function is continuous iff it is continuous at a point). $\square$

3.1.6. Exercises

1. Show that $\langle \bar{a}, \bar{b}\rangle = ||\bar{a}||\ ||\bar{b}|| \cos\theta$, where θ is the measure of the angle between $\bar{a}$ and $\bar{b}$, defines a scalar product on the real vector space of the free vectors.♣ Hw.

2. Show that the function defined by

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i, \quad \text{for all } x = (x_1, \ldots, x_n),\ y = (y_1, \ldots, y_n) \in \mathbb{R}^n$$

is a scalar product on $\mathbb{R}^n$.♣ Hw.

3. Show that the function defined by

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i \overline{y_i}, \quad \text{for all } x = (x_1, \ldots, x_n),\ y = (y_1, \ldots, y_n) \in \mathbb{C}^n$$

is a scalar product on $\mathbb{C}^n$.

Solution. We verify the axioms of the scalar product

$$\begin{aligned}
\langle x, y\rangle &= \sum_{i=1}^{n} x_i \bar{y}_i = \sum_{i=1}^{n} \overline{\bar{x}_i y_i} = \overline{\sum_{i=1}^{n} y_i \bar{x}_i} = \overline{\langle y, x\rangle}, \\
\langle x + y, z\rangle &= \sum_{i=1}^{n} (x_i + y_i)\bar{z}_i = \sum_{i=1}^{n} x_i \bar{z}_i + \sum_{i=1}^{n} y_i \bar{z}_i = \langle x, z\rangle + \langle y, z\rangle, \\
\alpha\langle x, y\rangle &= \alpha \sum_{i=1}^{n} x_i \bar{y}_i = \sum_{i=1}^{n} (\alpha x_i)\bar{y}_i = \langle \alpha x, y\rangle, \\
\langle x, x\rangle &= \sum_{i=1}^{n} x_i \bar{x}_i = \sum_{i=1}^{n} |x_i|^2 \geq 0, \text{and the equality holds} \\
&\quad \text{iff } \ x_i = 0,\ i = \overline{1, n}, \text{ that is, for } x = 0.
\end{aligned}$$

4. Consider the real vector space $V = \{f : [a, b] \to \mathbb{R} \mid f \text{ continuous}\}$. Show that V is a Euclidean space with respect to the application defined by

$$\langle f, g\rangle = \int_a^b f(t)g(t)dt, \quad \text{for all } f, g \in V.$$

Solution. The problem requires to show that $\langle f, g\rangle$ is a scalar product. The continuity ensures the existence of the integral, and if f, g, h are continuous functions with real values and r is a real number, then we have

$$\begin{aligned}
\langle f, g\rangle &= \int_a^b f(t)g(t)dt = \int_a^b g(t)f(t)dt = \langle g, f\rangle, \\
\langle f, g + h\rangle &= \int_a^b f(t)(g(t) + h(t))dt = \int_a^b f(t)g(t)dt + \int_a^b f(t)h(t)dt = \\
&= \langle f, g\rangle + \langle f, h\rangle, \\
r\langle f, g\rangle &= r\int_a^b f(t)g(t)dt = \int_a^b (rf)(t)g(t)dt = \langle rf, g\rangle.
\end{aligned}$$

For a continuous function $f \neq 0$, there exists $t_0 \in (a, b)$ such that $f(t_0) \neq 0$, i.e., $f(t_0) > 0$ or $f(t_0) < 0$. By continuity, there exists an open interval I such that $t_0 \in I \subset [a, b]$ and $f^2(t) > 0$, for any $t \in I$. Let $[c, d] \subset I,\ c < d$ and $m = \min\limits_{t\in[c,d]} f^2(t) > 0$. Then the positivity axiom follows

$$0 < m(d-c) \leq \int_c^d f^2(t)dt = \int_a^b f^2(t)dt = \langle f, f\rangle.$$

5. Consider the complex vector space $V = \{f : [a, b] \to \mathbb{C} \mid f \text{ continuous}\}$. Show that V is a Euclidean space with respect to the application defined by

$$\langle f, g\rangle = \int_a^b f(t)\overline{g(t)}dt, \quad \text{for all } f, g \in V.$$

6. Let V be the real vector space of the real sequences $\{x_n\}$ for which the series $\sum\limits_{n=0}^{\infty} x_n^2$ is convergent. Let $x = \{x_n\},\ y = \{y_n\}$ be two elements of V.

a) Show that the series $\sum\limits_{n=0}^{\infty} x_n y_n$ is absolutely convergent.

b) Verify that the function defined by $\langle x, y\rangle = \sum\limits_{n=0}^{\infty} x_n y_n$ is a scalar product on V.

c) Show that the sequence with the general term $x_n = e^{-n}$ belongs to the exterior of the sphere of centre the origin, of radius $\dfrac{3}{2\sqrt{2}}$ and to the open ball of centre the origin and radius $\dfrac{2}{\sqrt{3}}$.

d) Write the ball of centre $x_0 = \left\{\dfrac{1}{n+1}\right\}$, radius 1 and give examples of elements of V contained in this ball.

e) Calculate $\langle x, y\rangle$ for

$$x_n = \frac{2n-1}{2^{n/2}},\ y_n = \frac{1}{2^{n/2}},\ n \geq 1.$$

f) Let $x_n = 2^{1-n},\ y_n = 3^{1-n},\ n \geq 1$. Determine the "angle" between x and y.

Solution. a) Taking into account that for any real numbers a and b we have $2|ab| \leq a^2 + b^2$, it follows

$$2\sum_{i=0}^{\infty} |x_i y_i| \leq \sum_{i=0}^{\infty} x_i^2 + \sum_{i=0}^{\infty} y_i^2.$$

Using the comparison criterion from the positive terms series, the series $\sum_{i=0}^{\infty} |x_i y_i|$ is convergent, and then the series $\sum_{i=0}^{\infty} x_i y_i$ absolutely converges.

b) We have to verify only the positive definiteness of the associate quadratic form. Since $\langle x, x\rangle = \sum_{i=0}^{\infty} x_n^2 \geq 0$, for any $x \in V$, it is sufficient to show that $\langle x, x\rangle = 0 \Rightarrow x = 0$. Indeed, the relations

$$0 \leq x_0^2 \leq x_0^2 + x_1^2 \leq \ldots \leq x_0^2 + x_1^2 + \ldots x_n^2 \leq \ldots \leq \sum_{n=0}^{\infty} x_n^2 = 0$$

imply $0 = x_0 = x_1 = \ldots = x_n = \ldots$ (we used the fact that the limit of a convergent increasing sequence is the supremum of the set of values of the sequence).

c) Since $2 < e < 3$, we find $\frac{1}{3^n} \leq \frac{1}{e^n} \leq \frac{1}{2^n}$, $n \in \mathbf{N}$, and then

$$\sum_{n=0}^{\infty} \frac{1}{3^{2n}} < \sum_{n=0}^{\infty} \frac{1}{e^{2n}} < \sum_{n=0}^{\infty} \frac{1}{2^{2n}} \qquad \Leftrightarrow \qquad \frac{9}{8} < \sum_{n=0}^{\infty} \frac{1}{e^{2n}} < \frac{4}{3},$$

whence $\frac{9}{8} < ||e^{-n}||^2 < \frac{4}{3}$.

d)♣ Hw. e) We have subsequently

$$1 + 2x + \ldots + nx^{n-1} + \ldots = \frac{1}{(1-x)^2},\ |x| < 1 \Rightarrow \langle x, y\rangle = \sum_{n=1}^{\infty} \frac{2n-1}{2^n} = 3.$$

f) The angle is $\theta = \arccos \frac{2\sqrt{6}}{5}$.

7. On the real vector space

$$\mathcal{C}^0[1, e] = \{f : [1, e] \to \mathbb{R} \mid f \text{ continuous}\}$$

we define the scalar product

$$\langle f, g\rangle = \int_1^e (\ln x) f(x) g(x) dx, \quad \text{for all } f, g \in \mathcal{C}^0[1, e].$$

a) Calculate the angle between $f(x) = x$ and $g(x) = \sqrt{x}$.

b) Show that the functions f which satisfy the property $f^2(x) < \frac{2}{x}$ belong to the ball of the centre at origin and of radius one.

c) Calculate $||f||$ for $f(x) = \sqrt{x}$. Write the ball having as centre the function $f_0(x) = 1$, for all $x \in [1, e]$, of radius 1, and give examples of functions in this ball.

Solution. a) Integrating by parts and considering

$$u = \ln x,\ dv = x^{3/2}dx \quad \Rightarrow \quad du = \frac{1}{x}dx,\ v = \frac{2}{5}x^{5/2},$$

we have

$$\langle f, g\rangle = \int_1^e x^{3/2}\ln x = \frac{2}{5}x^{5/2}\ln x\Big|_1^e - \frac{2}{5}\int_1^e x^{3/2}dx = \frac{4+6e^{5/2}}{25},$$

and also

$$||f||^2 = \int_1^e x^2 lnx dx = \frac{x^3}{3}\ln x\Big|_1^e - \int_1^e \frac{x^3}{3}\frac{1}{x}dx = \frac{1+2e^3}{9}$$

$||g||^2 = \int_1^e x\ln x dx = \frac{1+e^2}{4}$,

whence we infer

$$\cos\theta = \frac{\langle f, g\rangle}{||f||\,||g||} = \frac{12}{25}\frac{2+3e^{5/2}}{\sqrt{(1+e^2)(1+2e^3)}}.$$

b) We remark that since $f^2(x) < \frac{2}{x}$, we get

$$||f||^2 = \int_1^e (\ln x)f^2(x)dx < \int_1^e \ln x\cdot\frac{2}{x}dx = \ln^2 x\Big|_1^e = 1.$$

8. Specify the reasons for which the following functions are *not* scalar products:

a) $\varphi : \mathbb{C}^n\times\mathbb{C}^n \to \mathbb{C}, \quad \varphi(x, y) = \sum_{i=1}^n x_i y_i$;

b) $\varphi : \mathbb{R}^n\times\mathbb{R}^n \to \mathbb{R}, \quad \varphi(x, y) = \sum_{i=1}^n |x_i y_i|$;

c) $\varphi : \mathcal{C}^0[0, 1]\times\mathcal{C}^0[0, 1] \to \mathbb{R}, \quad \varphi(f, g) = f(0)g(0)$. ↟ Hw.

9. In the canonical Euclidean space $\mathbb{R}^3$, consider the vector $v = (2, 1, -1)$ and the vector subspace $P : x - y + 2z = 0$. Find the orthogonal projection w of v on P and the vector $w^\perp$. ↟ Hw.

10. Let $P_n = \{p \in \mathbb{R}[X] \mid \deg p \le n\}$ and the mapping

$$\varphi : P_n \to P_n, \quad p \xrightarrow{\varphi} \varphi(p) = X^n\int_0^1 tp(t)dt.$$

Give an example of scalar product on P_n relative to which φ is a symmetric endomorphism.

Hint. Check that the scalar product $\langle p, q\rangle = \sum_{i=0}^n a_i b_i$, where

$$p = \sum_{i=0}^n a_i X^i,\ q = \sum_{i=0}^n b_i X^i \in P_n$$

satisfies the required property.

3.2 Orthonormal basis for a Hilbert space

Till now a basis of a vector space $\mathbf{V}$ was a purely algebraic concept, introduced using finite linear combinations and the idea of linear independence. Here we want to attach a meaning to infinite linear combinations of the form (series)

$$\sum_{i=1}^{\infty} \alpha_i x_i,\ x_i \in \mathbf{V},\ \alpha_i \in \mathbb{C}.$$

It will appear type of bases which involve topological as well as algebraic structure.

3.2.1 Definition. Let $\mathbf{V}$ be a Euclidean space.

1) The set $\{x_i\} \subset \mathbf{V}$ is called *orthogonal* if $x_i \perp x_j$, whenever $i \neq j$.

2) The set $\{x_i\} \subset \mathbf{V}$ is called *orthonormal* if

$$\langle x_i, x_j\rangle = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j, \end{cases}$$

where δ_{ij} is the Kronecker symbol.

The index i may range over a finite, countable infinite (when the set is a sequence) or uncountable index set.

3.2.2 Theorem. *If $\{x_i\}$ is an orthonormal set in $\mathbf{V}$, then the set $\{x_i\}$ is linearly independent.*

Proof. We start with

$$\alpha_1 x_1 + \ldots + \alpha_n x_n = 0.$$

Taking the scalar product with x_i, we obtain

$$\begin{aligned} 0 = \langle 0, x_i\rangle &= \alpha_1\langle x_1, x_i\rangle + \ldots + \alpha_i\langle x_i, x_i\rangle + \ldots + \alpha_n\langle x_n, x_i\rangle = \\ &= \alpha_i\langle x_i, x_i\rangle = \alpha_i,\quad i = \overline{1,n}, \end{aligned}$$

and hence the set is linearly independent. □

3.2.3 Definition. An orthonormal set $\{x_i\} \subset \mathbf{V}$ is called *maximal* if there is no unit vector $x_0 \in \mathbf{V}$ such that $\{x_i\} \cup \{x_0\}$ is an orthonormal set.

In other words, an orthonormal set $\{x_i\}$ is maximal if $x \perp x_i$, for all i, implies that $x = 0$.

3.2.4 Theorem. *Let $\{x_i\}$ be an orthonormal set. Then there is a maximal orthonormal set B which contains $\{x_i\}$.*

The proof of this theorem is a straightforward application of Zorn lemma.

3.2.5 Definition. A maximal orthonormal set B in a Hilbert space H is called an *orthonormal basis* of H.

Examples of orthonormal bases

1) $I = [0,1],\ H =$ the complex space $L_2(I)$ with the scalar product

$$\langle f, g\rangle = \int_0^1 f(t)\ \bar{g}(t)dt.$$

We claim that

$$\Phi_n(t) = e^{2\pi int},\ n = 0, \pm 1, \pm 2, \ldots$$

is an orthonormal set, i.e.,

$$\langle \Phi_n, \Phi_m\rangle = 0,\ n \neq m \text{ and } ||\Phi_n||^2 = 1.$$

Next is possible to show that $\{\Phi_n\}$ is a maximal orthonormal set, i.e.,

$$f \perp \Phi_n \Rightarrow f = 0.$$

Note. $L_2[a,b] = \overline{\mathcal{C}^0[a,b]}$, where the bar means closedness.

2) The *Laguerre functions* form an orthonormal set for $L_2[0,\infty)$. These are defined by

$$\Phi_n(t) = \frac{1}{n!}e^{-t/2}L_n(t),\ n = 0, 1, \ldots$$

where $L_n(t)$ are the *Laguerre polynomials*

$$L_n(t) = e^t \frac{d^n}{dt^n}(t^n e^{-t}).$$

3) The *Hermite functions*

$$\Phi_n(t) = \frac{e^{-t^2/2}}{[2^n n!\sqrt{\pi}]^{1/2}} H_n(t),\ n = 0, 1, \ldots$$

form an orthonormal basis for $L_2(-\infty,\infty)$, where

$$H_n(t) = (-1)^n e^{t^2} \frac{d^n}{dt^n}(e^{-t^2})$$

are the *Hermite polynomials*.

4) The *Legendre functions*

$$\Phi_n(t) = \left(\frac{2n+1}{2}\right)^{1/2} P_n(t),\ n = 0, 1, \ldots$$

form an orthonormal set for $L_2[-1,1]$, where $P_n(t)$ are the *Legendre polynomials*

$$P_n(t) = \frac{1}{2^n n!}\frac{d^n}{dt^n}(t^2-1)^n.$$

Note. Given any linearly independent sequence $\{x_n\}_{n\in\mathbb{N}}$ in a Euclidean space, it is always possible to construct an orthonormal set from it. The construction is known as *the Gram-Schmidt orthogonalization process.*

3.2.6. Exercises

1. Replace the following bases with orthogonal ones:

a) $\{v_1 = (1,1,0),\ v_2 = (1,0,1),\ v_3 = (0,1,1)\}$ in $\mathbb{R}^3$;

b) $\{1, x, \ldots, x^n, \ldots\}$ in the real vector space of all the real polynomial functions defined on $[-1,1]$.

Solution. We use the Gram-Schmidt orthogonalization method.

a) The canonical scalar product in $\mathbb{R}^3$ is $\langle x, y\rangle = x_1y_1 + x_2y_2 + x_3y_3$. We put $w_1 = v_1$. Considering now $w_2 = v_2 + \alpha v_1 = (1,0,1) + (\alpha,\alpha,0) = (1+\alpha,\alpha,1)$, the relation $\langle w_2, w_1\rangle = 0$ is equivalent to $1 + 2\alpha = 0$, whence $\alpha = -\frac{1}{2}$; it follows $w_2 = \left(\frac{1}{2}, -\frac{1}{2}, 1\right)$.

We search similarly for w_3 of the form

$$\begin{aligned} w_3 &= v_3 + \beta w_2 + \gamma w_1 = (0,1,1) + \left(\frac{\beta}{2}, -\frac{\beta}{2}, \beta\right) + (\gamma,\gamma,0) = \\ &= \left(\frac{\beta}{2} + \gamma, 1 - \frac{\beta}{2} + \gamma, 1 + \beta\right), \end{aligned}$$

satisfying the orthogonality conditions $\langle w_3, w_1\rangle = 0$ and $\langle w_3, w_2\rangle = 0$, which rewrite

$$\begin{cases} 1 + 2\gamma = 0 \\ 3\beta + 1 = 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} \gamma = -\frac{1}{2} \\ \beta = -\frac{1}{3}. \end{cases}$$

Hence $w_3 = \left(-\frac{2}{3}, \frac{2}{3}, \frac{2}{3}\right)$ and $\{w_1, w_2, w_3\}$ is an orthogonal basis. Remark that $\left\{\frac{w_1}{||w_1||}, \frac{w_2}{||w_2||}, \frac{w_3}{||w_3||}\right\}$ is an orthonormal basis.

b) We use the canonical scalar product $\langle p, q\rangle = \int_{-1}^{1} p(x)q(x)dx$ and the orthogonalization formulas. We denote $p_i(x) = x^i,\ i \in \mathbb{N}$. Set first $q_0(x) = p_0(x) = 1$. Computing $\langle q_0, q_0\rangle = \int_{-1}^{1} dx = 2$, and $\langle p_1, q_0\rangle = \int_{-1}^{1} xdx = 0$ we find

$$q_1(x) = p_1(x) - \frac{\langle p_1, q_0\rangle}{\langle q_0, q_0\rangle} q_0(x) = p_1(x) = x.$$

By analogy, the relations

$$\begin{cases} \langle p_2, q_0 \rangle = \int_{-1}^{1} x^2 dx = \frac{2}{3} \\ \langle p_2, q_1 \rangle = \int_{-1}^{1} x^3 dx = 0 \\ \langle q_1, q_1 \rangle = \int_{-1}^{1} x^2 dx = \frac{2}{3} \end{cases}$$

imply

$$q_2(x) = p_2(x) - \frac{\langle p_2, q_0 \rangle}{\langle q_0, q_0 \rangle} q_0(x) - \frac{\langle p_2, q_1 \rangle}{\langle q_1, q_1 \rangle} q_1(x) = x^2 - \frac{1}{3}.$$

Similarly we find

$$q_3(x) = x^3 - \frac{3}{5}x, \qquad q_4(x) = x^4 - \frac{6}{7}x^2 + \frac{3}{35}, \qquad q_5(x) = x^5 - \frac{10}{9}x^3 + \frac{5}{21}x.$$

By induction, it may be proved that the required polynomials q_n can be computed by the formula

$$q_n(x) = \frac{n!}{(2n)!} \frac{d^n}{dx^n} (x^2 - 1)^n, \quad n \in \mathbb{N}.$$

These polynomials are called *the Legendre polynomials.*

2. Show that equality holds in the *Cauchy-Schwartz inequality* $|\langle x, y \rangle| \leq ||x|| \, ||y||$ iff the set $\{x, y\}$ is linearly dependent.

3. Let $V = L_2[0, T]$ be endowed with the scalar product

$$\langle x, y \rangle = \int_0^T x(t)\overline{y(t)} dt, \quad \text{for all } x, y \in V$$

and let F be the filter mapping of V into itself defined by

$$(Fx)(t) = \int_0^t e^{-(t-\tau)} x(\tau) d\tau, \quad t \in [0, T].$$

Find an input $x \in V$ such that $\langle x, x \rangle = 1$ and $(Fx)(T)$ is maximal.

Solution. Let $y(t) = e^t$, then $(Fx)(T) = e^{-T}\langle y, x \rangle$. The *Cauchy-Schwartz inequality* implies

$$|(Fx)(T)| \leq e^{-T} ||x|| \cdot ||y||,$$

with equality when $x = cy$ (x and y are collinear). The solution is

$$x(t) = \sqrt{\frac{2}{e^{2T} - 1}} e^t, \quad t \in [0, T],$$

and

$$(Fx)(T) = \sqrt{\frac{1}{2}(1 - e^{-2T})}.$$

4. Describe all possible scalar products on the complex vector space $\mathbb{C}$.

Show that the real vector space $\mathbb{R}^2$ has a scalar product which is not derived from a complex scalar product.

5. Let $\langle x, y \rangle$ be a scalar product on the vector space V.

1) Show that the mappings $y \to \langle x, y \rangle$, $x \to \langle x, y \rangle$ are continuous.

2) Discuss the relations

$$\left\langle x, \sum_{n=1}^{\infty} y_n \right\rangle = \sum_{n=1}^{\infty} \langle x, y_n \rangle, \qquad \left\langle \sum_{n=1}^{\infty} x_n, y \right\rangle = \sum_{n=1}^{\infty} \langle x_n, y \rangle.$$

6. Let x, y be vectors in a Euclidean vector space V and assume that

$$||\lambda x + (1-\lambda)y|| = ||x||, \quad \text{for all } \lambda \in [0,1].$$

Show that $x = y$.

7. Let $\{x, y\}$ be a linearly independent set in a complex Euclidean vector space V. Define $f : \mathbb{C} \to \mathbb{R}$,

$$f(\alpha) = ||x - \alpha y||, \ \forall \alpha \in \mathbb{C}.$$

Where does f take on its minimum value ?

8. Let y be a fixed vector in a Hilbert space H. Describe the operator

$$L : H \to H, \quad L(x) = \langle x, y \rangle y, \quad \text{for all } x \in H.$$

9. Let V denote the vector subspace of $L_2[0, 2\pi]$ made up of all trigonometric polynomials of the form

$$x(t) = \sum_{k=-n}^{n} a_k e^{ikt}.$$

Let M be the subspace defined by

$$M = \left\{ x \in V \ \middle| \ \int_0^{2\pi} t x(t) dt = 0 \right\}.$$

Let x_0 in V with $x_0 \notin M$. Show that there is no vector y_0 in M with

$$||x_0 - y_0|| = \inf \{||x_0 - y||, \ y \in M\}$$

10. Let $z = x + y$, where $x \perp y$. Show that $\langle z, x \rangle$ is real.

11. Let $\{y, z_1, z_2, \ldots\}$ be an orthonormal set in a Hilbert space H. We construct the linearly independent set $\{x_n\}_{n \in \mathbb{N}^*}$ by

$$x_n = \left(\cos \frac{1}{n}\right) y + \left(\sin \frac{1}{n}\right) z_n, \ \ n \in \mathbb{N}^*.$$

Denote $M = L(\{x_n\}_{n\in\mathbb{N}^*})$ and $\overline{M}$ the closure of M.

1) Show that $y \in \overline{M}$.

2) Show that y cannot be expressed in the form

$$y = \alpha_1 x_1 + \alpha_2 x_2 + \ldots$$

Solution. 1) $y = \lim\limits_{n\to\infty} x_n \in \overline{M}$.

2) Suppose, $y = \alpha_1 x_1 + \alpha_2 x_2 + \ldots$. Then

$$y = \sum_{n=1}^{\infty} \alpha_n \left(\left(\cos\frac{1}{n}\right) y + \left(\sin\frac{1}{n}\right) z_n \right),$$

and since $\{y, z_1, z_2, \ldots\}$ is an orthonormal set, one has

$$y = \left(\sum_{n=1}^{\infty} \alpha_n \cos\frac{1}{n}\right) y + \sum_{n=1}^{\infty} \alpha_n \left(\sin\frac{1}{n}\right) z_n.$$

Since the two terms of the right-hand side are orthogonal to one another it follows

$$\sum_{n=1}^{\infty} \alpha_n \cos\frac{1}{n} = 1, \quad \sum_{n=0}^{\infty} \alpha_n \left(\sin\frac{1}{n}\right) z_n = 0.$$

But the theory implies

$$\sum_{n=1}^{\infty} |\alpha_n|^2 \sin^2\frac{1}{n} = 0^2 = 0,$$

and so $\alpha_1 = \alpha_2 = \ldots = 0$. This implies $y = 0$ which is a contradiction.

12. Let $\{x_n\}_{n\in\mathbb{N}}$ be an orthonormal sequence in a Euclidean space V, and also let $x = \sum\limits_{n=0}^{\infty} \alpha_n x_n$. Show that $\alpha_n = \langle x, x_n\rangle$ and $||x||^2 = \sum\limits_{n=0}^{\infty} |\langle x, x_n\rangle|^2$.

13. Let $\{x_n\}_{n\in\mathbb{N}}$ be an orthonormal sequence in a Euclidean space V, and let m be a natural number. Let $x \in V$ be fixed. Show that the set

$$B_m = \{x_n \mid ||x||^2 \leq m|\langle x, x_n\rangle|^2 \}$$

contains at most $m-1$ elements.

14. Let $\{x_1, x_2, x_3\}$ be linearly independent set in a complex Euclidean space V and assume that $\langle x_i, x_j\rangle = \delta_{ij},\ i, j = 1, 2$. Show that $f : \mathbb{C}^2 \to \mathbb{R}$,

$$f(\alpha_1, \alpha_2) = ||\alpha_1 x_1 + \alpha_2 x_2 - x_3||, \ \ \forall(\alpha_1, \alpha_2) \in \mathbb{C}^2,$$

attains its minimum when $\alpha_i = \langle x_3, x_i\rangle,\ i = 1, 2$.

15. Let $T_i : V \to W$ be two linear transformations on the Euclidean (complex or real) vector space. Assume that

$$\langle x, T_1 y\rangle = \langle x, T_2 y\rangle, \ \ \text{for all } x, y \in V.$$

Show that $T_1 = T_2$. Is linearity essential here ?

3.3 Fourier series

Let H be a Hilbert space, and $\{x_n\}_{n\in\mathbb{N}^*}$ be an orthonormal sequence. For $x \in H$, we compute the (projections) complex numbers $\langle x, x_n\rangle$, which are called *the Fourier coefficients* of the vector x.

The series

$$\sum_{n=1}^{\infty} \langle x, x_n\rangle x_n$$

is called *Fourier series* attached to x with respect to $\{x_n\}$.

The following questions arise: a) is the series convergent ? b) is the sum of the series equal to x ?

3.3.1 Fourier Series Theorem. *Let H be a Hilbert space and $\{x_n\}_{n\in\mathbb{N}^*} \subset H$ be an orthonormal set. The following statements are equivalent.*

1) The set $\{x_n\}$ is an orthonormal basis, i.e., it is a maximal orthonormal set.

*2) (*Fourier series expansion*) For any $x \in H$ we have $x = \sum_{n=1}^{\infty} \langle x, x_n\rangle x_n$.*

*3) (*Parseval equality*) For any $x, y \in H$, one has*

$$\langle x, y\rangle = \sum_{n=1}^{\infty} \langle x, x_n\rangle \overline{\langle y, x_n\rangle}.$$

4) For any $x \in H$, one has

$$||x||^2 = \sum_{n=1}^{\infty} |\langle x, x_n\rangle|^2.$$

5) Let M be any vector subspace of H that contains $\{x_n\}_{n\in\mathbb{N}^*}$. Then M is dense in H.

To prove this theorem we need three preliminary results:

- the Bessel inequality,
- discussion of the convergence of $\sum_{n=1}^{\infty} \alpha_n x_n$ when $\{x_n\}_{n\in\mathbb{N}^*}$ is an orthonormal sequence,
- a formula for computing orthogonal projections in terms of an orthonormal basis.

3.3.2 (The Bessel Inequality**) Lemma.**
If $\{x_n\}_{n\in\mathbb{N}^*}$ is an orthonormal set in a Euclidean space, then

$$\sum_{n=1}^{\infty} |\langle x, x_n\rangle|^2 \leq ||x||^2.$$

Proof. Let $\{x_1, \dots, x_k\} \subset \{x_n\}_{n\in\mathbb{N}^*}$. Then

$$\begin{aligned} 0 \quad &\leq ||x - \sum_{i=1}^{k} \langle x, x_i\rangle x_i||^2 = \langle x - \sum_{i=1}^{k} \langle x, x_i\rangle x_i, x - \sum_{j=1}^{k} \langle x, x_j\rangle x_j\rangle = \\ &= \langle x, x\rangle - \sum_{i=1}^{k} \langle x, x_i\rangle\langle x_i, x\rangle - \sum_{j=1}^{k} \overline{\langle x, x_j\rangle}\langle x, x_j\rangle + \sum_{i=1}^{k}\sum_{j=1}^{k} \langle x, x_i\rangle\overline{\langle x, x_j\rangle}\langle x_i, x_j\rangle = \\ &= ||x||^2 - \sum_{i=1}^{k} |\langle x, x_i\rangle|^2, \end{aligned}$$

since $\langle x_i, x_j\rangle = \delta_{ij}$, independent of k. □

Now let us take a careful look at the series of the form

$$\sum_{n=1}^{\infty} \alpha_n x_n, \quad \alpha \in \mathbb{C},$$

where $\{x_n\}_{n\in\mathbb{N}^*}$ is an orthonormal sequence.

3.3.3 Lemma. *Let H be a Hilbert space and $\{x_n\}_{n\in\mathbb{N}^*} \subset H$ be an orthonormal sequence. Then the following assertions are valid:*

1) $\sum_{n=1}^{\infty} \alpha_n x_n$ converges iff $\sum_{n=1}^{\infty} |\alpha_n|^2$ converges.

2) Assume $\sum_{n=1}^{\infty} \alpha_n x_n$ convergent. If

$$x = \sum_{n=1}^{\infty} \alpha_n x_n = \sum_{n=1}^{\infty} \beta_n x_n,$$

then $\alpha_n = \beta_n$ and $||x||^2 = \sum_{n=1}^{\infty} |\alpha_n|^2$.

Proof. 1) Suppose $\sum_{n=1}^{\infty} \alpha_n x_n$ is convergent and $x = \sum_{n=1}^{\infty} \alpha_n x_n$, i.e.,

$$\lim_{N\to\infty} ||x - \sum_{n=1}^{N} \alpha_n x_n||^2 = 0.$$

Since the scalar product is a continuous function, we have

$$\langle x, x_j\rangle = \langle\sum_{n=1}^{\infty} \alpha_n x_n, x_j\rangle = \sum_{n=1}^{\infty} \alpha_n \langle x_n, x_j\rangle = \alpha_j.$$

The Bessel inequality gives

$$\sum_{n=1}^{\infty} |\langle x, x_n\rangle|^2 = \sum_{n=1}^{\infty} |\alpha_n|^2 \leq ||x||^2$$

and so $\sum_{n=1}^{\infty} |\alpha_n|^2$ is convergent.

Next assume that $\sum_{n=1}^{\infty} |\alpha_n|^2$ is convergent, and denote $s_n = \sum_{i=1}^{n} \alpha_i x_i$ (partial sums). It follows

$$||s_n - s_m||^2 = \sum_{m+1}^{n} |\alpha_i|^2,$$

and hence $\{s_n\}_{n\in\mathbb{N}^*}$ is a Cauchy sequence. The completeness of H ensures the convergence of the sequence $\{s_n\}_{n\in\mathbb{N}^*}$.

2) Let us prove

$$||x||^2 = \sum_{n=1}^{\infty} |\alpha_n|^2.$$

For that, using the Cauchy-Schwartz inequality, we evaluate

$$\begin{aligned} ||x||^2 - \sum_{n=1}^{N} |\alpha_n|^2 &= \langle x, x - \sum_{n=1}^{N} \alpha_n x_n\rangle + \langle x - \sum_{n=1}^{N} \alpha_n x_n, \sum_{n=1}^{N} \alpha_n x_n\rangle \leq \\ &\leq ||x - \sum_{n=1}^{N} \alpha_n x_n||(||x|| + ||\sum_{n=1}^{N} \alpha_n x_n||) \leq \\ &\leq 2||x||\,||x - \sum_{n=1}^{N} \alpha_n x_n|| \to 0. \end{aligned}$$

Now suppose $x = \sum_{n=1}^{\infty} \alpha_n x_n = \sum_{n=1}^{\infty} \beta_n x_n$. This implies

$$0 = \lim_{N\to\infty} \left(\sum_{n=1}^{N} \alpha_n x_n - \sum_{n=1}^{N} \beta_n x_n\right) = \lim_{N\to\infty} \sum_{n=1}^{N} (\alpha_n - \beta_n) x_n.$$

Consequently

$$0 = \sum_{n=1}^{\infty} (\alpha_n - \beta_n)x_n \Rightarrow 0^2 = \sum_{n=1}^{\infty} |\alpha_n - \beta_n|^2 \Rightarrow \alpha_n = \beta_n. \qquad \square$$

3.3.4 Corollary. *Let H be a Hilbert space and $S = \{x_n\}_{n\in\mathbb{N}^*} \subset H$ be an orthonormal sequence. The series $\sum_{n=1}^{\infty} \alpha_n x_n$ is convergent iff it is unconditionally convergent (any rearrangement of this series is convergent).*

Now we wish to show that if $\{x_n\}_{n\in\mathbb{N}^*} \subset H$ is an orthonormal set, then the formula

$$\mathcal{P}x = \sum_{n=1}^{\infty} \langle x, x_n\rangle x_n$$

defines an orthonormal projection of H onto $M = L(S)$.

For that we need the following

3.3.5 Lemma. *Let $B = \{x_1, \ldots, x_n\}$ be a finite orthonormal set in a Hilbert space H and $M = L(B)$. Then:*

1) M is closed.

2) The orthogonal projection of H onto M is given by

$$\mathcal{P}x = \sum_{i=1}^{n} \langle x, x_i\rangle x_i.$$

Proof. 1) M is closed because is a finite dimensional subspace.

2) Obviously, $\mathcal{P} : H \to H$ is linear. Also,

$$\mathcal{P}x_j = \sum_{i=1}^{n} \langle x_j, x_i\rangle x_i = x_j.$$

Now let $x \in H$. Then

$$\mathcal{P}^2 x = \mathcal{P}\left(\sum_{i=1}^{n} \langle x, x_i\rangle x_i\right) = \sum_{i=1}^{n} \langle x, x_i\rangle \mathcal{P}x_i = \sum_{i=1}^{n} \langle x, x_i\rangle x_i = \mathcal{P}x.$$

Consequently $\mathcal{P}$ is a projection.

Obviously $Im\,\mathcal{P} \subset M$. Conversely, for any $x \in M$ we can write

$$x = \alpha_1 x_1 + \ldots + \alpha_n x_n.$$

Consequently $\mathcal{P}x = x$ and so $Im\,\mathcal{P} = M$.

Now we show that $\mathcal{P}$ is orthogonal. Let $y \in Ker\mathcal{P}$, $x \in Im\,\mathcal{P}$, i.e.,

$$\mathcal{P}y = 0,\ \mathcal{P}x = x.$$

We find

$$\begin{aligned}\langle y,x\rangle &= \langle y,\mathcal{P}x\rangle = \langle y,\sum_{i=1}^{n}\langle x,x_i\rangle x_i\rangle = \sum_{i=1}^{n}\langle y,x_i\rangle\overline{\langle x,x_i\rangle} = \\ &= \sum_{i=1}^{n}\langle y,x_i\rangle\langle x_i,x\rangle = \langle\sum_{i=1}^{n}\langle y,x_i\rangle x_i,x\rangle = \langle\mathcal{P}y,x\rangle = \langle 0,x\rangle = 0.\end{aligned}$$

This concludes the proof. □

3.3.6 Corollary (a minimum problem).

Let $S = \{x_1,\ldots,x_n\}$ be an orthonormal set in a Hilbert space H and $x \in H$. For any choice of complex numbers $\{c_1,\ldots,c_n\}$ one has

$$||x-\sum_{i=1}^{n}\langle x,x_i\rangle x_i|| \le ||x-\sum_{i=1}^{n} c_i x_i||.$$

Hint. $\mathcal{P}x = \sum_{i=1}^{n}\langle x,x_i\rangle x_i$ is the projection of x onto the subspace generated by S.

3.3.7 Lemma. *Let $B = \{x_n\}_{n\in\mathbb{N}^*}$ be an orthonormal sequence in H and M be the closed vector subspace generated by B (the closure of $L(B)$). Then any vector $x \in M$ can be uniquely written as*

$$x = \sum_{n=1}^{\infty}\langle x,x_n\rangle x_n$$

and moreover

$$\mathcal{P}x = \sum_{n=1}^{\infty}\langle x,x_n\rangle x_n$$

defines the orthogonal projection of H onto M.

Proof. Let y_N be a finite linear combination of vectors in B, i.e., $y_N = \sum_{n=1}^{K}\alpha_n x_n$, where both α_n and K depend on N. Suppose $K \ge N$. It follows

$$||x-\sum_{n=1}^{K}\langle x,x_n\rangle x_n|| \le ||x-\sum_{n=1}^{K}\alpha_n x_n|| = ||x-y_N||.$$

If we accept $x = \lim_{N\to\infty} y_N$, then $\lim_{N\to\infty}||x-y_N|| \to 0$. This implies $x = \sum_{n=1}^{\infty}\langle x,x_n\rangle x_n$. Then, via the *Bessel inequality*, we infer

$$||\mathcal{P}x||^2 = ||\sum_{n=1}^{\infty}\langle x,x_n\rangle x_n||^2 = \sum_{n=1}^{\infty}|\langle x,x_n\rangle|^2 \le ||x||^2.$$

Consequently $\mathcal{P}$ is continuous. The orthogonality is obtained easily since continuity of the scalar product justifies the interchange of summations. □

3.3.8 Proof of the Fourier Series Theorem

1) ⇒ 2). Hypotheses: $\{x_n\}_{n\in\mathbb{N}^*}$ is a maximal orthonormal set in H, M is a closed vectorial subspace of H generated by $\{x_n\}$. If $x \in M^\perp$, then $x \perp x_n$, and so $x = 0$. Hence $M^\perp = \{0\}$, i.e., $M = H$. Therefore the orthogonal projection is just the identity

$$Id(x) = x = \sum_{n=1}^{\infty} \langle x, x_n \rangle x_n.$$

2) ⇒ 3). For $x = \sum_{n=1}^{\infty} \langle x, x_n \rangle x_n,\ y = \sum_{n=1}^{\infty} \langle y, x_m \rangle x_m$ we find

$$\langle x, y \rangle = \langle \sum_{n=1}^{\infty} \langle x, x_n \rangle x_n, \sum_{n=1}^{\infty} \langle y, x_m \rangle x_m \rangle = \sum_{n=1}^{\infty} \langle x, x_n \rangle \overline{\langle y, x_n \rangle}.$$

3) ⇒ 4). Obvious.

4) ⇒ 1). If $\{x_n\}_{n\in\mathbb{N}^*}$ is not maximal, then there exists a unit vector x_0 such that $\{x_n\}_{n\in\mathbb{N}^*} \cup \{x_0\}$ is an orthonormal set. But this yields the contradiction

$$1 = ||x_0||^2 = \sum |\langle x_0, x_n \rangle|^2 = 0.$$

2) ⇔ 5). The statement 5) is equivalent to "the orthogonal projection onto $\bar{M}$ is the identity". Lemma 3.7 proves the equivalence with 2).

3.4 Continuous linear functionals

We want to show that every continuous linear functional on a Hilbert space has a representation via the scalar product.

Let $\mathbf{V}$ be a Euclidean vector space and $y \in \mathbf{V}$. Define

$$l : \mathbf{V} \to \mathbb{C},\ l(x) = \langle x, y \rangle.$$

This function is linear and

$$|l(x)| = |\langle x, y \rangle| \le ||y||\ ||x||.$$

Consequently

$$||l|| = \sup_{||x||=1} |l(x)| \le ||y||.$$

But $|l(y)| = ||y||\ ||y||$, so $||l|| \ge ||y||$. Finally $||l|| = ||y||$. Since l is bounded, automatically it is continuous.

Consequently, a continuous linear functional l is naturally associated to each vector y: any bounded (continuous) linear functional l in a Hilbert space can be written as $l(x) = \langle x, y\rangle,\ x \in H,\ y$ fixed in H.

3.4.1 Riesz Representation Theorem. *Let H be a Hilbert space and $l : H \to \mathbb{C}$ be a bounded (continuous) linear functional. Then there is one and only one vector $y \in H$ such that*

$$l(x) = \langle x, y\rangle, \quad \textit{for all } x \in H. \tag{66}$$

The fixed vector y is called *the representation of l.*

Proof. If $l = 0$, then we associate $y = 0$ to l. Assuming that $l \neq 0$ we remark that since l is continuous, the vector subspace $M = Ker\ l = \{x \in H | l(x) = 0\}$ is closed. Moreover, $M \neq H$, since $l \neq 0$. Therefore, there exist $z \in H$ satisfying $z \perp M, ||z|| = 1$.

We further show that $\dim M^\perp = 1$ and hence $\{z\}$ is a basis of $M^\perp$. We note that the mapping $l^\perp = l|_{M^\perp} : M^\perp \to \mathbb{C}$ is surjective. Indeed, since $l(z) \neq 0$, we have dim Im $l^\perp > 0$. But Im $l^\perp \subset \mathbb{C}$, hence dim Im $l^\perp \leq 1$. Then dim Im ${}^\perp l = 1$ and Im $l^\perp = \mathbb{C}$, $l^\perp$ is surjective. It can be easily checked as well that $l^\perp$ is injective, hence $l^\perp$ is an isomorphism, whence $\dim M^\perp = \dim \mathbb{C} = 1$.

Since $z \notin M$, one has $l(z) \neq 0$. First we show that $l(x) = \langle x, y\rangle,\ \forall x \in H$, for $y = \bar{l}(z)z$. The projection theorem shows

$$H = M + M^\perp,\ x = m + m^\perp.$$

Since $M^\perp = L(\{z\})$, it follows that $x = m + \beta z$. Using the linearity of l and $\langle z, z\rangle = ||z||^2 = 1$, we have

$$\begin{aligned} l(x) &= l(m) + \beta l(z) = 0 + \beta l(z)\langle z, z\rangle = \langle m, \overline{l(z)}z\rangle + \langle \beta z, \overline{l(z)}z\rangle = \\ &= \langle m + \beta z, \overline{l(z)}z\rangle = \langle x, y\rangle,\ y = \overline{l(z)}z. \end{aligned}$$

The vector y is unique with the property (66). Indeed, if $y' \in H$ satisfies (66) too, then

$$\langle x, y\rangle = \langle x, y'\rangle, \quad \text{for all } x \in H,$$

whence $\langle x, y - y'\rangle = 0$, for all $x \in H$. For $x = y - y'$ we infer $0 = \langle y - y', y - y'\rangle = ||y - y'||^2$. It follows $y = y'$. □

3.4.2. Exercise

On $\mathbb{R}^n$ we consider the scalar product

$$\langle x, y\rangle = \sum_{i,j=1}^{n} a_{ij} x_i y_j.$$

Find a representation for $l : \mathbb{R}^n \to \mathbb{R}$, when l is given by

1) $l(x) = x_1$;
2) $l(x) = x_1 + x_2$;
3) $l(x) = \sum_{i=1}^{n} b_i x_i$.

3.5 Trigonometric Fourier series

3.5.1 We consider $I = [-\pi, \pi]$, and the real space $H = L_2(I)$ with the scalar product

$$\langle f, g\rangle = \int_{-\pi}^{\pi} f(t)g(t)dt.$$

The set

$$\left\{\frac{1}{\sqrt{2\pi}}\right\} \bigcup \left\{\frac{1}{\sqrt{\pi}} \cos kt \;\middle|\; k \in \mathbb{N}^*\right\} \bigcup \left\{\frac{1}{\sqrt{\pi}} \sin kt \;\middle|\; k \in \mathbb{N}^*\right\}$$

is an orthonormal sequence. Given $f \in H$, we attach the trigonometric Fourier series

$$S(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left(a_k \cos kt + b_k \sin kt\right),$$

where

$$\begin{aligned} a_0 &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\, dt \\ a_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t) \cos kt\, dt, \quad k \in \mathbb{N}^* \\ b_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t) \sin kt\, dt, \quad k \in \mathbb{N}^*. \end{aligned} \tag{67}$$

Then we have the equality $f(t) = S(t), \quad \text{for all } t \in (-\pi, \pi)$.

The trigonometric Fourier series are used in the study of periodical phenomena, where appears the problem of representing a function as the sum of this kind of series. In this sense, we have two problems:

- if f is the sum of a trigonometric Fourier series, find the coefficients $a_0,\ a_k,\ b_k$;

- the finding of conditions under which a function can be expanded in a trigonometric Fourier series.

Some answers to these questions are given in §3.

3.5.2. Exercises

1. Determine the Fourier series of the function $f : (-\pi, \pi) \to \mathbb{R},\ f(x) = x$.

Solution. On the symmetric interval $(-\pi, \pi)$ we have $S(x) = f(x)$, where

$$S(x) = \frac{a_0}{2} + \sum_{n\geq 1}(a_n \cos nx + b_n \sin nx), \tag{68}$$

and the coefficients are provided by (67). In our case, as f is an odd function on the symmetric interval $(-\pi, \pi)$ as well as is $\sin nx$, it follows that a_0 and a_k are expected to be zero. Indeed, by computation, we find

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} t dt = \frac{1}{\pi} \cdot \frac{t^2}{2}\Big|_{-\pi}^{\pi} = \frac{1}{\pi}(\pi^2 - \pi^2) = 0$$

and also, using that *cos* is an even function, we have

$$\begin{aligned} a_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} t\cos kt dt = \frac{1}{\pi}\left(\frac{t\sin kt}{k}\Big|_{-\pi}^{\pi} - \frac{1}{k}\int_{-\pi}^{\pi}\sin kt dt\right) = \\ &= \frac{1}{\pi}\left(0 - \frac{1}{k}\cdot\frac{-\cos kt}{k}\Big|_{-\pi}^{\pi}\right) = \frac{1}{\pi}\cdot\frac{1}{k^2}(\cos k\pi - \cos(-k\pi)) = 0. \end{aligned}$$

On the other hand, using

$$\cos k\pi = \begin{cases} 1, \text{ for } k = 2p \\ -1, \text{ for } k = 2p+1 \end{cases} = (-1)^k, \text{ for all } k \in \mathbb{N},$$

we compute

$$\begin{aligned} b_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin kt dt = \frac{1}{\pi}\left(-\frac{t\cos kt}{k}\Big|_{-\pi}^{\pi} + \frac{1}{k}\int_{-\pi}^{\pi}\cos kt dt\right) = \\ &= \frac{1}{\pi}\left(\frac{-\pi\cos k\pi}{k} - \frac{\pi\cos(-k\pi)}{k} + \frac{1}{k}\frac{\sin kt}{k}\Big|_{-\pi}^{\pi}\right) = -\frac{2}{k}(-1)^k = \frac{2(-1)^{k+1}}{k}. \end{aligned}$$

Hence

$$S(x) = 0 + \sum_{k=1}^{\infty}\left(0 + \frac{2}{k}(-1)^{k+1}\sin kx\right) = 2\sum_{n=1}^{\infty}\frac{(-1)^{n+1}\sin nx}{n}.$$

2. Determine the Fourier series for the function

$$f : [-\pi, \pi] \to \mathbb{R}, \quad f(x) = \begin{cases} c_1, & x \in [-\pi, 0] \\ c_2, & x \in (0, \pi]. \end{cases}$$

Solution. The coefficients of the Fourier series are

$$a_0 = \frac{1}{\pi}\left(c_1\int_{-\pi}^{0} dt + c_2\int_{0}^{\pi} dt\right) = \frac{1}{\pi}\left(c_1 t\Big|_{-\pi}^{0} + c_2 t\Big|_{0}^{\pi}\right) = \frac{1}{\pi}(c_1\pi + c_2\pi) = c_1 + c_2,$$

and

$$\begin{aligned} a_k &= \tfrac{1}{\pi}\left(c_1 \int\limits_{-\pi}^{0} \cos kt dt + c_2 \int\limits_{0}^{\pi} \cos kt dt\right) = \\ &= \tfrac{1}{\pi}\left(c_1 \tfrac{\sin kt}{k}\Big|_{-\pi}^{0} + c_2 \tfrac{\sin kt}{k}\Big|_{0}^{\pi}\right) = \tfrac{1}{\pi}(0+0) = 0; \end{aligned}$$

For the odd part, the coefficients are

$$\begin{aligned} b_k &= \tfrac{1}{\pi}\left(c_1 \int\limits_{-\pi}^{0} \sin kt dt + c_2 \int\limits_{0}^{\pi} \sin kt dt\right) = \\ &= \tfrac{1}{\pi}\left(c_1(-1)\tfrac{\cos kt}{k}\Big|_{-\pi}^{0} + c_2(-1)\tfrac{\cos kt}{k}\Big|_{0}^{\pi}\right) = \\ &= \tfrac{1}{\pi}\left(-\tfrac{c_1}{k}(1-(-1)^k) - \tfrac{c_2}{k}((-1)^k - 1)\right) = (c_2 - c_1)(1-(-1)^k)\tfrac{1}{k\pi}, \end{aligned}$$

and hence

$$b_k = \begin{cases} 0, \text{ for } k = 2p \\ -\dfrac{2}{k\pi}(c_1 - c_2), \text{ for } k = 2p+1. \end{cases}$$

Then the Fourier series of the function f is

$$\begin{aligned} S(x) &= \frac{c_1+c_2}{2} + \sum_{k\geq 1;\ k=\text{odd}} \left(-\frac{2}{k\pi}(c_1-c_2)\sin kx\right) = \\ &= \frac{c_1+c_2}{2} + \sum_{p\geq 0}^{\infty}\left(-\frac{2}{(2p+1)\pi}(c_1-c_2)\sin(2p+1)x \right) = \\ &= \frac{c_1+c_2}{2} - \frac{2(c_1-c_2)}{\pi}\sum_{k=0}^{\infty}\frac{\sin(2k+1)x}{2k+1}. \end{aligned}$$

3. Find the Fourier series for the following functions:

a) $f : [-\pi, \pi] \to \mathbb{R},\ f(x) = \sin \frac{3x}{2}$;

b) $f : [-\pi, \pi] \to \mathbb{R},\ f(x) = |x|$;

c) $f : [-\pi, \pi] \to \mathbb{R},\ f(x) = x^2$;

d) $f : [-\pi, \pi] \to \mathbb{R},\ f(x) = e^{ax}$.

e) $f : [-\pi, \pi] \to \mathbb{R},\ f(x) = \begin{cases} \pi + x, & x \in [-\pi, 0) \\ \pi - x, & x \in [0, \pi]. \end{cases}$

Answer. a) $S(x) = \dfrac{8}{\pi} \sum_{n \geq 0} \dfrac{n(-1)^n}{4n^2 - 9} \sin nx;$

b) $S(x) = \dfrac{\pi}{2} - \dfrac{4}{\pi} \sum_{n \geq 1} \dfrac{\cos(2n-1)x}{(2n-1)^2};$

c) $S(x) = \dfrac{\pi^2}{3} + 4 \sum_{n \geq 1} \dfrac{(-1)^n}{n^2} \cos(nx);$

d) $S(x) = \dfrac{2}{\pi} \text{ sh } a\pi \left\{ \dfrac{1}{2a} + \sum_{n \geq 1} \dfrac{(-1)^n}{a^2 + n^2} (a \cos nx - n \sin nx) \right\};$

e) $S(x) = \dfrac{\pi}{2} + \dfrac{4}{\pi} \sum_{k \geq 1} \dfrac{\cos(2k-1)x}{(2k-1)^2}.$

Chapter 4

Numerical Methods in Linear Algebra

4.1 The norm of a matrix

Consider the n-dimensional real vector space $\mathbb{R}^n$.

4.1.1 Definition. We call *norm* in $\mathbb{R}^n$ a mapping $||\cdot|| : x \in \mathbb{R}^n \to ||x|| \in \mathbb{R}_+$, satisfying the following conditions

1. *positivity*: $||x|| \geq 0$, for all $x \in \mathbf{V}$ *and* $||x|| = 0 \Leftrightarrow x = 0$;
2. *homogeneity*: $||\alpha x|| = |\alpha|||x||$, for all $x \in \mathbf{V}$;
3. the *triangle inequality*: $||x + y|| \leq ||x|| + ||y||$, for all $x, y \in \mathbf{V}$.

Examples. The so-called L_p - *norms* on $\mathbb{R}^n$, for $p \in [1, \infty)$, which are provided by

$$||x||_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}, \quad \text{for all } x = {}^t(x_1, \ldots, x_n) \in \mathbb{R}^n,$$

satisfy the conditions 1-3 from above. As particular cases, we have:

a) for $p = 1$, we get $||x||_1 = |x_1| + |x_2| + \ldots + |x_n|$, (the *sky-scrapper norm*);

b) for $p = 2$, we have $||x||_2 = \sqrt{x_1^2 + \ldots + x_n^2} = \sqrt{\langle x, x\rangle}$, (the *Euclidean norm*);

c) for $p = \infty$, we find $||x||_\infty = \max_{i \in \overline{1,n}} |x_i|$, (*the "max" norm*).

Remark that, for $p = 2$, we have $\langle x, y\rangle = {}^t xy$, and the *Cauchy-Schwartz inequality* holds true

$$|\langle x, y\rangle| \leq ||x|| \cdot ||y||, \quad \text{for all } x, y \in \mathbb{R}^n.$$

4.1.2 Theorem. *Let* $||\cdot||_\diamondsuit$, $||\cdot||_\clubsuit$ *be two norms on* $\mathbb{R}^n$. *Then there exist* $a, b > 0$ *such that*

$$a||\cdot||_\diamondsuit \leq ||\cdot||_\clubsuit \leq b||\cdot||_\diamondsuit.$$

Practically, the theorem states that any two norms on $\mathbb{R}^n$ are equivalent.

Remark. The theorem is no longer valid for infinite dimensional vector spaces.

Let $A = (a_{ij})_{i=\overline{1,m}, j=\overline{1,n}} \in M_{m\times n}(\mathbb{R})$. The matrix A defines uniquely a linear operator

$$L_A : \mathbb{R}^n \to \mathbb{R}^m,\ L_A(x) = Ax, \text{ for all } x = {}^t(x_1, \ldots, x_n) \in \mathbb{R}^n.$$

Let $\mathbb{R}^n$ and $\mathbb{R}^m$ be endowed with the norms $||\cdot||_\diamondsuit$ and $||\cdot||_\clubsuit$, respectively.

4.1.3 Theorem. *The real function* $||\cdot|| : M_{m\times n}(\mathbb{R}) \to \mathbb{R}_+$ *given by*

$$||A|| = \sup_{||x||_\diamondsuit = 1} ||L_A(x)||_\clubsuit.$$

is a norm on the mn *- dimensional real vector space* $M_{m\times n}(\mathbb{R})$ *of matrices.*

Proof. The positivity. Since $||L_A(x)||_\clubsuit \geq 0$, for all $x \in \mathbb{R}^n$, it follows $||A|| \geq 0$. We have then

$$||A|| = 0 \Rightarrow ||L_A(x)||_\clubsuit = 0, \text{ for all } x \in \mathbb{R}^n,$$

whence

$$L_A(x) = 0, \text{ for all } x \in \mathbb{R}^n \Rightarrow L_A = 0 \Rightarrow A = 0.$$

If $A = 0$, obviously $L_A = 0 \Rightarrow L_A(x) = 0$, for all $x \in \mathbb{R}^n \Rightarrow ||A|| = 0$. The other two properties of the norm are left as exercises. □

Remarks. 1°. If $m = n$ and $||\cdot||_1 = ||\cdot||_2 = ||\cdot||$, then we have

$$||AB|| \leq ||A|| \cdot ||B||, \text{ for all } A, B \in M_n(\mathbb{R}).$$

2°. For $||\cdot||_1$ and $||\cdot||_2$, we can consider any of the norms described in 1.1. In the case presented above, if $||\cdot||_1 = ||\cdot||_2 = ||\cdot||_\bullet$, then we denote the norm of the matrix A simply by $||A||_\bullet$.

4.1.4 Theorem. *If* $\mathbb{R}^m$ *and* $\mathbb{R}^n$ *are endowed with* L_p *- norms,* $p \in \{1, 2, \infty\}$, *then the following equalities hold true*

$$||A||_1 \stackrel{def}{=} \sup_{||x||_1=1} ||Ax||_1 = \max_{j=\overline{1,n}} \sum_{i=1}^m |a_{ij}|$$

$$||A||_\infty \stackrel{def}{=} \sup_{||x||_\infty=1} ||Ax||_\infty = \max_{i=\overline{1,m}} \sum_{j=1}^n |a_{ij}|$$

$$||A||_2 \stackrel{def}{=} \sup_{||x||_2=1} ||Ax||_2 = \sqrt{\max\{\, |\lambda| \mid \lambda \in \sigma({}^tAA)\,\}}.$$

Proof. Denote $||A||_\diamondsuit = \max\limits_{j=\overline{1,n}} \sum\limits_{i=1}^{m} |a_{ij}|$. Rewriting in coordinates the equality $y = Ax$, for all $x = {}^t(x_1, \ldots, x_n) \in \mathbb{R}^n$ we have $y_i = \sum\limits_{j=1}^{n} a_{ij}x_j$, $i = \overline{1,m}$. Then

$$\begin{aligned} ||Ax||_1 = ||y||_1 &= \sum_{i=1}^{m} |y_i| = \sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij}x_j \right| \leq \\ &\leq \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}||x_j| = \sum_{j=1}^{n} |x_j| \left(\sum_{i=1}^{n} |a_{ij}| \right) \leq \\ &\leq \sum_{j=1}^{n} |x_j| \left(\max_{j=\overline{1,n}} \sum_{i=1}^{n} |a_{ij}| \right) = \sum_{j=1}^{n} |x_j| ||A||_\diamondsuit = ||A||_\diamondsuit ||x||_1. \end{aligned}$$

Then

$$||Ax||_1 \leq ||A||_\diamondsuit ||x||_1 \Rightarrow ||A\left(\frac{1}{||x||}x\right)||_1 \leq ||A||_\diamondsuit \Rightarrow \sup_{||v||_1=1} ||Av||_1 \leq ||A||_\diamondsuit, \quad (69)$$

whence $||A||_1 \leq ||A||_\diamondsuit$.

Denoting $e_k = \underset{(1 \text{ in the k-th pos.})}{{}^t(0, \ldots, 1, \ldots, 0)}$ and $Ae_k = {}^t(a_{k1}, \ldots, a_{kn})$, it follows that

$$||Ae_k||_1 = \sum_{i=1}^{m} |a_{ik}|, \quad k = \overline{1,n}.$$

Hence, if $\max\limits_{j=\overline{1,n}} \sum\limits_{i=1}^{m} |a_{ij}|$ is attended on index $j = k$ then, using the relation above, we find

$$||A||_\diamondsuit = \max_{j=\overline{1,n}} \sum_{i=1}^{m} |a_{ij}| = ||Ae_k||_1,$$

and since $||e_k||_1 = 1$, we have

$$||A||_1 = \sup_{||x||_1=1} ||Ax||_1 \geq ||Ae_k|| = ||A||_\diamondsuit \Rightarrow ||A||_1 \geq ||A||_\diamondsuit. \quad (70)$$

From (69), (70), the two norms coincide, i.e., $||A||_1 \equiv ||A||_\diamondsuit$.

For $||A||_\infty$ the proof is similar. For $||A||_2$, remark that

$$||Ax||_2 = \sqrt{{}^tx\ {}^tAAx} = \sqrt{{}^txBx},$$

where we denote $B = {}^tAA$. But B is symmetrical, hence it has real eigenvalues $\sigma(A) = \{\lambda_1, \ldots, \lambda_n\}$, and moreover, these are nonnegative; for $|\lambda_1| \leq |\lambda_2| \leq \ldots \leq |\lambda_n|$, we have

$$|\lambda_1|\ {}^txx \leq |{}^txBx| \leq |\lambda_n|\ {}^txx, \text{ for all } x \in \mathbb{R}^n.$$

Then we obtain

$$|\lambda_1| \leq \frac{|{}^t xBx|}{||x||_2^2} \leq |\lambda_n|, \quad \text{for all } x \in \mathbb{R}^n.$$

Denoting $v = \frac{1}{||x||_2} x$, and hence $||v||_2 = 1$, we find

$$||Av||_2 = \frac{||Ax||_2}{||x||_2} = \sqrt{\frac{{}^t xBx}{{}^t xx}} \in \left[\sqrt{|\lambda_1|}, \sqrt{|\lambda_n|}\right],$$

so that

$$||A||_2 = \sup_{||v||=1} ||Av||_2 \leq |\lambda_n| \Rightarrow ||A||_2 \leq \sqrt{\max\{|\lambda| \mid \lambda \in \sigma\,({}^t AA)\}}.$$

The converse inequality holds also true. □

Corollary. *If* $A = {}^tA \in M_n(\mathbb{R})$, *then* $||A||_2 = \max\{\, |\lambda_i| \mid \lambda_i \in \sigma(A) \,\}$.

Proof. In this case the eigenvalues of $B = {}^t AA$ are exactly the squares of those of the matrix A, whence the result. □

4.1.5 Definition. Let $A \in M_n(\mathbb{R})$, $\det A \neq 0$. The number

$$k(A) = ||A|| \cdot ||A^{-1}||$$

is called *the conditioning number of the matrix* A relative to the particular norm used in the definition.

Remark. Consider the norm $||\cdot||_2$ on

$$GL_n(\mathbb{R}) = \{A \mid A \in M_n(\mathbb{R}),\ \det A \neq 0\}.$$

Since

$$1 = ||I||_2 = ||AA^{-1}||_2 \leq ||A||_2 \cdot ||A^{-1}||_2 = k(A),$$

we get

$$k(A) \geq 1, \quad \text{for all } A \in GL_n(\mathbb{R}).$$

4.1.6. Exercises

1. a) Compute $||v||_p$, for $v = {}^t(1, 0, -2)$, and $p \in \{1, 2, \infty\}$.

b) Compute $||A||_p$ for $A = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}$, and $p \in \{1, 2, \infty\}$.

Solution. a) By straightforward computation, we obtain

$$\begin{cases} ||v||_1 = |1| + |0| + |-2| = 3 \\ ||v||_2 = \sqrt{1 + 0 + 4} = \sqrt{5} \\ ||v||_\infty = \max\{|1|, |0|, |-2|\} = 2. \end{cases}$$

b) Using the theorem which characterizes the p–norms of matrices $A \in M_{m\times n}(\mathbb{R})$, we infer

$$\begin{cases} ||A||_1 = \max\limits_{j=\overline{1,n}} \sum\limits_{i=1}^{m} |a_{ij}| = \max\{|1| + |-1|, |0| + |2|\} = 2 \\ ||A||_2 = \max\limits_{\lambda\in\sigma({}^tA\ A)} \sqrt{|\lambda|} = \max\limits_{\lambda\in\{3\pm\sqrt{5}\}} \sqrt{|\lambda|} = \sqrt{3+\sqrt{5}} \\ ||A||_\infty = \max\limits_{i=\overline{1,m}} \sum\limits_{j=1}^{n} |a_{ij}| = \max\{|1| + |-0|, |-1| + |2|\} = 3. \end{cases}$$

The last norm was obtained following the computations

$${}^tAA = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix},$$

whence

$$P_{{}^tAA}(\lambda) = \begin{vmatrix} 2-\lambda & -2 \\ -2 & 4-\lambda \end{vmatrix} = \lambda^2 - 6\lambda + 4 = 0 \Rightarrow \lambda_{1,2} = 3 \pm \sqrt{5}.$$

2. a) Check that the spectral radius $\rho(A) = \max\ \{|\lambda| \mid \lambda \in \sigma(A) \subset \mathbb{C}\}$ of the matrix $A = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}$ is less or equal than any of the norms of A of the previous problem.

b) Find the conditioning number of A with respect to the three considered norms.

c) Compute $\rho(B)$ and $||B||_2$, for $B = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$.

Solution. a) In this case $\sigma(A) = \{1, 2\}$, whence

$$\rho(A) = \max_{\lambda\in\sigma(A)} |\lambda| = 2.$$

We have ${}^tA\ A = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix}$ and the three norms of A are

$$||A||_1 = 2,\ ||A||_\infty = 3,\ ||A||_2 = \sqrt{3+\sqrt{5}};$$

remark that all are greater or equal to $\rho(A) = 2$.

b) The inverse is $A^{-1} = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$, ${}^t(A^{-1})A^{-1} = \begin{pmatrix} 5/4 & 1/4 \\ 1/4 & 1/4 \end{pmatrix}$, and the norms of A^{-1} are

$$||A^{-1}||_1 = \frac{3}{2},\ ||A^{-1}||_\infty = 1,\ ||A^{-1}||_2 = \frac{\sqrt{3+\sqrt{5}}}{2}.$$

Then the corresponding conditioning numbers are

$$k_1(A) = 3,\ k_\infty(A) = 3,\ k_2(A) = \frac{3+\sqrt{5}}{2}.$$

c) The *characteristic polynomial* of B is $P_B(\lambda) = \lambda^2 - 3\lambda + 1$, whence its roots are $\lambda_{1,2} = (3 \pm \sqrt{5})/2$ and hence $\rho(B) = (3 + \sqrt{5})/2$. As well, ${}^tBB = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}$, $\sigma({}^tBB) = \{(7 \pm 3\sqrt{5})/2\}$ whence

$$||B||_2 = \max_{\lambda \in \sigma({}^tBB)} \sqrt{|\lambda|} = \sqrt{(7 + 3\sqrt{5})/2}.$$

We note that $||B||_2 = \rho(B)$, since $B = {}^tB$.

4.2 The inverse of a matrix

Let $A \in M_n(\mathbb{R})$. We try to relate the spectrum of A, $\sigma(A) = \{\lambda_1, \ldots, \lambda_n\} \subset \mathbb{C}$ to the norms of the matrix A.

4.2.1 Definition. The real positive number $\rho(A) = \max_{i=\overline{1,n}} |\lambda_i|$ is called *the spectral radius of the matrix A.*

If $u \in \mathbb{R}^n \backslash \{0\}$ is an eigenvector associated to $\lambda \in \sigma(A)$, then, considering the associated unit vector $v = \frac{1}{||u||} u$, $(||v|| = 1)$, we find

$$Au = \lambda u \;\Rightarrow\; ||Au|| = |\lambda| \cdot ||u|| \;\Rightarrow\; \frac{||Au||}{||u||} = |\lambda| \;\Rightarrow\; |\lambda| = ||Av||.$$

Then $|\lambda| \leq \sup_{||w||=1} ||Aw|| = ||A||$. So that for all $\lambda \in \sigma(A)$, we have $|\lambda| \leq ||A||$, whence $\max_{\lambda_i \in \sigma(A)} | \lambda_i | \leq ||A||$, and thus

$$\rho(A) \leq ||A||. \tag{71}$$

4.2.2 Theorem. *a) For all $\varepsilon > 0$, there exists a norm $|| \cdot || : M_n(\mathbb{R}) \to \mathbb{R}_+$, such that*

$$||A|| \leq \rho(A) + \varepsilon, \quad \textit{for all } A \in M_n(\mathbb{R}).$$

b) Let $A \in M_n(\mathbb{R})$. We have $\lim_{k\to\infty} A^k = 0_{n\times n}$ if and only if $\rho(A) < 1$.

Proof. b) We prove first the direct implication by *reductio ad absurdum.* Let $\lim_{k\to\infty} A^k = 0$ and let $\lambda \in \sigma(A)$ be an eigenvalue such that $|\lambda| \geq 1$, and let $x \in \mathbb{R}^n \backslash \{0\}$ be a corresponding eigenvector, assumed unit vector. But $A^k x = \lambda^k x$, for all $k \in \mathbb{N}^*$, whence

$$\lim_{k\to\infty} ||A^k x|| = \lim_{k\to\infty} ||\lambda^k x|| = \lim_{k\to\infty} |\lambda|^k \geq 1,$$

since $|\lambda| \geq 1$. So that

$$||A^k|| = \sup_{||v||=1} ||A^k v|| \geq ||A^k x|| \geq 1,$$

Thus we have $\lim\limits_{k\to\infty} ||A^k|| \neq 0$, so $\lim\limits_{k\to\infty} A^k \neq 0_{n\times n}$, which contradicts the initial assertion.

The converse implication. If $\rho(A) < 1$, according to a) there exists a norm $||\cdot||_\diamondsuit$, such that $||A||_\diamondsuit < 1$. Then

$$0 \le ||A^k||_\diamondsuit \le {||A||_\diamondsuit}^k \underset{k\to\infty}{\longrightarrow} 0,$$

and it follows that $\lim\limits_{k\to\infty} ||A^k||_\diamondsuit = 0$, whence $\lim\limits_{k\to\infty} A^k = 0_{n\times n}$. □

4.2.3 Theorem. *Let $B \in M_n(\mathbb{R})$ be a matrix such that $\rho(B) < 1$. Then $I - B$ is invertible, and*

$$(I-B)^{-1} = \lim_{k\to\infty} \sum_{i=0}^{k} B^i = \sum_{i=0}^{\infty} B^i,$$

where we denoted $B^0 = I_n$.

Proof. If $\rho(B) < 1$, then $(I - B)$ has no vanishing eigenvalue, therefore

$$\det(I-B) = \lambda_1 \dots \lambda_n \neq 0,$$

i.e., is non degenerate, and hence invertible. Using the equalities

$$B^m B^n = B^n B^m = B^{m+n},$$

we find

$$(I-B)(I+B+\dots+B^{k-1}) = I - B^k,$$

whence for $k \to \infty$ we obtain the result,

$$(I-B) \lim_{k\to\infty} \sum_{i=0}^{k-1} B^i = I - \lim_{k\to\infty} B^k.$$

Corollary. If $||B|| < 1$, then the matrix $I - B$ is invertible, and we have

$$||(I-B)^{-1}|| \le \sum_{i=0}^{\infty} ||B||^i.$$

Proof. The invertibility follows from $\rho(B) \le ||B|| < 1$ and from the theorem above. Also, using the relation

$$||AB|| \le ||A|| \cdot ||B||,$$

we obtain

$$||(I-B)^{-1}|| \le \sum_{i=1}^{\infty} ||B^i|| \le \sum_{i=1}^{\infty} ||B||^i = \frac{1}{1-||B||}.$$ □

4.2.4 Definitions. a) Let A be an invertible matrix $A \in Gl(n, \mathbb{R})$ and let $b \in \mathbb{R}^n$. We consider the linear system

$$Ax = b, \tag{72}$$

in the unknown $x \in \mathbb{R}^n$, and call $x = A^{-1}b$ *the exact solution* of the system (72).

b) Let $\varepsilon \in M_n(\mathbb{R})$ be a matrix called *rounding error*, assumed to satisfy $||\varepsilon|| \ll ||A||$. Consider the linear attached system

$$(A + \varepsilon)\overline{x} = b; \tag{73}$$

if the matrix $A + \varepsilon$ is invertible, then we call $\overline{x} = (A + \varepsilon)^{-1}b$ *the approximate solution* of the initial system. In this case, the quantity

$$e = \frac{||\bar{x} - x||}{||x||},$$

where x is the solution of the initial equation, is called *the relative error.*

Remark. Denote $u = \overline{x} - x$ (and hence $\bar{x} = x + u$), and assume that $||u|| < ||x||$. Then the attached system (73) becomes: $(A + \varepsilon)(x + u) = b$. Using the relation $Ax = b$ fulfilled by x, the approximation $\varepsilon u \simeq 0$ which holds for ε, u small, and the fact that A is invertible, we obtain

$$Au + \varepsilon x = 0 \quad \Rightarrow \quad u = -A^{-1}\varepsilon x,$$

and hence $||u|| \leq ||A^{-1}|| \cdot ||\varepsilon|| \cdot ||x||$. The relative error becomes

$$e = \frac{||\overline{x} - x||}{||x||} = \frac{||u||}{||x||} \leq \frac{||A^{-1}|| \cdot ||\varepsilon|| \cdot ||x||}{||x||} = ||A|| \cdot ||A^{-1}|| \frac{||\varepsilon||}{||A||}.$$

The quantity

$$k(A) = ||A|| \cdot ||A^{-1}||$$

was called *conditioning number of the matrix A, with respect to the norm* $|| \cdot ||$.

For $k(A)$ having large values, the relative error increases and the computational methods give inaccurate results.

Example. Compute the exact solution, the approximate solution and the conditioning number of the following linear system, for the rounding error ε

$$\begin{cases} 2a + 1b = 3 \\ -2a + 3b = 1 \end{cases}, \qquad \varepsilon = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.1 \end{pmatrix}.$$

Solution. The given linear system rewrites

$$Ax = b \ \Leftrightarrow \begin{pmatrix} 2 & 1 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix},$$

and has the exact solution $x = {}^t(a, b) = {}^t(1, 1)$. The altered system

$$(A + \varepsilon)\bar{x} = b \quad \Leftrightarrow \quad \begin{cases} 2.5a + 1b = 3 \\ -2a + 3.1b = 1 \end{cases}$$

has the solution $\bar{x} = \left(\frac{166}{195}, \frac{34}{39}\right)$. Then $\bar{x}$ is close to x iff the number $k(A)$ is sufficiently small.

↟ Hw. Compute $k(A) = ||A|| \cdot ||A^{-1}||$ for $|| \cdot ||_2$.

4.3 Triangularization of a matrix

The idea of this section is to decompose a given matrix as a product between a lower-triangular and an upper-triangular matrix. Then the linear systems whose matrix of coefficients is the decomposed matrix become equivalent to triangular linear systems which can be solved in straightforward manner.

4.3.1 Definitions. a) A quadratic matrix in which all elements are zero above/below the diagonal is called *lower/upper triangular matrix*.

b) Let $A \in M_n(\mathbb{R}),\ A = (a_{ij})_{\overline{1,n} \times \overline{1,n}}$. Then the matrix blocks

$$A^{[k]} \in M_k(\mathbb{R}),\ A^{[k]} = (a_{ij})_{i,j=\overline{1,k}}, \quad k = \overline{1,n}$$

are called *principal submatrices of A*.

Example. The following matrix is upper triangular $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

Theorem. *If det $(A^{[k]}) \neq 0$, for all $k = \overline{1, n-1}$, then there exists a non-singular lower triangular matrix $M \in M_n(\mathbb{R})$, such that $MA = U$, where U is upper-triangular. Thus, for $L = M^{-1}$, the matrix A decomposes as $A = LU$ (the* LU decomposition of A).

Proof. We search for M of the form $M = M_{n-1} \dots M_1$, where $M_k = I_n - m_k\, {}^t e_k$, with

$$m_k = {}^t(0, \dots, 0, \mu_{k+1,k}, \dots, \mu_{n,k}), \quad e_k = {}^t \underbrace{(0, \dots, 0, 1, 0, \dots, 0)}_{1 \text{ in the } k\text{-th position}}.$$

We remark that:

1) $M_k^{-1} = I_n + m_k {}^t e_k$, for all $k = \overline{1, n-1}$.

2) If $x = {}^t(\zeta_1, \dots, \zeta_k, \zeta_{k+1}, .., \zeta_n)$ is the $k-$th column of the matrix

$$M_{k-1} M_{k-2} \dots M_1 A$$

(or A, in case that $k = 1$), and if ${}^t e_k x \equiv \zeta_k \neq 0$ (satisfied due to the requirements imposed on A), then there exists an upper-triangular matrix M_k such that in the vector $M_k x$ (i.e., in the $k-$th column of $M_k M_{k-1} \dots M_1 A$, the last $n-k$ components all vanish: $M_k x = {}^t(\zeta_1, \dots, \zeta_k, 0, \dots, 0)$.

In fact this determines recursively M_k. We find

$$M_k x = (I_n - m_k {}^t e_k) x = x - m_k \xi_k$$

of components

$$(M_k(x))_i = \zeta_i - \mu_{ik}\zeta_k = \zeta_i, \ i = \overline{1, n}.$$

Imposing that $\mu_{ik} = 0$ for $i = \overline{1, k}$, we get that the last $n-k$ components of the product $M_k x$ vanish iff

$$\mu_{ik} = \zeta_i / \zeta_k, \ i = \overline{k+1, n}.$$

Therefore

$$M_k = I_n - \underbrace{\underbrace{{}^t(0, \dots, 0, \frac{\zeta_{k+1}}{\zeta_k}, \dots, \frac{\zeta_n}{\zeta_k})}_{n\times 1} \underbrace{(0, \dots, 1, \dots 0)}_{1\times n}}_{n \times n \text{ matrix}}.$$

Remark that based on the properties above, the multiplication of M_k with the matrix $M_{k-1} \dots M_1 A$ provides a new matrix which has zeros below diagonal in its first k columns.

This is why denoting $M = M_{n-1} M_{n-2} \dots M_1$, the matrix $U = MA$ is upper-triangular, and both M and $L = M^{-1}$ are lower-triangular. Hence $A = LU$ is the required decomposition. □

Example. Let $A = \begin{pmatrix} 1 & -1 & 2 \\ 2 & 1 & 1 \\ 1 & 2 & -3 \end{pmatrix}$. Find the LU decomposition of A. Find the lower-triangular matrix M, such that MA is a upper-triangular matrix.

Solution. Consider first the pivot $\mu_{11} = \zeta_1 = 1$ in A, and

$$\begin{cases} \mu_{21} = -\dfrac{\zeta_2}{\zeta_1} = -\dfrac{2}{1} = -2 \\[2ex] \mu_{31} = -\dfrac{\zeta_3}{\zeta_1} = -\dfrac{1}{1} = -1, \end{cases}$$

whence the first lower-triangular multiplying matrix is $M_1 = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$. Then

$$M_1 \cdot A = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 3 & -3 \\ 0 & 3 & -5 \end{pmatrix}.$$

Also, the next pivot, taken from M_1A, is $\zeta_2 = 3$, whence the only nontrivial coefficient of the next multiplying matrix M_2 is

$$\mu_{32} = \frac{-\zeta_3}{\zeta_2} = -\frac{3}{3} = -1,$$

and the second lower-triangular multiplying matrix is $M_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$. Then

$$M_2M_1A = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 3 & -3 \\ 0 & 0 & -2 \end{pmatrix} = U.$$

Hence the overall lower-triangular multiplying matrix and its inverse are respectively

$$M = M_2M_1 = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \quad \text{and} \quad L = M^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

We have the relation $MA = U$, with U upper triangular matrix, and A obeys the $LU-$decomposition $A = LU$.

The triangularization algorithm. Using the Gauss elimination, we can describe the following *algorithm of triangularization*:

a) $A_1 = A$;

b) Let

$$A_k = M_{k-1}M_{k-2}\dots M_1A = (a_{ij}^{(k)})_{i,j=\overline{1,n}}, \quad k \geq 2$$

be upper triangular with respect to the first $k-1$ columns. Assume that at each step we have an nonvanishing *pivot element*, $a_{kk}^{(k)} \neq 0$. Then one determines M_k as above, using the condition of cancelling the elements below the diagonal on the k-th column. Then $A_{k+1} = M_kA_k$ becomes upper triangular with respect to the columns $\overline{1,k}$.

c) Repeat step b), for $k = \overline{1,n-1}$.

d) Denoting $M = M_{n-1} \cdot \ldots \cdot M_1$, we have that $U = A_n = MA$ is an upper triangular matrix, and for $L = M^{-1}$, we have the LU decomposition $A = LU$.

The algorithm proves to be extremely useful in solving linear systems of the form

$$Ax = b.$$

Namely, after performing the algorithm which provides M and U such that $MA = U$, and denoting $c = Mb$, this system becomes

$$MAx = Mb \quad \Leftrightarrow \quad Ux = c.$$

If $U = (r_{ij})_{i,j=\overline{1,n}}$ has the elements $r_{ii} \neq 0$, for all $i = \overline{1,n}$, then the equivalent system rewrites explicitly

$$\begin{cases} r_{11}x_1 + r_{12}x_2 + r_{13}x_3 + \ldots + r_{1n}x_n = c_1 \\ r_{22}x_2 + r_{23}x_3 + \ldots + r_{2n}x_n = c_2 \\ r_{33}x_3 + \ldots + r_{3n}x_n = c_3 \\ \ldots \\ r_{nn}x_n = c_n. \end{cases}$$

This can be solved easily by regression

$$x_n = \frac{c_n}{r_{nn}} \to x_{n-1} \to \ldots \to x_1,$$

by means of the relations

$$x_k = \frac{c_k - \sum\limits_{i=k+1}^{n} r_{ki}x_i}{r_{ii}}, \quad k \in \{n-1, n-2, \ldots, 2, 1\},$$

where we denoted $c = Mb = {}^t(c_1, \ldots, c_n)$.

4.3.2. Exercises

1. Given the matrices

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 5 \\ 2 & 1 & -3 \end{pmatrix} \text{ and } b = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix},$$

triangularize the matrix A and solve the system $Ax = b$.

Solution. The first pivot is $a_{11} = 1$ and

$$M_1 = \begin{pmatrix} \boxed{1} & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \quad \Rightarrow \quad M_1A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & \boxed{-1} & 5 \\ 0 & -3 & -5 \end{pmatrix}.$$

The next pivot is -1 and

$$M_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3 & 1 \end{pmatrix} \quad \Rightarrow \quad U = M_2M_1A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 5 \\ 0 & 0 & -20 \end{pmatrix},$$

and the multiplying lower-triangular matrix is

$$M = M_2M_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & -3 & 1 \end{pmatrix}.$$

Then

$$Mb = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & -3 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

and the system $Ax = b$ rewrites

$$MAx = Mb \Leftrightarrow \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 5 \\ 0 & 0 & 20 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

whence the solution is computed recursively, $x_3 = 0, x_2 = 0, x_1 = 1$, and hence $x = {}^t(1, 0, 0)$.

4.4 Iterative methods for solving linear systems

For given $A \in GL(n, \mathbb{R})$ and $b \in M_{n\times 1}(\mathbb{R})$, let be the associated linear system

$$Ax = b. \tag{74}$$

Note that the classical method of solving this system based on computing first the inverse A^{-1}, and afterwards the solution $x = A^{-1}b$ is inefficient because of cummulated errors and the large number of operations.

On the other hand, the iterative methods give, under certain defined conditions, more precise and less time-consuming results. We shall describe in the following a procedure of this type.

Let A be decomposed as $A = N - P$ with N invertible (usually diagonal, Jordan or triangular). Then the system rewrites $Nx = Px + b$, and one can build a sequence of vectors of given initial seed $x^{(0)} \in \mathbb{R}^n$, defined by

$$Nx^{(k+1)} = Px^{(k)} + b, \quad k \geq 0. \tag{75}$$

Note that the recurrence relation can be rewritten as

$$x^{(k+1)} = N^{-1}Px^{(k)} + N^{-1}b, \quad k \geq 0, \tag{76}$$

or, denoting $G = N^{-1}P$ and $c = N^{-1}b$,

$$x^{(k+1)} = Gx^{(k)} + c, \ \ k \geq 0.$$

We determine some cases in which $x^{(k)}$ converges to the exact solution $x_* = A^{-1}b$ of the system (74).

4.4.1 Theorem. *Given the sequence (76), we have* $\lim\limits_{k\to\infty} x^{(k)} = x_*$ *iff* $\rho(G) < 1$.

Proof. Let $\varepsilon^{(k)} = x - x^{(k)}$ be the error at step k, $k \geq 0$. We check that $\lim\limits_{k\to\infty} \varepsilon^{(k)} = 0$ iff $\rho(N^{-1}P) < 1$.

Indeed, since $x^{(k)} = x - \varepsilon^{(k)}$, by replacing in (75) and using that $Ax_* = b$ (i.e., equivalently, $Nx_* = Px_* + b$), we have

$$0 = Ne^{(k+1)} - Pe^{(k)} \Rightarrow e^{(k+1)} = Ge^{(k)}, \;\; k = \overline{0,n},$$

where we denoted $G = N^{-1}P$. Writing the last relation for $k = \overline{1,m}$, we infer $e^{(k+1)} = G^k e^{(0)}, \;\; k = \overline{0,n}$.

Hence we find

$$\rho(G) < 1 \Leftrightarrow \lim_{k\to\infty} G^k = 0 \Leftrightarrow \lim_{k\to\infty} e^{(k)} = 0 \Leftrightarrow \lim_{k\to\infty} x^{(k)} = x_*. \qquad \square$$

Application. Let $A = L + D + R$, where L is strictly lower triangular, D is diagonal and R is strictly upper triangular,

$$(a_{ij}) = \underbrace{\begin{pmatrix} 0 & & 0 \\ & \ddots & \\ a_{\alpha\beta} & & 0 \end{pmatrix}}_{L} + \underbrace{\begin{pmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}}_{D} + \underbrace{\begin{pmatrix} 0 & & a_{\mu v} \\ & \ddots & \\ 0 & & 0 \end{pmatrix}}_{R}.$$

Then we can use one of the following two methods of solving iteratively the linear system $Ax = b$.

1. The Jacobi method.

In this case, we choose

$$N = D, \quad P = N - A = -(L + R).$$

Provided that the condition in the theorem is fulfilled, the sequence of iterations $x^{(k)} = (x_i^{(k)})_{i=\overline{1,n}}$ which converge to the solution x_*, is given by

$$x_i^{(k+1)} = \left(b_i - \sum_{j=1,\, j\neq i}^{n} a_{ij} x_j^{(k)} \right) / a_{ii}, \;\; i = \overline{1,n}, \; k \in \mathbb{N}.$$

2. The Gauss-Seidel method.

In this case we have

$$N = L + D, \quad P = N - A = -R,$$

and the sequence is given by

$$x_i^{(k+1)} = \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} + \sum_{j=i+1}^{n} a_{ij} x_j^{(k)} \right) / a_{ii}, \; i = \overline{1,n}, \; k \in \mathbb{N}.$$

Note. For $A = {}^tA$ (hence A symmetric) and having positively defined associated quadratic form, the Gauss-Seidel algorithm converges.

Example. The system $Ax = b$, where $A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$, $b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ admits the solution $x_* = {}^t(1,1)$ to which converges the Gauss-Seidel sequence. Here we have

$$A = {}^tA, \quad \Delta_1 = |2| = 2 > 0, \quad \Delta_2 = \begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} = 3 > 0,$$

with $x^{(0)}$ given by, e.g., $x^{(0)} = {}^t(0,0)$.

4.4.2. Exercises

1. Solve the system $\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $x^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, using:

a) the Jacobi method;

b) the Gauss-Seidel method.

Solution. a) The Jacobi method. We have

$$N = D = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \quad P = N - A = -L - R = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad N^{-1} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix},$$

and

$$N^{-1}P = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix},$$

whence

$$\sigma(N^{-1}P) = \left\{-\frac{1}{2}, \frac{1}{2}\right\} \quad \Rightarrow \quad \rho(N^{-1}P) = \frac{1}{2} < 1;$$

consequently, the sequence converges to the solution of the system. The sequence is given by

$$x_i^{(k+1)} = \left(b_i - \sum_{j=1, j\neq i}^{n} a_{ij} x_j^{(k)} \right) \frac{1}{a_{ii}}, \quad i = \overline{1,n}, \; k \in \mathbb{N}.$$

These relations rewrite $Nx^{(k+1)} = Px^{(k)} + b$ and the next term of the sequence is provided by a system of the form

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

and rewrites

$$\begin{cases} 2x' = y + 1 \\ 2y' = x + 1 \end{cases} \Leftrightarrow \begin{cases} x' = (y+1)/2 \\ y' = (x+1)/2 \end{cases} \Rightarrow \begin{cases} x^{(n+1)} = (y^{(n)} + 1)/2 \\ y^{(n+1)} = (x^{(n)} + 1)/2. \end{cases}$$

Then the required iterations are

$x^{(0)}$	$x^{(1)}$	$x^{(2)}$	$x^{(3)}$	$x^{(4)}$
0	1/2	3/4	7/8	15/16
0	1/2	3/4	7/8	15/16

which provide at limit the solution $x^* = {}^t(1,1)$.

b) The Gauss-Seidel method. We have

$$N = L + D = \begin{pmatrix} 2 & 0 \\ -1 & 2 \end{pmatrix}, \ P = N - A = -R = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \ N^{-1} = \begin{pmatrix} 1/2 & 0 \\ 1/4 & 1/2 \end{pmatrix},$$

and

$$N^{-1}P = \begin{pmatrix} 1/2 & 0 \\ 1/4 & 1/2 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1/2 \\ 0 & 1/4 \end{pmatrix}.$$

We obtain $\sigma(N^{-1}P) = \{0, \frac{1}{4}\}$ and $\rho(N^{-1}P) = 1/4 < 1$ whence the sequence converges to the solution of the system. The sequence is given by

$$x_i^{(k+1)} = \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} + \sum_{j=i+1}^{n} a_{ij} x_j^{(k)} \right) \frac{1}{a_{ii}}, \quad i = \overline{1,n}, \ k \in \mathbb{N}.$$

These relations rewrite $Nx^{(k+1)} = Px^{(k)} + b$ and the next term of the sequence is provided by a system of the form

$$\begin{pmatrix} 2 & 0 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

and rewrites

$$\begin{cases} 2x' = y + 1 \\ -x' + 2y' = 1 \end{cases} \Leftrightarrow \begin{cases} x' = (y+1)/2 \\ y' = (y+3)/4 \end{cases} \Rightarrow \begin{cases} x^{(n+1)} = (y^{(n)} + 1)/2 \\ y^{(n+1)} = (y^{(n)} + 3)/4. \end{cases}$$

The corresponding first three iterations are

$x^{(0)}$	$x^{(1)}$	$x^{(2)}$	$x^{(3)}$
0	1/2	7/8	31/32
0	3/4	15/16	63/64.

which provide at limit the solution $x^* = {}^t(1,1)$.

4.5 Solving linear systems in the sense of least squares

Consider $A \in M_{m\times n}(\mathbb{R}),\ b \in M_{m\times 1}(\mathbb{R}) \equiv \mathbb{R}^m$ and the associated system

$$Ax = b, \tag{77}$$

with the unknown $x \in M_{n\times 1}(\mathbb{R}) \equiv \mathbb{R}^n$. The compatibility of the system holds iff, e.g., rank $A =$ rank $(A|b)$ (the *Kronecker-Capelli theorem*).

To find a pseudo-solution for (77) by using the method of least squares, means to find $x^* \in \mathbb{R}^n$, which obeys the condition

$$||Ax^* - b||^2 = \min_{x\in\mathbb{R}^n} ||Ax - b||^2,$$

where we use the Euclidean norm, $||x|| = \sqrt{\langle x, x\rangle}$.

4.5.1 Theorem. *Let rank $A = n$. Then for all $b \in \mathbb{R}^n$, the system $Ax = b$ has a unique pseudo-solution x^* given by*

$$x^* = \tilde{A}b,$$

where we denoted by $\tilde{A} = ({}^tA\cdot A)^{-1}$.

The matrix $\tilde{A}^{-1}$ is called the pseudo-inverse matrix of A.

Remark. Since $A \in M_{m\times n}(\mathbb{R})$, we get ${}^tA\,A \in M_n(\mathbb{R})$ and $\tilde{A} \in M_{n\times m}(\mathbb{R})$.

Proof. Let x^* be the requested pseudo-solution, and denote $r^* = b - Ax^* \in \mathbb{R}^m$. From the theory of least squares we know that r^* is orthogonal to the set of vectors

$$\text{Im } A = \{y \mid \exists x \in \mathbb{R}^n,\ y = Ax\} \subset \mathbb{R}^m.$$

Therefore

$$0 = \langle Ax, r^*\rangle_{\mathbb{R}^m} = {}^t(Ax)r^* = {}^tx\cdot{}^tA\cdot r^*, \text{ for all } x \in \mathbb{R}^n \ \Rightarrow\ {}^tAr^* = 0.$$

Hence

$$0 = {}^tAr^* = {}^tAb - {}^tAAx^* \ \Rightarrow\ ({}^tAA)x^* = {}^tAb. \tag{78}$$

We note that the matrix ${}^tA\,A$ is quadratic (see the remark), symmetrical and positive definite. Indeed

$${}^t({}^tAA) = {}^tA^t({}^tA) = {}^tAA,$$

and

$${}^tx^tAAx = {}^t(Ax)(Ax) = \langle Ax, Ax\rangle_{\mathbb{R}^m} = ||Ax||^2_{\mathbb{R}^m} \geq 0,$$

and the equality holds true only when $x = 0$. Also, ${}^tAA$ is invertible (since rank $A = n$). Hence from (78) we infer $x^* = ({}^tAA)^{-1\,t}Ab$. □

Remarks. 1°. Emerging from the system $Ax = b$ we obtain the *normal associated system*

$${}^tAAx^* = {}^tAb, \tag{79}$$

which has the solution x^*. Then

$$||b - Ax^*||^2 = \min_{x\in\mathbb{R}^n} ||b - Ax||^2,$$

and x^* approximates x, i.e., $x^* \simeq x$.

The system (79) can be solved using numerical analysis and computer techniques. Unfortunately, $B = {}^tAA$ is affected by rounding errors. For example, for

$$A = \begin{pmatrix} 1 & 1+\varepsilon \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$$

and $\varepsilon > \varepsilon_m$, where ε_m is the computer precision error, we obtain

$$B = {}^tAA = \begin{pmatrix} 3 & 3+\varepsilon \\ 3+\varepsilon & 3+\varepsilon+\varepsilon^2 \end{pmatrix},$$

and if $\varepsilon^2 < \varepsilon_m$ though rank $A = 2$, we can have ${}^tAA$ singular matrix (!) hence without inverse, and x^* cannot be determined.

2° The normal system (79) can still be solved by special numerical methods (avoiding the laborious calculations required for computing $({}^tAA)^{-1}$, for $n >> 0$).

4.5.2. Exercises

1. Find the *pseudo-solution* for the linear system

$$Ax = B \Leftrightarrow \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}.$$

Solution. We have

$$\tilde{A} = {}^tAA = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & 6 \end{pmatrix}, \quad \tilde{A}^{-1} = \begin{pmatrix} \frac{1}{5} & 0 \\ 0 & \frac{1}{6} \end{pmatrix}$$

and rank $\tilde{A} = 2$. Also, the pseudosolution is

$$\begin{aligned} x^* = ({}^tAA)^{-1}\, {}^tAb \quad &= \begin{pmatrix} \frac{1}{5} & 0 \\ 0 & \frac{1}{6} \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = \\ &= \begin{pmatrix} \frac{1}{5} & 0 \\ 0 & \frac{1}{6} \end{pmatrix} \begin{pmatrix} 5 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \end{aligned}$$

Remark that the initial system $\begin{cases} x + 2y = 1 \\ y = 0 \\ 2x - y = 2 \end{cases}$ has the exact solution $\bar{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, which coincides with the pseudosolution, $(\bar{x} = x^*)$.

♣ Hw. Same problem, for $b = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$. Remark that in this case the system is incompatible, hence it has no exact solution.

2. Find the pseudo-solution of the system $Ax = b$, where

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ \varepsilon & 0 & 0 & 0 \\ 0 & \varepsilon & 0 & 0 \\ 0 & 0 & \varepsilon & 0 \\ 0 & 0 & 0 & \varepsilon \end{pmatrix} \in M_{5\times 4}(\mathbb{R}),\ b = \begin{pmatrix} 10 \\ \varepsilon \\ 2\varepsilon \\ 3\varepsilon \\ 4\varepsilon \end{pmatrix}.$$

Solution. The system is obviously compatible for all $\varepsilon \in \mathbb{R}^*$, and its exact solution is $x = {}^t(1, 2, 3, 4)$. Still, in certain cases the pseudosolution of the system cannot be determined, as shown below.

We attach the normal associated system ${}^tAAx^* = {}^tAb$. We remark that though rank $(A) = 4$, and ${}^tAA \in M_4(\mathbb{R})$, we have

$$\det({}^tAA) = 4\varepsilon^6 + \varepsilon^8 \underset{\varepsilon\to 0}{\longrightarrow} 0.$$

That is why, for $\varepsilon \ll 1$, the matrix ${}^tAA$ is practically singular, and in our case, the matrix ${}^tAA$ is non-invertible; the pseudo-solution x^* can be determined only by special techniques.

3. Find the real numbers $\{x_k\}_{\overline{1,n}} \subset \mathbb{R}$ which are the closest to the corresponding numbers $\{a_k\}_{\overline{1,n}}$ respectively, and are closest each to the other.

Solution. We consider the quadratic form

$$Q(x) = \sum_{k=1}^{n} c_k(x_k - x_{k-1})^2 + \sum_{k=1}^{n} d_k(x_k - a_k)^2,$$

where $x_0 = x_n,\quad c_k, d_k \in \mathbb{R},\ k = \overline{1,n}$. This positively defined form attends its minimum at the solution point $x^* = (x_k^*)_{k=\overline{1,n}}$, which is obtained by solving the attached linear system

$$\frac{\partial Q}{\partial x^k}(x) = 0,\ k = \overline{1,n}.$$

♣ Hw. Apply the procedure for $n = 3$ and $a_1 = 1, a_2 = 0, a_3 = -1;\quad c_1 = 0, c_2 = c_3 = 1;\quad d_1 = d_2 = d_3 = 1$.

4.6 Numerical computation of eigenvectors and of eigenvalues

Let $A \in M_n(\mathbb{R})$, let $\lambda \in \sigma(A)$ be an eigenvalue of A and let $x \neq 0$ be an eigenvector associated to λ. These mathematical objects are connected by the relation

$$Ax = \lambda x \ \Leftrightarrow\ (A - \lambda I)x = 0.$$

The solutions of the *characteristic equation*

$$P_A(\lambda) \equiv \det(A - \lambda I) = 0, \quad \lambda \in \mathbb{C}$$

form the complex spectrum

$$\sigma(A) = \{\lambda | P_A(\lambda) = 0\} \subset \mathbb{C}.$$

The attempt of solving the algebraic equation $P_A(\lambda) = 0$ by numerical methods usually fails (these methods are unstable, since small variations in the coefficients of P infer large variations in the solution λ). We shall describe other ways of finding the eigenvalues.

4.6.1. The Schur canonic form.

Definition. The matrices $A, B \in M_n(\mathbb{R})$ are called *orthogonally similar* iff $\exists Q \in M_n(\mathbb{R})$, Q *orthogonal matrix*, (i.e., ${}^tQ = Q^{-1}$, equivalent to $Q\,{}^tQ = I$), such that $B = {}^tQ\,AQ$.

Remarks. a) If A, B are orthogonally similar then $\sigma(A) = \sigma(B)$, i.e., they have the same eigenvalues.

b) If A, B are orthogonally similar, then the corresponding eigenvectors are connected via $v_B = Qv_A$.

Theorem. *Let* $A \in M_n(\mathbb{R})$ *and let*

$$O(n) \equiv \{A | A \in M_n(\mathbb{R}) \mid {}^tA\,A = I\}$$

be the set of orthogonal matrices of order n. *Then there exists an orthogonal matrix* Q, *such that the matrix* $S = {}^tQ\,AQ$ *to be upper almost triangular matrix.*

In this case S is called the *canonical Schur form of* A, and its columns are called *Schur columns.*

Proof. Let $\lambda \in \sigma(A)$ and $x \in \mathbb{R}^n \backslash \{0\}$ an eigenvector associated to λ, i.e., $Ax = \lambda x$. We can assume $||x|| = 1$ (without loss of generality). Let $Q = [x, Y]$ be an orthogonal matrix, $Y \in M_{n\times(n-1)}(\mathbb{R})$ such that $\langle x, y\rangle = 0$, for all $y \in \text{Im}\,\tilde{Y}$, where the matrix Y defines a linear transformation $\tilde{Y} \in L(\mathbb{R}^{n-1}, \mathbb{R}^n)$, such that $[\tilde{Y}] = {}^tY$. Hence we have ${}^tY x = 0$. Then

$$AQ = A[x, Y] = [Ax, AY] = [\lambda x, AY]. \tag{80}$$

But since

$$AQ = IAQ = (Q\,{}^tQ)AQ = Q({}^tQAQ), \tag{81}$$

denoting $A' = {}^tQAQ$, we obtain from (80) and (81) that $[\lambda x, AY] = QA'$, whence, multiplying by $Q^{-1} = {}^tQ = \begin{pmatrix} {}^tx \\ {}^tY \end{pmatrix}$, we infer

$$A' = {}^tQAQ = \begin{pmatrix} \lambda & {}^txAY \\ 0 & B \end{pmatrix},$$

where we denoted $B = {}^tYAY \in M_{n-1}(\mathbb{R})$. Moreover, the reduction of the order of A occurs also for $\lambda_{1,2} = \alpha + i\beta$ eigenvalues with $\beta \neq 0$. If $x_{1,2} = u \pm iv \in \mathbb{C}^n$ are two conjugate associated eigenvectors, then consider $Q = [X, Y]$, where X is the matrix associated to an orthonormal basis of the subspace $L(u, v)$, and Y is a matrix of completion. Then $AX = XM$, where M is of order 2 with eigenvalues $\alpha \pm i\beta$, and hence we get

$$A' = {}^tQAQ = \begin{pmatrix} M & N \\ 0 & C \end{pmatrix}, \quad C \in M_{n-2}(\mathbb{R}).$$

Since the process can be iterated for $B \in M_{n-1}(\mathbb{R})$, using induction one obtains after at most $n-1$ steps *the Schur normal matrix* associated to A. □

4.6.2. The Power Method. Let $A \in M_n(\mathbb{R})$, having *simple* eigenvalues

$$\sigma(A) = \{\lambda_i\}_{\overline{1,n}} \subset \mathbb{R}.$$

Let x_i be eigenvectors associated to λ_i, $i = \overline{1,n}$ respectively, i.e., $Ax_i = \lambda_i x_i$. Then $B = \{x_i | i = \overline{1,n}\} \subset \mathbb{R}^n$ represents a basis of $\mathbb{R}^n$. Let $y \in \mathbb{R}^n$ be a unit vector ($||y|| = 1$), which decomposes w.r.t this basis as

$$y = \sum_{i=1}^{n} \beta_i x_i. \tag{82}$$

Then $Ay = \sum_{i=1}^{n} \beta_i A x_i = \sum_{i=1}^{n} \beta_i \lambda_i x_i$, whence

$$A^k y = \sum_{i=1}^{n} \beta_i \lambda_i^k x_i. \tag{83}$$

Theorem. *Let $\sigma(A) = \{\lambda_1, \ldots, \lambda_n\} \subset \mathbb{R}$, such that*

$$|\lambda_1| > |\lambda_2| > \ldots > |\lambda_n|,$$

and let the coefficient β_1 in (82) satisfy $\beta_1 \neq 0$. Then $A^k y$ converges in direction to x_1, i.e., denoting $y_* = \lim_{k\to\infty} A^k y$, there exists $\alpha \in \mathbb{R}\backslash\{0\}$ such that $y_* = \alpha x_1$.

Proof. Dividing the relation (83) by λ_1^k, and using

$$|\lambda_i/\lambda_1| < 1, \quad \text{for all } i \in \overline{2,n},$$

we infer

$$\frac{1}{\lambda_1^k} A^k y = \beta_1 x_1 + \sum_{i=2}^{n} \left(\frac{\lambda_i}{\lambda_1}\right)^k \beta_i x_i \quad \Rightarrow \quad \lim_{k\to\infty} \frac{1}{\lambda_1^k} A^k y = \beta_1 x_1. \tag{84}$$

Let now $y^{(0)} \in \mathbb{R}^n$ be the initial seed of the vector sequence

$$y^{(k)} = A^k y^{(0)}, \ \ k \geq 1.$$

This sequence converges also to a vector collinear to x_1, i.e., for some $\alpha \in \mathbb{R}$, we have

$$\lim_{k\to\infty} y^{(k)} = \alpha x_1.$$

The relation $Ax = \lambda x, \ x \neq 0$ infers $\lambda = \dfrac{\langle x, Ax\rangle}{\langle x, x\rangle}$. Hence, using (84) we have

$$\lim_{k\to\infty} \frac{{}^t y^{(k)} A y^{(k)}}{{}^t y^{(k)} y^{(k)}} = \lim_{k\to\infty} \frac{\langle y^{(k)}, Ay^{(k)}\rangle}{\langle y^{(k)}, y^{(k)}\rangle} = \frac{\langle \alpha x_1, A(\alpha x_1)\rangle}{\langle \alpha x_1, \alpha x_1\rangle} = \frac{\langle x_1, Ax_1\rangle}{\langle x_1, x_1\rangle} = \lambda_1,$$

where we denoted $\lim\limits_{k\to\infty} y^{(k)} = \alpha x_1, \ \alpha \in \mathbb{R}$, and

$$\lim_{k\to\infty} \frac{{}^t y^{(k)} A y^{(k)}}{{}^t y^{(k)} y^{(k)}} = \lambda_1. \qquad \square$$

Remark. When finding the *direction* of x_1 by (84) in concrete computer calculation, if $\alpha x_1 = 0$, then a change in choosing the initial vector $y^{(0)}$ must happen, and the process has to be re-iterated. If all the components of $y^{(k)}$ converge to 0, usually the new vector $y^{(0)}$ is chosen with one component strictly greater than unit.

4.6.3. Exercises

1. Let $A = \begin{pmatrix} 4 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}$. Find four iterations in determining the eigenvalue of maximal absolute value $\lambda \in \sigma(A)$ (i.e., such that $|\lambda| \geq |\mu|, \ \ \text{for all } \mu \in \sigma(A)$), and find the corresponding four estimates of the eigenvector x associated to λ. The seed vector of the iterations is $y^{(0)} = {}^t(1, 1, 1) \in \mathbb{R}^3$.

Solution. The approximative values of the eigenvalues of A are

$$\sigma(A) = \{\lambda_1 \cong 4.4605, \ \lambda_2 \cong 2.23912, \ \lambda_3 \cong .300372\}.$$

The sequence of vectors $y^{(k)} = A^k y^{(0)}, \ k \in \mathbb{N}$, which converges in direction to the eigenvector x_1 is

Component\vector	$y^{(0)}$	$y^{(1)}$	$y^{(2)}$	$y^{(3)}$	$y^{(4)}$	$\cdots$
1	1	5	24	111	504	$\cdots$
2	1	4	15	60	252	$\cdots$
3	1	2	6	21	81	$\cdots$

Denoting $v_k = y^{(k)} = A^k y^{(0)} = Ay^{(k-1)}$, the sequence $\eta_k = \dfrac{{}^t v_k A v_k}{{}^t v_k v_k}$ which converges to λ_1 has the first five iterations given by

$$\eta_0 = \frac{11}{3} \approx 3.6666, \ \eta_1 = \frac{64}{15} \approx 4.26666, \ \eta_2 = \frac{410}{93} \approx 4.40860,$$

$$\eta_3 = \frac{2695}{606} \approx 4.444719,\ \eta_4 = \frac{17833}{4001} \approx 4.45713.$$

Also, one can use the obtained limit value λ_1, and remark that the *normalized* sequence of vectors

$$z^{(k)} = \frac{v_k}{\lambda_1^k},\ k \in \mathbb{N}$$

converges also in direction to x_1, and the first four iterations of this sequence are

Component\vector	$z^{(0)}$	$z^{(1)}$	$z^{(2)}$	$z^{(3)}$	$z^{(4)}$	$\cdots$
1	1	1.12094	1.20626	1.25075	1.27319	$\cdots$
2	1	.896759	0.753916	.676081	.636596	$\cdots$
3	1	.448379	.301566	.236628	.204620	$\cdots$

2. Let $A = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}$. Find $\lambda \in \sigma(A)$, such that $|\lambda| = \max_{\gamma \in \sigma(A)} |\gamma|$. Find its eigenvectors.

Solution. The first seven iterations for the sequence $y^{(k)} = Ay^{(k)} = A^k y^{(0)}, k \geq 1$ are

Component\vector	$y^{(0)}$	$y^{(1)}$	$y^{(2)}$	$y^{(3)}$	$y^{(4)}$	$y^{(5)}$	$y^{(6)}$	$y^{(7)}$	$\ldots$
1	1	0	-2	-6	-14	-30	-62	-126	$\ldots$
2	1	2	4	8	16	32	64	128	$\ldots$

This converges to $v_{\max}$. Also, if $\lambda_{\max}$ is known, then the sequence $z^{(k)}$ below has the same property:

$$z^{(0)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},\quad z^{(k)} = \frac{1}{\lambda^k} A^k y^{(0)} \to v_{\max}.$$

The eigenvalue $\lambda_{\max}$ is also provided by the sequence

$$\eta_k = \frac{{}^t y^{(k)} A y^{(k)}}{{}^t y^{(k)} y^{(k)}} \to \lambda_{\max},$$

whose first three iterations look like

$$\begin{aligned} &\eta_0 = 1, && \eta_1 = 2, \\ &\eta_2 = \tfrac{11}{5} = 2.2, && \eta_3 = \tfrac{53}{25} = 2.12, \\ &\eta_4 = \tfrac{233}{113} \approx 2.06194, && \eta_5 = \tfrac{977}{481} \approx 2.03118, \\ &\eta_6 = \tfrac{4001}{1985} \approx 2.01561, && \eta_7 = \tfrac{16193}{8065} \approx 2.00781. \end{aligned}$$

We note that $\sigma(A) = \{1, 2\}$, and the largest in absolute value eigenvalue of A is $\lambda = 2$. Then it can be easily seen that the sequence η_k approaches at limit this eigenvalue.

Bibliography

[1] I.S.Antoniu, *Calculul matricial şi tensorial în electrotehnică* (in Romanian), Ed. Tehnică, Bucharest, 1962.

[2] T.M.Apostol, *Calculus*, vol. II, Blaisdell Publishing Company, 1969.

[3] V.Balan, *Algebră liniară, geometrie analitică* (in Romanian), Fair Partners Eds., 1999.

[4] V.Balan, I-R.Nicola, *Algebră liniară, geometrie analitică şi diferenţială, ecuaţii diferenţiale* (in Romanian), Bren Eds., Bucharest2004, 2006.

[5] I.Beju, O.Soos, P.P.Teodorescu, *Tehnici de calcul spinorial şi tensorial neeuclidian cu aplicaţii* (in Romanian), Ed. Tehnică, Bucharest, 1968.

[6] M.Berger, B.Gostiaux, *Géométrie différentielle*, Armand Colin, Paris, 1972.

[7] G.Berman, *Problèmes d'analyse mathématique*, Éditions de Moscou, 1976.

[8] R.M.Bowen, C.C.Wang, *Introduction to Vectors and Tensors*, vol. 1-2, Plenum Press, New York, 1976.

[9] F.Brickell, R.S.Clark, *Differentiable Manifolds*, Van Nostrand Reinhold Company, London, 1970.

[10] B.M.Budak, S.V.Fomin, *Multiple Integrals. Field Theory and Series*, Mir Publishers, Moskow, 1973.

[11] S.Chiriţă, *Probleme de matematici superioare* (in Romanian), Ed. Did. Ped., Bucharest, 1989.

[12] C.Drăguşin, L.Drăguşin, *Analiză matematică*, vol. II, (in Romanian), Ed. Matrix Rom, Bucharest, 1999.

[13] B.Dubrovin, S.Novikov, A.Fomenko, *Geometria Contemporanea*, Editori Riuniti, Edizioni Mir, 1987.

[14] G.M.Fihtenholţ, Curs de calcul diferenţial şi integral, vol. 3 (in Romanian), Ed. Tehnică, Bucharest, 1965.

[15] Gh. Gheorghiev, V. Oproiu, *Geometrie diferenţială* (in Romanian), Ed. Did. Ped., Bucharest, 1977.

[16] I.Glazman, Iu.Liubici, Analiză liniară pe spaţii finit dimensionale (in Romanian), Ed. Şt. Encicl., Bucharest, 1980; Nauka Eds (in Russian), 1969.

[17] J.N.Harris, H.Stocker, Handbook of Mathematics and Computational Sciences, Springer Verlag, 1998.

[18] S.Helgason, *Differential Geometry and Symmetric Spaces*, Academic Press, 1962.

[19] R.Hermann, *Topics in Physical Geometry*, Topics in Physical Geometry, v.26, Mathematical Science Press, Brooklin, 1988.

[20] S.Ianuş, *Geometrie diferenţială cu aplicaţii în teoria relativităţii* (in Romanian), Academy Eds., Bucharest, 1983.

[21] W.Klingenberg, *Lineare Algebra und Geometrie*, Springer-Verlag, Berlin, 1990.

[22] N.E.Kocin, *Calculul vectorial şi introducere în calculul tensorial* (in Romanian), Ed. Tehnică, Bucharest, 1954.

[23] I.A.Kostrikin, I.Yu.Manin, *Linear Algebra and Geometry*, Gordon and Breach Science Publishers, 1989.

[24] N.I.Kovantzov, G.M.Zrajevskaya, V.G.Kotcharovski, V.I.Mihailovski, *Problem Book in Differential Geometry and Topology* (in Russian), Visha Shkola Eds., Kiev, 1989.

[25] S.Lang, *Differential and Riemannian Manifolds*, GTM 160, Springer Verlag, 1998.

[26] A.Lichnerowicz, *Algebrè et analyse linéaire*, Masson & Cie, Paris, 1960.

[27] R.Millman, G.Parker, *Elements of Differential Geometry*, Prentice- Hall, Inc., New Jersey, 1977.

[28] A.S.Mischenko, Iu.P.Soloviev, A.T.Fomenko, *Problem book on Differential Geometry and Topology* (in Russian), University of Moskow, 1981.

[29] Gh.Munteanu, V.Balan, *Lecţii de teoria relativităţii* (in Romanian), Bren Eds., Bucharest, 2000.

[30] A.W.Naylor, G.R.Sell, *Linear operator theory in engineering and science*, Springer-Verlag, 1982.

[31] L.Nicolescu, *Lecţii de geometrie* (in Romanian), Univ. Bucureşti, Bucharest, 1990.

[32] V.Olariu, V.Prepeliţă, *Teoria distribuţiilor, funcţii complexe şi aplicaţii*, Ed. Şt. Encicl., Bucharest, 1986.

[33] B. O'Neill, *Elementary Differential Geometry*, Academic Press, Fourth Printing, 1970.

[34] P.V.O′Neill, Advanced Engineering Mathematics, Wadsworth Eds., 1991.

[35] D. Papuc, *Geometrie diferenţială* (in Romanian), Ed. Did. Ped., Bucharest, 1982.

[36] D. Papuc, I. D. Albu, *Elemente de geometrie diferenţială globală* (in Romanian), cap. I, II. 1, II. 4, VII. 2, Ed. Did. Ped., Bucharest, 1973,

[37] G.Pavel, F.I.Tomiţă, I.Gavrea, *Matematici speciale. Aplicaţii* (in Romanian), Ed. Dacia, Cluj-Napoca, 1981.

[38] I. Popescu, *Probleme de matematici superioare* (in Romanian), Ed. Did. Ped., Bucharest, 1964.

[39] C.Radu, C.Drăguşin, L.Drăguşin, *Algebră liniară, analiză matematică, geometrie analitică şi diferenţială (culegere de probleme)* (in Romanian), Ed.Fair Partners, Bucharest, 2000.

[40] C.Radu, C.Drăguşin, L.Drăguşin, *Aplicaţii de algebră, geometrie şi matematici speciale* (in Romanian), Ed. Did. Ped., Bucharest, 1991.

[41] L.D.Raigorodski, P.C.Stavrinos, V.Balan, *Introduction to the Physical Principles of Differential Geometry*, University of Athens, 1999.

[42] W.Rudin, *Analiză reală şi complexă* (in Romanian), Ed. Theta, Bucharest, 1999.

[43] V.Rudner, C.Nicolescu, *Probleme de matematici speciale* (in Romanian), Ed. Did. Ped., Bucharest, 1982.

[44] B.Spain, *Tensor Calculus*, Central Book Co., Dublin, 1953.

[45] Şt. Staicu, *Aplicaţii ale calculului matriceal în mecanica solidelor* (in Romanian), Academy Eds., Bucharest, 1986.

[46] Şt. Staicu, *Mecanică teoretică* (in Romanian), Ed. Did. Ped., Bucharest, 1998.

[47] L.Stoica, *Elements of Differentiable Manifolds*, Geometry Balkan Press, Bucureşti, Romania, 1998.

[48] K.Teleman, *Introducere în geometria diferenţială* (in Romanian), Univ. Bucureşti, Bucharest, 1986.

[49] J.A.Thorpe, *Elementary Topics in Differential Geometry*, Springer-Verlag, 1979.

[50] C.Udrişte, *Linii de câmp* (in Romanian), Ed. Tehnică, Bucharest, 1988.

[51] C.Udrişte, *Geometric Dynamics*, Kluwer Academic Publishers, 2000.

[52] C.Udrişte, *Problems of Algebra, Geometry and Differential Equations I, II*, University Politehnica of Bucharest, Bucharest, 1992.

[53] C.Udrişte, *Aplicaţii de algebră, geometrie şi ecuaţii diferenţiale* (in Romanian), Ed. Did. Ped., Bucharest, 1993.

[54] C.Udrişte, *Algebră liniară, geometrie analitică* (in Romanian), Geometry Balkan Press, Bucharest, 2000.

[55] C.Udrişte, *Geometrie diferenţială, ecuaţii diferenţiale* (in Romanian), Geometry Balkan Press, Bucharest, 1997.

[56] C.Udrişte, *Problems of Linear Algebra, Analytic and Differential Geometry and Differential Equations*, Geometry Balkan Press, Bucharest, 2000.

[57] C.Udrişte, V.Iftode, M.Postolache, *Metode numerice de calcul* (in Romanian), Ed. Tehnică, Bucharest, 1996.

[58] C.Udrişte, C.Radu, C.Dicu, O.Mălăncioiu, *Probleme de algebră, geometrie şi ecuaţii diferenţiale* (in Romanian), Ed. Did. Ped., Bucharest, 1981.

[59] C.Udrişte, C.Radu, C.Dicu, O.Mălăncioiu, *Algebră, geometrie şi ecuaţii diferenţiale* (in Romanian), Ed. Did. Ped., Bucharest, 1982.

[60] E.C.Young, *Vector and Tensor Analysis*, Pure and Applied Mathematics 172, M.Dekker, 1992.