Istituto di Scienza e Tecnologie
dell'Informazione "A. Faedo"
Consiglio Nazionale delle Ricerche

ISTI Technical Reports

Introduction to tensor calculus

Cristina Padovani, ISTI-CNR, Pisa, Italy

ISTI-TR-2021/003

Introduction to tensor calculus
Padovani C.
ISTI-TR-2021/003

Introduction to Tensor Calculus - Lecture Notes - International Doctorate in Civil and Environmental Engineering, Universities of Firenze, Pisa and Braunschweig, February 2021.

Keywords: Finite-dimensional vector spaces, Tensor algebra, Tensor analysis.

Citation

Padovani C., Introduction to tensor calculus. ISTI Technical Reports 2021/003. DOI: 10.32079/ISTI-TR-2021/003.

Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo"
Area della Ricerca CNR di Pisa
Via G. Moruzzi 1
56124 Pisa Italy
http://www.isti.cnr.it

Introduction to Tensor Calculus

Cristina Padovani [1]
International Doctorate in Civil and Environmental Engineering
Universities of Firenze, Pisa and Braunschweig

February 2021

[1]Istituto di Scienza e Tecnologie dell'Informazione "Alessandro Faedo" - Consiglio Nazionale delle Ricerche, Pisa.

Contents

Chapter 1

Finite-dimensional vector spaces

1.1 Notations

Let X be a set, we write $x \in X$ for the statement "x is an element of X" and $x \notin X$ for the statement "x is not an element of X". If Y is a subset of X we write $Y \subset X$. Given A and B subsets of X, we define the following subsets of X,

$$A \cup B = \{x \in X \mid x \in A \text{ or } x \in B\}, \tag{1.1}$$

$$A \cap B = \{x \in X \mid x \in A \text{ and } x \in B\}, \tag{1.2}$$

called the *union* and the *intersection* of A and B. We denote by $\emptyset$ the empty set; two sets A and B are *disjoint* if $A \cap B = \emptyset$.

If A is a subset of X, the difference

$$X - A = \{x \in X \mid x \in X \text{ and } x \notin A\}, \tag{1.3}$$

is the *complement* of A (in X) and is denoted by $\complement A$.

Let X_1 and X_2 be two sets. The set of the ordered pairs (x_1, x_2), with $x_1 \in X_1$ and $x_2 \in X_2$, is the *Cartesian product* of X_1 and X_2; it is denoted by $X_1 \times X_2$.

We denote by $\mathbb{N} = \{0, 1, 2, ...\}$ the set of natural numbers, by $\mathbb{Q}$ the set of rational numbers and by $\mathbb{R}$ the set of real numbers.

Let X and Y be two nonempty sets, a *function* T from X to Y (or *mapping* on X into Y) is a rule that assigns to each $x \in X$ a unique element $y \in Y$,

$$T : X \to Y \tag{1.4}$$

we denote by $T(x)$ the element y called the *image* of x under T.

Let A be a subset of X, the set

$$T(A) = \{v \in Y \mid v = T(u) \text{ for some } u \in A\} \tag{1.5}$$

is the *image* of A under T ($T(\emptyset) = \emptyset$).

Let B be a subset of Y, the set

$$T^{-1}(B) = \{u \in X \mid T(u) \in B\} \tag{1.6}$$

is the *inverse image* of B ($T^{-1}(\emptyset) = \emptyset$).

The function $T : X \to Y$ is *injective* (or one-to-one) if

$$u_1 \neq u_2 \Longrightarrow T(u_1) \neq T(u_2), \tag{1.7}$$

and is *surjective* (or onto Y) if for each $w \in Y$ there exists (at least) $u \in X$ such that $w = T(u)$, in this case $T(X) = Y$. A function T which is both injective and surjective is *bijective*.

Let X be a set. A *distance* (o *metric*) on X is a function d on the Cartesian product $X \times X$ with real values,

$$d : X \times X \to \mathbb{R} \tag{1.8}$$

such that for each $x, y, z \in X$:

d1. $d(x, y) \geq 0$,

d2. $d(x, y) = d(y, x)$,

d3. $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality)

d4. $d(x, y) = 0$ if and only if $x = y$.

The real number $d(x, y)$ is the *distance between x and y*. A set X with the distance d is called *metric space* and is usually denoted by (X, d). The elements of X are called *points*.

Conditions **d1** and **d4** are quite natural and intuitive. Condition **d3** generalizes the triangle inequality for the triangles in the Euclidean space and has important consequences, in particular it allows to prove that the limit of a convergent sequence in a metric space is unique.

Two different metrics d and d' on the same set X define different metric spaces (X, d) and (X, d').

Proposition 1. *Let X be a set with the metric d; for each $x, y, z \in X$ we have*

$$|d(x, z) - d(y, z)| \leq d(x, y). \tag{1.9}$$

Proof. From the triangle distance in **d3** it follows that

$$d(x, z) - d(y, z) \leq d(x, y),$$

Changing x with y and taking **d2** into account, we have

$$d(y, z) - d(x, z) \leq d(x, y),$$

and (1.9) follows. □

Let $\mathbb{R}^n$ be the set of the n-tuples of ordered real numbers $\mathbf{x} = (x_1, x_2, ..., x_n)$, the function

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}, \tag{1.10}$$

with $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^n$ is a metric called *Euclidean distance*. The conditions **d1**, **d2** and **d4** are easy to prove; relation **d3** follows from the inequality

$$\sqrt{\sum_{i=1}^{n} (a_i + b_i)^2} \leq \sqrt{\sum_{i=1}^{n} a_i^2} + \sqrt{\sum_{i=1}^{n} b_i^2}, \tag{1.11}$$

putting $a_i = x_i - y_i$ and $b_i = y_i - z_i$, $i = 1, ..., n$.

The inequality (1.11) is trivial if $a_i = 0$ or $b_i = 0$ for $i = 1, ..., n$; then let us assume that some a_i and some b_i are different from zero. For each $\lambda > 0$, from the inequalities

$$(\sqrt{\lambda} a_i + \frac{1}{\sqrt{\lambda}} b_i)^2 \geq 0, \quad (\sqrt{\lambda} a_i - \frac{1}{\sqrt{\lambda}} b_i)^2 \geq 0,$$

we get

$$2|a_i b_i| \leq \lambda a_i^2 + \frac{1}{\lambda} b_i^2, \quad i = 1, ..., n \tag{1.12}$$

and then

$$2|\sum_{i=1}^{n} a_i b_i| \leq \lambda \sum_{i=1}^{n} a_i^2 + \frac{1}{\lambda} \sum_{i=1}^{n} b_i^2 \tag{1.13}$$

follows. The two addends in the right-hand side are equal for $\lambda = \sqrt{\sum_{i=1}^{n} b_i^2} / \sqrt{\sum_{i=1}^{n} a_i^2}$ and for this value of λ equation (1.13) becomes

$$|\sum_{i=1}^{n} a_i b_i| \leq \sqrt{\sum_{i=1}^{n} a_i^2} \sqrt{\sum_{i=1}^{n} b_i^2}, \tag{1.14}$$

which is known as Cauchy-Schwarz inequality. Thus, in view of (1.14), we have

$$\sum_{i=1}^{n} (a_i + b_i)^2 = \sum_{i=1}^{n} a_i^2 + \sum_{i=1}^{n} b_i^2 + 2 \sum_{i=1}^{n} a_i b_i$$

$$\leq \sum_{i=1}^{n} a_i^2 + \sum_{i=1}^{n} b_i^2 + 2|\sum_{i=1}^{n} a_i b_i| \leq \sum_{i=1}^{n} a_i^2 + \sum_{i=1}^{n} b_i^2 + 2\sqrt{\sum_{i=1}^{n} a_i^2} \sqrt{\sum_{i=1}^{n} b_i^2}$$

$$= \left(\sqrt{\sum_{i=1}^{n} a_i^2} + \sqrt{\sum_{i=1}^{n} b_i^2} \right)^2$$

which coincides with (1.11).

Let X be a set ($X \neq \emptyset$), the function

$$d(x,y) = \begin{cases} 0 & x = y, \\ 1 & x \neq y, \end{cases} \tag{1.15}$$

is a metric called *discrete metric*.

Let X be the set of all possible sequences of k bits, each element of X is constituted by a string $x = x_1 x_2 ... x_k$ of k symbols with $x_i \in \{0,1\}$, $i = 1, ..., k$. We define the distance between two strings x and y of X as the number of positions at which the corresponding symbols are different. This distance, called *Hamming distance*, measures the number of substitutions needed to convert a string in the other, or, equivalently, the number of errors that have transformed a string in the other. For example, for $k = 6$, given $x = 001001$ and $y = 000011$, we have $d(x,y) = 2$.

The functions

$$d_1(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} |x_i - y_i|, \tag{1.16}$$

and

$$d_\infty(\mathbf{x}, \mathbf{y}) = \max_{i=1,\dots,n} |x_i - y_i|, \tag{1.17}$$

with $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^n$ are distances in $\mathbb{R}^n$

In the set

$$C[a,b] = \{f : [a,b] \to \mathbb{R} \mid \text{f is continuous in } [a,b]\}, \tag{1.18}$$

the functions

$$d_\infty(f,g) = \max_{t \in [a,b]} |f(t) - g(t)|, \tag{1.19}$$

and

$$d_1(f,g) = \int_a^b |f(t) - g(t)| dt, \tag{1.20}$$

with f, $g \in C[a,b]$ are distances.

Let A be a subset of $\mathbb{R}$, $b \in \mathbb{R}$ is an *upper bound* for A if $a \leq b$, for each $a \in A$. In this case A is *bounded from above*. We define the *least upper bound* or *supremum* of A, denoted by $\sup A$, as the minimum s of the upper bounds of A. The supremum s is characterized by the following properties,

$$a \leq s, \quad \text{for each } a \in A, \tag{1.21}$$

$$\text{for aech } \varepsilon > 0 \text{ there exists } a \in A \text{ such that } a > s - \varepsilon. \tag{1.22}$$

$c \in \mathbb{R}$ is a *lower bound* for A if $a \geq c$, for each $a \in A$. In this case A is *bounded from below*. We define the *greatest lower bound* or *infimum* of$A$, denoted by $\inf A$, as the maximum i of the lower bounds of A. The infimum$i$ is characterized by the following properties,

$$a \geq i, \quad \text{for each } a \in A, \tag{1.23}$$

$$\text{for each } \varepsilon > 0 \text{ there exists } a \in A \text{ such that } a < i + \varepsilon. \tag{1.24}$$

1.2 Vector spaces

A (real) *vector space* is a set $\mathcal{S}$ of elements called *vectors* satisfying the following axioms.

(A) To every pair, $\mathbf{a}$ and $\mathbf{b}$, of vectors in $\mathcal{S}$ there corresponds a vector $\mathbf{a}+\mathbf{b}$, called the *sum* of $\mathbf{a}$ and $\mathbf{b}$, in such a way that

1. addition is commutative, $\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$,
2. addition is associative, $\mathbf{a}+(\mathbf{b}+\mathbf{c})=(\mathbf{a}+\mathbf{b})+\mathbf{c}$,
3. there exists in $\mathcal{S}$ a unique vector $\mathbf{0}$ (called the origin) such that $\mathbf{a}+\mathbf{0}=\mathbf{a}$ per for every vector $\mathbf{a}$,
4. to every vector $\mathbf{a}$ in $\mathcal{S}$ there corresponds a unique vector $-\mathbf{a}$ such that $\mathbf{a}+(-\mathbf{a})=\mathbf{0}$.

(B) To every pair, α and $\mathbf{a}$, where α is a real number and $\mathbf{a}$ is a vector in $\mathcal{S}$, there corresponds a vector $\alpha\mathbf{a}$, called the *product* of α and $\mathbf{a}$, in such a way that

1. multiplication by scalars is associative, $\alpha(\beta\mathbf{a})=(\alpha\beta)\mathbf{a}$,
2. $1\mathbf{a}=\mathbf{a}$ for every vector $\mathbf{a}$.

(C) The following properties hold

1. multiplication by scalars is distributive with respect to vector addition, $\alpha(\mathbf{a}+\mathbf{b})=\alpha\mathbf{a}+\alpha\mathbf{b}$, for each $\mathbf{a}$, $\mathbf{b}\in\mathcal{S}$, $\alpha\in\mathbb{R}$,
2. multiplication by vectors is distributive with respect to scalar addition, $(\alpha+\beta)\mathbf{a}=\alpha\mathbf{a}+\beta\mathbf{a}$, for each $\mathbf{a}\in\mathcal{S}$, $\alpha,\beta\in\mathbb{R}$.

The sets

$$\mathbb{R}^n=\{\mathbf{x}=(x_1,...,x_n)\mid x_i\in\mathbb{R}, i=1,...,n\}, \tag{1.25}$$

constituted by the n-tuples of real numbers,

$$\mathcal{P}_n=\{p(x)=a_0+a_1x+...+a_nx^n\mid x\in[0,1], a_i\in\mathbb{R}, i=0,...,n\}, \tag{1.26}$$

constituted by the polynomials of degree less than or equal to n and real coefficients,

$$\mathcal{M}_{m,n}=\{A=[a_{ij}]\mid a_{ij}\in\mathbb{R}, i=1,...,m, j=1,...,n\}, \tag{1.27}$$

constituted by the matrices with real coefficients, m rows and n columns, are real vector spaces.

1.3 Norms on a vector space

Given the vector space $\mathcal{S}$, a *norm* is a function $\| \ \|$ on $\mathcal{S}$ into $\mathbb{R}$ such that

n1. $\|\mathbf{a}\| \geq 0$ for all $\mathbf{a} \in \mathcal{S}$,

n2. $\|\mathbf{a}\| = 0$ if and only if $\mathbf{a} = \mathbf{0}$,

n3. $\|\alpha\mathbf{a}\| = |\alpha| \ \|\mathbf{a}\|$ for all $\mathbf{a} \in \mathcal{S}$, $\alpha \in \mathbb{R}$,

n4. $\|\mathbf{a} + \mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|$ for all $\mathbf{a}, \mathbf{b} \in \mathcal{S}$ (triangle inequality).

The vector space $\mathcal{S}$ with the norm $\| \ \|$ is a *normed space*.

On $\mathbb{R}^n$ we can define the following norms

$$\|\mathbf{x}\|_\infty = \max_{i=1,\ldots,n} |x_i|, \tag{1.28}$$

$$\|\mathbf{x}\|_k = \left(\sum_{i=1}^{n} |x_i|^k \right)^{1/k}, \quad \text{with } k \text{ integer, } k \geq 1, \tag{1.29}$$

and on $\mathcal{P}_n$ we can consider the following norms

$$\|f\|_\infty = \max_{x\in[0,1]} |f(x)| \tag{1.30}$$

and

$$\|f\|_k = \left(\int_0^1 |f(x)|^k \ dx \right)^{1/k}, \quad \text{with } k \text{ integer, } k \geq 1. \tag{1.31}$$

Lastly

$$\|A\| = \max_{i=1,\ldots,m} \sum_{j=1}^{n} |a_{ij}|, \tag{1.32}$$

$$\|A\| = \max_{i,j} |a_{ij}|, \tag{1.33}$$

$$\|A\|_F = \left(\sum_{i,j} |a_{ij}|^2 \right)^{1/2}, \tag{1.34}$$

are norms on $\mathcal{M}_{m,n}$. The latter is called Frobenius norm.

From the property **n4**, it follows that

$$| \, \|\mathbf{a}\| - \|\mathbf{b}\| \, | \leq \|\mathbf{a} - \mathbf{b}\|, \quad \text{for each } \mathbf{a}, \mathbf{b} \in \mathcal{S}. \tag{1.35}$$

A normed space $\mathcal{S}$ is a metric space with the distance induced by the norm $\| \ \|$

$$d(\mathbf{a}, \mathbf{b}) = \|\mathbf{a} - \mathbf{b}\|, \quad \mathbf{a}, \mathbf{b} \in \mathcal{S}. \tag{1.36}$$

For example, on $\mathbb{R}^n$ the norm $\|\mathbf{x}\|_2 = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2}$ induces the Euclidean metric (1.10).

It is possible to prove that a distance d on a vector space $\mathcal{S}$ is induced by a norm if and only if

1. d is invariant with respect to translations,
$$d(\mathbf{a}+\mathbf{c}, \mathbf{b}+\mathbf{c}) = d(\mathbf{a}, \mathbf{b}), \quad \text{for each } \mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathcal{S}, \tag{1.37}$$

2. d is invariant with respect to homotheties,
$$d(\lambda\mathbf{a}, \mathbf{0}) = |\lambda| d(\mathbf{a}, \mathbf{0}), \quad \text{for each } \mathbf{a} \in \mathcal{S},\ \lambda \in \mathbb{R}. \tag{1.38}$$

There exist metrics on a vector space $\mathcal{S}$ that are not induced by any norm. For example, on $\mathbb{R}^n$ there is no norm that induces the metric d defined in (1.15), since d does not satisfy (1.38). On $\mathbb{R}^2$ let us consider the Euclidean distance d defined in (1.10), it is easy to prove that

$$d' = \frac{d}{1+d} \tag{1.39}$$

is a distance on $\mathbb{R}^2$, in fact properties **d1**, **d2** and **d4** are easy to prove and, as far as the triangle inequality is concerned, for each $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^2$ we have

$$d'(\mathbf{x},\mathbf{y}) + d'(\mathbf{y},\mathbf{z}) = \frac{d(\mathbf{x},\mathbf{y})}{1+d(\mathbf{x},\mathbf{y})} + \frac{d(\mathbf{y},\mathbf{z})}{1+d(\mathbf{y},\mathbf{z})} \geq \frac{d(\mathbf{x},\mathbf{y})}{1+d(\mathbf{x},\mathbf{y})+d(\mathbf{y},\mathbf{z})}$$
$$+ \frac{d(\mathbf{y},\mathbf{z})}{1+d(\mathbf{x},\mathbf{y})+d(\mathbf{y},\mathbf{z})} \geq \frac{d(\mathbf{x},\mathbf{z})}{1+d(\mathbf{x},\mathbf{z})},$$

because $f(b) = \frac{b}{1+b}$, $b \geq 0$, is an increasing function and d satisfies the triangle inequality. Since d' does not satisfy (1.38), there exists no norm that induces it.

1.4 Inner products

Let $\mathcal{S}$ be a vector space, an *inner product* (or *scalar product*) is a function $<,>$ on $\mathcal{S} \times \mathcal{S}$ into $\mathbb{R}$ such that

s1. $<\mathbf{a}, \mathbf{b}> = <\mathbf{b}, \mathbf{a}>$ for each $\mathbf{a}, \mathbf{b} \in \mathcal{S}$ (symmetry),

s2. $<\alpha_1\mathbf{a}_1 + \alpha_2\mathbf{a}_2, \mathbf{b}> = \alpha_1 <\mathbf{a}_1, \mathbf{b}> + \alpha_2 <\mathbf{a}_2, \mathbf{b}>$ for each $\mathbf{a}_1, \mathbf{a}_2, \mathbf{b} \in \mathcal{S}$, e $\alpha_1, \alpha_2 \in \mathbb{R}$ (bilinearity),

s3. $<\mathbf{a}, \mathbf{a}> \geq 0$ for each $\mathbf{a} \in \mathcal{S}$ (positivity),

s4. $<\mathbf{a}, \mathbf{a}> = 0$ if and only if $\mathbf{a} = \mathbf{0}$.

Vectors $\mathbf{a}, \mathbf{b} \in \mathcal{S}$ are *orthogonal* if $< \mathbf{a}, \mathbf{b} >= 0$. Vectors $\mathbf{u}_1, ..., \mathbf{u}_m \in \mathcal{S}$ are *orthonormal* if

$$< \mathbf{u}_i, \mathbf{u}_j >= \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases} \tag{1.40}$$

In this case we say that $\{\mathbf{u}_1, ..., \mathbf{u}_m\}$ is an *orthonormal set* of vectors in $\mathcal{S}$.

On $\mathbb{R}^n$ we can define the inner product

$$< \mathbf{x}, \mathbf{y} >= \sum_{i=1}^{n} x_i y_i, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n; \tag{1.41}$$

in the vector space $\mathcal{P}_n$ the product

$$< f, g >= \int_0^1 f(x) g(x) dx, \qquad f, g \in \mathcal{P}_n \tag{1.42}$$

is an inner product. On the space $\mathcal{M}_{m,n}$ the product

$$< A, B >= \sum_{\substack{i=1,...,m \\ j=1,...,n}} a_{ij} b_{ij}, \qquad A, B \in \mathcal{M}_{m,n} \tag{1.43}$$

is an inner product. In $\mathbb{R}^3$ the vectors

$$\mathbf{x}^1 = (1,0,0), \quad \mathbf{x}^2 = (0,1,0), \quad \mathbf{x}^3 = (0,0,1) \tag{1.44}$$

are orthonormal. In $\mathcal{P}_n$ the polynomials $f(x) = 1$, $g(x) = x - 1/2$, $x \in [0,1]$, are orthogonal with respect to the scalar product (1.42), in fact $< f, g >= \int_0^1 (x - 1/2) dx = 0$.

Given a scalar product in $\mathcal{S}$, the function that assigns to each vector $\mathbf{a}$ the quantity

$$\|\mathbf{a}\| = \sqrt{< \mathbf{a}, \mathbf{a} >} \tag{1.45}$$

satisfies the conditions **n1-n4** and thus is a norm on $\mathcal{S}$, called norm induced by the inner product $<,>$. The quantity (1.45) is called *length* (or norm) of the vector $\mathbf{a} \in \mathcal{S}$.

Proposition 2. *Let $\mathcal{S}$ be a vector space equipped with the scalar product $<,>$. Given the vectors $\mathbf{a}, \mathbf{b} \in \mathcal{S}$,, the Schwarz inequality*

$$| < \mathbf{a}, \mathbf{b} > | \leq \|\mathbf{a}\| \, \|\mathbf{b}\| , \tag{1.46}$$

the parallelogram law

$$\|\mathbf{a} + \mathbf{b}\|^2 + \|\mathbf{a} - \mathbf{b}\|^2 = 2 \|\mathbf{a}\|^2 + 2 \|\mathbf{b}\|^2 , \tag{1.47}$$

and the Pitagora theorem

$$\text{if } < \mathbf{a}, \mathbf{b} >= 0 \text{ then } \|\mathbf{a} + \mathbf{b}\|^2 = \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 , \tag{1.48}$$

hold

Proof. If $\mathbf{a} = 0$, (1.46) is trivially verified. Let us assume that $\mathbf{a} \neq 0$ and consider $\alpha \in \mathbb{R}$,

$$0 \leq < \alpha\mathbf{a} + \mathbf{b}, \alpha\mathbf{a} + \mathbf{b} > = \alpha^2 \|\mathbf{a}\|^2 + 2\alpha < \mathbf{a}, \mathbf{b} > + \|\mathbf{b}\|^2 =$$

$$\|\mathbf{a}\|^2 \left[\alpha^2 + \frac{2\alpha}{\|\mathbf{a}\|^2} < \mathbf{a}, \mathbf{b} > + \frac{< \mathbf{a}, \mathbf{b} >^2}{\|\mathbf{a}\|^4} \right] + \|\mathbf{b}\|^2 - \frac{< \mathbf{a}, \mathbf{b} >^2}{\|\mathbf{a}\|^2} =$$

$$\|\mathbf{a}\|^2 \left[\alpha + \frac{< \mathbf{a}, \mathbf{b} >}{\|\mathbf{a}\|^2} \right]^2 + \|\mathbf{b}\|^2 - \frac{< \mathbf{a}, \mathbf{b} >^2}{\|\mathbf{a}\|^2}. \tag{1.49}$$

If we put $\alpha = - < \mathbf{a}, \mathbf{b} > / \|\mathbf{a}\|^2$, from (1.49) we get

$$\|\mathbf{b}\|^2 \geq \frac{< \mathbf{a}, \mathbf{b} >^2}{\|\mathbf{a}\|^2}$$

and then (1.46). □

We have seen that given an inner product on $\mathcal{S}$, it is possible to define in a natural way the norm (1.45) on $\mathcal{S}$.

In $\mathbb{R}^n$ the norm $\|\mathbf{x}\|_2 = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2}$ is induced by the scalar product (1.41) and in $\mathcal{P}_n$, the norm $\|f\|_2 = \left(\int_0^1 |f(x)|^2 \, dx \right)^{1/2}$ is induced by the scalar product (1.42). In the space $\mathcal{M}_{m,n}$ the Frobenius norm defined in (1.34) is associated to the scalar product (1.43).

Nevertheless, it is possible define norms that are not induced by any inner product, For example, in $\mathcal{P}_1$, the norm (1.30) is not induced by any inner product. This follows from the fact that (1.30) does not satisfy the parallelogram law, as it is easy to prove choosing the polynomials $f_1(x) = 1$ and $f_2(x) = x$, $x \in [0, 1]$, for which $\|f_1\|_\infty = \|f_2\|_\infty = 1$, $\|f_1 - f_2\|_\infty = 1$ and $\|f_1 + f_2\|_\infty = 2$.

Analogously, in $\mathbb{R}^2$ the norm (1.28) does not satisfy the parallelogram law (for $\mathbf{x} = (1, 1)$ and $\mathbf{y} = (1, 0)$ we have $\|\mathbf{x}\|_\infty = \|\mathbf{y}\|_\infty = 1$, $\|\mathbf{x} - \mathbf{y}\|_\infty = 1$ and $\|\mathbf{x} + \mathbf{y}\|_\infty = 2$) and then it is not induced by any inner product.

It is possible to prove that if a norm $\| \ \|$ satisfies the parallelogram law, then $\| \ \|$ is induced by the following scalar product

$$< \mathbf{a}, \mathbf{b} > = \frac{1}{4} (\|\mathbf{a} + \mathbf{b}\|^2 - \|\mathbf{a} - \mathbf{b}\|^2). \tag{1.50}$$

1.5 Bases of a vector space

Let $\mathcal{S}$ be a vector space. Given the vectors $\mathbf{u}_1, ..., \mathbf{u}_m \in \mathcal{S}$ and the scalars $\alpha_1, ..., \alpha_m$, the vector $\alpha_1\mathbf{u}_1 + ... + \alpha_m\mathbf{u}_m$ is a *linear combination* of $\mathbf{u}_1, ..., \mathbf{u}_m$.

Vectors $\mathbf{u}_1, ..., \mathbf{u}_m \in \mathcal{S}$ are *linearly independent* if

$$\alpha_1 \mathbf{u}_1 + ... + \alpha_m \mathbf{u}_m = 0 \Longrightarrow \alpha_1 = ... = \alpha_m = 0. \tag{1.51}$$

If there exist α_i different from zero such that $\alpha_1 \mathbf{u}_1 + ... + \alpha_m \mathbf{u}_m = 0$, then vectors $\mathbf{u}_1, ..., \mathbf{u}_m$ are *linearly dependent*.

A *basis* of $\mathcal{S}$ is a set $\mathcal{B}$ of linearly independent vectors of $\mathcal{S}$ such that each vector in $\mathcal{S}$ is a (finite) linear combination of elements of $\mathcal{B}$. Of course, this combination is unique. It is possible to prove that every vector space has at least a basis.

A vector space has *finite dimension* if it has a finite basis. It is possible to prove that if $\mathcal{B}_1$ and $\mathcal{B}_2$ are two bases of the finite-dimensional vector space $\mathcal{S}$, then $\mathcal{B}_1$ and $\mathcal{B}_2$ have the same numbers of elements.

Thus, it is possible to define the *dimension* of a finite-dimensional vector space $\mathcal{S}$, which is the number of elements of a basis of $\mathcal{S}$.

Herein after we shall consider vector spaces $\mathcal{S}$ of finite dimension n and denote by $\{\mathbf{u}_1, ..., \mathbf{u}_n\}$ a *basis* of $\mathcal{S}$. For each $\mathbf{u} \in \mathcal{S}$ there exist (and are unique) $\beta_1, ..., \beta_n \in \mathbb{R}$ such that

$$\mathbf{u} = \sum_{i=1}^{n} \beta_i \mathbf{u}_i. \tag{1.52}$$

If $\mathcal{S}$ has an inner product and $\mathbf{u}_1, ..., \mathbf{u}_n$ are orthonormal, then $\{\mathbf{u}_1, ..., \mathbf{u}_n\}$ is an *orthonormal basis*.

In $\mathbb{R}^n$ let us consider the vectors

$$\begin{aligned} \mathbf{x}^1 &= (1, 0, ..., 0), \\ \mathbf{x}^2 &= (0, 1, ..., 0), \\ &... \\ \mathbf{x}^n &= (0, 0, ..., 1), \end{aligned}$$

$\{\mathbf{x}^1, ..., \mathbf{x}^n\}$ is an orthonormal basis, called canonic basis. The dimension of $\mathbb{R}^n$ is n.

In $\mathbb{R}^2$ let us consider the vectors $\mathbf{x}^1 = (1, 0)$ and $\mathbf{x}^2 = (1, 1)$, $\{\mathbf{x}^1, \mathbf{x}^2\}$ is a basis of $\mathbb{R}^2$, which is not orthonormal.

In $\mathcal{P}_1$ let us consider the polynomials $f_1(x) = 1$, $f_2(x) = x$, $g_2(x) = \sqrt{3}(1 - 2x)$, $x \in [0, 1]$, $\{f_1, f_2\}$ is a basis and $\{f_1, g_2\}$ is an orthonormal basis of $\mathcal{P}_1$. The dimension of the vector space $\mathcal{P}_n$ is $n + 1$.

In $\mathcal{M}_{m,n}$ the matrices $\{A^{ij}\}_{\substack{i=1,...,m \\ j=1,...,n}}$ with coefficients

$$a^{ij}_{kl} = \begin{cases} 1, & k = i, l = j \\ 0, & \text{otherwise.} \end{cases} \tag{1.53}$$

are an orthonormal basis and the dimension of $\mathcal{M}_{m,n}$ is $m \times n$.

Let $\mathcal{U}$ and $\mathcal{W}$ be two vector spaces, a function $T : \mathcal{U} \to \mathcal{W}$ is *linear* if it is *homogenous*

$$T(\alpha\mathbf{a}) = \alpha T(\mathbf{a}), \quad \text{for each } \mathbf{a} \in \mathcal{U},\ \alpha \in \mathbb{R}, \tag{1.54}$$

and *additive*

$$T(\mathbf{a} + \mathbf{b}) = T(\mathbf{a}) + T(\mathbf{b}), \quad \text{for each } \mathbf{a}, \mathbf{b} \in \mathcal{U}. \tag{1.55}$$

In particular, if T is linear then $T(\mathbf{0}) = \mathbf{0}$.

A bijective linear function is called *isomorphism* and two vector spaces $\mathcal{U}$ and $\mathcal{W}$ are *isomorphic* if there exists an isomorphism $T : \mathcal{U} \to \mathcal{W}$.

Vector spaces with the same dimension are isomorphic. In fact, the following theorem holds.

Theorem 1. *Every vector space $\mathcal{S}$ of dimension n is isomorphic to $\mathbb{R}^n$.*

Proof. Let $\{\mathbf{u}_1, ..., \mathbf{u}_n\}$ be a basis of $\mathcal{S}$. Then, each $\mathbf{u} \in \mathcal{S}$ can be written in the form $\mathbf{u} = \sum_{i=1}^{n} \beta_i \mathbf{u}_i$, with the scalars $\beta_1, ..., \beta_n$ being uniquely determined. The bijective function

$$\mathbf{u} \mapsto (\beta_1, ..., \beta_n) \tag{1.56}$$

from $\mathcal{S}$ to $\mathbb{R}^n$ is the required isomorphism. □

Vice versa, two isomorphic vector spaces $\mathcal{U}$ and $\mathcal{W}$ have the same dimension.

Theorem 2. *If the vector spaces $\mathcal{U}$ and $\mathcal{W}$ are isomorphic, then they have the same dimension.*

Proof. If $\{\mathbf{u}_1, ..., \mathbf{u}_n\}$ is a basis of $\mathcal{U}$, then $\{T\mathbf{u}_1, ..., T\mathbf{u}_n\}$ is a basis of $\mathcal{W}$. Firstly, we prove that the vectors $T\mathbf{u}_1, ..., T\mathbf{u}_n$ are linearly independent. In fact,

$$\alpha_1 T\mathbf{u}_1 + ... + \alpha_n T\mathbf{u}_n = \mathbf{0}$$

implies

$$T(\alpha_1 \mathbf{u}_1 + ... + \alpha_n \mathbf{u}_n) = \mathbf{0}$$

and then, in view of the fact that T is injective, we have

$$\alpha_1 \mathbf{u}_1 + ... + \alpha_n \mathbf{u}_n = \mathbf{0}$$

from which we deduce

$$\alpha_1 = ... = \alpha_n = 0,$$

because $\mathbf{u}_1, ..., \mathbf{u}_n$ are linearly independent. To prove that each $\mathbf{w} \in \mathcal{W}$ can be written as a unique linear combination of vectors $T\mathbf{u}_1, ..., T\mathbf{u}_n$, we proceed in the following way. Since T is bijective, given $\mathbf{w} \in \mathcal{W}$ there exists $\mathbf{u} \in \mathcal{U}$ such that $T\mathbf{u} = \mathbf{w}$. From the relation $\mathbf{u} = \alpha_1 \mathbf{u}_1 + ... + \alpha_n \mathbf{u}_n$ it follows that $\mathbf{w} = \alpha_1 T\mathbf{u}_1 + ... + \alpha_n T\mathbf{u}_n$. □

1.6 Subspaces

A non-empty subset $\mathcal{M}$ of the vector space $\mathcal{S}$ is a *subspace* if for each $\mathbf{a}, \mathbf{b} \in \mathcal{M}$, $\alpha, \beta \in \mathbb{R}$, the vector $\alpha\mathbf{a} + \beta\mathbf{b}$ belongs to $\mathcal{M}$.

Let $\mathcal{D}$ be a non-empty set of vectors of $\mathcal{S}$, the intersection of all subspaces containing $\mathcal{D}$ is a subspace of $\mathcal{S}$, called *subspace spanned by* $\mathcal{D}$ and denoted by Span($\mathcal{D}$). Span($\mathcal{D}$) contains all possible (finite) linear combinations of elements of $\mathcal{D}$.

If $\mathcal{M}_1$ and $\mathcal{M}_2$ are two subspaces of $\mathcal{S}$, Span($\mathcal{M}_1$,$\mathcal{M}_2$) is the subspace of $\mathcal{S}$ constituted by all the vectors $\mathbf{a} + \mathbf{b}$ with $\mathbf{a} \in \mathcal{M}_1, \mathbf{b} \in \mathcal{M}_2$ and is denoted by $\mathcal{M}_1 + \mathcal{M}_2$.

A subspace $\mathcal{M}_2$ of $\mathcal{S}$ is a *complement of a subspace* $\mathcal{M}_1$ if

$$\mathcal{M}_1 \cap \mathcal{M}_2 = \{\mathbf{0}\} \text{ and } \mathcal{S} = \mathcal{M}_1 + \mathcal{M}_2. \tag{1.57}$$

In this case, we say that $\mathcal{S}$ is the *direct sum* of the subspaces $\mathcal{M}_1$ and $\mathcal{M}_2$ and we write

$$\mathcal{S} = \mathcal{M}_1 \oplus \mathcal{M}_2. \tag{1.58}$$

If that is the case, every vector $\mathbf{s} \in \mathcal{S}$ can be written in a unique way as

$$\mathbf{s} = \mathbf{a} + \mathbf{b}, \text{ with } \mathbf{a} \in \mathcal{M}_1, \mathbf{b} \in \mathcal{M}_2. \tag{1.59}$$

In fact, (1.59) follows from the definition of $\mathcal{M}_1 + \mathcal{M}_2$, and to prove the uniqueness of $\mathbf{a}$ and $\mathbf{b}$, let us assume to have

$$\mathbf{s} = \mathbf{a}_1 + \mathbf{b}_1 = \mathbf{a}_2 + \mathbf{b}_2, \ \mathbf{a}_1, \mathbf{a}_2 \in \mathcal{M}_1, \mathbf{b}_1, \mathbf{b}_2 \in \mathcal{M}_2. \tag{1.60}$$

Then,

$$\mathbf{a}_1 - \mathbf{a}_2 = \mathbf{b}_2 - \mathbf{b}_1,$$

and, from $(1.57)_1$, we get $\mathbf{a}_1 = \mathbf{a}_2$ and $\mathbf{b}_1 = \mathbf{b}_2$.

In $\mathbb{R}^2$, given $\mathbf{x}^1 = (1, 0)$ and $\mathbf{x}^2 = (1, 1)$, let us consider $\mathcal{M}_1$ =Span($\mathbf{x}^1$) and $\mathcal{M}_2$ =Span($\mathbf{x}^2$), we have

$$\mathcal{M}_1 \cap \mathcal{M}_2 = \{(0, 0)\} \text{ and } \mathcal{M}_1 + \mathcal{M}_2 = \mathbb{R}^2,$$

thus $\mathbb{R}^2 = \mathcal{M}_1 \oplus \mathcal{M}_2$.

In $\mathbb{R}^3$ let us consider the vectors $\mathbf{x}^1 = (1, 0, 0), \mathbf{x}^2 = (0, 1, 0)$ and $\mathbf{x}^3 = (0, 0, 1)$; for $\mathcal{M}_1$ =Span($\mathbf{x}^1, \mathbf{x}^2$) and $\mathcal{M}_2$ =Span($\mathbf{x}^2, \mathbf{x}^3$), we have $\mathcal{M}_1 + \mathcal{M}_2 = \mathbb{R}^3$, but $\mathcal{M}_1 \cap \mathcal{M}_2$ =Span($\mathbf{x}^2$), then $\mathbb{R}^3$ is not direct sum of $\mathcal{M}_1$ and $\mathcal{M}_2$.

The subspace $\mathcal{M}$ of $\mathcal{S}$ has dimension m if it is spanned by m linear independent vectors of $\mathcal{S}$.

Given a subspace $\mathcal{M}$ of dimension m of a vector space $\mathcal{S}$ of dimension n ($m \leq n$), there exists a basis of $\mathcal{S}$ which contains a basis of $\mathcal{M}$.

Let $\mathcal{S}$ be a vector space with the scalar product $<,>$. Two subspaces $\mathcal{M}_1$ and $\mathcal{M}_2$ of $\mathcal{S}$ are *orthogonal* if each vector of the former is orthogonal to each vector of the latter.

Proposition 3. *Let $\mathcal{S}$ be a vector space with the scalar product $<,>$. Let $\{\mathbf{u}_1, ..., \mathbf{u}_m\}$ be an orthonormal set of vectors of $\mathcal{S}$. For each $\mathbf{u} \in \mathcal{S}$, putting $\alpha_i =< \mathbf{u}, \mathbf{u}_i >$, the following Bessel inequality holds*

$$\sum_{i=1}^{m} |\alpha_i|^2 \leq \|\mathbf{u}\|^2 . \tag{1.61}$$

Moreover, the vector $\mathbf{u}' = \mathbf{u} - \sum_{i=1}^{m} \alpha_i \mathbf{u}_i$ is orthogonal to Span$(\mathbf{u}_1, ..., \mathbf{u}_m)$.

Proof. We have

$$0 \leq \|\mathbf{u}'\|^2 =< \mathbf{u} - \sum_{i=1}^{m} \alpha_i \mathbf{u}_i, \mathbf{u} - \sum_{i=1}^{m} \alpha_i \mathbf{u}_i >=$$

$$\|\mathbf{u}\|^2 - \sum_{i=1}^{m} |\alpha_i|^2 - \sum_{i=1}^{m} |\alpha_i|^2 + \sum_{i=1}^{m} |\alpha_i|^2 =$$

$$\|\mathbf{u}\|^2 - \sum_{i=1}^{m} |\alpha_i|^2,$$

from which (1.61) follows. Moreover, we have

$$< \mathbf{u}', \mathbf{u}_j >=< (\mathbf{u}, \mathbf{u}_j > - \sum_{i=1}^{m} \alpha_i < \mathbf{u}_i, \mathbf{u}_j >= \alpha_j - \alpha_j = 0.$$

□

1.7 Orthonormal bases

An orthonormal set $\{\mathbf{u}_1, ..., \mathbf{u}_m\}$ of vectors of $\mathcal{S}$ is *complete* if it not contained in any larger orthonormal set. In particular, a complete orthonormal set of $\mathcal{S}$ is an orthonormal basis of $\mathcal{S}$, in fact, the following proposition holds.

Proposition 4. *Let $\mathcal{O} = \{\mathbf{u}_1, ..., \mathbf{u}_m\}$ be an orthonormal set of vector in the vector space $\mathcal{S}$ equipped with the inner product $<,>$. The following conditions are equivalent to each other.*

(1) *The orthonormal set $\mathcal{O}$ is complete.*

(2) *If $< \mathbf{u}, \mathbf{u}_j >= 0$ for $j = 1, ..., m$ then $\mathbf{u} = \mathbf{0}$.*

(3) *The subspace Span($\mathcal{O}$) coincides with $\mathcal{S}$.*

(4) *If $\mathbf{u} \in \mathcal{S}$, we have $\mathbf{u} = \sum_{i=1}^{m} < \mathbf{u}, \mathbf{u}_i > \mathbf{u}_i$.*

(5) *If $\mathbf{u}, \mathbf{v} \in \mathcal{S}$, the Parseval identity holds,*

$$< \mathbf{u}, \mathbf{v} >= \sum_{i=1}^{m} < \mathbf{u}, \mathbf{u}_i >< \mathbf{v}, \mathbf{u}_i > . \tag{1.62}$$

(6) *If* $\mathbf{u} \in \mathcal{S}$*, then we have*

$$\|\mathbf{u}\|^2 = \sum_{i=1}^{m} < \mathbf{u}, \mathbf{u}_i >^2 . \tag{1.63}$$

Proof. (1) $\Longrightarrow$ (2) If $< \mathbf{u}, \mathbf{u}_j >= 0$ for $j = 1, ..., m$ with $\mathbf{u} \neq \mathbf{0}$, the union of $\mathcal{O}$ and the vector $\mathbf{u}/\|\mathbf{u}\|$, is an orthonormal set of $\mathcal{S}$ containing $\mathcal{O}$.

(2) $\Longrightarrow$ (3) If there were $\mathbf{u} \in \mathcal{S}$ which is not a linear combination of vectors $\mathbf{u}_i$, then, in view of proposition 3 the vector $\mathbf{u}' = \mathbf{u} - \sum_{i=1}^{m} < \mathbf{u}, \mathbf{u}_i > \mathbf{u}_i$ would be different from zero and orthogonal to each $\mathbf{u}_i$.

(3) $\Longrightarrow$ (4) If each $\mathbf{u} \in \mathcal{S}$ had the expression $\mathbf{u} = \sum_{j=1}^{m} \alpha_j \mathbf{u}_j$, the, for each $i = 1, ..., m$, it would be

$$< \mathbf{u}, \mathbf{u}_i >= \sum_{j=1}^{m} \alpha_j < \mathbf{u}_j, \mathbf{u}_i >= \alpha_i.$$

(4) $\Longrightarrow$ (5) If $\mathbf{u} = \sum_{i=1}^{m} < \mathbf{u}, \mathbf{u}_i > \mathbf{u}_i$, $\mathbf{v} = \sum_{j=1}^{m} < \mathbf{v}, \mathbf{u}_j > \mathbf{u}_j$, then

$$< \mathbf{u}, \mathbf{v} >= \sum_{i,j=1}^{m} < \mathbf{u}, \mathbf{u}_i >< \mathbf{v}, \mathbf{u}_j >< \mathbf{u}_i, \mathbf{u}_j >= \sum_{i=1}^{m} < \mathbf{u}, \mathbf{u}_i >< \mathbf{v}, \mathbf{u}_i > .$$

(5) $\Longrightarrow$ (6) Put $\mathbf{u} = \mathbf{v}$ in (1.62).

(6) $\Longrightarrow$ (1) Let $\mathbf{u}_0 \in \mathcal{S}$ be orthogonal to all $\mathbf{u}_i$. Then,

$$\|\mathbf{u}_0\|^2 = \sum_{i=1}^{m} < \mathbf{u}_0, \mathbf{u}_i >^2 = 0$$

which implies $\mathbf{u}_0 = \mathbf{0}$. □

Let $\mathcal{M}$ be a subspace of $\mathcal{S}$; the set

$$\mathcal{M}^\perp = \{\mathbf{u} \in \mathcal{S} \mid < \mathbf{v}, \mathbf{u} >= 0 \quad \text{for each } \mathbf{v} \in \mathcal{M}\} \tag{1.64}$$

is a subspace of $\mathcal{S}$ called *orthogonal complement* of $\mathcal{M}$. The vector space $\mathcal{S}$ is the direct sum of $\mathcal{M}$ and $\mathcal{M}^\perp$,

$$\mathcal{S} = \mathcal{M} \oplus \mathcal{M}^\perp. \tag{1.65}$$

In fact, if $\{\mathbf{e}_1, ..., \mathbf{e}_m\}$ is an orthonormal basis of $\mathcal{M}$, for each $\mathbf{v} \in \mathcal{S}$ we have

$$\mathbf{v} = \overline{\mathbf{v}} + \mathbf{v}_o, \tag{1.66}$$

where $\overline{\mathbf{v}} = \sum_{i=1}^{m} < \mathbf{e}_i, \mathbf{v} > \mathbf{e}_i \in \mathcal{M}$ e $\mathbf{v}_o = \mathbf{v} - \sum_{i=1}^{m} < \mathbf{e}_i, \mathbf{v} > \mathbf{e}_i \in \mathcal{M}^\perp$, in view of proposition 3.

Let $\mathcal{S}_1$ and $\mathcal{S}_2$ two vector spaces with inner products $< .,. >_{\mathcal{S}_1}$ and $< .,. >_{\mathcal{S}_2}$. A linear function $T : \mathcal{S}_1 \to \mathcal{S}_2$ which satisfies

$$< \mathbf{u}, \mathbf{v} >_{\mathcal{S}_1} = < T(\mathbf{u}), T(\mathbf{v}) >_{\mathcal{S}_2}, \quad \text{for each } \mathbf{u}, \mathbf{v} \in \mathcal{S}_1, \tag{1.67}$$

is an *isometry*. An isomorphism satisfying (1.67) is called *isometric isomorphism* and the space $\mathcal{S}_1$ and $\mathcal{S}_2$ are called *isometrically isomorphic*.

An isometry preserves the scalar product and then preserves the norm

$$||\mathbf{u}||_{\mathcal{S}_1} = ||T(\mathbf{u})||_{\mathcal{S}_2}, \quad \text{for each } \mathbf{u} \in \mathcal{S}_1. \tag{1.68}$$

From (1.68) it follows that if $|T(\mathbf{u})||_{\mathcal{S}_2} = 0$ then $||\mathbf{u}||_{\mathcal{S}_1} = 0$, thus an isometry is injective.

Proposition 5. *Every vector space $\mathcal{S}$ of dimension n with inner product $< \ , >_{\mathcal{S}}$ is isometrically isomorphic to $\mathbb{R}^n$.*

Proof. Let $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ be an orthonormal basis of $\mathcal{S}$, the function T on $\mathcal{S}$ into $\mathbb{R}^n$ defined by

$$T(\mathbf{u}) = (u_1, ..., u_n), \quad u_i = < \mathbf{e}_i, \mathbf{u} >_{\mathcal{S}}, \quad i = 1, ..., n \tag{1.69}$$

is an isometric isomorphism, in fact, for each $\mathbf{u}, \mathbf{v} \in S$ we have

$$< T(\mathbf{u}), T(\mathbf{v}))_{\mathbb{R}^n} = \sum_{i=1}^{n} u_i v_i = \sum_{i=1}^{n} < \mathbf{e}_i, \mathbf{u} >_{\mathcal{S}} < \mathbf{e}_i, \mathbf{v} >_{\mathcal{S}} = < \mathbf{u}, \mathbf{v} >_{\mathcal{S}}, \tag{1.70}$$

where the last equality follows from (1.62). □

1.8 Convergence of vectors

Let us now introduce the notion of convergence of a sequence of vectors in a vector space equipped with a scalar product.

A sequence $\{\mathbf{v}^{(k)}\}_{k \in \mathbb{N}}$ of vectors in $\mathcal{S}$ *converges* to a vector $\mathbf{v} \in \mathcal{S}$ if for each $\varepsilon > 0$ there exists $\overline{k} > 0$ such that

$$\left\| \mathbf{v}^{(k)} - \mathbf{v} \right\| < \varepsilon \quad \text{for each } \ k \geq \overline{k}. \tag{1.71}$$

In that case, the sequence $\{\mathbf{v}^{(k)}\}_{k \in \mathbb{N}}$ is *convergent* and the vector $\mathbf{v}$ is the *limit* of $\{\mathbf{v}^{(k)}\}_{k \in \mathbb{N}}$ for k going to infinity,

$$\lim_{k \to \infty} \mathbf{v}^{(k)} = \mathbf{v}, \text{ or } \quad \mathbf{v}^{(k)} \to \mathbf{v}, \text{ for } k \to \infty. \tag{1.72}$$

In $\mathbb{R}^n$ the sequence $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}}$ converges to $\mathbf{x}$ if the sequence of real numbers $\left\| \mathbf{x}^{(k)} - \mathbf{x} \right\|_2^2 = \sum_{i=1}^{n} |\mathbf{x}_i^{(k)} - x_i|^2$ converges to 0 for k going to infinity and, in particular, if

$$x_i^{(k)} \to x_i \text{ for } k \to \infty, i = 1, ..., n. \tag{1.73}$$

The convergence defined above is also called *strong convergence*. It is easy to prove that the limit of a convergent sequence $\{\mathbf{v}^{(k)}\}_{k\in\mathbb{N}}$ is unique. In fact, let us assume that

$$\mathbf{v}^{(k)} \to \mathbf{v} \quad \text{and} \quad \mathbf{v}^{(k)} \to \mathbf{w}, \quad \text{for } k \to \infty. \tag{1.74}$$

Then we have

$$\|\mathbf{v} - \mathbf{w}\| \le \left\|\mathbf{v} - \mathbf{v}^{(k)}\right\| + \left\|\mathbf{v}^{(k)} - \mathbf{w}\right\|, \tag{1.75}$$

from which, in view of (1.71), we get $\|\mathbf{v} - \mathbf{w}\| = 0$.

For each $\mathbf{w} \in \mathcal{S}$, from the Schwarz inequality it follows that

$$| < \mathbf{v}^{(k)} - \mathbf{v}, \mathbf{w} > | \le \left\|\mathbf{v}^{(k)} - \mathbf{v}\right\| \; ||\mathbf{w}||$$

therefore, if $\{\mathbf{v}^{(k)}\}_{k\in\mathbb{N}}$ converges to $\mathbf{v}$ we have that

$$\lim_{k\to\infty} < \mathbf{v}^{(k)}, \mathbf{w} > = < \mathbf{v}, \mathbf{w} > \quad \text{for each } \mathbf{w} \in \mathcal{S}. \tag{1.76}$$

If condition (1.76) is satisfied, we say that the sequence $\{\mathbf{v}^{(k)}\}_{k\in\mathbb{N}}$ in $\mathcal{S}$ *converges weakly* to $\mathbf{v} \in \mathcal{S}$ and we write

$$\mathbf{v}^{(k)} \rightharpoonup \mathbf{v}, \quad \text{for } k \to \infty. \tag{1.77}$$

Of course, if a sequence is strongly convergent, then it is weakly convergent. Unlike infinite dimensional vector spaces, where strong and weak convergence do not coincide, in finite-dimensional vector spaces each weakly convergent sequence is (strongly) convergent. In fact, let us assume that (1.76) holds and let $\{\mathbf{u}_1, ..., \mathbf{u}_n\}$ an orthonormal basis of $\mathcal{S}$. Then, we have that $< \mathbf{v}^{(k)} - \mathbf{v}, \mathbf{u}_i > \to 0$ when $k \to \infty$, for each $i = 1, ..., n$. In view of relation (1.63) of the Proposition 4 we have

$$\left\|\mathbf{v}^{(k)} - \mathbf{v}\right\|^2 = \sum_{i=1}^{n} < \mathbf{v}^{(k)} - \mathbf{v}, \mathbf{u}_i >^2,$$

thus, $\lim_{k\to\infty} \mathbf{v}^{(k)} = \mathbf{v}$.

Let $\mathcal{S}$ be a normed vector space, a sequence $\{\mathbf{v}^{(k)}\}_{k\in\mathbb{N}} \subset \mathcal{S}$ is a *Cauchy sequence* if for each $\varepsilon > 0$ there is $\overline{q} \in \mathbb{N}$ such that $\left\|\mathbf{v}^{(p)} - \mathbf{v}^{(q)}\right\| < \varepsilon$ when $p, q > \overline{q}$ or, equivalently, if $\left\|\mathbf{v}^{(p)} - \mathbf{v}^{(q)}\right\| \to 0$, for $p, q \to \infty$.

If $\{\mathbf{v}^{(k)}\}_{k\in\mathbb{N}}$ is a convergent sequence, with limit $\mathbf{v}$, then it is a Cauchy sequence, in fact

$$\left\|\mathbf{v}^{(p)} - \mathbf{v}^{(q)}\right\| \le \left\|\mathbf{v}^{(p)} - \mathbf{v}\right\| + \left\|\mathbf{v} - \mathbf{v}^{(q)}\right\|, \tag{1.78}$$

and $\left\|\mathbf{v}^{(p)} - \mathbf{v}^{(q)}\right\|$ converges to 0 for $p, q \to \infty$.

A normed vector space $\mathcal{S}$ is *complete* if for each Cauchy sequence $\{\mathbf{v}^{(k)}\}_{k\in\mathbb{N}} \subset \mathcal{S}$ there is a unique vector $\mathbf{v} \in \mathcal{S}$ such that $\mathbf{v}_k \to \mathbf{v}$ when $k \to \infty$.

Proposition 6. *Each finite-dimensional vector space $\mathcal{S}$ with inner product $<,>$ is complete.*

Proof. From Proposition 5 it follows that if $\mathcal{S}$ has dimension n then it is isometrically isomorphic to $\mathbb{R}^n$. Let T be be isometric isomorphism definite in (1.69) and $\{\mathbf{v}^{(k)}\}_{k\in\mathbb{N}} \subset \mathcal{S}$ a Cauchy sequence, we have

$$||T(\mathbf{v}^{(p)}) - T(\mathbf{v}^{(q)})||^2_{\mathbb{R}^n} =< T(\mathbf{v}^{(p)}) - T(\mathbf{v}^{(q)}), T(\mathbf{v}^{(p)}) - T(\mathbf{v}^{(q)}) >_{\mathbb{R}^n}$$
$$=< \mathbf{v}^{(p)} - \mathbf{v}^{(q)}, \mathbf{v}^{(p)} - \mathbf{v}^{(q)} >_{\mathcal{S}} = ||\mathbf{v}^{(p)} - \mathbf{v}^{(q)}||^2_{\mathcal{S}}, \tag{1.79}$$

therefore, $\{T(\mathbf{v}^{(k)})\}_{k\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}^n$. Since $\mathbb{R}^n$ is complete, there is $\mathbf{x} \in \mathbb{R}^n$ such that $T(\mathbf{v}^{(k)}) \to \mathbf{x}$ for $k \to \infty$, and in view of the fact that T is surjective, there exists $\mathbf{v} \in \mathcal{S}$ such that $T(\mathbf{v}) = \mathbf{x}$. Then we have,

$$||\mathbf{v} - \mathbf{v}^{(k)}||_{\mathcal{S}} = ||T(\mathbf{v}) - T(\mathbf{v}^{(k)})||_{\mathbb{R}^n} = ||\mathbf{x} - T(\mathbf{v}^{(k)})||_{\mathbb{R}^n}, \tag{1.80}$$

from which it follows that $\mathbf{v}^{(k)} \to \mathbf{v}$ when $k \to \infty$. □

Two norms $\| \|_1$ and $\| \|_2$ on a vector space $\mathcal{U}$ are *equivalent* if there are two positive constants λ and μ such that

$$\lambda \|\mathbf{u}\|_1 \leq \|\mathbf{u}\|_2 \leq \mu \|\mathbf{u}\|_1 \quad \text{for each } \mathbf{u} \in \mathcal{U}. \tag{1.81}$$

The following theorem holds.

Theorem 3. *In a finite-dimensional vector spaces all the norms are equivalent.*

1.9 Open and closed sets, neighborhoods

Let $\mathcal{S}$ be a vector space with the norm $\| \|$. Given $\mathbf{a} \in \mathcal{S}$, $r > 0$, the sets

$$B(\mathbf{a}, r) = \{\mathbf{b} \in \mathcal{S} \mid \|\mathbf{a} - \mathbf{b}\| < r\}, \tag{1.82}$$

$$B'(\mathbf{a}, r) = \{\mathbf{b} \in \mathcal{S} \mid \|\mathbf{a} - \mathbf{b}\| \leq r\}, \tag{1.83}$$

$$S(\mathbf{a}, r) = \{\mathbf{b} \in \mathcal{S} \mid \|\mathbf{a} - \mathbf{b}\| = r\}, \tag{1.84}$$

are the *open ball*, *closed ball* e *sphere* of center $\mathbf{a}$ and radius r.

For example, in $\mathcal{S} = \mathbb{R}^2$ with the inner product (1.41) we have

$$B(\mathbf{0}, r) = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 < r^2\}, \tag{1.85}$$
$$B'(\mathbf{0}, r) = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq r^2\}, \tag{1.86}$$
$$S(\mathbf{0}, r) = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = r^2\}. \tag{1.87}$$

A subset A of $\mathcal{S}$ is *open* if for each $\mathbf{a} \in A$ there exists $r > 0$ such that $B(\mathbf{a}, r) \subset A$. A subset C of $\mathcal{S}$ is *closed* if its complement $\mathcal{S} - C$ is open. Given $\mathbf{a} \in \mathcal{S}$, a subset U_a of $\mathcal{S}$ which contains an open ball with center $\mathbf{a}$ is a *neighborhood* of $\mathbf{a}$.

Open balls, closed balls and spheres can be defined in a set X with metric d, they have been introduced in a normed vector space because in these notes we are interested in focusing on normed vector spaces.

The following propositions holds.

Proposition 7. *Sets $\mathcal{S}$ and $\emptyset$ are closed and open.*
The union of an arbitrary family of open sets is open.
The intersection of a finite family of open sets is open.
The intersection of an arbitrary family of closed sets is closed.
The union of a finite family of closed sets is closed.

A subset $\mathcal{K}$ of $\mathcal{S}$ is *convex* if given $\mathbf{a}, \mathbf{b} \in \mathcal{K}$, we have $\alpha\mathbf{a} + (1-\alpha)\mathbf{b} \in \mathcal{K}$, for each $\alpha \in [0, 1]$. A subset $\mathcal{K}$ of $\mathcal{S}$ is *bounded* if there exists $\kappa > 0$ such that $||\mathbf{a}|| \leq \kappa$ for every $\mathbf{a} \in \mathcal{K}$.

The balls $B(\mathbf{a}, r)$ and $B'(\mathbf{a}, r)$ are convex. The subset $\mathcal{K} = \{\mathbf{v}^{(k)}\}_{k\in\mathbb{N}}$ of $\mathcal{S}$ constituted by the elements of the convergent sequence $\{\mathbf{v}^{(k)}\}_{k\in\mathbb{N}}$ is bounded and not convex, on the contrary, $\text{Span}(\mathcal{K})$ is convex but not bounded.

1.10 Mappings on vector spaces

Let $\mathcal{U}$ and $\mathcal{W}$ be normed vector spaces and $T : \mathcal{U} \to \mathcal{W}$ a mapping (or function).

T is *continuous* at $\mathbf{a}_0 \in \mathcal{U}$ if for each $\varepsilon > 0$ there is $\delta > 0$ such that for each $\mathbf{a} \in \mathcal{U}$ satisfying $||\mathbf{a} - \mathbf{a}_0||_{\mathcal{U}} < \delta$, we have$||T(\mathbf{a}) - T(\mathbf{a}_0)||_{\mathcal{W}} < \varepsilon$. T is continuous on $\mathcal{U}$ if it is continuous at each $\mathbf{a}_0 \in \mathcal{U}$.

In other words, T is *continuous* at $\mathbf{a}_0$ if for every open ball $B(T(\mathbf{a}_0), \varepsilon)$ with center $T(\mathbf{a}_0)$ and radius ε there is an open ball $B(\mathbf{a}_0, \delta)$ with center $\mathbf{a}_0$ and radius δ such that $T(B(\mathbf{a}_0, \delta)) \subset B(T(\mathbf{a}_0), \varepsilon)$.

An alternative formulation of continuity can be expressed in terms of open and closed sets.

Proposition 8. *Let $\mathcal{U}$ and $\mathcal{W}$ be normed vector spaces and $T : \mathcal{U} \to \mathcal{W}$ a mapping. T is continuous on $\mathcal{U}$ (that is at each $\mathbf{a}_0 \in \mathcal{U}$) if and only is for each open (closed) set $\mathcal{A}$ in $\mathcal{W}$, the inverse image $T^{-1}(\mathcal{A})$ of $\mathcal{A}$ under T, is an open (closed) set in $\mathcal{U}$.*

The notion of convergence of a sequence of vectors can be used to characterize closed sets and continuous functions. In fact, the following propositions hold.

Proposition 9. *Let $\mathcal{C}$ be a non-empty set of the normed vector space $\mathcal{S}$. $\mathcal{C}$ is closed if and only if each convergent sequence constituted by vectors in $\mathcal{C}$, converges to a vector of $\mathcal{C}$.*

The following result generalizes the well known relationship between continuity of real functions and convergence of sequences.

Proposition 10. *Let $\mathcal{U}$ and $\mathcal{W}$ be normed vector spaces and $T : \mathcal{U} \to \mathcal{W}$ a mapping. T is continuous at $\mathbf{a}_0 \in \mathcal{U}$ if and only if for each sequence $\{\mathbf{a}^{(k)}\}_{k\in\mathbb{N}} \subset \mathcal{U}$ such that $\lim_{k\to\infty} \mathbf{a}^{(k)} = \mathbf{a}_0$, we have $\lim_{k\to\infty} T(\mathbf{a}^{(k)}) = T(\mathbf{a}_0)$.*

A mapping $T : \mathcal{U} \to \mathcal{W}$ is *linear* if the properties (1.54) and (1.55) are satisfied.

Example 1. *Let $\{\mathbf{e}_1, ..., \mathbf{e}_m\}$ be an orthonormal set $\mathcal{S}$, function L defined from $\mathcal{S}$ into Span$(\mathbf{e}_1, ..., \mathbf{e}_m)$ such that*

$$L(\mathbf{u}) = \sum_{i=1}^{m} < \mathbf{u}, \mathbf{e}_i > \mathbf{e}_i, \quad \forall \mathbf{u} \in \mathcal{S}, \tag{1.88}$$

is linear, on the contrary, the function that assigns to each vector $\mathbf{u}$ in $\mathcal{S}$ the constant vector $\overline{\mathbf{u}}$ is not linear.

A bijective map T is *invertible* and the function $T^{-1} : \mathcal{W} \to \mathcal{U}$ defined by $T^{-1}(\mathbf{v}) = \mathbf{u}$, if and only if $T(\mathbf{u}) = \mathbf{v}$ is called *inverse* of T. If T is linear and invertible, then T^{-1} is linear. In fact, for $\mathbf{z}$, $\mathbf{w} \in \mathcal{W}$, let $\mathbf{u}$, $\mathbf{v}$ the unique vectors of $\mathcal{U}$ such that $T(\mathbf{u}) = \mathbf{z}$ and $T(\mathbf{v}) = \mathbf{w}$. For α and $\beta \in \mathbb{R}$ we have

$$\alpha\mathbf{z} + \beta\mathbf{w} = \alpha T(\mathbf{u}) + \beta T(\mathbf{v}) = T(\alpha\mathbf{u} + \beta\mathbf{v}) \tag{1.89}$$

and then

$$T^{-1}(\alpha\mathbf{z} + \beta\mathbf{w}) = \alpha\mathbf{u} + \beta\mathbf{v} = \alpha T^{-1}(\mathbf{z}) + \beta T^{-1}(\mathbf{w}). \tag{1.90}$$

Let $T : \mathcal{U} \to \mathcal{W}$ be a linear mapping, T is *bounded* if there is $\kappa > 0$ such that

$$||T(\mathbf{a})||_{\mathcal{W}} \leq \kappa ||\mathbf{a}||_{\mathcal{U}}, \quad \text{for each } \mathbf{a} \in \mathcal{U}. \tag{1.91}$$

All linear mappings on finite-dimensional vector spaces are bounded. The following proposition holds.

Proposition 11. *Let $\mathcal{U}$ and $\mathcal{W}$ be finite-dimensional normed vector spaces. Every linear mapping $L : \mathcal{U} \to \mathcal{W}$ is bounded.*

Proof. For the sake of simplicity, let us limit ourselves to prove the proposition in the case in which the norm on $\mathcal{U}$ is induced by the scalar product $<, >_{\mathcal{U}}$. Let $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ be an orthonormal basis of $\mathcal{U}$, for each $\mathbf{u} \in \mathcal{U}$ we have

$$\mathbf{u} = \sum_{i=1}^{n} u_i \mathbf{e}_i, \tag{1.92}$$

with $u_i =< \mathbf{u}, \mathbf{e}_i >_{\mathcal{U}}$, $i = 1, ..., n$, and

$$L(\mathbf{u}) = \sum_{i=1}^{n} u_i L(\mathbf{e}_i). \tag{1.93}$$

Therefore, in view of the properties **n3** and **n2** of the norm, from (1.93) we get

$$||L(\mathbf{u})||_{\mathcal{W}} \leq \sum_{i=1}^{n} |u_i| \; ||L(\mathbf{e}_i)||_{\mathcal{W}} \leq \beta \sum_{i=1}^{n} |u_i|, \tag{1.94}$$

where $\beta = \max_{i=1,\dots,n} ||L(\mathbf{e}_i)||_{\mathcal{W}}$. Finally, from (1.94), using the Schwarz inequality we obtain that

$$||L(\mathbf{u})||_{\mathcal{W}} \leq \beta \sum_{i=1}^{n} | < \mathbf{u}, \mathbf{e}_i >_{\mathcal{U}} | \leq \beta \sum_{i=1}^{n} ||\mathbf{e}_i||_{\mathcal{U}} \; ||\mathbf{u}||_{\mathcal{U}} = \beta n ||\mathbf{u}||_{\mathcal{U}}, \tag{1.95}$$

for each $\mathbf{u} \in \mathcal{U}$, and then L is bounded. □

Proposition 12. *Let $T : \mathcal{U} \to \mathcal{W}$ be a linear mapping; T is continuous on $\mathcal{U}$ if and only if it is continuous at $\mathbf{0} \in \mathcal{U}$.*

Proof. Let us assume that T is continuous at $\mathbf{0} \in \mathcal{U}$, then for each $\varepsilon > 0$ there is $\delta > 0$ such that if $||\mathbf{a}||_{\mathcal{U}} < \delta$ then $||T(\mathbf{a})||_{\mathcal{V}} < \varepsilon$ (in view of the linearity of T, $T(\mathbf{0}) = \mathbf{0}$). Now consider $\mathbf{a}_0 \in \mathcal{U}$, for each $\mathbf{a} \in \mathcal{U}$ such that $||\mathbf{a} - \mathbf{a}_0||_{\mathcal{U}} < \delta$ we have $||T(\mathbf{a}) - T(\mathbf{a}_0)||_{\mathcal{V}} = ||T(\mathbf{a} - \mathbf{a}_0)||_{\mathcal{V}} < \varepsilon$, therefore if T is continuous at $\mathbf{a}_0$. □

Proposition 13. *Let $T : \mathcal{U} \to \mathcal{W}$ be a linear mapping. T is continuous on $\mathcal{U}$ if and only if is bounded on $\mathcal{U}$.*

Proof. Let us assume that T is bounded, then from (1.91) it follows that for each $\varepsilon > 0$ putting $\delta = \varepsilon/\kappa$ we have

$$||T(\mathbf{u})||_{\mathcal{W}} \leq \kappa ||\mathbf{u}||_{\mathcal{U}} \leq \kappa\delta = \varepsilon \tag{1.96}$$

and then T is continuous at $\mathbf{0} \in \mathcal{U}$; the thesis follows from Proposition 12. Vice versa let us assume that T is continuous but not bounded (reductio ad absurdum): then, for each $k \in \mathbb{N}$ there is $\mathbf{u}^{(k)} \in \mathcal{U}$ such that

$$||T(\mathbf{u}^{(k)})||_{\mathcal{W}} > k ||\mathbf{u}^{(k)}||_{\mathcal{U}}. \tag{1.97}$$

In particular, we have $\mathbf{u}^{(k)} \neq \mathbf{0}$, then, we can put

$$\mathbf{v}^{(k)} = \frac{\mathbf{u}^{(k)}}{k||\mathbf{u}^{(k)}||_{\mathcal{U}}}. \tag{1.98}$$

The sequence $\{\mathbf{v}^{(k)}\}_{k\in\mathbb{N}}$ converges to $\mathbf{0}$, but

$$||T(\mathbf{v}^{(k)})||_{\mathcal{W}} = \frac{||T(\mathbf{u}^{(k)})||_{\mathcal{W}}}{k||\mathbf{u}^{(k)}||_{\mathcal{U}}} > 1 \tag{1.99}$$

which is in contrast with the continuity of T. □

Remark 1. *In infinite dimensional vector spaces there are linear functions that are not continuous. Let us consider the set $\mathcal{P}$ of the polynomials on $[0,1]$ with the scalar product $< f, g >= \int_0^1 f(t)g(t)dt$ and let $L : \mathcal{P} \to \mathcal{P}$ be the mapping that assigns to each polynomial its derivative. L is linear, but not bounded. Given the polynomials $f_k(t) = t^k$, $k \in \mathbb{N}$, we have*

$$||f_k||^2 = \frac{1}{2k+1} < 1$$

and

$$||L(f_k)||^2 = k^2 \int_0^1 t^{2k-2}dt = \frac{k^2}{2k-1},$$

therefore, $||L(f_k)||^2 \to +\infty$ when $k \to \infty$ and L is not bounded.

Proposition 14. *Every subspace $\mathcal{M}$ of a vector space $\mathcal{S}$ with inner product is closed.*

Proof. The proof is based on Proposition 9. Let $\mathcal{M}$ be a subspace of the vector space $\mathcal{S}$ and $\{\mathbf{v}^{(k)}\}_{k\in\mathbb{N}} \subset \mathcal{M}$ a sequence converging to $\mathbf{v} \in \mathcal{S}$. For $\{\mathbf{u}_1, ..., \mathbf{u}_m\}$ an orthonormal basis of $\mathcal{M}$, we have

$$\mathbf{v}^{(k)} = \sum_{i=1}^{m} < \mathbf{v}^{(k)}, \mathbf{u}_i > \mathbf{u}_i, \quad k \in \mathbb{N},$$

and

$$\left\|\mathbf{v}^{(k)} - \mathbf{v}\right\| \to 0 \quad \text{when} \;\; k \to \infty.$$

From the inequality

$$\left\|\sum_{i=1}^{m} < \mathbf{v}, \mathbf{u}_i > \mathbf{u}_i - \mathbf{v}\right\| \le \left\|\sum_{i=1}^{m} < \mathbf{v}, \mathbf{u}_i > \mathbf{u}_i - \sum_{i=1}^{m} < \mathbf{v}^{(k)}, \mathbf{u}_i > \mathbf{u}_i\right\| + \left\|\mathbf{v}^{(k)} - \mathbf{v}\right\|$$

taking into account that $\{\mathbf{v}^{(k)}\}_{k\in\mathbb{N}}$ converges t $\mathbf{v}$ and then converges weakly to a $\mathbf{v}$, we get that $\mathbf{v} \in \mathcal{M}$. □

Let $\mathcal{U}$ and $\mathcal{W}$ be vector spaces with inner product and dimension n and m, respectively. Let us denote by $\mathcal{L}(\mathcal{U}, \mathcal{W})$ the set of all linear mappings on $\mathcal{U}$ into $\mathcal{W}$

$$\mathcal{L}(\mathcal{U}, \mathcal{W}) = \{L : \mathcal{U} \to \mathcal{W} \mid L \text{ is linear}\}. \tag{1.100}$$

If we define the sum of two mappings and the product by a scalar in the following natural way

$$(L_1 + L_2)(\mathbf{u}) = L_1(\mathbf{u}) + L_2(\mathbf{u}), \tag{1.101}$$

$$(\alpha L_1)(\mathbf{u}) = \alpha L_1(\mathbf{u})$$

for each $\mathbf{u} \in \mathcal{U}$, $\alpha \in \mathbb{R}$, $\mathcal{L}(\mathcal{U}, \mathcal{W})$ turns out to be a vector space.

Let $\{\mathbf{u}_1, ..., \mathbf{u}_n\}$ and $\{\mathbf{w}_1, ..., \mathbf{w}_m\}$ be two orthonormal bases of $\mathcal{U}$ and $\mathcal{W}$, respectively. The $m \times n$ linear mappings L_{ij} defined by

$$L_{ij}(\mathbf{u}_k) = \left\{ \begin{array}{ll} \mathbf{w}_i & k = j \\ \mathbf{0} & k \neq j \end{array} \right. , \quad i = 1, ..., m, \ j = 1, ..., n \tag{1.102}$$

are linear independent. In fact, from the condition $\sum_{i,j} \alpha_{ij} L_{ij} = 0$ we get that $\sum_{i,j} \alpha_{ij} L_{ij}(\mathbf{u}_k) = \mathbf{0}$, $k = 1, ..., n$ and then, in view of (1.102), $\sum_{i} \alpha_{ik} \mathbf{w}_i = \mathbf{0}$, $k = 1, ..., n$. From the linear independence of vectors $\mathbf{w}_1, ..., \mathbf{w}_m$ it follows that the coefficients α_{ij} are zero. Moreover, for each $L \in \mathcal{L}(\mathcal{U}, \mathcal{W})$ we have

$$L = \sum_{\substack{i=1,...,m \\ j=1,...,n}} (L(\mathbf{u}_j), \mathbf{w}_i)_{\mathcal{W}} \ L_{ij}. \tag{1.103}$$

In fact, for each $\mathbf{u} \in \mathcal{U}$, we have $\mathbf{u} = \sum_{k=1}^{n} (\mathbf{u}, \mathbf{u}_k)_{\mathcal{U}} \ \mathbf{u}_k$, from which

$$\sum_{\substack{i=1,...,m \\ j=1,...,n}} (L(\mathbf{u}_j), \mathbf{w}_i)_{\mathcal{W}} \ L_{ij}(\mathbf{u}) = \sum_{\substack{i=1,...,m \\ j=1,...,n}} (L(\mathbf{u}_j), \mathbf{w}_i)_{\mathcal{W}} \ L_{ij} \left(\sum_{k=1}^{n} (\mathbf{u}, \mathbf{u}_k)_{\mathcal{U}} \ \mathbf{u}_k \right) =$$

$$\sum_{\substack{i=1,...,m \\ j=1,...,n}} (L(\mathbf{u}_j), \mathbf{w}_i)_{\mathcal{W}} \ (\mathbf{u}, \mathbf{u}_j)_{\mathcal{U}} \ \mathbf{w}_i =$$

$$\sum_{i=1,...,m} (L \left(\sum_{j=1,...,n} (\mathbf{u}, \mathbf{u}_j)_{\mathcal{U}} \ \mathbf{u}_j \right), \mathbf{w}_i)_{\mathcal{W}} \ \mathbf{w}_i =$$

$$\sum_{i=1,...,m} (L(\mathbf{u}), \mathbf{w}_i)_{\mathcal{W}} \ \mathbf{w}_i = L(\mathbf{u}),$$

and then (1.103) is proved. Thus the linear mappings defined in (1.102) are a basis of the vector space $\mathcal{L}(\mathcal{U}, \mathcal{W})$ and the dimension of $\mathcal{L}(\mathcal{U}, \mathcal{W})$ is $m \times n$. The vector spaces $\mathcal{L}(\mathcal{U}, \mathcal{W})$ and $\mathcal{M}_{m,n}$ are isomorphic.

For each $L \in \mathcal{L}(\mathcal{U}, \mathcal{W})$ let us consider the quantity

$$||L||_N = \sup_{\mathbf{u} \in \mathcal{U}, \ \mathbf{u} \neq \mathbf{0}} \frac{||L(\mathbf{u})||_{\mathcal{W}}}{||\mathbf{u}||_{\mathcal{U}}}. \tag{1.104}$$

It is easy to verify that (1.104) is a norm on $\mathcal{L}(\mathcal{U}, \mathcal{W})$. Firstly, if $L \in \mathcal{L}(\mathcal{U}, \mathcal{W})$ then L is bounded (Proposition 11), then there is $\kappa > 0$ such that $||L(\mathbf{u})||_{\mathcal{V}} \leq \kappa ||\mathbf{u}||_{\mathcal{U}}$ for each $\mathbf{u} \in \mathcal{U}$, and $\sup_{\mathbf{u} \in \mathcal{U}, \ \mathbf{u} \neq \mathbf{0}} ||L\mathbf{u}||_{\mathcal{W}} / ||\mathbf{u}||_{\mathcal{U}}$ exists and is finite. Moreover, from the linearity of L it follows that

$$\sup_{\mathbf{u} \in \mathcal{U}, \ \mathbf{u} \neq \mathbf{0}} \frac{||L(\mathbf{u})||_{\mathcal{W}}}{||\mathbf{u}||_{\mathcal{U}}} = \sup_{||\mathbf{u}||_{\mathcal{U}} = 1} ||L(\mathbf{u})||_{\mathcal{W}} \tag{1.105}$$

and

$$||L_1 + L_2||_N = \sup_{||\mathbf{u}||_{\mathcal{U}}=1} ||L_1(\mathbf{u}) + L_2(\mathbf{u})||_{\mathcal{W}} \leq \sup_{||\mathbf{u}||_{\mathcal{U}}=1} (||L_1(\mathbf{u})||_{\mathcal{W}} + ||L_2(\mathbf{u})||_{\mathcal{W}})$$

$$\leq \sup_{||\mathbf{u}||_{\mathcal{U}}=1} ||L_1(\mathbf{u})||_{\mathcal{W}} + \sup_{||\mathbf{u}||_{\mathcal{U}}=1} ||L_2(\mathbf{u})||_{\mathcal{W}} = ||L_1||_N + ||L_2||_N. \qquad (1.106)$$

In a similar way we can prove that $||\alpha L||_N = |\alpha|\ ||L||_N$ for each $L \in \mathcal{L}(\mathcal{U}, \mathcal{W})$ and $\alpha \in \mathbb{R}$. Finally we have $||L||_N = 0$ if and only if $\sup_{||\mathbf{u}||_{\mathcal{U}}=1} ||L(\mathbf{u})||_{\mathcal{W}} = 0$ if and only if $||L(\mathbf{u})||_{\mathcal{W}} \leq 0$ for each $\mathbf{u} \in \mathcal{U}$, $||\mathbf{u}||_{\mathcal{U}} = 1$, if and only if $L(\mathbf{u}) = 0$ for each $\mathbf{u} \in \mathcal{U}$. We can the conclude that (1.104) is a norm on $\mathcal{L}(\mathcal{U}, \mathcal{W})$ called *natural norm*.

The norm (1.104) is not induced by any scalar product. To prove this, let us put $\mathcal{U} = \mathcal{W}$ and let $\{\mathbf{u}_1, ..., \mathbf{u}_n\}$ be an orthonormal basis of $\mathcal{U}$. For the linear mappings

$$L_1(\mathbf{u}) = \mathbf{u},\ L_2(\mathbf{u}) = (\mathbf{u}, \mathbf{u}_1)\mathbf{u}_1;$$

we have $||L_1 + L_2||_N = 2$, $||L_1 - L_2||_N = 1$, $||L_1||_N = ||L_2||_N = 1$, then (1.104) does not satisfy the parallelogram law.

Example 2. *Let $T : \mathcal{U} \to \mathcal{W}$ be an isometry, we have*

$$||T||_N = \sup_{\mathbf{u} \in \mathcal{U},\ \mathbf{u} \neq \mathbf{0}} \frac{||T(\mathbf{u})||_{\mathcal{W}}}{||\mathbf{u}||_{\mathcal{U}}} = \sup_{\mathbf{u} \in \mathcal{U},\ \mathbf{u} \neq \mathbf{0}} \frac{||\mathbf{u}||_{\mathcal{U}}}{||\mathbf{u}||_{\mathcal{U}}} = 1. \qquad (1.107)$$

For the mapping (1.88) defined in the example 1, we have

$$||L||_N = \sup_{\mathbf{u} \in \mathcal{S},\ \mathbf{u} \neq \mathbf{0}} \frac{||L(\mathbf{u})||_{\mathcal{S}}}{||\mathbf{u}||_{\mathcal{S}}} = \sup_{\mathbf{u} \in \mathcal{S},\ \mathbf{u} \neq \mathbf{0}} \frac{\sqrt{\sum_{i=1}^{k} |(\mathbf{u}, \mathbf{e}_i)|^2}}{||\mathbf{u}||_{\mathcal{S}}} \leq 1,$$

choosing $\mathbf{u} \in$Span$(\mathbf{e}_1, ..., \mathbf{e}_k)$, we get $||L||_N = 1$.

1.11 Functionals

Let $\mathcal{S}$ be a vector space with inner product $<, >$. A function ψ on $\mathcal{S}$ into $\mathbb{R}$ is called *functional.* ψ is a *linear* functional if the properties (1.54) and (1.55)are satisfied,

1. $\psi(\alpha \mathbf{a}) = \alpha \psi(\mathbf{a})$, for each $\mathbf{a} \in \mathcal{S}$, $\alpha \in \mathbb{R}$ (homogeneity).
2. $\psi(\mathbf{a} + \mathbf{b}) = \psi(\mathbf{a}) + \psi(\mathbf{b})$, for each $\mathbf{a}, \mathbf{b} \in \mathcal{S}$ (additivity),

Given $\mathbf{b} \in \mathcal{S}$, the functional $\psi(\mathbf{u}) =< \mathbf{u}, \mathbf{b} >$, $\mathbf{u} \in \mathcal{S}$ is linear, on the contrary $\psi(\mathbf{u}) = ||\mathbf{u}||$, $\mathbf{u} \in \mathcal{S}$, is not linear, in fact,in general we have $||\alpha \mathbf{u}|| = |\alpha|\ ||\mathbf{u}|| \neq \alpha ||\mathbf{u}||$.

The following theorem is known are representation theorem for linear functionals.

Theorem 4. *Let $\mathcal{S}$ be a finite-dimensional vector space with inner product $<,>$ and $\psi : \mathcal{S} \to \mathbb{R}$ a linear functional. There exists a unique $\mathbf{a} \in \mathcal{S}$ such that*

$$\psi(\mathbf{u}) =< \mathbf{a}, \mathbf{u} >, \quad \textit{for each } \mathbf{u} \in \mathcal{S}. \tag{1.108}$$

Proof. If $\psi = 0$, (1.108) is verified by $\mathbf{a} = \mathbf{0}$. Them let us assume that $\psi \neq 0$ and consider the subspaces of $\mathcal{S}$,

$$\mathcal{M} = \{\mathbf{v} \in \mathcal{S} \ \mid \psi(\mathbf{v}) = 0\} \tag{1.109}$$

and

$$\mathcal{M}^{\perp} = \{\mathbf{u} \in \mathcal{S} \ \mid \ < \mathbf{v}, \mathbf{u} >= 0 \ \text{ for each } \mathbf{v} \in \mathcal{M}\}. \tag{1.110}$$

Since $\psi \neq 0$, $\mathcal{M}^{\perp}$ contains at least an element $\mathbf{z}$ different from zero, and then we can put $\mathbf{a} = \psi(\mathbf{w})\mathbf{w}$, where $\mathbf{w} = \mathbf{z}/||\mathbf{z}||$. We have

$$< \mathbf{a}, \mathbf{w} >= \psi(\mathbf{w}) < \mathbf{w}, \mathbf{w} >= \psi(\mathbf{w}) \tag{1.111}$$

and, if $\mathbf{u} \in \mathcal{M}$, $0 = \psi(\mathbf{u}) =< \mathbf{a}, \mathbf{u} >$.

Let us fix $\mathbf{u} \in \mathcal{S}$; for each $\lambda \in \mathbb{R}$ we have

$$\mathbf{u} = \lambda\mathbf{w} + \mathbf{u} - \lambda\mathbf{w}, \tag{1.112}$$

where $\lambda\mathbf{w} \in \mathcal{M}^{\perp}$, and if we choose $\lambda = \psi(\mathbf{u})/\psi(\mathbf{w})$ we have

$$\psi(\mathbf{u} - \lambda\mathbf{w}) = \psi(\mathbf{u}) - \frac{\psi(\mathbf{u})}{\psi(\mathbf{w})}\psi(\mathbf{w}) = 0, \tag{1.113}$$

and then $\mathbf{u} - \lambda\mathbf{w} \in \mathcal{M}$. Taking into account the linearity of ψ, the choice of λ and (1.111), we have

$$< \mathbf{a}, \mathbf{u} >=< \mathbf{a}, \lambda\mathbf{w} + \mathbf{u} - \lambda\mathbf{w} >=< \mathbf{a}, \lambda\mathbf{w} >= \frac{\psi(\mathbf{u})}{\psi(\mathbf{w})} < \mathbf{a}, \mathbf{w} >= \psi(\mathbf{u}), \tag{1.114}$$

which proves the existence of $\mathbf{a}$. As far as the uniqueness is concerned, let us assume that there exist $\mathbf{a}_1, \mathbf{a}_2 \in \mathcal{S}$ such that

$$\psi(\mathbf{u}) =< \mathbf{a}_1, \mathbf{u} >=< \mathbf{a}_2, \mathbf{u} >, \quad \text{for each } \mathbf{u} \in \mathcal{S}. \tag{1.115}$$

Chosen $\mathbf{u} = \mathbf{a}_1 - \mathbf{a}_2$ from (1.115) we get

$$< \mathbf{a}_1 - \mathbf{a}_2, \mathbf{a}_1 - \mathbf{a}_2 >= 0, \tag{1.116}$$

from which, in view of the property **s3.** of the inner product, the equality $\mathbf{a}_1 = \mathbf{a}_2$ follows □

The subspace $\mathcal{M}$ is called the *kernel* of ψ. From the theorem above it follows that if $\mathcal{S}$ has dimension n and ψ is different from 0, then the dimension of $\mathcal{M}$ is $n - 1$.

A functional $\varphi : \mathcal{S} \to \mathbb{R}$ is *continuous* at $\mathbf{a}_0 \in \mathcal{S}$ if for each $\varepsilon > 0$ there is $\delta > 0$ such that for each $\mathbf{a} \in \mathcal{S}$ satisfying $||\mathbf{a} - \mathbf{a}_0|| < \delta$, we have $|\varphi(\mathbf{a}) - \varphi(\mathbf{a}_0)| < \varepsilon$.

φ is continuous on $\mathcal{S}$ if it is continuous at each $\mathbf{a}_0 \in \mathcal{S}$.

The functional $||\ || : \mathcal{S} \to \mathbb{R}$ which assigns to each vector its norm, is continuous at each $\mathbf{a}_0 \in \mathcal{S}$, in fact, in view of the inequality (1.35) we have

$$|||\mathbf{a}|| - ||\mathbf{a}_0||| \leq ||\mathbf{a} - \mathbf{a}_0||.$$

Analogously, from the inequality (1.35) it follows that for each given $\overline{\mathbf{a}} \in \mathcal{S}$, the functional $\varphi : \mathcal{S} \to \mathbb{R}$ defined by $\varphi(\mathbf{v}) = ||\mathbf{v} - \overline{\mathbf{a}}||$ with $\mathbf{v} \in \mathcal{S}$ is continuous.

Remark 2. *From theorem 4 it follows that every linear functional $\psi : \mathcal{S} \to \mathbb{R}$ is bounded and then continuous on $\mathcal{S}$. In fact, from both (1.108) and the Schwarz inequality (1.46) we get*

$$|\psi(\mathbf{v})| \leq ||\mathbf{a}||\ ||\mathbf{v}|| \quad \textit{for each } \mathbf{v} \in \mathcal{S}, \tag{1.117}$$

then ψ is bounded

The vector space $\mathcal{S}^* = \mathcal{L}(\mathcal{S}, \mathbb{R})$ constituted by all linear functionals on $\mathcal{S}$ is called the *dual space* of $\mathcal{S}$. If the vector space $\mathcal{S}$ has dimension n, then $\mathcal{S}^*$ has dimension n. For $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ an orthonormal basis of $\mathcal{S}$, the n linear functionals $\varphi_i \in \mathcal{S}^*$ with $\varphi_i(\mathbf{u}) =< \mathbf{e}_i, \mathbf{u} >$, for each $\mathbf{u} \in \mathcal{S}$ are a basis of $\mathcal{S}^*$. In fact, the linear independence of φ_i, $i = 1, ..., n$ follows from the linear independence of vectors $\mathbf{e}_i$, $i = 1, ..., n$, moreover, given $\varphi \in \mathcal{S}^*$ in virtue of Theorem 4 there exist $\mathbf{a} \in \mathcal{S}$ such that $\varphi(\mathbf{u}) = (\mathbf{a}, \mathbf{u})$, for each $\mathbf{u} \in \mathcal{S}$ and

$$\varphi = \sum_{i=1}^{n} < \mathbf{a}, \mathbf{e}_i > \varphi_i.$$

In view of (1.104) and (1.117) we have

$$||\varphi||_N = ||\mathbf{a}||. \tag{1.118}$$

The vector spaces $\mathcal{S}$ and $\mathcal{S}^*$, having the same dimension, are isomorphic.

1.12 Projections

Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be subspaces of $\mathcal{S}$ with $\mathcal{M}_2$ complement of $\mathcal{M}_1$. Then each $\mathbf{s} \in \mathcal{S}$ can be written in a unique way as $\mathbf{s} = \mathbf{s}_1 + \mathbf{s}_2$, with $\mathbf{s}_1 \in \mathcal{M}_1$ and $\mathbf{s}_2 \in \mathcal{M}_2$. The *projection on $\mathcal{M}_1$ along $\mathcal{M}_2$* is the mapping $P_{\mathcal{M}_1}$ defined by $P_{\mathcal{M}_1}(\mathbf{s}) = \mathbf{s}_1$. $P_{\mathcal{M}_1}$ is linear and idempotent, $(P_{\mathcal{M}_1})^2 = P_{\mathcal{M}_1}$.

Example 3. *In $\mathbb{R}^2$ given $\mathbf{x}^1 = (1, 0)$ and $\mathbf{x}^2 = (1, 1)$, consider the subspaces $\mathcal{M}_1 = Span(\mathbf{x}^1)$ and $\mathcal{M}_2 = Span(\mathbf{x}^2)$. For each $\mathbf{v} = (v_1, v_2)$, we have $\mathbf{v} = \mathbf{v}^1 + \mathbf{v}^2$, with $\mathbf{v}^1 = (v_1 - v_2, 0) \in \mathcal{M}_1$, $\mathbf{v}^2 = (v_2, v_2) \in \mathcal{M}_2$, and the projection on $\mathcal{M}_1$ along $\mathcal{M}_2$ is defined by $P_{\mathcal{M}_1}(\mathbf{v}) = \mathbf{v}^1$. For $\mathbf{x}^3 = (1, 2)$ and $\mathcal{M}_3 = Span(\mathbf{x}^3)$, we have $\mathbf{v} = \mathbf{u}^1 + \mathbf{u}^2$, with $\mathbf{u}^1 = (\frac{2v_1 - v_2}{2}, 0) \in \mathcal{M}_1$, $\mathbf{u}^2 = (\frac{v_2}{2}, v_2) \in \mathcal{M}_3$, and the projection on $\mathcal{M}_1$ along $\mathcal{M}_3$ is defined by $P_{\mathcal{M}_1}(\mathbf{v}) = \mathbf{u}^1$.*

Let $\mathcal{M}$ be a subspace of $\mathcal{S}$ with an inner product, the *orthogonal projection* on $\mathcal{M}$ is the linear mapping $P_{\mathcal{M}}$ that assigns to each vector $\mathbf{v} \in S$ the vector $\overline{\mathbf{v}} = P_{\mathcal{M}}(\mathbf{v}) \in \mathcal{M}$ defined in (1.66). $\overline{\mathbf{v}}$ satisfies the condition

$$< \mathbf{v} - \overline{\mathbf{v}}, \mathbf{w} >= 0 \ \text{ for each } \mathbf{w} \in \mathcal{M}. \tag{1.119}$$

Now we are in the position to extend the notion of projection to a non-empty closed convex set of $\mathcal{S}$.

Given a subset $\mathcal{A}$ of $\mathcal{S}$ and $\mathbf{u}_0 \in \mathcal{S}$, the *distance* of $\mathbf{u}_0$ from $\mathcal{A}$ is the scalar

$$dist((\mathbf{u}_0, \mathcal{A}) = \inf_{\mathbf{v}\in\mathcal{A}} ||\mathbf{u}_0 - \mathbf{v}||. \tag{1.120}$$

The following result is known as minimum norm theorem.

Theorem 5. *Let $\mathcal{S}$ be a finite-dimensional real vector space with the inner product $<,>$ and let $\mathcal{K} \subset \mathcal{S}$ be a non-empy closed convex subset of $\mathcal{S}$. For each $\mathbf{f} \in \mathcal{S}$ there is a unique $\mathbf{u} \in \mathcal{K}$ which satisfies the following equivalent conditions*

$$\|\mathbf{f} - \mathbf{u}\| = \min_{\mathbf{v}\in\mathcal{K}} \|\mathbf{f} - \mathbf{v}\| = dist(\mathbf{f}, \mathcal{K}), \tag{1.121}$$

$$< \mathbf{f} - \mathbf{u}, \mathbf{v} - \mathbf{u} >\leq 0 \quad \textit{for each } \mathbf{v} \in \mathcal{K}. \tag{1.122}$$

The vector $\mathbf{u} = P_{\mathcal{K}}(\mathbf{f})$ is called *projection* of $\mathbf{f}$ onto the closed convex set $\mathcal{K}$.

Proof. First of all, let us prove that there exists $\mathbf{u} \in \mathcal{K}$ which satisfies (1.121), then we prove the equivalence of (1.121) and (1.122) and finally the uniqueness of $\mathbf{u} \in \mathcal{K}$ satisfying (1.122).

If $\mathbf{f} \in \mathcal{K}$, then $\mathbf{u} = \mathbf{f}$; if $\mathbf{f} \notin \mathcal{K}$, we set $d = dist(\mathbf{f}, \mathcal{K})$. From the definition of infimum it follows that there is a sequence $\{\mathbf{u}^{(k)}\}_{k\in\mathbb{N}} \subset \mathcal{K}$ such that $d_k = \left\|\mathbf{u}^{(k)} - \mathbf{f}\right\| \to d$, for $k \to \infty$. $\{\mathbf{u}^{(k)}\}_{k\in\mathbb{N}}$ is a Cauchy sequence, in fact, by using the parallelogram law (1.47) with $\mathbf{a} = \mathbf{f} - \mathbf{u}^{(p)}$, $\mathbf{b} = \mathbf{f} - \mathbf{u}^{(q)}$, recalling that $\mathcal{K}$ is convex, we have

$$||\mathbf{u}^{(p)} - \mathbf{u}^{(q)}||^2 \leq 2d_p^2 + 2d_q^2 - 4d^2, \tag{1.123}$$

from which we get

$$||\mathbf{u}^{(p)} - \mathbf{u}^{(q)}|| \to 0, \quad \text{when } \ p, q \to \infty. \tag{1.124}$$

Thus, $\mathbf{u}^{(p)} \to \mathbf{u} \in \mathcal{K}$ for $p \to \infty$ (see Proposition 9) and $d = \|\mathbf{u} - \mathbf{f}\|$ because the norm is a continuous functional.

Now we have to prove the equivalence of (1.121) and (1.122). Let us assume that $\mathbf{u} \in \mathcal{K}$ satisfies (1.121), for each $\mathbf{w} \in \mathcal{K}$ we have

$$\mathbf{v} = (1 - t)\mathbf{u} + t\mathbf{w} \in \mathcal{K} \ \text{ for each } t \in [0, 1]$$

and then

$$\|\mathbf{f} - \mathbf{u}\| \leq \|\mathbf{f} - (1 - t)\mathbf{u} - t\mathbf{w}\| = \|\mathbf{f} - \mathbf{u} - t(\mathbf{w} - \mathbf{u})\|.$$

As a consequence, for $t \in (0,1]$, we have

$$\|\mathbf{f}-\mathbf{u}\|^2 \leq \|\mathbf{f}-\mathbf{u}\|^2 - 2t(\mathbf{f}-\mathbf{u},\mathbf{w}-\mathbf{u}) + t^2\|\mathbf{w}-\mathbf{u}\|^2$$

and then $2 < \mathbf{f}-\mathbf{u},\mathbf{w}-\mathbf{u} > \leq t\|\mathbf{w}-\mathbf{u}\|^2$ which implies (1.122) when $t \to 0$.

Vice versa, let us assume that $\mathbf{u} \in \mathcal{K}$ satisfies (1.122), we have

$$\|\mathbf{u}-\mathbf{f}\|^2 - \|\mathbf{v}-\mathbf{f}\|^2 = 2 < \mathbf{f}-\mathbf{u},\mathbf{v}-\mathbf{u} > - \|\mathbf{u}-\mathbf{v}\|^2 \leq 0 \quad \text{for each } \mathbf{v} \in \mathcal{K},$$

from which (1.121) follows.

In order to prove the uniqueness of $\mathbf{u}$, let $\mathbf{u}_1$,$\mathbf{u}_2 \in \mathcal{K}$ satisfy (1.122). We have

$$< \mathbf{f}-\mathbf{u}_1,\mathbf{v}-\mathbf{u}_1 > \leq 0 \quad \text{for each } \mathbf{v} \in \mathcal{K}, \tag{1.125}$$

$$< \mathbf{f}-\mathbf{u}_2,\mathbf{v}-\mathbf{u}_2 > \leq 0 \quad \text{for each } \mathbf{v} \in \mathcal{K}. \tag{1.126}$$

Setting $\mathbf{v} = \mathbf{u}_2$ in (1.125) and $\mathbf{v} = \mathbf{u}_1$ in (1.126), and summing we get

$$||\mathbf{u}_2 - \mathbf{u}_1||^2 \leq 0. \tag{1.127}$$

□

Example 4. *Consider $\mathcal{S} = \mathbb{R}^2$ with the scalar product defined in (1.41), $\mathcal{K} = \{(x_1,x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}$. If $\mathbf{f} \notin \mathcal{K}$, we have $P_{\mathcal{K}}(\mathbf{f}) = \mathbf{f}/||\mathbf{f}||$. In fact, for each $\mathbf{v} \in \mathcal{K}$, we have*

$$||\mathbf{f}-\mathbf{v}|| \geq |||\mathbf{f}|| - ||\mathbf{v}||| = ||\mathbf{f}|| - ||\mathbf{v}|| \geq ||\mathbf{f}|| - 1 = |||\mathbf{f}|| - 1| = ||\mathbf{f} - \frac{\mathbf{f}}{||\mathbf{f}||}||.$$

If $\mathcal{K}$ is not convex, the uniqueness of the projection is not guaranteed (for $\mathcal{S} = \mathbb{R}^2$ and $\mathcal{K} = S(0,1)$, $P_{\mathcal{K}}(\mathbf{0}) = S(0,1)$) and if $\mathcal{K}$ is not closed the existence of the projection is guaranteed (see for example $\mathcal{K} = B(0,1)$).

The mapping $P_{\mathcal{K}} : \mathcal{S} \to \mathcal{K}$ defined in the preceding theorem is continuous. In fact, the following proposition holds.

Proposition 15. *Under the hypotheses of theorem 5 we have*

$$||P_{\mathcal{K}}(\mathbf{f}_1) - P_{\mathcal{K}}(\mathbf{f}_2)|| \leq ||\mathbf{f}_1 - \mathbf{f}_2||, \quad \textit{for each } \mathbf{f}_1,\mathbf{f}_2 \in \mathcal{S}. \tag{1.128}$$

Proof. Setting $\mathbf{u}_1 = P_{\mathcal{K}}(\mathbf{f}_1)$ and $\mathbf{u}_2 = P_{\mathcal{K}}(\mathbf{f}_2)$, in view of (1.122) we have

$$< \mathbf{f}_1-\mathbf{u}_1,\mathbf{v}-\mathbf{u}_1 > \leq 0 \quad \text{for each } \mathbf{v} \in \mathcal{K}, \tag{1.129}$$

$$< \mathbf{f}_2-\mathbf{u}_2,\mathbf{v}-\mathbf{u}_2 > \leq 0 \quad \text{for each } \mathbf{v} \in \mathcal{K}. \tag{1.130}$$

Putting $\mathbf{v} = \mathbf{u}_2$ in (1.129) and $\mathbf{v} = \mathbf{u}_1$ in (1.130), and summing we get

$$||\mathbf{u}_1 - \mathbf{u}_2||^2 \leq < \mathbf{f}_1 - \mathbf{f}_2, \mathbf{u}_1 - \mathbf{u}_2 >,$$

that, by taking the Schwarz inequality into account, implies

$$||\mathbf{u}_1 - \mathbf{u}_2|| \leq ||\mathbf{f}_1 - \mathbf{f}_2||.$$

□

Since a subspace is closed and convex, the orthogonal projection onto a subspace defined in (1.119) can be obtained as particular case of the minimum norm theorem and the following proposition holds. Unlike the projection onto a convex closed set, the projection onto a subspace is linear.

Proposition 16. *If $\mathcal{K} = \mathcal{M}$, with $\mathcal{M}$ subspace of $\mathcal{S}$, for each $\mathbf{f} \in \mathcal{S}$ the projection $\mathbf{u} = P_{\mathcal{M}}(\mathbf{f})$ of $\mathbf{f}$ onto $\mathcal{M}$ is characterized by*

$$\mathbf{u} \in \mathcal{M}, \quad < \mathbf{f} - \mathbf{u}, \mathbf{v} >= 0, \quad \textit{for each } \mathbf{v} \in \mathcal{M}, \tag{1.131}$$

and $P_{\mathcal{M}}$ is a linear mapping.

Proof. From (1.122) we get

$$< \mathbf{f} - \mathbf{u}, \mathbf{v} - \mathbf{u} > \leq 0 \quad \text{ for each } \mathbf{v} \in \mathcal{M},$$

and then

$$< \mathbf{f} - \mathbf{u}, t\mathbf{v} - \mathbf{u} > \leq 0 \quad \text{ for each } \mathbf{v} \in \mathcal{M}, \text{ for each } t \in \mathbb{R},$$

thus, it follows that

$$< \mathbf{f} - \mathbf{u}, \mathbf{v} >= 0, \quad \text{for each} \mathbf{v} \in \mathcal{M}.$$

Moreover, if $\mathbf{u}$ satisfies (1.131) we have

$$< \mathbf{f} - \mathbf{u}, \mathbf{v} - \mathbf{u} >= 0 \quad \text{ for each } \mathbf{v} \in \mathcal{M}.$$

Given $\mathbf{f}_1, \mathbf{f}_2 \in \mathcal{S}$, putting $\mathbf{u}_1 = P_{\mathcal{M}}(\mathbf{f}_1)$ and $\mathbf{u}_2 = P_{\mathcal{M}}(\mathbf{f}_2)$, from (1.131) it follows that

$$< \mathbf{f}_1 + \mathbf{f}_2 - \mathbf{u}_1 - \mathbf{u}_2, \mathbf{v} >=< \mathbf{f}_1 - \mathbf{u}_1, \mathbf{v} > + < \mathbf{f}_2 - \mathbf{u}_2, \mathbf{v} >= 0, \tag{1.132}$$

for each $\mathbf{v} \in \mathcal{M}$, then, $P_{\mathcal{M}}(\mathbf{f}_1 + \mathbf{f}_2) = P_{\mathcal{M}}(\mathbf{f}_1) + P_{\mathcal{M}}(\mathbf{f}_2)$. Analogously, we prove that $P_{\mathcal{M}}(\alpha \mathbf{f}) = \alpha P_{\mathcal{M}}(\mathbf{f})$ for each $\alpha \in \mathbb{R}$, $\mathbf{f} \in \mathcal{S}$. □

Example 5. *Let $A \in \mathcal{M}_{m,n}$ and $\mathbf{y} \in \mathbb{R}^m$ be given. For $m > n$ the linear system*

$$A\mathbf{x} = \mathbf{y}, \quad \mathbf{x} \in \mathbb{R}^n \tag{1.133}$$

may be overdetermined and can be solved via the least squares approach, which consists in minimizing the functional

$$\phi(\mathbf{x}) = ||A\mathbf{x} - \mathbf{y}||_2, \quad \mathbf{x} \in \mathbb{R}^n. \tag{1.134}$$

For $i = 1, ..., n$, the vectors $\mathbf{a}^{(i)}$, constituted by the columns of A, belongs to $\mathbb{R}^m$. Assuming that they are linearly independent, the subspace $\mathcal{M} = span(\mathbf{a}^{(1)}, ..., \mathbf{a}^{(n)})$ of $\mathbb{R}^m$ has dimension n and $A\mathbf{x} \in \mathcal{M}$ for each $\mathbf{x} \in \mathbb{R}^n$. Then, minimize (1.134) is equivalent to calculate

$$\min_{\mathbf{v} \in \mathcal{M}} ||\mathbf{v} - \mathbf{y}||_2. \tag{1.135}$$

In view of the minimum norm theorem, there is a unique $\mathbf{u} \in \mathbb{R}^m$ *such that*

$$||\mathbf{u} - \mathbf{y}||_2 = \min_{\mathbf{v} \in \mathcal{M}} ||\mathbf{v} - \mathbf{y}||_2 \tag{1.136}$$

and

$$< \mathbf{y} - \mathbf{u} > \cdot \mathbf{v} = 0, \text{ for each } \mathbf{v} \in \mathcal{M}. \tag{1.137}$$

Vector $\mathbf{u}$ *is the projection* $P_{\mathcal{M}}(\mathbf{y})$ *of* $\mathbf{y}$ *onto* $\mathcal{M}$ *and the unique* $\mathbf{x} \in \mathbb{R}^n$ *such that* $A\mathbf{x} = \mathbf{u}$ *is the minimum point of (1.134).*

1.13 Differentiation

Here we introduce the notion of differentiation of functions on normed vector spaces. Let $\mathcal{U}$ and $\mathcal{W}$ be two vector spaces with inner product and let T be a function defined on a neighborhood of $\mathbf{0} \in \mathcal{U}$ with values in $\mathcal{W}$. We say that $T(\mathbf{u})$ *approaches zero faster than* $\mathbf{u}$ and we write

$$T(\mathbf{u}) = o(\mathbf{u}) \quad \text{as } \mathbf{u} \to \mathbf{0} \tag{1.138}$$

if

$$\lim_{\mathbf{u} \neq \mathbf{0},\ \mathbf{u} \to \mathbf{0}} \frac{||T(\mathbf{u})||_{\mathcal{W}}}{||\mathbf{u}||_{\mathcal{U}}} = 0^1. \tag{1.139}$$

If T_1 and T_2 are two functions, $T_1(\mathbf{u}) = T_2(\mathbf{u}) + o(\mathbf{u})$ means that $T_1(\mathbf{u}) - T_2(\mathbf{u}) = o(\mathbf{u})$.

For example, for $\mathcal{U} = \mathcal{W} = \mathbb{R}$ and $T(t) = t^\alpha$ with $\alpha > 1$, we have $T(t) = o(t)$ as $t \to 0$.

Let $\mathbf{g}$ be a function defined on the open set $\mathcal{D} \subset \mathbb{R}$ into the vector space $\mathcal{W}$, the *derivative* of $\mathbf{g}$ at t, if it exists, is defined by

$$\dot{\mathbf{g}}(t) = \frac{d}{dt}\mathbf{g}(t) = \lim_{s \to 0} \frac{\mathbf{g}(t+s) - \mathbf{g}(t)}{s}. \tag{1.140}$$

In that case we say that $\mathbf{g}$ is *differentiable* at t. The function $\mathbf{g} : \mathcal{D} \to \mathcal{W}$ is of class C^1 (or smooth) if $\dot{\mathbf{g}}(t)$ exists at each $t \in \mathcal{D}$ and if the function $\dot{\mathbf{g}}$ is continuous on $\mathcal{D}$.

Let $\mathbf{g}$ be differentiable at t, then we have

$$\lim_{s \to 0} \frac{\mathbf{g}(t+s) - \mathbf{g}(t) - s\dot{\mathbf{g}}(t)}{s} = 0, \tag{1.141}$$

or equivalently

$$\mathbf{g}(t+s) = \mathbf{g}(t) + s\dot{\mathbf{g}}(t) + o(s), \quad s \to 0. \tag{1.142}$$

[1] In other words, for every $k > 0$ there is $k' > 0$ such that $||T(\mathbf{u})||_{\mathcal{W}} < k||\mathbf{u}||_{\mathcal{U}}$ if $||\mathbf{u}||_{\mathcal{U}} < k'$.

Since $s\dot{\mathbf{g}}(t)$ is linear in s, $\mathbf{g}(t+s)-\mathbf{g}(t)$ is equal to a function linear in s plus a term that approaches zero faster than s.

Let $\mathcal{U}$ and $\mathcal{W}$ be normed vector spaces, $\mathcal{D}$ an open subset of $\mathcal{U}$ and $T : \mathcal{D} \to \mathcal{W}$ a function. We say that T is *(Fréchet) differentiable* at $\mathbf{u} \in \mathcal{D}$ if the difference $T(\mathbf{u}+\mathbf{h})-T(\mathbf{u})$ is equal to a linear function of $\mathbf{h}$ plus a term that approaches zero faster than $\mathbf{h}$. More precisely, if there exists a linear mapping

$$DT(\mathbf{u}) : \mathcal{U} \to \mathcal{W} \tag{1.143}$$

such that

$$T(\mathbf{u}+\mathbf{h}) = T(\mathbf{u}) + DT(\mathbf{u})[\mathbf{h}] + o(\mathbf{h}), \quad \text{as } \mathbf{h} \to \mathbf{0}. \tag{1.144}$$

If $DT(\mathbf{u})$ exists, it is unique. In fact, for each $\mathbf{h} \in \mathcal{U}$ we have

$$DT(\mathbf{u})[\mathbf{h}] = \lim_{\alpha \to 0} \frac{T(\mathbf{u}+\alpha\mathbf{h}) - T(\mathbf{u})}{\alpha} = \frac{d}{d\alpha} T(\mathbf{u}+\alpha\mathbf{h})|_{\alpha=0}. \tag{1.145}$$

We call $DT(\mathbf{u})$ the *(Fréchet) derivative* of T at $\mathbf{u}$.

If T is differentiable at each $\mathbf{u} \in \mathcal{D}$, then DT is a function from $\mathcal{D}$ to the space $\mathcal{L}(\mathcal{U},\mathcal{W})$ of linear mappings from $\mathcal{U}$ to $\mathcal{W}$, introduced in (1.100),

$$DT : \mathcal{D} \to \mathcal{L}(\mathcal{U},\mathcal{W}), \tag{1.146}$$

which assigns to each $\mathbf{u} \in \mathcal{D}$ the linear mapping $DT(\mathbf{u})$.

A function $T : \mathcal{D} \to \mathcal{W}$ is of class C^1 (or smooth) if T is differentiable at each $\mathbf{u} \in \mathcal{D}$ and DT is continuous.

If $\mathcal{D}$ is an open subset of $\mathbb{R}$ and $\mathbf{g}$ a function from $\mathcal{D}$ to $\mathcal{W}$, from (1.142) it follows that $D\mathbf{g}(t)[s] = s\dot{\mathbf{g}}(t)$.

The following theorem holds.

Theorem 6. *Let $T : \mathcal{D} \to \mathcal{W}$ be a function, with $\mathcal{D}$ open subset of $\mathcal{U}$. If T is (Fréchet) differentiable at $\mathbf{u}_0 \in \mathcal{D}$ then T is continuous at $\mathbf{u}_0$.*

Proof. For $\mathbf{u} \in \mathcal{U}$, we have

$$T(\mathbf{u}_0+\mathbf{u}) - T(\mathbf{u}_0) = T(\mathbf{u}_0+\mathbf{u}) - T(\mathbf{u}_0) - DT(\mathbf{u}_0)[\mathbf{u}] + DT(\mathbf{u}_0)[\mathbf{u}], \tag{1.147}$$

since the ratio

$$\frac{||T(\mathbf{u}_0+\mathbf{u}) - T(\mathbf{u}_0) - DT(\mathbf{u}_0)[\mathbf{u}]||}{||\mathbf{u}||} \tag{1.148}$$

converges to 0 as $\mathbf{u} \to \mathbf{0}$, there is $\delta_0 > 0$ such that if $0 < ||\mathbf{u}|| < \delta_0$ we have

$$||T(\mathbf{u}_0+\mathbf{u}) - T(\mathbf{u}_0) - DT(\mathbf{u}_0)[\mathbf{u}]|| \le ||\mathbf{u}||. \tag{1.149}$$

Moreover, since $DT(\mathbf{u}_0)$ is linear, from Proposition 11, it is bounded as well, the there exists $\kappa > 0$ such that

$$||DT(\mathbf{u}_0)[\mathbf{u}]|| \leq \kappa||\mathbf{u}||, \quad \text{for each } \mathbf{u} \in \mathcal{U}. \tag{1.150}$$

Then, for $||\mathbf{u}|| < \delta_0$ we have

$$||T(\mathbf{u}_0 + \mathbf{u}) - T(\mathbf{u}_0)|| \leq ||T(\mathbf{u}_0 + \mathbf{u}) - T(\mathbf{u}_0) - DT(\mathbf{u}_0)[\mathbf{u}]|| + ||DT(\mathbf{u}_0)[\mathbf{u}]|| \leq (1+\kappa)||\mathbf{u}||. \tag{1.151}$$

Finally, for each $\varepsilon > 0$, putting $\delta = \min(\delta_0, \varepsilon/1+\kappa)$, the continuity of T at $\mathbf{u}_0$ follows. □

Example 6. *Let $L : \mathcal{U} \to \mathcal{W}$ be a linear application. For $\mathbf{u}_0, \mathbf{u} \in \mathcal{U}$ we have*

$$L(\mathbf{u}_0 + \mathbf{u}) = L(\mathbf{u}_0) + L(\mathbf{u}), \tag{1.152}$$

therefore, $DL(\mathbf{u}_0) = L$, that is DL is constant.

Let $\mathcal{M}$ be a subspace of the vector space $\mathcal{S}$ with inner product, $\{\mathbf{e}_1, ..., \mathbf{e}_k\}$ an orthonormal basis $\mathcal{M}$ and $P_{\mathcal{M}}$ the projection onto $\mathcal{M}$, $P_{\mathcal{M}}(\mathbf{v}) = \sum_{i=1}^{k} < \mathbf{v}, \mathbf{e}_i > \mathbf{e}_i$. We have

$$DP_{\mathcal{M}}(\mathbf{v})[\mathbf{h}] = \sum_{i=1}^{k} < \mathbf{h}, \mathbf{e}_i > \mathbf{e}_i = P_{\mathcal{M}}(\mathbf{h}), \quad \textit{for each } \mathbf{h} \in \mathcal{S}. \tag{1.153}$$

Example 7. *Let $\mathcal{U}$ a vector space with inner product $<,>$, $\phi : \mathcal{U} \to \mathbb{R}$ the nonlinear functional defined by $\phi(\mathbf{u}) =< \mathbf{u}, \mathbf{u} >$, $\mathbf{u} \in \mathcal{U}$. We have*

$$\phi(\mathbf{u} + \mathbf{h}) =< \mathbf{u} + \mathbf{h}, \mathbf{u} + \mathbf{h} >= \phi(\mathbf{u}) + 2 < \mathbf{u}, \mathbf{h} > + \phi(\mathbf{h}), \tag{1.154}$$

with $\phi(\mathbf{h}) = o(\mathbf{h}) \quad$ as $\mathbf{h} \to \mathbf{0}$, in fact,

$$\frac{< \mathbf{h}, \mathbf{h} >}{||\mathbf{h}||} = ||\mathbf{h}|| \to 0 \ \textit{as } \mathbf{h} \to \mathbf{0}. \tag{1.155}$$

Since $2(\mathbf{u}, \mathbf{h})$ is linear in $\mathbf{h}$, we have that

$$D\phi(\mathbf{u})[\mathbf{h}] = 2 < \mathbf{u}, \mathbf{h} >, \quad \textit{for each } \mathbf{h} \in \mathcal{U}. \tag{1.156}$$

Let $\mathcal{X}$, $\mathcal{X}_1$, $\mathcal{X}_2$ and $\mathcal{Y}$ be finite-dimensional vector spaces with inner product; let $\mathcal{D}$ be an open subset of $\mathcal{X}$.

Let us consider the bilinear mapping [2] $\pi : \mathcal{X}_1 \times \mathcal{X}_2 \to \mathcal{Y}$ which assigns to each $\mathbf{f}_0 \in \mathcal{X}_1$ and $\mathbf{g}_0 \in \mathcal{X}_2$ the product $\pi(\mathbf{f}_0, \mathbf{g}_0) \in \mathcal{Y}$. Within this framework, the product $P = \pi(F, G)$ of two functions $F : \mathcal{D} \to \mathcal{X}_1$ and $G : \mathcal{D} \to \mathcal{X}_2$ is the function $P : \mathcal{D} \to \mathcal{Y}$ defined by

$$P(\mathbf{u}) = \pi(F(\mathbf{u}), G(\mathbf{u})), \quad \text{for each } \mathbf{u} \in \mathcal{D}. \tag{1.157}$$

Let us state the following fundamental proposition.

[2] A mapping $\pi : \mathcal{X}_1 \times \mathcal{X}_2 \to \mathcal{Y}$ is bilinear if $\pi(\alpha\mathbf{f}_1 + \beta\mathbf{f}_2, \mathbf{g}) = \alpha\pi(\mathbf{f}_1, \mathbf{g}) + \beta\pi(\mathbf{f}_2, \mathbf{g})$ and $\pi(\mathbf{f}, \alpha\mathbf{g}_1 + \beta\mathbf{g}_2) = \alpha\pi(\mathbf{f}, \mathbf{g}_1) + \beta\pi(\mathbf{f}, \mathbf{g}_2)$ for each $\mathbf{f}, \mathbf{f}_1, \mathbf{f}_2 \in \mathcal{X}_1$, $\mathbf{g}, \mathbf{g}_1, \mathbf{g}_2 \in \mathcal{X}_2$, $\alpha, \beta \in \mathbb{R}$.

Proposition 17. *(Product Rule) Let F and G be differentiable at $\mathbf{u} \in \mathcal{D}$. Then their product $P = \pi(F, G)$ is differentiable at $\mathbf{u}$ and*

$$DP(\mathbf{u})[\mathbf{h}] = \pi(DF(\mathbf{u})[\mathbf{h}], G(\mathbf{u})) + \pi(F(\mathbf{u}), DG(\mathbf{u})[\mathbf{h}]) \tag{1.158}$$

for all $\mathbf{h} \in \mathcal{X}$.

Remark 3. *If $\mathcal{X} = \mathbb{R}$, by replacing $\mathbf{u}$ with t in (1.158) we have*

$$\dot{P}(t) = \pi(\dot{F}(t), G(t)) + \pi(F(t), \dot{G}(t)). \tag{1.159}$$

Let $\mathcal{G}$ be an open subset of $\mathcal{X}_1$, $F : \mathcal{D} \to \mathcal{X}_1$ and $G : \mathcal{G} \to \mathcal{Y}$, with $F(\mathcal{D}) = \{\mathbf{v} \in \mathcal{X}_1 : \mathbf{v} = F(\mathbf{u}), \mathbf{u} \in \mathcal{D}\} \subset \mathcal{G}$.

Proposition 18. *(Chain Rule)*

Let F be differentiable at $\mathbf{u} \in \mathcal{D}$ and G be differentiable at $\mathbf{v} = F(\mathbf{u})$. The the composition $C = G \circ F$ is differentiable at $\mathbf{u}$ and

$$DC(\mathbf{u})[\mathbf{h}] = DG(F(\mathbf{u}))[DF(\mathbf{u})[\mathbf{h}]] \tag{1.160}$$

for every $\mathbf{h} \in \mathcal{X}$.

Remark 4. *If $\mathcal{X} = \mathbb{R}$, writing t in place of $\mathbf{u}$ in (1.160) we have*

$$\frac{d}{dt}C(t) = DG(F(t)[\dot{F}(t)]. \tag{1.161}$$

Example 8. *Let us consider the functional $\psi : \mathcal{U} - \mathbf{0} \to \mathbb{R}$ defined by $\psi(\mathbf{u}) = \sqrt{< \mathbf{u}, \mathbf{u} >}$, $\mathbf{u} \in \mathcal{U}$. ψ if the composition of the function $f : \mathbb{R}^+ \to \mathbb{R}$ defined by $f(s) = \sqrt{s}$, for each $s \in \mathbb{R}^+$ and the functional ϕ given in Example 7,*

$$\psi(\mathbf{u}) = f(\phi(\mathbf{u})), \quad \mathbf{u} \in \mathcal{U}. \tag{1.162}$$

From Proposition 18, by taking (1.156) into account, we get that the derivative of ψ at $\mathbf{u}$ is given by

$$D\psi(\mathbf{u})[\mathbf{h}] = Df(\phi(\mathbf{u}))[D\phi(\mathbf{u})[\mathbf{h}]] \tag{1.163}$$

$$= \frac{1}{2}(\phi(\mathbf{u}))^{-1/2}\ 2 < \mathbf{u}, \mathbf{h} >= \frac{1}{||\mathbf{u}||} < \mathbf{u}, \mathbf{h} >, \quad \textit{for every } \mathbf{h} \in \mathcal{U}. \tag{1.164}$$

Chapter 2

Tensor calculus

This chapter is devoted to some results of tensor algebra and analysis. The term tensor stands for a linear function from an inner product space to itself.

Let $\mathcal{V}$ be a real vector space of dimension $n \geq 2$ equipped with the scalar product $\cdot$. Denoted by $\{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n\}$ an orthonormal basis of $\mathcal{V}$, for every $\mathbf{u} \in \mathcal{V}$ the quantities

$$u_i = \mathbf{u} \cdot \mathbf{e}_i, \qquad i = 1, ..., n \tag{2.1}$$

are the (Cartesian) *components* of $\mathbf{u}$ and we have

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{n} u_i v_i \qquad \text{e} \qquad \|\mathbf{u}\| = \sqrt{\sum_{i=1}^{n} u_i^2}. \tag{2.2}$$

If $n = 3$ it is possible to prove via geometric considerations that $\mathbf{u} \cdot \mathbf{v} = ||\mathbf{u}||\, ||\mathbf{v}|| \cos\theta$, where $\theta \in [0, \pi]$ is the angle between the vectors $\mathbf{u}$ and $\mathbf{v}$.

2.1 Second-order tensors

A *(second-order) tensor* $\mathbf{A}$ is a linear mapping from $\mathcal{V}$ into $\mathcal{V}$,

$$\mathbf{A}(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha\mathbf{A}\mathbf{u} + \beta\mathbf{A}\mathbf{v}, \quad \text{for each } \alpha, \beta \in \mathbb{R},\ \mathbf{u}, \mathbf{v} \in \mathcal{V}. \tag{2.3}$$

The set

$$\text{Lin} = \{\mathbf{A} : \mathcal{V} \to \mathcal{V} \mid \mathbf{A} \text{ is linear}\} \tag{2.4}$$

of all tensors is a vector space. Given $\mathbf{A}, \mathbf{B} \in$Lin, $\alpha \in \mathbb{R}$, the tensors $\mathbf{A} + \mathbf{B}$ and $\alpha\mathbf{A}$ are defined as in the following

$$(\mathbf{A} + \mathbf{B})\mathbf{v} = \mathbf{A}\mathbf{v} + \mathbf{B}\mathbf{v}, \quad \text{for all } \mathbf{v} \in \mathcal{V}, \tag{2.5}$$

$$(\alpha\mathbf{A})\mathbf{v} = \alpha\mathbf{A}\mathbf{v}, \quad \text{for all } \mathbf{v} \in \mathcal{V}. \tag{2.6}$$

The zero tensor in Lin is the tensor $\mathbf{0}$ defined by

$$\mathbf{0v} = \mathbf{0}, \quad \text{for all } \mathbf{v} \in \mathcal{V}, \tag{2.7}$$

and the identity tensor $\mathbf{I}$ is defined by

$$\mathbf{Iv} = \mathbf{v}, \quad \text{for all } \mathbf{v} \in \mathcal{V}. \tag{2.8}$$

For $\alpha \in \mathbb{R}$, the mapping that assigns to each $\mathbf{v} \in \mathcal{V}$ the vector $\alpha\mathbf{v}$ is a tensor, on the contrary the function that assigns to each $\mathbf{v}$ the vector $(\mathbf{v} \cdot \mathbf{v})\mathbf{v}$ is not a tensor, because it is not linear.

If $\mathbf{A}, \mathbf{B} \in$ Lin, then the *product* $\mathbf{AB} \in$ Lin is defined by $(\mathbf{AB})\mathbf{u} = \mathbf{A}(\mathbf{Bu})$ for all $\mathbf{u} \in \mathcal{V}$. In general, $\mathbf{AB} \neq \mathbf{BA}$; if $\mathbf{AB} = \mathbf{BA}$ then we say that $\mathbf{A}$ and $\mathbf{B}$ *commute*. Given $\mathbf{A} \in$ Lin and the integer $k \geq 0$, we define the following powers of $\mathbf{A}$

$$\mathbf{A}^k = \begin{cases} \mathbf{I} & \text{if } k = 0, \\ \mathbf{A}^{k-1}\mathbf{A} & \text{if } k \geq 1. \end{cases} \tag{2.9}$$

Proposition 19. *For every tensor $\mathbf{A} \in$ Lin there is a unique tensor $\mathbf{A}^T$ such that*

$$\mathbf{A}^T\mathbf{v} \cdot \mathbf{u} = \mathbf{v} \cdot \mathbf{Au} \quad \textit{for all } \mathbf{u}, \mathbf{v} \in \mathcal{V}. \tag{2.10}$$

Tensor $\mathbf{A}^T$ is called transpose of $\mathbf{A}$.

Proof. Let us first prove that for each $\mathbf{A} \in$ Lin there is a tensor $\mathbf{A}^T$ which satisfies (2.10). To this end, for a fixed $\mathbf{v} \in \mathcal{V}$, let us consider the linear functional $\psi : \mathcal{V} \to \mathbb{R}$ defined by $\psi(\mathbf{u}) = \mathbf{Au} \cdot \mathbf{v}$, $\mathbf{u} \in \mathcal{V}$. From the theorem of representation of linear functionals it follows that there is a unique $\mathbf{a_v} \in \mathcal{V}$ such that

$$\psi(\mathbf{u}) = \mathbf{Au} \cdot \mathbf{v} = \mathbf{a_v} \cdot \mathbf{u}, \quad \text{for aech } \mathbf{u} \in \mathcal{V}. \tag{2.11}$$

Now let us consider the function $\mathbf{B}$ from $\mathcal{V}$ to $\mathcal{V}$ defined by

$$\mathbf{Bv} = \mathbf{a_v}, \quad \text{for each } \mathbf{v} \in \mathcal{V}; \tag{2.12}$$

$\mathbf{B}$ is linear, in fact if $\mathbf{v}, \mathbf{w} \in \mathcal{V}$ and $\alpha \in \mathbb{R}$, we have

$$\begin{aligned} \mathbf{B}(\mathbf{v} + \mathbf{w}) \cdot \mathbf{u} = \mathbf{a_{v+w}} \cdot \mathbf{u} = \mathbf{Au} \cdot (\mathbf{v} + \mathbf{w}) = \\ \mathbf{Au} \cdot \mathbf{v} + \mathbf{Au} \cdot \mathbf{w} = \mathbf{a_v} \cdot \mathbf{u} + \mathbf{a_w} \cdot \mathbf{u} = \\ \mathbf{Bv} \cdot \mathbf{u} + \mathbf{Bw} \cdot \mathbf{u}, \quad \text{for each } \mathbf{u} \in \mathcal{V}; \end{aligned} \tag{2.13}$$

$$\begin{aligned} \mathbf{B}(\alpha\mathbf{v}) \cdot \mathbf{u} = \mathbf{a_{\alpha v}} \cdot \mathbf{u} = \mathbf{Au} \cdot (\alpha\mathbf{v}) = \\ \alpha\mathbf{a_v} \cdot \mathbf{u} = \alpha\mathbf{Bv} \cdot \mathbf{u}, \quad \text{for each } \mathbf{u} \in \mathcal{V}. \end{aligned} \tag{2.14}$$

Let us put $\mathbf{A}^T = \mathbf{B}$, we have

$$\mathbf{Au} \cdot \mathbf{v} = \mathbf{a_v} \cdot \mathbf{u} = \mathbf{Bv} \cdot \mathbf{u} = \mathbf{A}^T\mathbf{v} \cdot \mathbf{u} \quad \text{for each } \mathbf{u}, \mathbf{v} \in \mathcal{V}. \tag{2.15}$$

To prove the uniqueness, let as assume that there exist two tensors $\mathbf{B}$ and $\mathbf{C}$ such that

$$\mathbf{v} \cdot \mathbf{A}\mathbf{u} = \mathbf{B}\mathbf{v} \cdot \mathbf{u} = \mathbf{C}\mathbf{v} \cdot \mathbf{u} \quad \text{for all } \mathbf{u}, \mathbf{v} \in \mathcal{V}, \tag{2.16}$$

then

$$(\mathbf{B} - \mathbf{C})\mathbf{v} \cdot \mathbf{u} = 0 \tag{2.17}$$

for each $\mathbf{u}, \mathbf{v} \in \mathcal{V}$. Setting $\mathbf{u} = (\mathbf{B} - \mathbf{C})\mathbf{v}$ from (2.17) we get $||(\mathbf{B} - \mathbf{C})\mathbf{v}|| = 0$, from which we obtain $(\mathbf{B} - \mathbf{C})\mathbf{v} = \mathbf{0}$ for each $\mathbf{v} \in \mathcal{V}$, and then $\mathbf{B} - \mathbf{C} = \mathbf{0}$. □

Proposition 20. *For $\mathbf{A}, \mathbf{B} \in$Lin, the following properties hold,*

$$(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T, \tag{2.18}$$

$$(\mathbf{A}\mathbf{B})^T = \mathbf{B}^T \mathbf{A}^T, \tag{2.19}$$

$$(\mathbf{A}^T)^T = \mathbf{A}. \tag{2.20}$$

Proof. For each $\mathbf{u}, \mathbf{v} \in \mathcal{V}$ we have

$$\mathbf{u} \cdot (\mathbf{A} + \mathbf{B})^T \mathbf{v} = (\mathbf{A} + \mathbf{B})\mathbf{u} \cdot \mathbf{v} = (\mathbf{A}\mathbf{u} + \mathbf{B}\mathbf{u}) \cdot \mathbf{v} = \tag{2.21}$$

$$\mathbf{u} \cdot \mathbf{A}^T \mathbf{v} + \mathbf{u} \cdot \mathbf{B}^T \mathbf{v} = \mathbf{u} \cdot (\mathbf{A}^T + \mathbf{B}^T)\mathbf{v}, \tag{2.22}$$

which proves (2.18). Properties (2.19) and (2.20) follows directly from the following equalities,

$$\mathbf{u} \cdot (\mathbf{A}\mathbf{B})^T \mathbf{v} = (\mathbf{A}\mathbf{B})\mathbf{u} \cdot \mathbf{v} = \mathbf{A}(\mathbf{B}\mathbf{u}) \cdot \mathbf{v} = \mathbf{B}\mathbf{u} \cdot \mathbf{A}^T \mathbf{v} = \tag{2.23}$$

$$\mathbf{u} \cdot \mathbf{B}^T \mathbf{A}^T \mathbf{v}, \tag{2.24}$$

$$\mathbf{u} \cdot (\mathbf{A}^T)^T \mathbf{v} = \mathbf{A}^T \mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{A}\mathbf{v}. \tag{2.25}$$

□

2.2 Symmetric and skew-symmetric tensors

A tensor $\mathbf{A} \in \text{Lin}$ is *symmetric* if $\mathbf{A}^T = \mathbf{A}$ and is *skew-symmetric* if $\mathbf{A}^T = -\mathbf{A}$. Let us denote by

$$\text{Sym} = \{\mathbf{A} \in \text{Lin} : \mathbf{A} = \mathbf{A}^T\} \tag{2.26}$$

the subspace of Lin of all symmetric tensors and by

$$\text{Skw} = \{\mathbf{W} \in \text{Lin} : \mathbf{W} = -\mathbf{W}^T\} \tag{2.27}$$

the subspace of Lin of all skew-symmetric tensors. Every $\mathbf{A} \in \text{Lin}$ can be written in a unique way as the sum of $(\mathbf{A} + \mathbf{A}^T)/2 \in \text{Sym}$ and $(\mathbf{A} - \mathbf{A}^T)/2 \in \text{Skw}$. Moreover, since $\text{Sym} \cap \text{Skw} = \{\mathbf{0}\}$, Lin is the direct sum of Sym and Skw,

$$\text{Lin} = \text{Sym} \oplus \text{Skw}. \tag{2.28}$$

Tensors $(\mathbf{A}+\mathbf{A}^T)/2$ e $(\mathbf{A}-\mathbf{A}^T)/2$ are called *symmetric part* and *skew-symmetric part* of $\mathbf{A}$.

2.3 Dyads

For $\mathbf{a}, \mathbf{b} \in \mathcal{V}$, $\mathbf{a} \otimes \mathbf{b}$ is the element of Lin defined by,

$$(\mathbf{a} \otimes \mathbf{b})\mathbf{u} = (\mathbf{u} \cdot \mathbf{b})\mathbf{a} \quad \text{for all } \mathbf{u} \in \mathcal{V}. \tag{2.29}$$

Tensor $\mathbf{a} \otimes \mathbf{b}$ is also called *dyad* and the symbol $\otimes$ denotes the *tensor product*. The relation (2.29) defines a tensor, in fact,

$$\mathbf{a} \otimes \mathbf{b}(\mathbf{u} + \mathbf{v}) = [(\mathbf{u} + \mathbf{v}) \cdot \mathbf{b}]\mathbf{a} = (\mathbf{u} \cdot \mathbf{b} + \mathbf{v} \cdot \mathbf{b})\mathbf{a} =$$

$$(\mathbf{a} \otimes \mathbf{b})\mathbf{u} + (\mathbf{a} \otimes \mathbf{b})\mathbf{v}, \tag{2.30}$$

$$\mathbf{a} \otimes \mathbf{b}(\alpha \mathbf{u}) = \alpha(\mathbf{u} \cdot \mathbf{b})\mathbf{a} = \alpha(\mathbf{a} \otimes \mathbf{b})\mathbf{u} \tag{2.31}$$

for all $\mathbf{u}, \mathbf{v} \in \mathcal{V}$, $\alpha \in \mathbb{R}$.

Proposition 21. *Consider* $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d} \in \mathcal{V}$ *and and let* $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ *be an orthonormal basis of* $\mathcal{V}$*.*

(i) *The following properties hold*

$$(\mathbf{a} \otimes \mathbf{b})^T = (\mathbf{b} \otimes \mathbf{a}), \tag{2.32}$$

$$(\mathbf{a} \otimes \mathbf{b})(\mathbf{c} \otimes \mathbf{d}) = (\mathbf{b} \cdot \mathbf{c})(\mathbf{a} \otimes \mathbf{d}), \tag{2.33}$$

$$(\mathbf{e}_i \otimes \mathbf{e}_j)(\mathbf{e}_k \otimes \mathbf{e}_l) = \begin{cases} 0, & j \neq k, \\ \mathbf{e}_i \otimes \mathbf{e}_l, & j = k, \end{cases} \tag{2.34}$$

$$\sum_{i=1}^{n} \mathbf{e}_i \otimes \mathbf{e}_i = \mathbf{I}. \tag{2.35}$$

(ii) *The dyad* $\mathbf{a} \otimes \mathbf{b}$ *is symmetric if and only if* $\mathbf{b} = \alpha \mathbf{a}$, $\alpha \in \mathbb{R}$ *and id skew-symmetric if and only if* $\mathbf{a} = \mathbf{b} = \mathbf{0}$.

Proof. The proof of (2.32) follows from the equalities

$$\mathbf{u} \cdot (\mathbf{a} \otimes \mathbf{b})^T \mathbf{v} = (\mathbf{a} \otimes \mathbf{b})\mathbf{u} \cdot \mathbf{v} = (\mathbf{u} \cdot \mathbf{b})(\mathbf{a} \cdot \mathbf{v}) =$$

$$\mathbf{u} \cdot (\mathbf{a} \cdot \mathbf{v})\mathbf{b} = \mathbf{u} \cdot (\mathbf{b} \otimes \mathbf{a})\mathbf{v}, \quad \text{for each } \mathbf{u}, \mathbf{v} \in \mathcal{V}. \tag{2.36}$$

From the relations

$$(\mathbf{a} \otimes \mathbf{b})(\mathbf{c} \otimes \mathbf{d})\mathbf{u} = (\mathbf{a} \otimes \mathbf{b})(\mathbf{d} \cdot \mathbf{u})\mathbf{c} = (\mathbf{b} \cdot \mathbf{c})(\mathbf{d} \cdot \mathbf{u})\mathbf{a} =$$

$$(\mathbf{b} \cdot \mathbf{c})(\mathbf{a} \otimes \mathbf{d})\mathbf{u}, \quad \text{for all } \mathbf{u} \in \mathcal{V}, \tag{2.37}$$

condition (2.33) follows, moreover (2.34) follows directly from (2.33). As far as (2.35) is concerned, we have

$$(\mathbf{e}_1 \otimes \mathbf{e}_1 + ... + \mathbf{e}_n \otimes \mathbf{e}_n)\mathbf{u} =$$

$$u_1 \mathbf{e}_1 + ... + u_n \mathbf{e}_n = \mathbf{I}\mathbf{u}, \quad \text{for each} \mathbf{u} \in \mathcal{V}. \tag{2.38}$$

□

Next exercise summarizes some properties of the tensor product.

Exercise 1. *Prove that given* $\mathbf{A} \in$ *Lin and* $\mathbf{a}, \mathbf{b} \in \mathcal{V}$*, we have*

$$\mathbf{A}(\mathbf{a} \otimes \mathbf{b}) = \mathbf{A}\mathbf{a} \otimes \mathbf{b}, \tag{2.39}$$

$$(\mathbf{a} \otimes \mathbf{b})\mathbf{A} = \mathbf{a} \otimes \mathbf{A}^T\mathbf{b}. \tag{2.40}$$

Solution. For each $\mathbf{u} \in \mathcal{V}$, we have

$$\mathbf{A}(\mathbf{a} \otimes \mathbf{b})\mathbf{u} = \mathbf{A}(\mathbf{b} \cdot \mathbf{u})\mathbf{a} = (\mathbf{b} \cdot \mathbf{u})\mathbf{A}\mathbf{a} = (\mathbf{A}\mathbf{a} \otimes \mathbf{b})\mathbf{u}, \tag{2.41}$$

$$(\mathbf{a} \otimes \mathbf{b})\mathbf{A}\mathbf{u} = (\mathbf{b} \cdot \mathbf{A}\mathbf{u})\mathbf{a} = (\mathbf{A}^T\mathbf{b} \cdot \mathbf{u})\mathbf{a} = (\mathbf{a} \otimes \mathbf{A}^T\mathbf{b})\mathbf{u}. \tag{2.42}$$

Consider $\mathbf{e} \in \mathcal{V}$ with $||\mathbf{e}|| = 1$, for each $\mathbf{v} \in \mathcal{V}$ the vector $(\mathbf{e} \otimes \mathbf{e})\mathbf{v} = (\mathbf{v} \cdot \mathbf{e})\mathbf{e}$ is the projection of $\mathbf{v}$ onto Span($\mathbf{e}$); the vector $(\mathbf{I} - \mathbf{e} \otimes \mathbf{e})\mathbf{v}$ is the projection of $\mathbf{v}$ onto the subspace orthogonal to $\mathbf{e}$,

$$P_{\text{Span}(\mathbf{e})} = \mathbf{e} \otimes \mathbf{e}, \quad P_{\text{Span}(\mathbf{e})^\perp} = \mathbf{I} - \mathbf{e} \otimes \mathbf{e}.$$

2.4 Components of a tensor

Given an orthonormal basis $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ of $\mathcal{V}$, the *Cartesian components* of a tensor $\mathbf{A} \in$ Lin are

$$A_{ij} = \mathbf{e}_i \cdot \mathbf{A}\mathbf{e}_j, \quad i, j = 1, ..., n. \tag{2.43}$$

For $\mathbf{u} \in \mathcal{V}$, we have

$$\mathbf{u} = \sum_{j=1}^{n} u_j \mathbf{e}_j, \tag{2.44}$$

putting $\mathbf{v} = \mathbf{A}\mathbf{u}$, for each $i = 1, ..., n$ we have

$$v_i = \mathbf{v} \cdot \mathbf{e}_i = \mathbf{A}\mathbf{u} \cdot \mathbf{e}_i = \mathbf{e}_i \cdot \sum_{j=1}^{n} \mathbf{A}(u_j \mathbf{e}_j) = \sum_{j=1}^{n} \mathbf{e}_i \cdot \mathbf{A}\mathbf{e}_j u_j = \sum_{j=1}^{n} A_{ij} u_j. \tag{2.45}$$

Proposition 22. *Let* $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ *be an orthonormal basis of* $\mathcal{V}$*, the dyads* $\{\mathbf{e}_i \otimes \mathbf{e}_j\}_{i,j=1,...,n}$ *are a basis of Lin. In particular, for each* $\mathbf{A} \in$ *Lin, we have*

$$\mathbf{A} = \sum_{i,j=1}^{n} A_{ij} (\mathbf{e}_i \otimes \mathbf{e}_j). \tag{2.46}$$

Proof. Let us start by proving that the dyads $\{\mathbf{e}_i \otimes \mathbf{e}_j\}_{i,j=1,\dots,n}$ are linearly independent tensors. Let us consider the linear combination of the dyads $\{\mathbf{e}_i \otimes \mathbf{e}_j\}_{i,j=1,\dots,n}$ with coefficients α_{ij}, we have

$$\sum_{i,j=1}^{n} \alpha_{ij} \mathbf{e}_i \otimes \mathbf{e}_j = \mathbf{0} \tag{2.47}$$

if and only if

$$\sum_{i,j=1}^{n} \alpha_{ij} (\mathbf{e}_i \otimes \mathbf{e}_j)\mathbf{u} = \sum_{i,j=1}^{n} \alpha_{ij} u_j \mathbf{e}_i = \mathbf{0}, \quad \text{for each } \mathbf{u} \in \mathcal{V}. \tag{2.48}$$

Putting

$$\beta_i = \sum_{j=1}^{n} \alpha_{ij} u_j, \quad i = 1, ..., n \tag{2.49}$$

from (2.48) we get $\beta_1 = ... = \beta_n = 0$ and then

$$\left(\sum_{j=1}^{n} \alpha_{ij} \mathbf{e}_j \right) \cdot \mathbf{u} = 0, \quad \text{for each } \mathbf{u} \in \mathcal{V},\ i = 1, ..., n. \tag{2.50}$$

The relations in (2.50) are equivalent to

$$\sum_{j=1}^{n} \alpha_{ij} \mathbf{e}_j = \mathbf{0},\ i = 1, ..., n, \tag{2.51}$$

that, in their turn, taking into account the linear independence of vectors $\mathbf{e}_1, ..., \mathbf{e}_n$, imply the equalities $\alpha_{ij} = 0$, $i, j = 1, ..., n$.

For each $\mathbf{u} \in \mathcal{V}$ we have

$$\mathbf{A}\mathbf{u} = \sum_{i=1}^{n} (\mathbf{A}\mathbf{u})_i \mathbf{e}_i = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} u_j \mathbf{e}_i = \sum_{i,j=1}^{n} A_{ij} (\mathbf{u} \cdot \mathbf{e}_j) \mathbf{e}_i = \sum_{i,j=1}^{n} A_{ij} (\mathbf{e}_i \otimes \mathbf{e}_j)\mathbf{u},$$

which proves (2.46) and allows to conclude that $\{\mathbf{e}_i \otimes \mathbf{e}_j\}_{i,j=1,\dots,n}$ is a basis of the vector space Lin that, therefore, has dimension n^2. □

Proposition 23. *Given* $\mathbf{A} \in$ *Lin, we have*

$$\mathbf{A} = \sum_{j=1}^{n} (\mathbf{A}\mathbf{e}_j \otimes \mathbf{e}_j). \tag{2.52}$$

Proof. To prove (2.52) we use the representations (2.46) and (2.43) along with the relation (2.33), from which we get

$$\mathbf{A} = \sum_{i,j=1}^{n} A_{ij} (\mathbf{e}_i \otimes \mathbf{e}_j) = \sum_{i,j=1}^{n} (\mathbf{e}_i \cdot \mathbf{A}\mathbf{e}_j)(\mathbf{e}_i \otimes \mathbf{e}_j) = \sum_{i,j=1}^{n} (\mathbf{e}_i \otimes \mathbf{e}_i)(\mathbf{A}\mathbf{e}_j \otimes \mathbf{e}_j) =$$

$$\left(\sum_{i=1}^{n} \mathbf{e}_i \otimes \mathbf{e}_i\right)\left(\sum_{j=1}^{n} \mathbf{A}\mathbf{e}_j \otimes \mathbf{e}_j\right) = \sum_{j=1}^{n} \mathbf{A}\mathbf{e}_j \otimes \mathbf{e}_j,$$

where the latest equality follows from (2.35). □

For $\mathbf{a}, \mathbf{b} \in \mathcal{V}$, we have

$$(\mathbf{a} \otimes \mathbf{b})_{ij} = a_i b_j, \quad i, j = 1, ..., n, \tag{2.53}$$

in fact,

$$(\mathbf{a} \otimes \mathbf{b})_{ij} = \mathbf{e}_i \cdot (\mathbf{a} \otimes \mathbf{b})\mathbf{e}_j = (\mathbf{a} \cdot \mathbf{e}_i)(\mathbf{b} \cdot \mathbf{e}_j) = a_i b_j.$$

Moreover the components of the identity tensor $\mathbf{I}$ are

$$I_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases} \tag{2.54}$$

and, for $\mathbf{S}$ a symmetric tensor and $\mathbf{W}$ a skew-symmetric tensor, we have

$$S_{ij} = \mathbf{e}_i \cdot \mathbf{S}\mathbf{e}_j = \mathbf{S}\mathbf{e}_i \cdot \mathbf{e}_j = S_{ji}, \quad i, j = 1, ..., n, \tag{2.55}$$

$$W_{ij} = \mathbf{e}_i \cdot \mathbf{W}\mathbf{e}_j = -\mathbf{W}\mathbf{e}_i \cdot \mathbf{e}_j = -W_{ji}, \quad i, j = 1, ..., n, \tag{2.56}$$

Given a tensor $\mathbf{A} \in$ Lin, the matrix

$$[\mathbf{A}] = \begin{bmatrix} A_{11} & A_{12} & . & A_{1n} \\ A_{21} & A_{22} & . & A_{2n} \\ . & . & . & . \\ A_{n1} & A_{n2} & & A_{nn} \end{bmatrix} \tag{2.57}$$

is the *matrix of the components* of $\mathbf{A}$ with respect to $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$.

Given the tensors $\mathbf{A}, \mathbf{B} \in$Lin, we have

$$[\mathbf{A}^T] = [\mathbf{A}]^T, \tag{2.58}$$

$$[\mathbf{A}\mathbf{B}] = [\mathbf{A}][\mathbf{B}], \tag{2.59}$$

$$[\mathbf{I}] = \begin{bmatrix} 1 & 0 & . & 0 \\ 0 & 1 & . & 0 \\ . & . & . & . \\ 0 & 0 & . & 1 \end{bmatrix}. \tag{2.60}$$

2.5 Inner product and norm on Lin

The *trace* is the linear functional on Lin that assigns to each tensor $\mathbf{A}$ the scalar $tr\mathbf{A}$ and satisfies

$$tr(\mathbf{a}\otimes\mathbf{b}) = \mathbf{a}\cdot\mathbf{b}, \quad \text{for each } \mathbf{a},\mathbf{b}\in\mathcal{V} \tag{2.61}$$

From the relation (2.46) and the linearity of tr we have

$$tr\mathbf{A} = tr\left(\sum_{i,j=1}^{n} A_{ij}(\mathbf{e}_i\otimes\mathbf{e}_j)\right) = \sum_{i,j=1}^{n} A_{ij}tr(\mathbf{e}_i\otimes\mathbf{e}_j) =$$

$$\sum_{i,j=1}^{n} A_{ij}(\mathbf{e}_i\cdot\mathbf{e}_j) = \sum_{i=1}^{n} A_{ii}. \tag{2.62}$$

Proposition 24. *The trace has the following properties*

$$tr\mathbf{A} = tr\mathbf{A}^T, \tag{2.63}$$

$$tr(\mathbf{AB}) = tr(\mathbf{BA}), \tag{2.64}$$

for each $\mathbf{A},\mathbf{B}\in$*Lin.*

Proof. We have

$$tr\mathbf{A} = \sum_{i=1}^{n} A_{ii} = \sum_{i=1}^{n}\mathbf{e}_i\cdot\mathbf{A}\mathbf{e}_i = \sum_{i=1}^{n}\mathbf{e}_i\cdot\mathbf{A}^T\mathbf{e}_i = tr\mathbf{A}^T,$$

and (2.63) is proved. As far as (2.64) is concerned, we remark that

$$\mathbf{AB} = \left(\sum_{i,j=1}^{n} A_{ij}(\mathbf{e}_i\otimes\mathbf{e}_j)\right)\left(\sum_{l,m=1}^{n} B_{lm}(\mathbf{e}_l\otimes\mathbf{e}_m)\right) =$$

$$\sum_{i,j,l,m=1}^{n} A_{ij}B_{lm}(\mathbf{e}_i\otimes\mathbf{e}_j)(\mathbf{e}_l\otimes\mathbf{e}_m) = \sum_{i,j,m=1}^{n} A_{ij}B_{jm}(\mathbf{e}_i\otimes\mathbf{e}_m), \tag{2.65}$$

$$\mathbf{BA} = \left(\sum_{l,m=1}^{n} B_{lm}(\mathbf{e}_l\otimes\mathbf{e}_m)\right)\left(\sum_{i,j=1}^{n} A_{ij}(\mathbf{e}_i\otimes\mathbf{e}_j)\right) =$$

$$\sum_{i,j,l,m=1}^{n} A_{ij}B_{lm}(\mathbf{e}_l\otimes\mathbf{e}_m)(\mathbf{e}_i\otimes\mathbf{e}_j) = \sum_{i,j,l=1}^{n} A_{ij}B_{li}(\mathbf{e}_l\otimes\mathbf{e}_j). \tag{2.66}$$

From (2.65) we get

$$tr(\mathbf{AB}) = \sum_{i,j,m=1}^{n} A_{ij}B_{jm}(\mathbf{e}_i\cdot\mathbf{e}_m) = \sum_{i,j=1}^{n} A_{ij}B_{ji},$$

and from (2.66) we have

$$tr(\mathbf{BA}) = \sum_{i,j,l=1}^{n} A_{ij}B_{li}(\mathbf{e}_l \cdot \mathbf{e}_j) = \sum_{i,j=1}^{n} A_{ij}B_{ji},$$

and the thesis follows. □

In particular, from the preceding proof, it follows that the components of the product **AB** are given by

$$(AB)_{ml} = \sum_{j=1}^{n} A_{mj}B_{jl}, \quad m,l = 1,...,n. \tag{2.67}$$

The vector space Lin can be equipped with the inner product

$$\mathbf{A} \cdot \mathbf{B} = tr(\mathbf{A}^T\mathbf{B}), \quad \mathbf{A}, \mathbf{B} \in \text{Lin}. \tag{2.68}$$

Let us verify that (2.68) is scalar product. The symmetry is satisfied, in fact

$$\mathbf{A} \cdot \mathbf{B} = tr(\mathbf{A}^T\mathbf{B}) = tr(\mathbf{B}^T\mathbf{A}) = \mathbf{B} \cdot \mathbf{A},$$

as for the bilinearity, given $\mathbf{A}, \mathbf{B}, \mathbf{C} \in$Lin we have

$$\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = tr(\mathbf{A}^T(\mathbf{B} + \mathbf{C})) = tr(\mathbf{A}^T\mathbf{B}) + tr(\mathbf{A}^T\mathbf{C}) =$$

$$\mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C},$$

moreover, for each $\alpha \in \mathbb{R}$,

$$\mathbf{A} \cdot (\alpha\mathbf{B}) = tr(\alpha\mathbf{A}^T\mathbf{B}) = \alpha tr(\mathbf{A}^T\mathbf{B}) = \alpha\mathbf{A} \cdot \mathbf{B}.$$

Finally, as for the positivity, we have

$$\mathbf{A} \cdot \mathbf{A} = tr(\mathbf{A}^T\mathbf{A}) = tr\left[\left(\sum_{i,j=1}^{n} A_{ij}(\mathbf{e}_j \otimes \mathbf{e}_i)\right) \sum_{l,m=1}^{n} A_{lm}(\mathbf{e}_l \otimes \mathbf{e}_m)\right] =$$

$$tr\left(\sum_{i,j,l,m=1}^{n} A_{ij}A_{lm}(\mathbf{e}_j \otimes \mathbf{e}_i)(\mathbf{e}_l \otimes \mathbf{e}_m)\right) = tr\left(\sum_{i,j,m=1}^{n} A_{ij}A_{im}(\mathbf{e}_j \otimes \mathbf{e}_m)\right) =$$

$$\sum_{i,j=1}^{n} A_{ij}^2 \geq 0,$$

moreover, $\mathbf{A} \cdot \mathbf{A} = 0$ if and only if $A_{ij} = 0$ for $i,j = 1,...,n$.

The inner product of **A** and **B** in terms of components is given by

$$\mathbf{A} \cdot \mathbf{B} = \sum_{i,j=1}^{n} A_{ij}B_{ij}, \tag{2.69}$$

in fact,

$$\mathbf{A}\cdot\mathbf{B} = tr(\mathbf{A}^T\mathbf{B}) = tr(\sum_{i,j=1}^{n} A_{ij}(\mathbf{e}_j\otimes\mathbf{e}_i)\sum_{k,l=1}^{n} B_{kl}(\mathbf{e}_k\otimes\mathbf{e}_l))$$

$$= tr(\sum_{i,j,l=1}^{n} A_{ij}B_{il}(\mathbf{e}_j\otimes\mathbf{e}_l)) = \sum_{i,j=1}^{n} A_{ij}B_{ij}. \quad (2.70)$$

In the vector space Lin the norm induced by the inner product (2.68) is

$$||\mathbf{A}|| = \sqrt{\mathbf{A}\cdot\mathbf{A}} = \sqrt{tr(\mathbf{A}^T\mathbf{A})}, \quad \mathbf{A}\in \text{Lin}. \quad (2.71)$$

In particular, we have

$$||\mathbf{A}|| = ||\mathbf{A}^T||, \quad (2.72)$$

in fact,

$$||\mathbf{A}||^2 = \mathbf{A}\cdot\mathbf{A} = tr(\mathbf{A}^T\mathbf{A}) = tr(\mathbf{A}\mathbf{A}^T)$$
$$= tr((\mathbf{A}^T)^T\mathbf{A}^T) = \mathbf{A}^T\cdot\mathbf{A}^T = ||\mathbf{A}^T||^2. \quad (2.73)$$

Exercise 2. *The norm (2.71) is submultiplicative,*

$$||\mathbf{A}\mathbf{B}|| \leq ||\mathbf{A}||\ ||\mathbf{B}|| \qquad \textit{for each } \mathbf{A},\mathbf{B}\in \textit{Lin}. \quad (2.74)$$

Proposition 25. *For $\mathbf{A},\mathbf{B},\mathbf{C}\in$Lin, $\mathbf{u},\mathbf{v},\mathbf{a},\mathbf{b}\in\mathcal{V}$, the following relations hold*

$$\mathbf{I}\cdot\mathbf{A} = tr\mathbf{A}, \quad (2.75)$$

$$\mathbf{C}\cdot(\mathbf{A}\mathbf{B}) = (\mathbf{A}^T\mathbf{C})\cdot\mathbf{B} = (\mathbf{C}\mathbf{B}^T)\cdot\mathbf{A}, \quad (2.76)$$

$$\mathbf{u}\cdot\mathbf{A}\mathbf{v} = \mathbf{A}\cdot(\mathbf{u}\otimes\mathbf{v}), \quad (2.77)$$

$$(\mathbf{a}\otimes\mathbf{b})\cdot(\mathbf{u}\otimes\mathbf{v}) = (\mathbf{a}\cdot\mathbf{u})(\mathbf{b}\cdot\mathbf{v}), \quad (2.78)$$

$$||\mathbf{u}\otimes\mathbf{u}|| = ||\mathbf{u}||^2. \quad (2.79)$$

Proof. (2.75) is trivial; to prove (2.76) we remark that

$$\mathbf{C}\cdot(\mathbf{A}\mathbf{B}) = tr(\mathbf{C}^T\mathbf{A}\mathbf{B}) = tr((\mathbf{A}^T\mathbf{C})^T\mathbf{B}) = (\mathbf{A}^T\mathbf{C})\cdot\mathbf{B} =$$

$$tr(\mathbf{B}\mathbf{C}^T\mathbf{A}) = tr((\mathbf{C}\mathbf{B}^T)^T\mathbf{A}) = (\mathbf{C}\mathbf{B}^T)\cdot\mathbf{A}.$$

Moreover,

$$\mathbf{u}\cdot\mathbf{A}\mathbf{v} = \sum_{i=1}^{n} u_i\sum_{j=1}^{n} A_{ij}v_j = \sum_{i,j=1}^{n} A_{ij}u_iv_j = \sum_{i,j=1}^{n} A_{ij}(\mathbf{u}\otimes\mathbf{v})_{ij} = \mathbf{A}\cdot(\mathbf{u}\otimes\mathbf{v}),$$

then (2.77) is proved. Finally,

$$(\mathbf{a}\otimes\mathbf{b})\cdot(\mathbf{u}\otimes\mathbf{v}) = \sum_{i,j=1}^{n} a_ib_ju_iv_j = \left(\sum_{i=1}^{n} a_iu_i\right)\left(\sum_{i=1}^{n} b_jv_j\right) = (\mathbf{a}\cdot\mathbf{u})(\mathbf{b}\cdot\mathbf{v}),$$

and (2.78) is proved as well. Finally,

$$||\mathbf{u}\otimes\mathbf{u}||^2 = tr((\mathbf{u}\otimes\mathbf{u})(\mathbf{u}\otimes\mathbf{u})) = (\mathbf{u}\cdot\mathbf{u})^2 = ||\mathbf{u}||^4.$$

□

From (2.78) and (2.46) it follows that $\{\mathbf{e}_i\otimes\mathbf{e}_j\}_{i,j=1,\dots,n}$ is an orthonormal basis of Lin,

$$(\mathbf{e}_i\otimes\mathbf{e}_j)\cdot(\mathbf{e}_k\otimes\mathbf{e}_l) = (\mathbf{e}_i\cdot\mathbf{e}_k)(\mathbf{e}_j\cdot\mathbf{e}_l) = \begin{cases} 1 & i=k,\ j=l, \\ 0 & \text{otherwise.}\end{cases}$$

Given $\mathbf{A}\in$Lin, for each $\mathbf{u}\in\mathcal{V}$, we have

$$||\mathbf{A}\mathbf{u}||^2 = \mathbf{A}\mathbf{u}\cdot\mathbf{A}\mathbf{u} = \mathbf{u}\cdot\mathbf{A}^T\mathbf{A}\mathbf{u} = \mathbf{A}^T\mathbf{A}\cdot(\mathbf{u}\otimes\mathbf{u}) \le ||\mathbf{A}^T\mathbf{A}||\ ||\mathbf{u}\otimes\mathbf{u}||$$

$$\le ||\mathbf{A}^T||\ ||\mathbf{A}||\ ||\mathbf{u}\otimes\mathbf{u}|| \le ||\mathbf{A}||^2\ ||\mathbf{u}||^2,$$

then,

$$||\mathbf{A}\mathbf{u}|| \le ||\mathbf{A}||\ ||\mathbf{u}||, \quad \text{for each } \mathbf{u}\in\mathcal{V}. \tag{2.80}$$

In particular, in agreement with the fact that $\mathbf{A}$ is linear, $\mathbf{A}$ is bounded (Proposition 11)

Proposition 26. *Given $\mathbf{A},\mathbf{B}\in Lin$, the following properties hold,*

(1) If $\mathbf{A}$ is symmetric, we have

$$\mathbf{A}\cdot\mathbf{C} = \mathbf{A}\cdot\mathbf{C}^T = \mathbf{A}\cdot\frac{1}{2}(\mathbf{C}+\mathbf{C}^T), \quad \textit{for each } \mathbf{C}\in Lin. \tag{2.81}$$

(2) If $\mathbf{B}$ is skew-symmetric, we have

$$\mathbf{B}\cdot\mathbf{C} = -\mathbf{B}\cdot\mathbf{C}^T = \mathbf{B}\cdot\frac{1}{2}(\mathbf{C}-\mathbf{C}^T), \quad \textit{for each } \mathbf{C}\in Lin. \tag{2.82}$$

(3) If $\mathbf{A}$ is symmetric and $\mathbf{B}$ is skew-symmetric, we have $\mathbf{A}\cdot\mathbf{B}=0$.

(4) If $\mathbf{A}\cdot\mathbf{C}=0$ for every symmetric tensor $\mathbf{C}$, then $\mathbf{A}$ is skew-symmetric.

(5) If $\mathbf{A}\cdot\mathbf{C}=0$ for every skew-symmetric tensor $\mathbf{C}$, then $\mathbf{A}$ is symmetric.

Proof. (1) If $\mathbf{A}=\mathbf{A}^T$, then

$$\mathbf{A}\cdot\mathbf{C} = tr(\mathbf{A}^T\mathbf{C}) = tr(\mathbf{A}\mathbf{C}) = \mathbf{A}\cdot\mathbf{C}^T,$$

moreover,

$$\mathbf{A}\cdot\mathbf{C} = \frac{1}{2}(\mathbf{A}\cdot\mathbf{C}+\mathbf{A}\cdot\mathbf{C}) = \frac{1}{2}(\mathbf{A}\cdot\mathbf{C}+\mathbf{A}\cdot\mathbf{C}^T) =$$

$$\mathbf{A} \cdot \frac{1}{2}(\mathbf{C} + \mathbf{C}^T).$$

(2) On the contrary, if $\mathbf{A} = -\mathbf{A}^T$, we have

$$\mathbf{A} \cdot \mathbf{C} = tr(\mathbf{A}^T\mathbf{C}) = -tr(\mathbf{A}\mathbf{C}) = -\mathbf{A} \cdot \mathbf{C}^T,$$

and

$$\mathbf{A} \cdot \mathbf{C} = \frac{1}{2}(\mathbf{A} \cdot \mathbf{C} + \mathbf{A} \cdot \mathbf{C}) = \frac{1}{2}(\mathbf{A} \cdot \mathbf{C} - \mathbf{A} \cdot \mathbf{C}^T) =$$

$$\mathbf{A} \cdot \frac{1}{2}(\mathbf{C} - \mathbf{C}^T),$$

(3) If $\mathbf{A}$ is symmetric and $\mathbf{B}$ is skew-symmetric, then

$$\mathbf{A} \cdot \mathbf{B} = tr(\mathbf{A}\mathbf{B}) = \mathbf{B}^T \cdot \mathbf{A} = -\mathbf{B} \cdot \mathbf{A},$$

then $\mathbf{A} \cdot \mathbf{B} = 0$.

(4) Let us assume that $\mathbf{A}$ is not skew-symmetric, then $\mathbf{A} = \mathbf{S} + \mathbf{W}$, with $\mathbf{S} \in$Sym, $\mathbf{W} \in$Skw; in view of (3) we have

$$0 = \mathbf{A} \cdot \mathbf{C} = \mathbf{S} \cdot \mathbf{C} + \mathbf{W} \cdot \mathbf{C} = \mathbf{S} \cdot \mathbf{C}, \quad \text{for each } \mathbf{C} \in \text{Sym};$$

in particular, choosing $\mathbf{C} = \mathbf{S}$, we have $\mathbf{S} \cdot \mathbf{S} = 0$ and then $\mathbf{S} = \mathbf{0}$.

(5) The proof is analogous to that of point (4). □

We know that Lin=Sym⊕Skw, from the previous proposition it follows that Skw is the orthogonal complement of Sym and Sym is the orthogonal complement of Skw,

$$\text{Sym}^{\perp} = \text{Skw}, \quad \text{Skw}^{\perp} = \text{Sym} \tag{2.83}$$

and that $P_{\text{Sym}}(\mathbf{A}) = \frac{\mathbf{A}+\mathbf{A}^T}{2}$ e $P_{\text{Skw}}(\mathbf{A}) = \frac{\mathbf{A}-\mathbf{A}^T}{2}$ are the orthogonal projections of $\mathbf{A}$ onto the subspaces Sym and Skw, respectively.

2.6 Invertible tensors

A tensor $\mathbf{A}$ is called *invertible* if it is injective

(i) if $\mathbf{u}_1 \neq \mathbf{u}_2$ then $\mathbf{A}\mathbf{u}_1 \neq \mathbf{A}\mathbf{u}_2$,

and surjective,

(ii) for each $\mathbf{v} \in \mathcal{V}$ there exists (at least) $\mathbf{u} \in \mathcal{V}$ such that $\mathbf{A}\mathbf{u} = \mathbf{v}$.

If $\mathbf{A}$ is invertible, the tensor $\mathbf{A}^{-1}$, called *inverse* of $\mathbf{A}$, is defined as follows. Given $\mathbf{v}_0 \in \mathcal{V}$ there is a unique (in view of (i)) $\mathbf{u}_0 \in \mathcal{V}$ such that $\mathbf{A}\mathbf{u}_0 = \mathbf{v}_0$, then, we can define $\mathbf{A}^{-1}\mathbf{v}_0 = \mathbf{u}_0$. $\mathbf{A}^{-1}$ is linear, in fact, given $\mathbf{v}_1, \mathbf{v}_2$ with

$\mathbf{Au}_1 = \mathbf{v}_1, \mathbf{Au}_2 = \mathbf{v}_2$, from the linearity of $\mathbf{A}$ it follows that $\mathbf{A}(\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2) = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2$, then

$$\mathbf{A}^{-1}(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2) = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 =$$

$$\alpha_1\mathbf{A}^{-1}\mathbf{v}_1 + \alpha_2\mathbf{A}^{-1}\mathbf{v}_2. \tag{2.84}$$

From the definition it follows that if $\mathbf{A}$ is invertible then

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}. \tag{2.85}$$

These relations characterize $\mathbf{A}^{-1}$, in fact, the following theorem holds.

Theorem 7. *Let* $\mathbf{A}$ *be a tensor. If there exist two tensors* $\mathbf{B}, \mathbf{C} \in Lin$ *such that*

$$\mathbf{AB} = \mathbf{CA} = \mathbf{I}, \tag{2.86}$$

then $\mathbf{A}$ *is invertible and* $\mathbf{B} = \mathbf{C} = \mathbf{A}^{-1}$.

Proof. If $\mathbf{Au}_1 = \mathbf{Au}_2$ then $\mathbf{CAu}_1 = \mathbf{CAu}_2$ and $\mathbf{u}_1 = \mathbf{u}_2$, then $\mathbf{A}$ has the property (i). For each $\mathbf{v} \in \mathcal{V}$ put $\mathbf{u} = \mathbf{Bv}$, then $\mathbf{Au} = \mathbf{ABv} = \mathbf{v}$ and $\mathbf{A}$ satisfies (ii). From $\mathbf{AB} = \mathbf{I}$, multiplying (left) by $\mathbf{A}^{-1}$ we get $\mathbf{B} = \mathbf{A}^{-1}$, and from $\mathbf{CA} = \mathbf{I}$, multiplying (right) by $\mathbf{A}^{-1}$ we have $\mathbf{C} = \mathbf{A}^{-1}$. □

Theorem 8. $\mathbf{A} \in Lin$ *is injective if and only is surjective.*

Proof. Let us assume that $\mathbf{A}$ is injective, that is $\mathbf{Au} = \mathbf{0}$ implies $\mathbf{u} = \mathbf{0}$. Let $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ be a basis of $\mathcal{V}$, then $\{\mathbf{Ae}_1, ..., \mathbf{Ae}_n\}$ is also a basis of $\mathcal{V}$. In fact, from

$$\mathbf{0} = \sum_{i=1}^{n} \alpha_i \mathbf{Ae}_i = \mathbf{A}\left(\sum_{i=1}^{n} \alpha_i \mathbf{e}_i\right) \tag{2.87}$$

we get $\sum_{i=1}^{n} \alpha_i \mathbf{e}_i = \mathbf{0}$ and then $\alpha_1 = ... = \alpha_n = 0$. Therefore each $\mathbf{v} \in \mathcal{V}$ can be written as $\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{Ae}_i = \mathbf{A}\left(\sum_{i=1}^{n} \alpha_i \mathbf{e}_i\right) = \mathbf{Au}$ and $\mathbf{A}$ is surjective.

Now let us assume that $\mathbf{A}$ is surjective, that is each $\mathbf{v} \in \mathcal{V}$ can be written as $\mathbf{v} = \mathbf{Au}$. If $\{\mathbf{f}_1, ..., \mathbf{f}_n\}$ is a basis of $\mathcal{V}$, let $\mathbf{e}_1, ..., \mathbf{e}_n \in \mathcal{V}$ be vectors such that $\mathbf{f}_i = \mathbf{Ae}_i$, $i = 1, ..., n$, then $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ is a basis of $\mathcal{V}$, in fact $\sum_{i=1}^{n} \alpha_i \mathbf{e}_i = \mathbf{0}$ implies $\mathbf{A}\left(\sum_{i=1}^{n} \alpha_i \mathbf{e}_i\right) = \sum_{i=1}^{n} \alpha_i \mathbf{f}_i = \mathbf{0}$ and then $\alpha_1 = ... = \alpha_n = 0$. Thus, we have proved that if $\mathbf{Au} = \mathbf{0}$ then $\mathbf{u} = \mathbf{0}$, and then $\mathbf{A}$ is injective. □

In a infinite dimensional vector space theorem 8 does not hold. Let $\mathcal{P}[0,1]$ be the vector space of polynomials with real coefficients, $p(x) = a_0 + a_1 x + ... + a_k x^k$, with k integer and x belonging to the interval $[0,1]$. Consider the linear function $T : \mathcal{P}[0,1] \to \mathcal{P}[0,1]$ defined by $T(p(x)) = xp(x)$. T is injective, but not surjective, in fact given the polynomial $p(x) = a_0$, there is no $q \in \mathcal{P}[0,1]$ such that $T(q) = p$. Moreover the linear function $D : \mathcal{P}[0,1] \to \mathcal{P}[0,1]$ defined by $T(p(x)) = p'(x)$, is surjective, but not injective, in fact for $p(x) = a_0$ and $q(x) = b_0$, with $a_0 \neq b_0$ we have $D(p) = D(q)$.

Theorem 9. *(1) Let* $\mathbf{A}, \mathbf{B} \in Lin$ *be invertible tensors, then* $\mathbf{AB}$ *is invertible and*

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}. \tag{2.88}$$

(2) Let $\mathbf{A} \in Lin$ *be invertible and* $\alpha \neq 0$, *then* $\alpha\mathbf{A}$ *is invertible and*

$$(\alpha\mathbf{A})^{-1} = \frac{1}{\alpha}\mathbf{A}^{-1}.$$

(3) Let $\mathbf{A} \in Lin$ *be invertible, then* $\mathbf{A}^{-1}$ *is invertible and*

$$(\mathbf{A}^{-1})^{-1} = \mathbf{A}. \tag{2.89}$$

(4) Let $\mathbf{A} \in Lin$ *be invertible, then* $\mathbf{A}^T$ *is invertible and*

$$(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T \tag{2.90}$$

(5) Let $\mathbf{A} \in Lin$ *be invertible, then* $\mathbf{A}^k$ *is invertible for each* $k \in \mathbb{N}$ *and*

$$(\mathbf{A}^k)^{-1} = (\mathbf{A}^{-1})^k. \tag{2.91}$$

Let us consider the functional det :Lin$\to \mathbb{R}$ that assigns to each tensor $\mathbf{A}$ the determinant of the matrix $[\mathbf{A}]$ of the Cartesian components of $\mathbf{A}$ with respect to the orthonormal basis $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ of $\mathcal{V}$

$$\det \mathbf{A} = \det[\mathbf{A}]. \tag{2.92}$$

$\det \mathbf{A}$ is called *determinant* of the tensor $\mathbf{A}$. In the following, we will prove that the definition does not depend on the choice of the basis $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ di $\mathcal{V}$.

The following properties of the determinant of a tensor are a direct consequence of the analogous properties of the determinant of a matrix. For each $\mathbf{A}, \mathbf{B} \in$ Lin we have

$$\det(\mathbf{AB}) = \det \mathbf{A} \det \mathbf{B}, \tag{2.93}$$

$$\det(\mathbf{A}^T) = \det \mathbf{A}, \tag{2.94}$$

$$\det(\alpha\mathbf{A}) = \alpha^n \det \mathbf{A}, \quad \alpha \in \mathbb{R}, \tag{2.95}$$

$$\det(\mathbf{I}) = 1. \tag{2.96}$$

The following proposition holds.

Proposition 27. *A tensor* $\mathbf{A}$ *is invertible if and only if* $\det \mathbf{A} \neq 0$, *in this case*

$$\det(\mathbf{A}^{-1}) = (\det \mathbf{A})^{-1}. \tag{2.97}$$

Proof. If $\mathbf{A}$ is invertible, from (2.85), taking (2.93) into account, it follows that

$$1 = \det \mathbf{A} \det(\mathbf{A}^{-1}),$$

from which $\det \mathbf{A} \neq 0$ and (2.97) follow. Vice versa, Let us assume that $\det \mathbf{A}$ is different from 0, Then $\mathbf{A}$ is injective. In fact, the relation $\mathbf{A}\mathbf{u} = \mathbf{0}$ is equivalent to the linear system $\sum_{j=1}^{n} A_{ij}u_j = 0, i = 1, ..., n$ whose unique solution is $u_1 = ... = u_n = 0$. Then, in virtue of the theorem 8 we conclude that $\mathbf{A}$ is invertible. □

Example 9. *Given* $\mathbf{e} \in \mathcal{V}$ *with* $||\mathbf{e}|| = 1$*, the tensor* $\mathbf{e} \otimes \mathbf{e}$ *which assigns to each* $\mathbf{v} \in \mathcal{V}$ *the vector* $(\mathbf{e} \cdot \mathbf{v})\mathbf{e}$ *is not invertible since it maps the subspace orthogonal to* $\mathbf{e}$ *in the vector* $\mathbf{0}$*.*

2.7 Orthogonal tensors

A tensor $\mathbf{Q}$ is *orthogonal* if it preserves the inner product $\cdot$ on $\mathcal{V}$,

$$\mathbf{Q}\mathbf{u} \cdot \mathbf{Q}\mathbf{v} = \mathbf{u} \cdot \mathbf{v}, \quad \text{for each } \mathbf{u}, \mathbf{v} \in \mathcal{V}. \tag{2.98}$$

In particular, an orthogonal tensor is invertible, in fact, from(2.98) for $\mathbf{v} = \mathbf{u}$ we get

$$||\mathbf{Q}\mathbf{u}|| = ||\mathbf{u}||, \tag{2.99}$$

thus, if $\mathbf{Q}\mathbf{u} = \mathbf{0}$, then $\mathbf{u} = \mathbf{0}$. An orthogonal tensor is an isometry (see (1.67)).

Condition (2.99) expresses the fact that $\mathbf{Q}$ preserves the norm of vectors.

Proposition 28. $\mathbf{Q} \in Lin$ *is orthogonal if and only if*

$$\mathbf{Q}\mathbf{Q}^T = \mathbf{Q}^T\mathbf{Q} = \mathbf{I}. \tag{2.100}$$

Proof. Let us assume that condition (2.100) is satisfied, then

$$\mathbf{Q}\mathbf{u} \cdot \mathbf{Q}\mathbf{v} = \mathbf{u} \cdot \mathbf{Q}^T\mathbf{Q}\mathbf{v} = \mathbf{u} \cdot \mathbf{v}, \quad \text{for each } \mathbf{u}, \mathbf{v} \in \mathcal{V},$$

and $\mathbf{Q}$ is orthogonal. Vice versa, let us assume that $\mathbf{Q}$ is orthogonal,

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{Q}\mathbf{u} \cdot \mathbf{Q}\mathbf{v} = \mathbf{u} \cdot \mathbf{Q}^T\mathbf{Q}\mathbf{v}, \quad \text{for each } \mathbf{u}, \mathbf{v} \in \mathcal{V},$$

then we have that $\mathbf{u} \cdot (\mathbf{v} - \mathbf{Q}^T\mathbf{Q}\mathbf{v}) = 0$ for each $\mathbf{u} \in \mathcal{V}$, therefore, $\mathbf{v} - \mathbf{Q}^T\mathbf{Q}\mathbf{v} = \mathbf{0}$ for each $\mathbf{v} \in \mathcal{V}$ and finally

$$\mathbf{Q}^T\mathbf{Q} = \mathbf{I}. \tag{2.101}$$

If we right multiply (2.101) by $\mathbf{Q}^T$ we get

$$\mathbf{Q}^T\mathbf{Q}\mathbf{Q}^T = \mathbf{Q}^T,$$

from which, left multiplying by $\mathbf{Q}^{-T}$, we deduce that $\mathbf{Q}\mathbf{Q}^T = \mathbf{I}$. □

From the preceding proposition, we get that $\mathbf{Q}$ is orthogonal if and only if $\mathbf{Q}^T = \mathbf{Q}^{-1}$, moreover, if $\mathbf{Q}$ is orthogonal then $\det \mathbf{Q} = \pm 1$. An orthogonal tensor $\mathbf{R}$ with $\det \mathbf{R} = 1$ is called *rotation*.

We have seen that if $\mathbf{A}$ is an invertible tensor and $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ is a basis of $\mathcal{V}$, then $\{\mathbf{A}\mathbf{e}_1, ..., \mathbf{A}\mathbf{e}_n\}$ is a basis of $\mathcal{V}$. If $\mathbf{A}$ is an orthogonal tensor the following proposition holds.

Proposition 29. *If $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ is an orthonormal basis of $\mathcal{V}$ and $\mathbf{Q}$ is an orthogonal tensor, then $\{\mathbf{Q}\mathbf{e}_1, ..., \mathbf{Q}\mathbf{e}_n\}$ is an orthonormal basis of $\mathcal{V}$. Vice versa, if $\mathbf{Q}$ is a tensor such that if $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ is an orthonormal basis then $\{\mathbf{Q}\mathbf{e}_1, ..., \mathbf{Q}\mathbf{e}_n\}$ is an orthonormal basis, $\mathbf{Q}$ is orthogonal.*

Proof. Let $\mathbf{Q}$ be an orthogonal tensor, we have

$$\mathbf{Q}\mathbf{e}_i \cdot \mathbf{Q}\mathbf{e}_j = \mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij},$$

then $\{\mathbf{Q}\mathbf{e}_1, ..., \mathbf{Q}\mathbf{e}_n\}$ is an orthonormal basis of $\mathcal{V}$.

Now let us assume that $\{\mathbf{Q}\mathbf{e}_1, ..., \mathbf{Q}\mathbf{e}_n\}$ is an orthonormal basis of $\mathcal{V}$. Since

$$\mathbf{Q}\mathbf{e}_i \cdot \mathbf{Q}\mathbf{e}_j = \delta_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j,$$

it is an easy matter to verify that

$$\mathbf{Q}\mathbf{u} \cdot \mathbf{Q}\mathbf{v} = \mathbf{u} \cdot \mathbf{v}, \quad \text{for each } \mathbf{u}, \mathbf{v} \in \mathcal{V}.$$

□

Let $E = \{\mathbf{e}_1, ..., \mathbf{e}_n\}$ and $F = \{\mathbf{f}_1, ..., \mathbf{f}_n\}$ be two orthonormal bases of $\mathcal{V}$. The tensor

$$\mathbf{Q} = \sum_{i=1}^{n} \mathbf{f}_i \otimes \mathbf{e}_i \tag{2.102}$$

is orthogonal and

$$\mathbf{f}_i = \mathbf{Q}\mathbf{e}_i, \quad i = 1, ..., n. \tag{2.103}$$

Given $\mathbf{u} \in \mathcal{V}$ we have

$$\mathbf{u} = \sum_{i=1}^{n} \xi_i \mathbf{e}_i, \qquad \mathbf{u} = \sum_{i=1}^{n} \eta_i \mathbf{f}_i,$$

with

$$\xi_i = \mathbf{u} \cdot \mathbf{e}_i = \sum_{j=1}^{n} \eta_j \mathbf{f}_j \cdot \mathbf{e}_i$$

$$= \sum_{j=1}^{n} \eta_j \mathbf{Q}\mathbf{e}_j \cdot \mathbf{e}_i = \sum_{j=1}^{n} Q_{ij} \eta_j \quad i = 1, ..., n, \tag{2.104}$$

where $Q_{ij} = \mathbf{e}_i \cdot \mathbf{Q}\mathbf{e}_j$ are the components of $\mathbf{Q}$ with respect to the basis E.

Given $(\xi_1, ..., \xi_n) \in \mathbb{R}^n$, the vectors $\mathbf{u} = \sum_{i=1}^n \xi_i \mathbf{e}_i$ and $\mathbf{v} = \sum_{i=1}^n \xi_i \mathbf{f}_i$, in view of (2.103), are linked by the following relation

$$\mathbf{v} = \mathbf{Q}\mathbf{u}. \tag{2.105}$$

Thus, tensor $\mathbf{Q}$ (or more precisely the matrix with components Q_{ij}) can be considered a coordinate transformation as in (2.104) or as a vector transformation, as in (2.105). In this case, $\mathbf{Q}$ represents a change of basis, from basis E to basis F.

Given the tensor $\mathbf{B}$, we wonder what is the relationship between the matrix of its components B_{ij} with respect to E and the matrix of its components B'_{ij} with respect to F

$$\mathbf{B} = \sum_{i,j=1}^n B_{ij} \mathbf{e}_i \otimes \mathbf{e}_j, \tag{2.106}$$

$$\mathbf{B} = \sum_{i,j=1}^n B'_{ij} \mathbf{f}_i \otimes \mathbf{f}_i. \tag{2.107}$$

We have

$$B'_{ij} = \mathbf{f}_i \cdot \mathbf{B}\mathbf{f}_j = \mathbf{Q}\mathbf{e}_i \cdot \mathbf{B}\mathbf{Q}\mathbf{e}_j = \mathbf{e}_i \cdot \mathbf{Q}^T\mathbf{B}\mathbf{Q}\mathbf{e}_j, \tag{2.108}$$

and

$$[\mathbf{B}'] = [\mathbf{Q}]^T[\mathbf{B}][\mathbf{Q}], \tag{2.109}$$

where $[\mathbf{B}]$ and $[\mathbf{Q}]$ are the matrices of the components of $\mathbf{B}$ and $\mathbf{Q}$ with respect to E.

Finally, if B_{ij} are the components of a matrix, we want to determine the relationship between the tensors $\mathbf{B}$ and $\mathbf{C}$ defined, respectively, by

$$\mathbf{B} = \sum_{i,j=1}^n B_{ij} \mathbf{e}_i \otimes \mathbf{e}_j, \tag{2.110}$$

and

$$\mathbf{C} = \sum_{i,j=1}^n B_{ij} \mathbf{f}_i \otimes \mathbf{f}_j. \tag{2.111}$$

Fom (2.110), (2.111) and (2.103) we get

$$\mathbf{C} = \mathbf{Q}\mathbf{B}\mathbf{Q}^T. \tag{2.112}$$

Relation (2.112) expresses the link that must exist between a tensor $\mathbf{B}$ and a tensor $\mathbf{C}$ such that if $\mathbf{B}\mathbf{u} = \mathbf{v}$, then $\mathbf{C}\mathbf{Q}\mathbf{u} = \mathbf{Q}\mathbf{v}$, for each $\mathbf{u} \in V$,

$$\begin{array}{ccc} \mathbf{u} & \xrightarrow{\mathbf{B}} & \mathbf{v} \\ \downarrow & & \downarrow \\ \mathbf{Q}\mathbf{u} & \xrightarrow[\mathbf{C}]{} & \mathbf{Q}\mathbf{v} \end{array} \tag{2.113}$$

Tensor $\mathbf{Q}\mathbf{B}\mathbf{Q}^T$ is called *orthogonal conjugate* of $\mathbf{B}$ with respect to $\mathbf{Q}$.

Example 10. *Let us put $n = 3$ and consider the change of basis*

$$\mathbf{f}_1 = \cos\theta\mathbf{e}_1 + \sin\theta\mathbf{e}_2, \tag{2.114}$$

$$\mathbf{f}_2 = -\sin\theta\mathbf{e}_1 + \cos\theta\mathbf{e}_2, \tag{2.115}$$

$$\mathbf{f}_3 = \mathbf{e}_3, \tag{2.116}$$

corresponding to a positive (anticlockwise) rotation of an angle θ about $\mathbf{e}_3$. The rotation

$$\mathbf{R} = \mathbf{e}_3 \otimes \mathbf{e}_3 + \sin\theta(\mathbf{e}_2 \otimes \mathbf{e}_1 - \mathbf{e}_1 \otimes \mathbf{e}_2) + \cos\theta(\mathbf{e}_1 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_2)$$

is such that $\mathbf{R}\mathbf{e}_i = \mathbf{f}_i$ and the matrix of its components $R_{ij} = \mathbf{e}_i \cdot \mathbf{R}\mathbf{e}_j$ with respect to E is

$$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{2.117}$$

The orthogonal tensor $\mathbf{Q} = -\mathbf{I}$ is called *central reflection* in the space of vectors.

For $n = 3$, the orthogonal tensor $\mathbf{Q}$ whose matrix of components is given by

$$[\mathbf{Q}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \tag{2.118}$$

is a *reflection* with respect to the subspace spanned by vectors $\mathbf{e}_1$ and $\mathbf{e}_2$.

Now, we can prove the following result.

Proposition 30. *The definition of determinant given in (2.92) does not depend on the choice of the basis of $\mathcal{V}$.*

Proof. Let us start by noting that if $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ and $\{\mathbf{f}_1, ..., \mathbf{f}_n\}$ are two orthonormal bases of $\mathcal{V}$, there is an orthogonal tensor $\mathbf{Q}$ such that

$$\mathbf{Q}\mathbf{e}_i = \mathbf{f}_i, \quad i = 1, ..., n. \tag{2.119}$$

In fact, the tensor

$$\mathbf{Q} = \mathbf{f}_1 \otimes \mathbf{e}_1 + \mathbf{f}_2 \otimes \mathbf{e}_2 + ... + \mathbf{f}_n \otimes \mathbf{e}_n \tag{2.120}$$

satisfies (2.119) and is orthogonal in view of proposition 29. From definition (2.92) it follows that $\det\mathbf{A} = \det[\mathbf{A}]$, where the matrix $[\mathbf{A}]$ has components $A_{ij} = \mathbf{e}_i \cdot \mathbf{A}\mathbf{e}_j$, $i, j = 1, ..., n$. Let $[\mathbf{A}']$ be the matrix of the components of $\mathbf{A}$ with respect to the basis $\{\mathbf{f}_1, ..., \mathbf{f}_n\}$,

$$A'_{ij} = \mathbf{f}_i \cdot \mathbf{A}\mathbf{f}_j, \quad i, j = 1, ..., n \tag{2.121}$$

and let $[\mathbf{Q}^T\mathbf{AQ}]$ be the matrix of the components of the tensor $\mathbf{Q}^T\mathbf{AQ}$ with respect to the basis $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$,

$$A''_{ij} = \mathbf{e}_i \cdot \mathbf{Q}^T\mathbf{AQe}_j, \quad i, j = 1, ..., n, \tag{2.122}$$

in view of(2.119) we have

$$A'_{ij} = \mathbf{Qe}_i \cdot \mathbf{AQe}_j = A''_{ij}, \quad i, j = 1, ..., n. \tag{2.123}$$

Still from definition (2.92) we have

$$\det \mathbf{Q}^T\mathbf{AQ} = \det[\mathbf{Q}^T\mathbf{AQ}] = \det[\mathbf{A}'], \tag{2.124}$$

from the relation (2.226) we finally get that

$$\det[\mathbf{A}] = \det[\mathbf{A}']. \tag{2.125}$$

□

2.8 Some subsets of Lin

A tensor $\mathbf{A}$ is *positive semidefinite* if

$$\mathbf{v} \cdot \mathbf{Av} \geq 0, \quad \text{for each } \mathbf{v} \in \mathcal{V}, \tag{2.126}$$

is *positive definite* if $\mathbf{v} \cdot \mathbf{Av} > 0, \quad \text{for each } \mathbf{v} \neq \mathbf{0}$.

A tensor $\mathbf{A}$ is *negative semidefinite* if

$$\mathbf{v} \cdot \mathbf{Av} \leq 0, \quad \text{for each } \mathbf{v} \in \mathcal{V}, \tag{2.127}$$

is *negative definite* if $\mathbf{v} \cdot \mathbf{Av} < 0, \quad \text{for each } \mathbf{v} \neq \mathbf{0}$.

Let us consider the following subsets of Lin,

$$\text{Lin}^+ = \{\mathbf{A} \in \text{Lin} : \ \det \mathbf{A} > 0\}, \tag{2.128}$$

$$\text{Psym} = \{\mathbf{A} \in \text{Sym} : \ \mathbf{A} \text{ is positive definite}\}, \tag{2.129}$$

$$\text{Sym}^+ = \{\mathbf{A} \in \text{Sym} : \ \mathbf{A} \text{ is positive semidefinite}\}, \tag{2.130}$$

$$\text{Nsym} = \{\mathbf{A} \in \text{Sym} : \ \mathbf{A} \text{ is negative definite}\}, \tag{2.131}$$

$$\text{Sym}^- = \{\mathbf{A} \in \text{Sym} : \ \mathbf{A} \text{ is negative semidefinite}\}, \tag{2.132}$$

$$\text{Orth} = \left\{\mathbf{Q} \in \text{Lin} : \ \mathbf{QQ}^T = \mathbf{Q}^T\mathbf{Q} = \mathbf{I}\right\}, \tag{2.133}$$

$$\text{Orth}^+ = \{\mathbf{R} \in \text{Orth} : \ \det \mathbf{R} = 1\}. \tag{2.134}$$

Lin$^+$, Orth e Orth$^+$ are groups[1] with respect to the multiplication by tensors; Orth is called *orthogonal group*, Orth$^+$ is called *rotation group*. Psym, Nsym, Sym$^+$ and Sym$^-$ are convex cones[2]. Sets Sym and Skw defined in (2.26) and (2.27) are vector spaces of dimension $\frac{n(n+1)}{2}$ and $\frac{n(n-1)}{2}$, respectively. For $n = 3$ the sets

$$\Big\{\mathbf{e}_1 \otimes \mathbf{e}_1, \mathbf{e}_2 \otimes \mathbf{e}_2, \mathbf{e}_3 \otimes \mathbf{e}_3, \frac{1}{\sqrt{2}}(\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1),$$
$$\frac{1}{\sqrt{2}}(\mathbf{e}_1 \otimes \mathbf{e}_3 + \mathbf{e}_3 \otimes \mathbf{e}_1), \frac{1}{\sqrt{2}}(\mathbf{e}_2 \otimes \mathbf{e}_3 + \mathbf{e}_3 \otimes \mathbf{e}_3)\Big\}, \tag{2.135}$$

and

$$\Big\{\frac{1}{\sqrt{2}}(\mathbf{e}_1 \otimes \mathbf{e}_2 - \mathbf{e}_2 \otimes \mathbf{e}_1),$$
$$\frac{1}{\sqrt{2}}(\mathbf{e}_1 \otimes \mathbf{e}_3 - \mathbf{e}_3 \otimes \mathbf{e}_1), \frac{1}{\sqrt{2}}(\mathbf{e}_2 \otimes \mathbf{e}_3 - \mathbf{e}_3 \otimes \mathbf{e}_3)\Big\}, \tag{2.136}$$

are an orthonormal basis of Sym and Skw, respectively.

A tensor is called *spherical* if $\mathbf{A} = \alpha\mathbf{I}$, with $\alpha \in \mathbb{R}$. Given $\mathbf{A} \in$ Lin, the tensor

$$\mathbf{A}_0 = \mathbf{A} - \frac{1}{n}(tr\mathbf{A})\mathbf{I}, \tag{2.137}$$

is called *deviatoric part* of $\mathbf{A}$. From (2.137) it follows that $tr\mathbf{A}_0 = 0$. Let

$$\text{Dev} = \{\mathbf{A} \in \text{Lin} : tr\mathbf{A} = 0\} \tag{2.138}$$

be the set of deviatoric part of all tensors and

$$\text{Sph} = \{\alpha\mathbf{I} : \alpha \in \mathbb{R}\} \tag{2.139}$$

be the set of all spherical tensors. It is an easy matter to prove that Dev and Sph are subspaces of Lin with dimension $n^2 - 1$ and 1, respectively, that Dev is orthogonal to Sph and that

$$\text{Lin} = \text{Dev} + \text{Sph}. \tag{2.140}$$

Thus, it holds that

$$\text{Lin} = \text{Dev} \oplus \text{Sph}$$

and the orthogonal projections P_{Dev} and P_{Sph} of Lin onto Dev and Sph are defined by

$$P_{\text{Dev}}(\mathbf{A}) = \mathbf{A}_0, \quad P_{\text{Sph}}(\mathbf{A}) = \frac{1}{n}(tr\mathbf{A})\mathbf{I}, \quad \mathbf{A} \in \text{Lin}. \tag{2.141}$$

Exercise 3. *For* $\mathbf{D} \in$ *Psym,* $\mathbf{Q} \in$ *Orth, show that* $\mathbf{QDQ}^T \in$ *Psym.*

[1] A *group* $\mathcal{G}$ is a set of elements with the operation $*$ which satisfies the following properties:

1. If $a, b \in \mathcal{G}$, then $a * b \in \mathcal{G}$,
2. Fir each $a, b, c \in \mathcal{G}$, we have $(a * b) * c = a * (b * c)$,
3. There exists the identity element 1 such that $1 * a = a * 1 = a$, for each $a \in \mathcal{G}$,
4. For each $a \in \mathcal{G}$, there is an element $a^{-1} \in \mathcal{G}$, such that $a^{-1} * a = a * a^{-1} = 1$.

[2] A subset $\mathcal{C}$ of a vector space $\mathcal{S}$ is a *cone* if $\lambda\mathbf{u} \in \mathcal{C}$ for each $\lambda > 0$ and $\mathbf{u} \in \mathcal{C}$.

2.9 Vector product

In this section we fix $n = 3$. Given $\mathbf{u}, \mathbf{v} \in \mathcal{V}$ let us denote by $\mathbf{u} \wedge \mathbf{v}$ the *vector product* of $\mathbf{u}$ and $\mathbf{v}$.

Let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be a right orthonormal basis, the components of $\mathbf{u} \wedge \mathbf{v}$ with respect to this basis are

$$u_2 v_3 - u_3 v_2, \quad u_3 v_1 - u_1 v_3, \quad u_1 v_2 - u_2 v_1. \tag{2.142}$$

The vector product $\wedge$ has the following properties,

$$(\alpha \mathbf{u} + \beta \mathbf{v}) \wedge \mathbf{w} = \alpha \mathbf{u} \wedge \mathbf{w} + \beta \mathbf{v} \wedge \mathbf{w}, \quad \text{(bilinearity)} \tag{2.143}$$

$$\mathbf{u} \wedge \mathbf{v} = -\mathbf{v} \wedge \mathbf{u}, \quad \text{(skew-symmetry)} \tag{2.144}$$

$$\mathbf{u} \wedge \mathbf{u} = \mathbf{0}, \tag{2.145}$$

$$\mathbf{u} \cdot (\mathbf{v} \wedge \mathbf{w}) = \mathbf{w} \cdot (\mathbf{u} \wedge \mathbf{v}) = \mathbf{v} \cdot (\mathbf{w} \wedge \mathbf{u}) \tag{2.146}$$

for each $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V}$, $\alpha, \beta \in \mathbb{R}$.

Moreover, if $\mathbf{u} \neq \mathbf{0}$, then $\mathbf{u} \wedge \mathbf{v} = \mathbf{0}$ if and only if $\mathbf{v} = \alpha \mathbf{u}$ with $\alpha \in \mathbb{R}$. In fact, from (2.142) the following relations follow

$$u_2 v_3 = u_3 v_2, \ u_3 v_1 = u_1 v_3, \ u_1 v_2 = u_2 v_1. \tag{2.147}$$

From (2.147) assuming, for example, that $u_1 \neq 0$, we get

$$v_2 = \frac{u_2}{u_1} v_1, \quad v_3 = \frac{u_3}{u_1} v_1, \tag{2.148}$$

and then $\mathbf{v} = \frac{v_1}{u_1} \mathbf{u}$.

The vector $\mathbf{u} \wedge \mathbf{v}$ is orthogonal to the subspace spanned by $\mathbf{u}$ and $\mathbf{v}$ and we can prove that

$$||\mathbf{u} \wedge \mathbf{v}|| = ||\mathbf{u}|| \ ||\mathbf{v}|| \ \sin\theta, \tag{2.149}$$

where $\theta \in [0, \pi]$ is the angle between the vectors $\mathbf{u}$ and $\mathbf{v}$. Moreover, the mixed product $\mathbf{u} \cdot (\mathbf{v} \wedge \mathbf{w})$ is equal to zero if and only if $\mathbf{u}, \mathbf{v}$ and $\mathbf{w}$ are linearly dependent; in fact, if $\mathbf{u} \cdot (\mathbf{v} \wedge \mathbf{w}) = 0$ then $\mathbf{u} = \mathbf{0}$, or $\mathbf{v} \wedge \mathbf{w} = 0$, that is $\mathbf{w} = \alpha \mathbf{v}$ for some $\alpha \in \mathbb{R}$, or $\mathbf{u}$ is orthogonal to $\mathbf{v} \wedge \mathbf{w}$ and then belongs to the subspace spanned by $\mathbf{v}$ and $\mathbf{w}$.

The further properties hold,

$$||\mathbf{u} \wedge \mathbf{v}||^2 = ||\mathbf{u}||^2 ||\mathbf{v}||^2 - (\mathbf{u} \cdot \mathbf{v})^2, \quad \text{for each } \mathbf{u}, \mathbf{v} \in \mathcal{V}, \tag{2.150}$$

$$||\mathbf{u} \wedge \mathbf{v}||^2 + (\mathbf{u} \cdot \mathbf{v})^2 = 1, \quad \text{for each } \mathbf{u}, \mathbf{v} \in \mathcal{V} \text{ con } ||\mathbf{u}|| = ||\mathbf{v}|| = 1. \tag{2.151}$$

Proposition 31. *For* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V}$*, we have*

$$(\mathbf{u} \wedge \mathbf{v}) \wedge \mathbf{w} = (\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{v} \cdot \mathbf{w})\mathbf{u} = (\mathbf{v} \otimes \mathbf{u} - \mathbf{u} \otimes \mathbf{v})\mathbf{w}. \tag{2.152}$$

Exercise 4. *For* $\mathbf{a}, \mathbf{b} \in \mathcal{V}$*, show that the unique solution to the linear equation*

$$\mathbf{x} + \mathbf{a} \wedge \mathbf{x} = \mathbf{b} \tag{2.153}$$

is

$$\mathbf{x} = \frac{1}{1 + ||\mathbf{a}||^2}[\mathbf{b} + (\mathbf{a} \cdot \mathbf{b})\mathbf{a} + \mathbf{b} \wedge \mathbf{a}]. \tag{2.154}$$

Now we are in the position to prove that (for $n = 3$) there exists a linear bijective mapping from Skw to $\mathcal{V}$, which therefore, are isomorphic. For a vector $\mathbf{w}$ with components w_1, w_2, w_3, let us consider the skew-symmetric tensor

$$\begin{aligned} \mathbf{W} = -w_3(\mathbf{e}_1 \otimes \mathbf{e}_2 - \mathbf{e}_2 \otimes \mathbf{e}_1) + w_2(\mathbf{e}_1 \otimes \mathbf{e}_3 - \mathbf{e}_3 \otimes \mathbf{e}_1) \\ -w_1(\mathbf{e}_2 \otimes \mathbf{e}_3 - \mathbf{e}_3 \otimes \mathbf{e}_3). \end{aligned} \tag{2.155}$$

It is easy to verify that

$$\mathbf{W}\mathbf{a} = \mathbf{w} \wedge \mathbf{a}, \quad \text{for each } \mathbf{a} \in \mathcal{V}. \tag{2.156}$$

$\mathbf{w}$ is called *axial vector* of $\mathbf{W}$.

Vice versa let $\mathbf{W}$ be a skew-symmetric tensor

$$\mathbf{W} = \sum_{\substack{i,j=1 \\ i<j}}^{3} W_{ij}(\mathbf{e}_i \otimes \mathbf{e}_j - \mathbf{e}_j \otimes \mathbf{e}_i). \tag{2.157}$$

For every $\mathbf{a} \in \mathcal{V}$ we have

$$(\mathbf{e}_i \otimes \mathbf{e}_j - \mathbf{e}_j \otimes \mathbf{e}_i)\mathbf{a} = (\mathbf{e}_j \cdot \mathbf{a})\mathbf{e}_i - (\mathbf{e}_i \cdot \mathbf{a})\mathbf{e}_j = (\mathbf{e}_j \wedge \mathbf{e}_i) \wedge \mathbf{a}, \tag{2.158}$$

then

$$\mathbf{W}\mathbf{a} = \sum_{\substack{i,j=1 \\ i<j}}^{3} W_{ij}(\mathbf{e}_j \wedge \mathbf{e}_i) \wedge \mathbf{a}, \tag{2.159}$$

and

$$\mathbf{w} = \sum_{\substack{i,j=1 \\ i<j}}^{3} W_{ij}(\mathbf{e}_j \wedge \mathbf{e}_i) \tag{2.160}$$

is the axial vector of $\mathbf{W}$.

From (2.250) it follows that (for $\mathbf{W} \neq \mathbf{0}$) the subspace of $\mathcal{V}$

$$Ker\mathbf{W} = \{\mathbf{v} \in \mathcal{V} \mid \mathbf{W}\mathbf{v} = \mathbf{0}\} \tag{2.161}$$

has dimension 1 and is spanned by $\mathbf{w}$. $Ker\mathbf{W}$ is called *axis* of $\mathbf{W}$.

Given $\mathbf{e}_1, \mathbf{e}_2$ orthonormal vectors, the vector $\mathbf{e}_3 = \mathbf{e}_1 \wedge \mathbf{e}_2$ is the axial vector of the skew-symmetric tensor

$$\mathbf{W} = \mathbf{e}_2 \otimes \mathbf{e}_1 - \mathbf{e}_1 \otimes \mathbf{e}_2. \tag{2.162}$$

Proposition 32. *Consider* $\mathbf{Q} \in$*Orth, if* $\det \mathbf{Q} = 1$, *then there exists* $\mathbf{e} \in \mathcal{V}$ *such that* $\mathbf{Qe} = \mathbf{e}$. *If, on the contrary,* $\det \mathbf{Q} = -1$, *then there exists* $\mathbf{e} \in \mathcal{V}$ *such that* $\mathbf{Qe} = -\mathbf{e}$.

Proof. We have

$$\det(\mathbf{Q} - \mathbf{I}) = \det[\mathbf{Q}(\mathbf{I} - \mathbf{Q}^T) = \det \mathbf{Q} \det(\mathbf{I} - \mathbf{Q}^T) =$$

$$- \det \mathbf{Q} \det(\mathbf{Q} - \mathbf{I}) = - \det(\mathbf{Q} - \mathbf{I}),$$

then $\det(\mathbf{Q} - \mathbf{I}) = 0$ and in virtue of theorem 8, there is $\mathbf{e} \in \mathcal{V}$, $\mathbf{e} \neq 0$ such that $(\mathbf{Q} - \mathbf{I})\mathbf{e} = \mathbf{0}$. If $\det \mathbf{Q} = -1$, the proof is analogous. □

Exercise 5. *Let* $\mathbf{Q} \in$*Orth and* $\mathbf{e} \in \mathcal{V}$ *be vectors such that*$\mathbf{Qe} = \mathbf{e}$.

1. *Prove that* $\mathbf{Q}^T\mathbf{e} = \mathbf{e}$.
2. *Let* $\mathbf{w}$ *be the axial vector of the skew-symmetric part of* $\mathbf{Q}$, *prove that* $\mathbf{e} \in$*Span(*$\mathbf{w}$*).*

Solution. 1. $\mathbf{Qe} = \mathbf{e} \Longrightarrow \mathbf{Q}^T\mathbf{Qe} = \mathbf{Q}^T\mathbf{e} \Longrightarrow \mathbf{Q}^T\mathbf{e} = \mathbf{e}$.
2. Let $\mathbf{W} = (\mathbf{Q} - \mathbf{Q}^T)/2$ be the skew-symmetric part of $\mathbf{Q}$, we have

$$\mathbf{Wv} = \frac{1}{2}(\mathbf{Q} - \mathbf{Q}^T)\mathbf{v} = \mathbf{w} \wedge \mathbf{v}, \quad \mathbf{v} \in \mathcal{V}, \tag{2.163}$$

in particular,

$$\mathbf{w} \wedge \mathbf{e} = \mathbf{We} = \mathbf{0}, \tag{2.164}$$

then $\mathbf{e} \in$Span($\mathbf{w}$).

From the proposition 32 we get that the subspace $\mathcal{A}(\mathbf{Q}) = \{\mathbf{e} \in \mathcal{V} : \mathbf{Qe} = \mathbf{e}\}$ contains non-zero elements. $\mathcal{A}(\mathbf{Q})$ is called *axis* of $\mathbf{Q}$ and, in view of the exercise 5, has dimension 1.

Exercise 6. *Given* $\mathbf{W}, \mathbf{Z} \in$*Skw, let* $\mathbf{w}, \mathbf{z} \in \mathcal{V}$ *be the corresponding axial vectors. Prove that*

$$\mathbf{WZ} = \mathbf{z} \otimes \mathbf{w} - (\mathbf{z} \cdot \mathbf{w})\mathbf{I}; \tag{2.165}$$

thus, in particular

$$\mathbf{WZ} - \mathbf{ZW} = \mathbf{z} \otimes \mathbf{w} - \mathbf{w} \otimes \mathbf{z}, \tag{2.166}$$

$$\mathbf{Z} \cdot \mathbf{W} = 2(\mathbf{z} \cdot \mathbf{w}),$$

and

$$||\mathbf{w}|| = \frac{1}{\sqrt{2}}||\mathbf{W}||. \tag{2.167}$$

Solution. In view of (2.250) and (2.152) we have

$$\mathbf{WZv} = \mathbf{w} \wedge (\mathbf{z} \wedge \mathbf{v}) = (\mathbf{v} \cdot \mathbf{w})\mathbf{z} - (\mathbf{z} \cdot \mathbf{w})\mathbf{v}$$

$$= [\mathbf{z} \otimes \mathbf{w} - (\mathbf{z} \cdot \mathbf{w})\mathbf{I}]\mathbf{v}, \quad \text{for each } \mathbf{z} \in \mathcal{V}.$$

If the vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V}$ are linearly independent, then the scalar $|\mathbf{u} \cdot (\mathbf{v} \wedge \mathbf{w})|$ is the volume of the parallelepiped $\mathcal{P}$ determined by $\mathbf{u}, \mathbf{v}, \mathbf{w}$.

Proposition 33. *Given the linearly independent vectors* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V}$ *and* $\mathbf{A} \in Lin$ *we have*

$$\det \mathbf{A} = \frac{\mathbf{Au} \cdot (\mathbf{Av} \wedge \mathbf{Aw})}{\mathbf{u} \cdot (\mathbf{v} \wedge \mathbf{w})}. \tag{2.168}$$

In particular, from (2.168) we get the relation

$$|\det \mathbf{A}| = \frac{Vol(\mathbf{A}(\mathcal{P}))}{Vol(\mathcal{P})}, \tag{2.169}$$

which gives a geometrical interpretation of the determinant. In (2.169) $\mathbf{A}(\mathcal{P})$ is the image of $\mathcal{P}$ under $\mathbf{A}$ and Vol designates the volume.

Relation (2.168) comes from the following propositions.

Proposition 34. *For* $\mathbf{A} \in Lin$*, let* $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ *and* $\{\mathbf{u}', \mathbf{v}', \mathbf{w}'\}$ *two sets of linearly independent vectors of* $\mathcal{V}$*. We have*

$$\frac{\mathbf{Au} \cdot (\mathbf{Av} \wedge \mathbf{Aw})}{\mathbf{u} \cdot (\mathbf{v} \wedge \mathbf{w})} = \frac{\mathbf{Au}' \cdot (\mathbf{Av}' \wedge \mathbf{Aw}')}{\mathbf{u}' \cdot (\mathbf{v}' \wedge \mathbf{w}')}. \tag{2.170}$$

Proof. Since $\mathbf{u}, \mathbf{v}, \mathbf{w}$ are linearly independent, we have that $\mathbf{u} \cdot (\mathbf{v} \wedge \mathbf{w}) \neq 0$. Let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be an orthonormal basis of $\mathcal{V}$, with $\mathbf{e}_3 = \mathbf{e}_1 \wedge \mathbf{e}_2$. To prove (2.170) it is sufficient to prove that for each set of linearly independent vectors $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ we have

$$\mathbf{Au} \cdot (\mathbf{Av} \wedge \mathbf{Aw}) = \mathbf{u} \cdot (\mathbf{v} \wedge \mathbf{w})[\mathbf{Ae}_1 \cdot (\mathbf{Ae}_2 \wedge \mathbf{Ae}_3)]. \tag{2.171}$$

The following relations hold

$$\mathbf{u} = u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + u_3\mathbf{e}_3, \tag{2.172}$$

$$\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + v_3\mathbf{e}_3, \tag{2.173}$$

$$\mathbf{w} = w_1\mathbf{e}_1 + w_2\mathbf{e}_2 + w_3\mathbf{e}_3, \tag{2.174}$$

from which we get

$$\mathbf{Au} \cdot (\mathbf{Av} \wedge \mathbf{Aw}) = (u_1\mathbf{Ae}_1 + u_2\mathbf{Ae}_2 + u_3\mathbf{Ae}_3) \cdot [(v_1\mathbf{Ae}_1 + v_2\mathbf{Ae}_2 +$$

$$v_3\mathbf{Ae}_3) \wedge (w_1\mathbf{Ae}_1 + w_2\mathbf{Ae}_2 + w_3\mathbf{Ae}_3)] =$$

$$(u_1\mathbf{Ae}_1 + u_2\mathbf{Ae}_2 + u_3\mathbf{Ae}_3) \cdot [v_1w_2\mathbf{Ae}_1 \wedge \mathbf{Ae}_2 + v_1w_3\mathbf{Ae}_1 \wedge \mathbf{Ae}_3 +$$

$$v_2 w_1 \mathbf{Ae_2} \wedge \mathbf{Ae}_1 + v_2 w_3 \mathbf{Ae_2} \wedge \mathbf{Ae}_3 + v_3 w_1 \mathbf{Ae_3} \wedge \mathbf{Ae}_1 + v_3 w_2 \mathbf{Ae_3} \wedge \mathbf{Ae}_2] =$$
$$u_1 v_2 w_3 \mathbf{Ae}_1 \cdot (\mathbf{Ae_2} \wedge \mathbf{Ae}_3) + u_1 v_3 w_2 \mathbf{Ae}_1 \cdot (\mathbf{Ae_3} \wedge \mathbf{Ae}_2) +$$
$$u_2 v_1 w_3 \mathbf{Ae}_2 \cdot (\mathbf{Ae_1} \wedge \mathbf{Ae}_3) + u_2 v_3 w_1 \mathbf{Ae}_2 \cdot (\mathbf{Ae_3} \wedge \mathbf{Ae}_1) +$$
$$u_3 v_1 w_2 \mathbf{Ae}_3 \cdot (\mathbf{Ae_1} \wedge \mathbf{Ae}_2) + u_3 v_2 w_1 \mathbf{Ae}_3 \cdot (\mathbf{Ae_2} \wedge \mathbf{Ae}_1) =$$
$$\mathbf{Ae}_1 \cdot (\mathbf{Ae_2} \wedge \mathbf{Ae}_3)[u_1 v_2 w_3 - u_1 v_3 w_2 - u_2 v_1 w_3 + u_2 v_3 w_1 +$$
$$u_3 v_1 w_2 - u_3 v_2 w_1] = [\mathbf{u} \cdot (\mathbf{v} \wedge \mathbf{w})] \mathbf{Ae}_1 \cdot (\mathbf{Ae_2} \wedge \mathbf{Ae}_3).$$

□

Proposition 35. *Let* $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ *be an orthonormal basis of* $\mathcal{V}$*, with* $\mathbf{e}_3 = \mathbf{e}_1 \wedge \mathbf{e}_2$*; for each tensor* $\mathbf{A}$ *we have*

$$\det \mathbf{A} = \mathbf{Ae}_1 \cdot (\mathbf{Ae}_2 \wedge \mathbf{Ae}_3). \tag{2.175}$$

Proof. We have

$$\mathbf{Ae}_k = \sum_{i=1}^{3} A_{ik} \mathbf{e}_i, \quad k = 1, 2, 3. \tag{2.176}$$

then

$$\mathbf{Ae}_2 \wedge \mathbf{Ae}_3 = A_{12} A_{23} \mathbf{e}_3 - A_{12} A_{33} \mathbf{e}_2 - A_{13} A_{22} \mathbf{e}_3 + \tag{2.177}$$
$$A_{22} A_{33} \mathbf{e}_1 - A_{13} A_{32} \mathbf{e}_2 - A_{32} A_{23} \mathbf{e}_1, \tag{2.178}$$
$$\mathbf{Ae}_1 \cdot (\mathbf{Ae}_2 \wedge \mathbf{Ae}_3) = A_{11}(A_{22} A_{33} - A_{32} A_{32}) + A_{21}(A_{13} A_{32} - A_{12} A_{33}) + \tag{2.179}$$
$$A_{31}(A_{12} A_{23} - A_{13} A_{22}). \tag{2.180}$$

On the other hand

$$\det \mathbf{A} = A_{11}(A_{22} A_{33} - A_{32} A_{32}) + A_{21}(A_{13} A_{32} - A_{12} A_{33}) + \tag{2.181}$$
$$A_{31}(A_{12} A_{23} - A_{13} A_{22}), \tag{2.182}$$

from which the thesis follows. □

Exercise 7. *Let* $\varphi : \mathcal{V} \times \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ *be a skew-symmetric trilinear functional, that is linear in each argument and*

$$\varphi(\mathbf{u}, \mathbf{v}, \mathbf{w}) = -\varphi(\mathbf{v}, \mathbf{u}, \mathbf{w}) = -\varphi(\mathbf{u}, \mathbf{w}, \mathbf{v}) = -\varphi(\mathbf{w}, \mathbf{v}, \mathbf{u}), \tag{2.183}$$

for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V}$*. Let* $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ *be an orthonormal basis of* $\mathcal{V}$*, for* $\mathbf{A} \in$ *Lin, prove that*

$$\varphi(\mathbf{Ae}_1, \mathbf{e}_2, \mathbf{e}_3) + \varphi(\mathbf{e}_1, \mathbf{Ae}_2, \mathbf{e}_3) + \varphi(\mathbf{e}_1, \mathbf{e}_2, \mathbf{Ae}_3) = (tr\mathbf{A})\varphi(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3). \tag{2.184}$$

Solution.

$$\varphi(\mathbf{A}\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = \varphi(\sum_{i=1}^{3} A_{i1}\mathbf{e}_i, \mathbf{e}_2, \mathbf{e}_3) = \tag{2.185}$$

$$\varphi(A_{11}\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) + \varphi(A_{21}\mathbf{e}_2, \mathbf{e}_2, \mathbf{e}_3) + \varphi(A_{31}\mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_3). \tag{2.186}$$

On the other hand we have

$$\varphi(A_{21}\mathbf{e}_2, \mathbf{e}_2, \mathbf{e}_3) = A_{21}\varphi(\mathbf{e}_2, \mathbf{e}_2, \mathbf{e}_3) = -A_{21}\varphi(\mathbf{e}_2, \mathbf{e}_2, \mathbf{e}_3), \tag{2.187}$$

then $\varphi(A_{21}\mathbf{e}_2, \mathbf{e}_2, \mathbf{e}_3) = 0$. In a similar way we prove that $\varphi(A_{31}\mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_3) = 0$ and from (2.186) we get

$$\varphi(\mathbf{A}\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = \varphi(A_{11}\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3). \tag{2.188}$$

Using the same arguments for $\varphi(\mathbf{e}_1, \mathbf{A}\mathbf{e}_2, \mathbf{e}_3)$ and $\varphi(\mathbf{e}_1, \mathbf{e}_2, \mathbf{A}\mathbf{e}_3)$ the thesis follows.

2.10 Cofactor of a second-order tensor

Put $n = 3$. Given $\mathbf{A} \in \text{Lin}$, the *cofactor* $\mathbf{A}^*$ of $\mathbf{A}$ is the unique element of Lin such that for each $\mathbf{w} \in \mathcal{V}$, $\mathbf{W} \in \text{Skw}$ linked by the relation

$$\mathbf{W}\mathbf{v} = \mathbf{w} \wedge \mathbf{v}, \quad \text{for each } \mathbf{v} \in \mathcal{V}, \tag{2.189}$$

vector $\mathbf{A}^*\mathbf{w}$ and tensor $\mathbf{A}\mathbf{W}\mathbf{A}^T$ in their turn, satisfy the relation

$$\mathbf{A}\mathbf{W}\mathbf{A}^T\mathbf{v} = (\mathbf{A}^*\mathbf{w}) \wedge \mathbf{v}, \quad \text{for each } \mathbf{v} \in \mathcal{V}. \tag{2.190}$$

Proposition 36. *Consider $\mathbf{A} \in Lin$, its cofactor $\mathbf{A}^*$, and $\mathbf{c} \in \mathcal{V}$. The following properties hold.*

(1) For each $\mathbf{a}, \mathbf{b} \in \mathcal{V}$, we have

$$\mathbf{A}^*(\mathbf{a} \wedge \mathbf{b}) = (\mathbf{A}\mathbf{a}) \wedge (\mathbf{A}\mathbf{b}). \tag{2.191}$$

(2) If $\mathbf{A}$ is invertible, then $\mathbf{A}^*$ is invertible

$$\mathbf{A}^* = (\det \mathbf{A})\mathbf{A}^{-T}. \tag{2.192}$$

(3) The rotation group $Orth^+$ coincides with the set

$$\mathcal{C} = \{\mathbf{R} \in Lin - \{0\} \;\; : \;\; \mathbf{R} = \mathbf{R}^*\}. \tag{2.193}$$

(4) The dyad $\mathbf{c} \otimes \mathbf{c}$ satisfies the relation

$$(\mathbf{I} - \mathbf{c} \otimes \mathbf{c})^* = (1 - \mathbf{c} \cdot \mathbf{c})\mathbf{I} + \mathbf{c} \otimes \mathbf{c}. \tag{2.194}$$

(5) Let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ *be an orthonormal basis of* $\mathcal{V}$*, we have*

$$\sum_{i=1}^{3} (\mathbf{I} - \mathbf{e}_i \otimes \mathbf{e}_i)^* = \mathbf{I}. \tag{2.195}$$

Proof. (1) Given $\mathbf{w} = \mathbf{a} \wedge \mathbf{b}$ and $\mathbf{W} = \mathbf{b} \otimes \mathbf{a} - \mathbf{a} \otimes \mathbf{b}$, the skew-symmetric tensor associated to $\mathbf{w}$, in view of (2.190) we have

$$(\mathbf{A}^*\mathbf{w}) \wedge \mathbf{v} = \mathbf{A}(\mathbf{b} \otimes \mathbf{a} - \mathbf{a} \otimes \mathbf{b})\mathbf{A}^T\mathbf{v} \tag{2.196}$$

$$= (\mathbf{A}\mathbf{b} \otimes \mathbf{A}\mathbf{a} - \mathbf{A}\mathbf{a} \otimes \mathbf{A}\mathbf{b})\mathbf{v} = (\mathbf{A}\mathbf{a} \wedge \mathbf{A}\mathbf{b}) \wedge \mathbf{v} \quad \text{for all } \mathbf{v} \in \mathcal{V}, \tag{2.197}$$

from which we deduce (2.191).

(2) Let $\mathbf{A}$ be an invertible tensor and assume that there is $\mathbf{v} \in \mathcal{V}$, $\mathbf{v} \neq \mathbf{0}$ such that $\mathbf{A}^*\mathbf{w} = \mathbf{0}$. Take $\mathbf{a}, \mathbf{b} \in \mathcal{V}$, both differente from zero such that $\mathbf{v} = \mathbf{a} \wedge \mathbf{b}$. In view of (2.191) we have

$$\mathbf{0} = \mathbf{A}^*(\mathbf{a} \wedge \mathbf{b}) = \mathbf{A}\mathbf{a} \wedge \mathbf{A}\mathbf{b}, \tag{2.198}$$

from which, taking into account the properties of the vector product, we deduce that $\mathbf{A}\mathbf{a} = \alpha\mathbf{A}\mathbf{b}$, then $\mathbf{a} - \alpha\mathbf{b} = \mathbf{0}$ and finally $\mathbf{v} = \mathbf{0}$. Now, let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be an orthonormal basis of $\mathcal{V}$ with $\mathbf{e}_3 = \mathbf{e}_1 \wedge \mathbf{e}_2$. We have

$$\mathbf{A}^*\mathbf{e}_1 = \mathbf{A}\mathbf{e}_2 \wedge \mathbf{A}\mathbf{e}_3, \tag{2.199}$$

$$\mathbf{A}^*\mathbf{e}_2 = \mathbf{A}\mathbf{e}_3 \wedge \mathbf{A}\mathbf{e}_1, \tag{2.200}$$

$$\mathbf{A}^*\mathbf{e}_3 = \mathbf{A}\mathbf{e}_1 \wedge \mathbf{A}\mathbf{e}_2, \tag{2.201}$$

and, in view of Proposition 35,the following relations hold

$$\mathbf{A}^T\mathbf{A}^*\mathbf{e}_1 \cdot \mathbf{e}_1 = \mathbf{A}^*\mathbf{e}_1 \cdot \mathbf{A}\mathbf{e}_1 = \det \mathbf{A}, \tag{2.202}$$

$$\mathbf{A}^T\mathbf{A}^*\mathbf{e}_2 \cdot \mathbf{e}_2 = \mathbf{A}^*\mathbf{e}_2 \cdot \mathbf{A}\mathbf{e}_2 = \det \mathbf{A}, \tag{2.203}$$

$$\mathbf{A}^T\mathbf{A}^*\mathbf{e}_3 \cdot \mathbf{e}_3 = \mathbf{A}^*\mathbf{e}_3 \cdot \mathbf{A}\mathbf{e}_3 = \det \mathbf{A}, \tag{2.204}$$

and

$$\mathbf{A}^T\mathbf{A}^*\mathbf{e}_i \cdot \mathbf{e}_j = \mathbf{A}^*\mathbf{e}_i \cdot \mathbf{A}\mathbf{e}_j = 0 \quad \text{if } i \neq j, \tag{2.205}$$

we can thus conclude that

$$\mathbf{A}^T\mathbf{A}^* = (\det \mathbf{A})\mathbf{I}, \tag{2.206}$$

from which (2.192) follows.

(3) Consider $\mathbf{R} \in$Orth$^+$, from (2.192) taking account that $\det \mathbf{R} = 1$ and $\mathbf{R}^T = \mathbf{R}^{-1}$, we get that $\mathbf{R}^* = \mathbf{R}$ and then $\mathbf{R} \in \mathcal{C}$. Vice versa, let us assume that $\mathbf{R} \in \mathcal{C}$, then

$$\mathbf{R}(\mathbf{a} \wedge \mathbf{b}) = \mathbf{R}^*(\mathbf{a} \wedge \mathbf{b}) = (\mathbf{R}\mathbf{a}) \wedge (\mathbf{R}\mathbf{b}), \quad \mathbf{a}, \mathbf{b} \in \mathcal{V}. \tag{2.207}$$

Let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be an orthonormal basis of $\mathcal{V}$ with $\mathbf{e}_3 = \mathbf{e}_1 \wedge \mathbf{e}_2$, in view of (2.207) we have

$$\mathbf{R}\mathbf{e}_3 = \mathbf{R}\mathbf{e}_1 \wedge \mathbf{R}\mathbf{e}_2, \tag{2.208}$$

$$\mathbf{R}\mathbf{e}_2 = \mathbf{R}\mathbf{e}_3 \wedge \mathbf{R}\mathbf{e}_1, \tag{2.209}$$

$$\mathbf{R}\mathbf{e}_1 = \mathbf{R}\mathbf{e}_2 \wedge \mathbf{R}\mathbf{e}_3, \tag{2.210}$$

then

$$\mathbf{R}\mathbf{e}_i \cdot \mathbf{R}\mathbf{e}_j = \begin{cases} \det \mathbf{R} & i = j, \\ 0 & i \neq j, \end{cases} \tag{2.211}$$

and finally

$$\mathbf{R}^T\mathbf{R} = (\det \mathbf{R})\mathbf{I}. \tag{2.212}$$

From (2.212) it follows that

$$(\det \mathbf{R})^2 = \det \mathbf{R}, \tag{2.213}$$

from which we obtain that $\det \mathbf{R} = 0$ or $\det \mathbf{R} = 1$. If $\det \mathbf{R} = 0$, from (2.212) we get $\mathbf{R}^T\mathbf{R} = \mathbf{0}$ and then $\mathbf{R} = \mathbf{0}$ which is excluded b that fact that $\mathbf{R} \in \mathcal{C}$. Therefore, we have that $\det \mathbf{R} = 1$, which, along with (2.212), allows to conclude that $\mathbf{R}^T = \mathbf{R}^{-1}$.

(4) For each $\mathbf{u}, \mathbf{v} \in \mathcal{V}$, we have

$$(\mathbf{I} - \mathbf{c} \otimes \mathbf{c})^*(\mathbf{u} \wedge \mathbf{v}) = (\mathbf{I} - \mathbf{c} \otimes \mathbf{c})\mathbf{u} \wedge (\mathbf{I} - \mathbf{c} \otimes \mathbf{c})\mathbf{v} \tag{2.214}$$

$$= \mathbf{u} \wedge \mathbf{v} - (\mathbf{c} \cdot \mathbf{u})\mathbf{c} \wedge \mathbf{v} + (\mathbf{c} \cdot \mathbf{v})\mathbf{c} \wedge \mathbf{u}, \tag{2.215}$$

and

$$[(1 - \mathbf{c} \cdot \mathbf{c})\mathbf{I} + \mathbf{c} \otimes \mathbf{c}](\mathbf{u} \wedge \mathbf{v}) = \mathbf{u} \wedge \mathbf{v} - (\mathbf{c} \cdot \mathbf{c})\mathbf{u} \wedge \mathbf{v} + [\mathbf{c} \cdot (\mathbf{u} \wedge \mathbf{v})]\mathbf{c} \tag{2.216}$$

$$= \mathbf{u} \wedge \mathbf{v} + \mathbf{c} \wedge [\mathbf{c} \wedge (\mathbf{u} \wedge \mathbf{v})], \tag{2.217}$$

where the latest equality comes from (2.152). On the other hand,

$$\mathbf{c} \wedge [\mathbf{c} \wedge (\mathbf{u} \wedge \mathbf{v})] = -\mathbf{c} \wedge [(\mathbf{u} \wedge \mathbf{v}) \wedge \mathbf{c}] \tag{2.218}$$

$$= -\mathbf{c} \wedge [(\mathbf{u} \cdot \mathbf{c})\mathbf{v} - (\mathbf{v} \cdot \mathbf{c})\mathbf{u}] = -(\mathbf{u} \cdot \mathbf{c})\mathbf{c} \wedge \mathbf{v} + (\mathbf{v} \cdot \mathbf{c})\mathbf{c} \wedge \mathbf{u}, \tag{2.219}$$

substituting (2.219) in (2.217) and comparing the obtained expression with (2.217), we get (2.194).

(5) To prove (2.195) it is sufficient to note that in view of (2.194) we have

$$\sum_{i=1}^{3} (\mathbf{I} - \mathbf{e}_i \otimes \mathbf{e}_i)^* = \sum_{i=1}^{3} \mathbf{e}_i \otimes \mathbf{e}_i = \mathbf{I}. \tag{2.220}$$

□

2.11 Principal invariants

For each $\mathbf{A} \in$ Lin, let us introduce the following scalar quantities,

$$I_1(\mathbf{A}) = tr\mathbf{A}, \tag{2.221}$$

$$I_2(\mathbf{A}) = \frac{1}{2}[(tr\mathbf{A})^2 - tr(\mathbf{A}^2)] \tag{2.222}$$

$$I_3(\mathbf{A}) = \det \mathbf{A}. \tag{2.223}$$

$I_1(\mathbf{A}), I_2(\mathbf{A}), I_3(\mathbf{A})$ are called *principal invariants* of $\mathbf{A}$. For $i = 1, 2, 3$ we have

$$I_i(\mathbf{Q}\mathbf{A}\mathbf{Q}^T) = I_i(\mathbf{A}), \quad \text{for each } \mathbf{Q} \in \text{Orth}. \tag{2.224}$$

In fact, for $\mathbf{Q} \in$Orth we have

$$I_1(\mathbf{A}) = tr\mathbf{A} = tr(\mathbf{Q}^T\mathbf{Q}\mathbf{A}) = tr(\mathbf{Q}\mathbf{A}\mathbf{Q}^T) = I_1(\mathbf{Q}\mathbf{A}\mathbf{Q}^T). \tag{2.225}$$

Moreover, let us note that

$$tr(\mathbf{A}^2) = tr(\mathbf{Q}^T\mathbf{Q}\mathbf{A}\mathbf{Q}^T\mathbf{Q}\mathbf{A}) = tr(\mathbf{Q}\mathbf{A}\mathbf{Q}^T\mathbf{Q}\mathbf{A}\mathbf{Q}^T) = tr[(\mathbf{Q}\mathbf{A}\mathbf{Q}^T)^2],$$

and then, in view of (2.225)

$$I_2(\mathbf{A}) = I_2(\mathbf{Q}\mathbf{A}\mathbf{Q}^T).$$

Finally, in view of (2.93) we have

$$\begin{aligned} I_3(\mathbf{A}) &= \det \mathbf{A} = (\det \mathbf{Q})(\det \mathbf{A})(\det \mathbf{Q}^T) \\ &= \det(\mathbf{Q}\mathbf{A}\mathbf{Q}^T) = I_3(\mathbf{Q}\mathbf{A}\mathbf{Q}^T). \end{aligned} \tag{2.226}$$

It is easy to verify that if $n = 3$ and A_{ij}, $i, j = 1, 2, 3$ are the components of $\mathbf{A}$ with respect to an orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ of $\mathcal{V}$, we have

$$I_1(\mathbf{A}) = A_{11} + A_{22} + A_{33}, \tag{2.227}$$

$$I_2(\mathbf{A}) = A_{11}A_{22} + A_{22}A_{33} + A_{11}A_{33} - A_{12}A_{21} - A_{13}A_{31} - A_{23}A_{32}, \tag{2.228}$$

$$\begin{aligned} I_3(\mathbf{A}) = A_{11}(A_{22}A_{33} - A_{32}A_{32}) + A_{21}(A_{13}A_{32} - A_{12}A_{33}) + \\ A_{31}(A_{12}A_{23} - A_{13}A_{22}), \end{aligned} \tag{2.229}$$

or equivalently,

$$I_1(\mathbf{A}) = \mathbf{I} \cdot \mathbf{A}, \tag{2.230}$$

$$I_2(\mathbf{A}) = \mathbf{e}_1 \cdot (\mathbf{A}\mathbf{e}_2 \wedge \mathbf{A}\mathbf{e}_3) + \mathbf{e}_2 \cdot (\mathbf{A}\mathbf{e}_3 \wedge \mathbf{A}\mathbf{e}_1) + \mathbf{e}_3 \cdot (\mathbf{A}\mathbf{e}_1 \wedge \mathbf{A}\mathbf{e}_2), \tag{2.231}$$

$$I_3(\mathbf{A}) = \mathbf{A}\mathbf{e}_1 \cdot (\mathbf{A}\mathbf{e}_2 \wedge \mathbf{A}\mathbf{e}_3), \tag{2.232}$$

Let $\eta(\mathbf{A}) = \{I_1(\mathbf{A}), I_2(\mathbf{A}), I_3(\mathbf{A})\}$ be denote the list of the principal invariants of $\mathbf{A}$.

2.12 Eigenvalues and eigenvectors

A real number a is an *eigenvalue* of $\mathbf{A} \in \mathrm{Lin}$ if there is a vector $\mathbf{u} \in \mathcal{V}$, $||\mathbf{u}|| = 1$, such that

$$\mathbf{A}\mathbf{u} = a\mathbf{u}, \tag{2.233}$$

$\mathbf{u}$ is called *eigenvector* of $\mathbf{A}$. Given an eigenvalue a of $\mathbf{A}$,

$$\mathcal{M}(a) = \{\mathbf{u} \in \mathcal{V} \mid \mathbf{A}\mathbf{u} = a\mathbf{u}\} \tag{2.234}$$

is a subspace of $\mathcal{V}$ called *characteristic space* of $\mathbf{A}$ corresponding to the eigenvalue a. If $\mathcal{M}(a)$ has dimension m, then the eigenvalue a is said to have *multiplicity* m. The set $\sigma(\mathbf{A})$ of the eigenvalues of $\mathbf{A}$, each repeated a number of times equal to its multiplicity is called *spectrum* of $\mathbf{A}$.

a is an eigenvalue of $\mathbf{A}$ if and only if the tensor $\mathbf{A} - a\mathbf{I}$ is not invertible and then a is a real root of the *characteristic polynomial* of $\mathbf{A}$,

$$p(a) = \det(\mathbf{A} - a\mathbf{I}). \tag{2.235}$$

The eigenvalues of a tensor are also called *principal components* (of the tensor), and the eigenvectors are called *principal vectors.*

Proposition 37. *If $n = 3$ the characteristic polynomial (2.235) of $\mathbf{A} \in$ Lin ha the following expression*

$$p(a) = -a^3 + I_1(\mathbf{A})a^2 - I_2(\mathbf{A})a + I_3(\mathbf{A}). \tag{2.236}$$

Proof. To prove (2.236) take into account that if $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is an orthonormal basis of $\mathcal{V}$, in view of proposition 35 we have,

$$\det(\mathbf{A} - a\mathbf{I}) = (\mathbf{A} - a\mathbf{I})\mathbf{e}_1 \cdot [(\mathbf{A} - a\mathbf{I})\mathbf{e}_2 \wedge (\mathbf{A} - a\mathbf{I})\mathbf{e}_3].$$

□

The third degree polynomial (with real coefficients) (2.236) has at least a real root.

2.13 Spectral theorem

Proposition 38. *The following properties hold.*

(a) *The characteristic spaces of a tensor $\mathbf{S} \in$ Sym are mutually orthogonal.*

(b) *The eigenvalues of a tensor $\mathbf{P} \in$ Psym are positive, the eigenvalues of a tensor $\mathbf{S} \in \mathrm{Sym}^+$ are non-negative.*

(c) *The eigenvalues of a tensor $\mathbf{N} \in$ Nsym are negative, the eigenvalues of a tensor $\mathbf{S} \in \mathrm{Sym}^-$ are non-positive.*

Proof. (a) Let a and b be the eigenvalues of $\mathbf{S} \in \mathrm{Sym}$ and let $\mathbf{u}$, $\mathbf{v}$ the corresponding eigenvectors,

$$\mathbf{Su} = a\mathbf{u}, \quad \mathbf{Sv} = b\mathbf{v}.$$

We have

$$a\mathbf{u} \cdot \mathbf{v} = \mathbf{Su} \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{Sv} = b\mathbf{u} \cdot \mathbf{v},$$

then

$$(a - b)\mathbf{u} \cdot \mathbf{v} = 0,$$

therefore, if $a \neq b$ then $\mathbf{u} \cdot \mathbf{v} = 0$.

(b) Given $\mathbf{P} \in \mathrm{Psym}$ let a be an eigenvalue and $\mathbf{v}$ the corresponding eigenvector of $\mathbf{P}$, we have

$$a = a\mathbf{v} \cdot \mathbf{v} = \mathbf{Pv} \cdot \mathbf{v} > 0.$$

□

Theorem 10. *(Spectral theorem). Let* $\mathbf{S}$ *be a symmetric tensor. There exist an orthonormal basis of* $\mathcal{V}$ *constituted by eigenvectors* $\mathbf{g}_1, ..., \mathbf{g}_n$ *of* $\mathbf{S}$*, and* n *eigenvalues* $s_1, ..., s_n$ *of* $\mathbf{S}$*,*

$$\mathbf{S}\mathbf{g}_i = s_i \ \mathbf{g}_i, \quad i = 1, ..., n, \tag{2.237}$$

such that

$$\mathbf{S} = \sum_{i=1}^{n} s_i \ \mathbf{g}_i \otimes \mathbf{g}_i. \tag{2.238}$$

In particular, for $n = 3$, one of the following cases hold:

1. $\mathbf{S}$ has three distinct eigenvalues, then the characteristic spaces of $\mathbf{S}$ are $\mathrm{Span}(\mathbf{g}_1)$, $\mathrm{Span}(\mathbf{g}_2)$ and $\mathrm{Span}(\mathbf{g}_3)$.

2. $\mathbf{S}$ has two distinct eigenvalues $s_1 \neq s_2$, $s_2 = s_3$, then (2.238) reduces to

$$\mathbf{S} = s_1\mathbf{g}_1 \otimes \mathbf{g}_1 + s_2(\mathbf{I} - \mathbf{g}_1 \otimes \mathbf{g}_1). \tag{2.239}$$

$\mathrm{Span}(\mathbf{g}_1)$ is the characteristic space corresponding to s_1 and $\mathrm{Span}(\mathbf{g}_1)^\perp$ the characteristic space corresponding to s_2.

3. $\mathbf{S}$ has only one eigenvalue $s_1 = s_2 = s_3 = s$,

$$\mathbf{S} = s\mathbf{I}, \tag{2.240}$$

in this case $\mathcal{V}$ is the only characteristic space of $\mathbf{S}$.

The relation (2.238) is the *spectral decomposition* of $\mathbf{S}$.

If $\mathcal{M}_i$ $(i = 1, ..., k \leq n)$ are the characteristic spaces of $\mathcal{S}$, then each vector $\mathbf{v}$ can be written in the form

$$\mathbf{v} = \sum_{i=1}^{k} \mathbf{v}_i, \quad \mathbf{v}_i \in \mathcal{M}_i, \tag{2.241}$$

and

$$\mathcal{S} = \mathcal{M}_1 \oplus ... \oplus \mathcal{M}_k. \tag{2.242}$$

The matrix $[\mathbf{S}]$ of the components of $\mathbf{S}$ with respect to the basis $\{\mathbf{g}_1, ..., \mathbf{g}_n\}$ of eigenvectors has the form

$$[\mathbf{S}] = \begin{bmatrix} s_1 & 0 & . & 0 \\ 0 & s_2 & . & 0 \\ . & . & . & . \\ 0 & 0 & . & s_n \end{bmatrix}. \tag{2.243}$$

For $n = 3$, from the spectral theorem it follows that if $\mathbf{S} \in$ Sym, $\eta(\mathbf{S})$ is completely characterized by the spectrum $\sigma(\mathbf{S})$; it is easy to prove that

$$I_1(\mathbf{S}) = s_1 + s_2 + s_3, \tag{2.244}$$

$$I_2(\mathbf{S}) = s_1 s_2 + s_1 s_3 + s_2 s_3, \tag{2.245}$$

$$I_2(\mathbf{S}) = s_1 s_2 s_3. \tag{2.246}$$

If $\mathbf{S} \in$ Sym, the multiplicity of an eigenvalue s coincides with the multiplicity of s as root of the characteristic equation $\det(\mathbf{S} - s\mathbf{I}) = 0$. The following proposition follows directly from the previous remark.

Proposition 39. *For $n = 3$ consider $\mathbf{S}, \mathbf{T} \in$ Sym such that $\eta(\mathbf{S}) = \eta(\mathbf{T})$, then $\mathbf{S}$ and $\mathbf{T}$ have the same spectrum, $\sigma(\mathbf{S}) = \sigma(\mathbf{T})$.*

We point out that this result hold only if $\mathbf{S}$ and $\mathbf{T}$ are symmetric. Let us consider the tensors

$$\mathbf{S} = \mathbf{I} + \mathbf{e}_3 \otimes \mathbf{e}_3, \quad \mathbf{T} = \mathbf{I} + \mathbf{e}_3 \otimes \mathbf{e}_3 + \mathbf{e}_1 \otimes \mathbf{e}_2,$$

we have $I_1(\mathbf{S}) = I_1(\mathbf{T}) = 4$, $I_2(\mathbf{S}) = I_2(\mathbf{T}) = 5$, $I_3(\mathbf{S}) = I_3(\mathbf{T}) = 2$, but $\sigma(\mathbf{S}) = \{1, 1, 2\}$ and $\sigma(\mathbf{T}) = \{1, 2\}$.

For non-symmetric tensors, eigenvectors corresponding to distinct eigenvalues are not necessarily orthogonal. For example, the spectrum of tensor $\mathbf{A} \in$ Lin

$$\mathbf{A} = \mathbf{e}_1 \otimes \mathbf{e}_1 + 2\mathbf{e}_2 \otimes \mathbf{e}_2 + 3\mathbf{e}_3 \otimes \mathbf{e}_3 + \mathbf{e}_1 \otimes \mathbf{e}_2, \tag{2.247}$$

is $\sigma(\mathbf{A}) = \{1, 2, 3\}$, the corresponding eigenvectors are $\mathbf{e}_1$, $\frac{1}{\sqrt{2}}(\mathbf{e}_1 + \mathbf{e}_2)$ e $\mathbf{e}_3$, and we have that $\mathbf{e}_1 \cdot (\mathbf{e}_1 + \mathbf{e}_2) = 1$.

Exercise 8. *For $n = 3$, $\mathbf{D} \in$ Sym, $\mathbf{Q} \in$ Orth, show that $\sigma(\mathbf{D}) = \sigma(\mathbf{Q}\mathbf{D}\mathbf{Q}^T)$.*

Solution. It is sufficient to remark that $\mathbf{Q}\mathbf{D}\mathbf{Q}^T$ is symmetric and that $\eta(\mathbf{D}) = \eta(\mathbf{Q}\mathbf{D}\mathbf{Q}^T)$, the desired result follows from proposition 39.

Consider $n = 3$. A skew-symmetric tensor $\mathbf{W} \neq 0$ has only one eigenvalue equal to zero a 0, the remaining roots of the characteristic polynomial are two conjugate imaginary numbers. The principal invariants of $\mathbf{W}$are

$$I_1(\mathbf{W}) = 0, \quad I_2(\mathbf{W}) = W_{23}^2 + W_{12}^2 + W_{13}^2, \quad I_3(\mathbf{W}) = 0, \tag{2.248}$$

where W_{ij} are the components of $\mathbf{W}$ with respect to the orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. The characteristic polynomial of $\mathbf{W}$ is therefore

$$a^3 + I_2(\mathbf{W})a = 0, \tag{2.249}$$

since $I_2(\mathbf{W}) > 0$, $\mathbf{W}$ has the only zero eigenvalue $a = 0$. The eigenvector corresponding to the null eigenvalue is the axial vector $\mathbf{w}$ of $\mathbf{W}$. In fact, from the relation $\mathbf{Wa} = \mathbf{w} \wedge \mathbf{a}$, $\mathbf{a} \in \mathcal{V}$, we have that $\mathbf{w}$ is the only eigenvector of $\mathbf{W}$ and $\mathbf{Ww} = \mathbf{0}$.

For $\mathbf{v} \in \mathcal{V}$ we have

$$\mathbf{W}^2\mathbf{v} = \mathbf{w} \wedge (\mathbf{w} \wedge \mathbf{v}) = (\mathbf{v} \wedge \mathbf{w}) \wedge \mathbf{w} = (\mathbf{v} \cdot \mathbf{w})\mathbf{w} - ||\mathbf{w}||^2\mathbf{v}, \tag{2.250}$$

and $\mathbf{W}^2$ has the expression

$$\mathbf{W}^2 = \mathbf{w} \otimes \mathbf{w} - ||\mathbf{w}||^2\mathbf{I}. \tag{2.251}$$

Tensor $\mathbf{W}^2$ turns out to be symmetric and its spectral decomposition is

$$\mathbf{W}^2 = -||\mathbf{w}||^2(\mathbf{I} - \frac{\mathbf{w}}{||\mathbf{w}||} \otimes \frac{\mathbf{w}}{||\mathbf{w}||}). \tag{2.252}$$

Tensor $\mathbf{W}^3$ is instead skew-symmetric and

$$\mathbf{W}^3 = -||\mathbf{w}||^2\mathbf{W}. \tag{2.253}$$

A dyad $\mathbf{a} \otimes \mathbf{b}$ with $\mathbf{a}, \mathbf{b} \in \mathcal{V}$ has a null eigenvalue with multiplicity $n-1$ and the corresponding characteristic space is the subspace orthogonal to $\mathbf{b}$. Dyad $\mathbf{a} \otimes \mathbf{b}$ has also the eigenvalue $\mathbf{a} \cdot \mathbf{b}$ whose characteristic space is Span($\mathbf{a}$). In general, these characteristic spaces are not orthogonal, they are orthogonal if and only if $\mathbf{a} = \alpha\mathbf{b}$.

Exercise 9. *Consider $n = 3$. Determine spectrum, characteristic spaces and spectral decomposition of the following symmetric tensors*

$$\mathbf{A} = \alpha\mathbf{I} + \beta\mathbf{m} \otimes \mathbf{m}, \quad \mathbf{B} = \mathbf{m} \otimes \mathbf{n} + \mathbf{n} \otimes \mathbf{m}, \tag{2.254}$$

con $\alpha, \beta \in \mathbb{R}$, $\mathbf{m}, \mathbf{n} \in \mathcal{V}$, $\mathbf{m} \cdot \mathbf{n} = 0$, $||\mathbf{m}|| = ||\mathbf{n}|| = 1$.

Solution. Putting $\mathbf{q} = \mathbf{m} \wedge \mathbf{n}$ we have

$$\mathbf{An} = \alpha\mathbf{n}, \ \mathbf{Am} = (\alpha + \beta)\mathbf{m}, \ \mathbf{Aq} = \alpha\mathbf{q}, \tag{2.255}$$

then $\sigma(\mathbf{A}) = \{\alpha, \alpha, \alpha+\beta\}$, the characteristic spaces are Span$(\mathbf{n}, \mathbf{q})$ and Span$(\mathbf{m})$ and the spectral decomposition of $\mathbf{A}$ is

$$\begin{aligned}\mathbf{A} &= \alpha(\mathbf{n} \otimes \mathbf{n} + \mathbf{q} \otimes \mathbf{q}) + (\alpha + \beta)\mathbf{m} \otimes \mathbf{m} \\ &= (\alpha + \beta)\mathbf{m} \otimes \mathbf{m} + \alpha(\mathbf{I} - \mathbf{m} \otimes \mathbf{m}).\end{aligned} \tag{2.256}$$

Moreover, we have

$$\mathbf{B}\mathbf{q} = \mathbf{0},\ \mathbf{B}(\mathbf{m} + \mathbf{n}) = \mathbf{m} + \mathbf{n},\ \mathbf{B}(\mathbf{m} - \mathbf{n}) = \mathbf{n} - \mathbf{m}, \tag{2.257}$$

then $\sigma(\mathbf{B}) = \{-1, 0, 1\}$, the characteristic spaces are Span$(\mathbf{m} - \mathbf{n})$, Span$(\mathbf{q})$ e Span$(\mathbf{m} + \mathbf{n})$ and the spectral decomposition of $\mathbf{B}$ is

$$\mathbf{B} = \frac{\mathbf{m} + \mathbf{n}}{\sqrt{2}} \otimes \frac{\mathbf{m} + \mathbf{n}}{\sqrt{(2)}} - \frac{\mathbf{m} - \mathbf{n}}{\sqrt{2}} \otimes \frac{\mathbf{m} - \mathbf{n}}{\sqrt{(2)}}.$$

Exercise 10. *Put $n = 3$. A tensor $\mathbf{P}$ is called orthogonal projection if $\mathbf{P} \in$ Sym and $\mathbf{P}^2 = \mathbf{P}$.*

(a) *For $\mathbf{n} \in \mathcal{V}$, $||\mathbf{n}|| = 1$, show that the following tensors are orthogonal projections,*

$$\mathbf{0},\ \mathbf{I},\ \mathbf{n} \otimes \mathbf{n},\ \mathbf{I} - \mathbf{n} \otimes \mathbf{n}. \tag{2.258}$$

(b) *Show that if $\mathbf{P}$ is an orthogonal projection, then $\mathbf{P}$ admits one of the representations (2.258).*

Solution. (a) It is easy to verify that tensors in (2.258) are orthogonal projections.

(b) If $\mathbf{P}$ is an orthogonal projection, let us calculate its eigenvalues. Let λ be an eigenvalue and $\mathbf{v}$ the corresponding eigenvector,

$$\lambda\mathbf{v} = \mathbf{P}\mathbf{v} = \mathbf{P}^2\mathbf{v} = \lambda^2\mathbf{v},$$

from which we have $\lambda = 0$ or $\lambda = 1$. the following four cases are possible,

- $\sigma(\mathbf{P}) = \{0, 0, 0\}$, $\mathbf{P} = \mathbf{0}$,
- $\sigma(\mathbf{P}) = \{1, 1, 1\}$, $\mathbf{P} = \mathbf{I}$,
- $\sigma(\mathbf{P}) = \{0, 1, 1\}$, $\mathbf{P} = \mathbf{I} - \mathbf{n} \otimes \mathbf{n}$, with $\mathbf{n}$ eigenvector corresponding to the 0 eigenvalue.
- $\sigma(\mathbf{P}) = \{0, 0, 1\}$, $\mathbf{P} = \mathbf{n} \otimes \mathbf{n}$, with $\mathbf{n}$ eigenvector corresponding to the eigenvalue 1.

Exercise 11. *Put $n = 3$. Given $\mathbf{R} \in Orth^+$ let $\mathbf{e} \in \mathcal{V}$ be such that $\mathbf{R}\mathbf{e} = \mathbf{e}$. For $\mathbf{W}$ the skew-symmetric tensor associated to $\mathbf{e}$, prove that $\mathbf{R}$ has the following representation*

$$\mathbf{R} = \mathbf{I} + \sin\theta\mathbf{W} + (1 - \cos\theta)\mathbf{W}^2, \tag{2.259}$$

with $\theta \in (-\pi, \pi)$.

The representation formula (2.259) proves that each rotation is completely characterized by an axis and an angle.

Some properties of the product of tensors are collected in the following propositions.

Proposition 40. *Put* $n = 3$. *A tensor* $\mathbf{A} \in$ *Lin commutes with each tensor* $\mathbf{W} \in$ *Skw if and only if* $\mathbf{A} = \omega \mathbf{I}$.

Proof. Let us assume that

$$\mathbf{AW} = \mathbf{WA} \quad \text{for each } \mathbf{W} \in \text{Skw}. \tag{2.260}$$

For $\mathbf{w} \in \mathcal{V}$ fixed, let $\mathbf{W}$ the the skew-symmetric tensor associated to $\mathbf{w}$, we have

$$\mathbf{W}(\mathbf{Aw}) = \mathbf{A}(\mathbf{Ww}) = \mathbf{A}(\mathbf{w} \wedge \mathbf{w}) = \mathbf{0},$$

then $\mathbf{Aw}$ belongs to the characteristic space of the null eigenvalue of $\mathbf{W}$,

$$\mathbf{Aw} = \lambda \mathbf{w}, \quad \text{with } \lambda = \widetilde{\lambda}(\mathbf{w}) \in \mathbb{R}. \tag{2.261}$$

Let $\mathbf{w}_1, \mathbf{w}_2$ be two linearly independent vectors in $\mathcal{V}$, in view of the linearity of $\mathbf{A}$ we have

$$\begin{aligned} &\widetilde{\lambda}(\mathbf{w}_1)\mathbf{w}_1 + \widetilde{\lambda}(\mathbf{w}_2)\mathbf{w}_2 = \mathbf{Aw}_1 + \mathbf{Aw}_2 = \\ &\mathbf{A}(\mathbf{w}_1 + \mathbf{w}_2) = \widetilde{\lambda}(\mathbf{w}_1 + \mathbf{w}_2)(\mathbf{w}_1 + \mathbf{w}_2), \end{aligned} \tag{2.262}$$

from which we get

$$[\widetilde{\lambda}(\mathbf{w}_1) - \widetilde{\lambda}(\mathbf{w}_1 + \mathbf{w}_2)]\mathbf{w}_1 + [\widetilde{\lambda}(\mathbf{w}_2) - \widetilde{\lambda}(\mathbf{w}_1 + \mathbf{w}_2)]\mathbf{w}_2 = \mathbf{0}, \tag{2.263}$$

and then $\widetilde{\lambda}(\mathbf{w}_1) = \widetilde{\lambda}(\mathbf{w}_2) = \omega$. □

Proposition 41. *A tensor* $\mathbf{S} \in$ *Sym commutes with each tensor* $\mathbf{Q} \in$ *Orth*$^+$ *if and only if* $\mathbf{S} = \omega \mathbf{I}$.

Proof. Let us assume that

$$\mathbf{SQ} = \mathbf{QS} \quad \text{for each } \mathbf{Q} \in \text{Orth}^+$$

and that $\mathbf{S}$ has two distinct eigenvalues ω and λ and let $\mathbf{u}$ and $\mathbf{v}$ be the corresponding orthogonal eigenvectors (of norm 1). Let $\{\mathbf{u}, \mathbf{v}, \mathbf{f}_3, ..., \mathbf{f}_n\}$ be an orthonormal basis of $\mathcal{V}$, put

$$\overline{\mathbf{Q}} = \mathbf{v} \otimes \mathbf{u} - \mathbf{v} \otimes \mathbf{u} + \sum_{i=3}^{n} \mathbf{f}_i \otimes \mathbf{f}_i,$$

we have that $\overline{\mathbf{Q}}\mathbf{u} = \mathbf{v}$ and $\overline{\mathbf{Q}} \in$Orth$^+$ since it transforms the orthonormal basis $\{\mathbf{u}, \mathbf{v}, \mathbf{f}_3, ..., \mathbf{f}_n\}$ into the orthonormal basis $\{\mathbf{v}, -\mathbf{u}, \mathbf{f}_3, ..., \mathbf{f}_n\}$ and $\det \overline{\mathbf{Q}} = 1$. Then, we have

$$\overline{\mathbf{Q}}\mathbf{S}\mathbf{u} = \omega \overline{\mathbf{Q}}\mathbf{u} = \omega \mathbf{v}, \quad \overline{\mathbf{Q}}\mathbf{S}\mathbf{u} = \mathbf{S}\overline{\mathbf{Q}}\mathbf{u} = \mathbf{S}\mathbf{v} = \lambda \mathbf{v},$$

from which we get $(\omega - \lambda)\mathbf{v} = \mathbf{0}$ and then $\omega = \lambda$. □

Theorem 11. *(Commutation theorem). Consider* $\mathbf{S}$, $\mathbf{A} \in Lin$ *such that* $\mathbf{SA} = \mathbf{AS}$. *Then, if* $\mathbf{v} \in \mathcal{V}$ *belongs to a characteristic space of* $\mathbf{S}$, $\mathbf{Av}$ *belongs to the same characteristic space. Vice versa if* $\mathbf{A}$ *leaves each characteristic space of a symmetric tensor* $\mathbf{S} \in Sym$ *invariant, then* $\mathbf{S}$ *and* $\mathbf{A}$ *commute.*

Proof. If $\mathbf{S}$ and $\mathbf{A}$ commute, let $\mathbf{v}$ be an eigenvector of $\mathbf{S}$ corresponding to the eigenvalue ω, $\mathbf{Sv} = \omega\mathbf{v}$, then

$$\mathbf{S}(\mathbf{Av}) = \mathbf{ASv} = \omega\mathbf{Av},$$

that is $\mathbf{Av}$ belongs to the characteristic space of $\mathbf{S}$ corresponding to ω.

Vice versa let $\mathcal{M}_i$, $i = 1, ..., k$ be the characteristic spaces of $\mathbf{S}$. Each $\mathbf{v} \in \mathcal{V}$ has the representation (2.241) and $\mathbf{Av}_i \in \mathcal{M}_i$ for each i, then

$$\mathbf{SAv}_i = \omega_i\mathbf{Av}_i = \mathbf{A}(\omega_i\mathbf{v}_i) = \mathbf{ASv}_i,$$

from which it follows that $\mathbf{SAv} = \mathbf{ASv}$. □

Tensors $\mathbf{A}$ and $\mathbf{B} \in$ Sym are called *coaxial* if there is at least one orthonormal basis of common eigenvectors.

Proposition 42. *The tensors* $\mathbf{A}, \mathbf{B} \in Sym$ *commute if and only if are coaxial.*

Proof. For the sake of simplicity, take $n = 3$. Let us assume that $\mathbf{A}$ and $\mathbf{B}$ are coaxial, let $\{\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3\}$ be a common basis of eigenvectors

$$\mathbf{A} = \sum_{i=1}^{3} a_i\ \mathbf{g}_i \otimes \mathbf{g}_i, \quad \mathbf{B} = \sum_{i=1}^{3} b_i\ \mathbf{g}_i \otimes \mathbf{g}_i. \tag{2.264}$$

Then, in view of (2.34) we have $\mathbf{AB} = \mathbf{BA}$. Vice versa, let us assume that $\mathbf{AB} = \mathbf{BA}$, we can consider the following cases.

(i) If $\mathbf{A} = a\mathbf{I}$, the coaxiality is evident.

(ii) If $\mathbf{A}$ has three distinct eigenvalues a_1, a_2, a_3, let us consider its spectral decomposition

$$\mathbf{A} = \sum_{i=1}^{3} a_i\ \mathbf{g}_i \otimes \mathbf{g}_i, \tag{2.265}$$

and put

$$\mathbf{B} = \sum_{i,j=1}^{3} B_{ij}\ \mathbf{g}_i \otimes \mathbf{g}_j. \tag{2.266}$$

Then, we have

$$\mathbf{0} = \mathbf{AB} - \mathbf{BA} = \sum_{\substack{i,j=1 \\ i \neq j}}^{3} a_i B_{ij}\ (\mathbf{g}_i \otimes \mathbf{g}_j - \mathbf{g}_j \otimes \mathbf{g}_i); \tag{2.267}$$

since the skew-symmetric tensors $\mathbf{g}_i \otimes \mathbf{g}_j - \mathbf{g}_j \otimes \mathbf{g}_i$ are linearly independent, from (2.267) we get

$$(a_1 - a_2)B_{12} = 0, \tag{2.268}$$

$$(a_1 - a_3)B_{13} = 0, \tag{2.269}$$

$$(a_2 - a_3)B_{23} = 0, \tag{2.270}$$

from which it follows that $B_{12} = B_{13} = B_{23} = 0$, thus, $\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3$ are eigenvectors of $\mathbf{B}$ as well.

(iii) Now let us consider the case in which $\mathbf{A}$ has two distinct eigenvalues $a_1 \neq a_2 = a_3$,

$$\mathbf{A} = a_1 \mathbf{g}_1 \otimes \mathbf{g}_1 + a_2(\mathbf{I} - \mathbf{g}_1 \otimes \mathbf{g}_1), \quad \mathbf{B} = \sum_{i,j=1}^{3} B_{ij}\ \mathbf{g}_i \otimes \mathbf{g}_j, \tag{2.271}$$

with $\mathbf{g}_2$ and $\mathbf{g}_3$ belonging to $\mathrm{Span}(\mathbf{g}_1)^{\perp}$. Considering once again the relation $\mathbf{0} = \mathbf{AB} - \mathbf{BA}$ we get that $B_{12} = B_{13} = 0$, then $\mathbf{B}\mathbf{g}_1 = b_1 \mathbf{g}_1$. Let $\mathbf{f}_2$ and $\mathbf{f}_3$ be the remaining two eigenvectors of $\mathbf{B}$ such that $\{\mathbf{g}_1, \mathbf{f}_2, \mathbf{f}_3\}$ is an orthonormal basis of $\mathcal{V}$, we conclude that $\{\mathbf{g}_1, \mathbf{f}_2, \mathbf{f}_3\}$ is a basis of eigenvectors for both tensors $\mathbf{B}$ and $\mathbf{A}$. □

Proposition 43. *The following properties hold.*

(1) Given $\mathbf{A} \in$ *Sym*$^-$ *(Sym*$^+$*), if there is* $\mathbf{u} \in \mathcal{V}$ *such that* $\mathbf{u} \cdot \mathbf{A}\mathbf{u} = 0$*, then* $\mathbf{A}\mathbf{u} = \mathbf{0}$.

(2) Consider $\mathbf{A}, \mathbf{B} \in$ *Sym. If* $\mathbf{A} \cdot \mathbf{B} \geq 0$ *for each* $\mathbf{B} \in$ *Sym*$^+$ *(Sym*$^-$*) then* $\mathbf{A} \in$ *Sym*$^+$ *(Sym*$^-$*).*

(3) Consider $\mathbf{A} \in$ *Sym*$^+$*. For each* $\mathbf{B} \in$ *Sym*$^+$ *(Sym*$^-$*) we have* $\mathbf{A} \cdot \mathbf{B} \geq 0$ *(*≤ 0*).*

(4) Consider $\mathbf{A} \in$ *Sym*$^+$*,* $\mathbf{B} \in$ *Sym*$^+$ *(Sym*$^-$*). If* $\mathbf{A} \cdot \mathbf{B} = 0$ *then* $\mathbf{AB} = \mathbf{BA} = \mathbf{0}$.

Proof. (1) Let $\mathbf{A} = \sum_{i=1}^{n} a_i \mathbf{q}_i \otimes \mathbf{q}_i$ with $a_i \leq 0$ ($a_i \geq 0$) be the spectral decomposition of $\mathbf{A}$. We have

$$\mathbf{A}\mathbf{u} = \sum_{i=1}^{n} a_i (\mathbf{q}_i \cdot \mathbf{u}) \mathbf{q}_i, \tag{2.272}$$

therefore

$$0 = \mathbf{u} \cdot \mathbf{A}\mathbf{u} = \sum_{i=1}^{n} a_i (\mathbf{q}_i \cdot \mathbf{u})^2 \tag{2.273}$$

if and only if

$$a_i (\mathbf{q}_i \cdot \mathbf{u})^2 = 0, \quad i = 1, ..., n \tag{2.274}$$

since a_i are non positive (non negative). (2.274) is verified if and only if

$$a_i (\mathbf{q}_i \cdot \mathbf{u}) = 0, \quad i = 1, ..., n, \tag{2.275}$$

which is equivalent to the condition $\mathbf{Au} = \mathbf{0}$.

(2) Let us assume that $\mathbf{A}$ does not belong to Sym^+ (Sym^-), then $\mathbf{A}$ has an eigenvalue $\lambda < 0$ ($\lambda > 0$), $\mathbf{Av} = \lambda\mathbf{v}$. The tensor $\overline{\mathbf{B}} = -\lambda\mathbf{v}\otimes\mathbf{v}$ belongs to Sym^+ (Sym^-) and

$$0 \leq \mathbf{A}\cdot\overline{\mathbf{B}} = -\lambda tr(\mathbf{Av}\otimes\mathbf{v}) = -\lambda^2 < 0.$$

(3) Let

$$\mathbf{A} = \sum_{i=1}^{n} a_i\mathbf{q}_i\otimes\mathbf{q}_i \quad \text{with } a_i \geq 0 \tag{2.276}$$

be the spectral decomposition of $\mathbf{A}$. Moreover, let

$$\mathbf{B} = \sum_{i=1}^{n} b_i\mathbf{p}_i\otimes\mathbf{p}_i \quad \text{con } b_i \geq 0 \;\; (b_i \leq 0) \tag{2.277}$$

be the spectral decomposition of $\mathbf{B}$. We have

$$\mathbf{AB} = \sum_{i,j=1}^{n} a_i b_j(\mathbf{q}_i\otimes\mathbf{q}_i)(\mathbf{p}_j\otimes\mathbf{p}_j) = \sum_{i,j=1}^{n} a_i b_j(\mathbf{q}_i\cdot\mathbf{p}_j)(\mathbf{q}_i\otimes\mathbf{p}_j), \tag{2.278}$$

and

$$\mathbf{A}\cdot\mathbf{B} = \sum_{i,j=1}^{n} a_i b_j(\mathbf{q}_i\cdot\mathbf{p}_j)^2 \geq 0 \;\; (\leq 0). \tag{2.279}$$

(4) Let (2.276) and (2.277) be the spectral decompositions of $\mathbf{A}$ and $\mathbf{B}$, respectively. In view of (2.279) the condition $\mathbf{A}\cdot\mathbf{B} = 0$ is equivalent to the conditions

$$a_i b_j(\mathbf{q}_i\cdot\mathbf{p}_j) = 0, \quad i, j = 1, ..., n, \tag{2.280}$$

therefore, from (2.278) we get that $\mathbf{AB} = \mathbf{0}$, in a similar way, we prove that $\mathbf{BA} = \mathbf{0}$. □

For $\mathbf{A} \in \mathrm{Sym}$, the function

$$q_A(\mathbf{u}) = \frac{\mathbf{u}\cdot\mathbf{Au}}{\mathbf{u}\cdot\mathbf{u}}, \quad \mathbf{u} \in \mathcal{V}, \mathbf{u} \neq \mathbf{0}, \tag{2.281}$$

is called *Rayleigh ratio.*

Proposition 44. *Given* $\mathbf{A} \in$ *Sym, its the Rayleigh ratio* q_A *satisfies the inequalities*

$$a_1 \leq q_A(\mathbf{u}) \leq a_n, \quad \textit{for each } \mathbf{u} \in \mathcal{V}, \mathbf{u} \neq \mathbf{0}, \tag{2.282}$$

where a_1 *and* a_n *are the minimum and the maximum eigenvalue of* $\mathbf{A}$*.*

Proof. Let

$$\mathbf{A} = \sum_{i=1}^{3} a_i \; \mathbf{g}_i\otimes\mathbf{g}_i, \tag{2.283}$$

be the spectral decomposition of $\mathbf{A}$, with

$$a_1 \leq a_2 \leq ... \leq a_n. \tag{2.284}$$

For $\mathbf{u} \in \mathcal{V}$, we have

$$\mathbf{u} = \sum_{i=1}^{3} \alpha_i \mathbf{g}_i, \text{ with } \alpha_i = \mathbf{u} \cdot \mathbf{g}_i, \tag{2.285}$$

and

$$\mathbf{Au} = \sum_{i=1}^{3} \alpha_i a_i \mathbf{g}_i. \tag{2.286}$$

Putting (2.285) and (2.286) in (2.281), we get

$$q_A(\mathbf{u}) = \frac{\sum\limits_{i=1}^{3} \alpha_i^2 a_i}{\sum\limits_{i=1}^{3} \alpha_i^2}, \tag{2.287}$$

From (2.287), taking (2.284) into account, we get (2.282). □

Exercise 12. *Given the tensor* $\mathbf{A}$*, with* $\mathbf{A} \neq \alpha\mathbf{I}$ *for each real number* α*, compute the orthogonal projection onto Span*$(\mathbf{I}, \mathbf{A})$*.*

Solution. From the minimum norm theorem, it follows that given $\mathbf{U} \in \text{Sym}$, there is a unique $\widehat{\mathbf{U}} \in \text{Span}(\mathbf{I}, \mathbf{A})$ such that $(\mathbf{U} - \widehat{\mathbf{U}}) \cdot \mathbf{V} \leq 0$ for each $\mathbf{V} \in \text{Span}(\mathbf{I}, \mathbf{A})$. Let us start by determining an orthonormal basis of $\text{Span}(\mathbf{I}, \mathbf{A})$. For

$$\mathbf{A}_0 = \mathbf{A} - \frac{tr\mathbf{A}}{n}\mathbf{I}, \tag{2.288}$$

the deviatoric part of $\mathbf{A}$, the tensors

$$\mathbf{A}_1 = \frac{\mathbf{I}}{\sqrt{n}}, \tag{2.289}$$

and

$$\mathbf{A}_2 = \frac{\mathbf{A}_0}{||\mathbf{A}_0||}, \tag{2.290}$$

with

$$||\mathbf{A}_0|| = \sqrt{||\mathbf{A}_0||^2 - \frac{(tr\mathbf{A})^2}{n}}, \tag{2.291}$$

are orthonormal and then are a basis of $\text{Span}(\mathbf{I}, \mathbf{A})$. Thus,

$$\begin{aligned} \widehat{\mathbf{U}} = P_{\text{Span}(\mathbf{I},\mathbf{A})}(\mathbf{U}) &= (\mathbf{U} \cdot \mathbf{A}_1)\mathbf{A}_1 + (\mathbf{U} \cdot \mathbf{A}_2)\mathbf{A}_2 \\ &= \frac{tr\mathbf{U}}{n}\mathbf{I} + (\mathbf{U} \cdot \mathbf{A}_2)\mathbf{A}_2 \end{aligned} \tag{2.292}$$

2.14 Square root theorem, polar decomposition theorem

Theorem 12. *(Square root theorem) For every* $\mathbf{C} \in$ *Psym, there exists a unique tensor* $\mathbf{U} \in$ *Psym such that*

$$\mathbf{U}^2 = \mathbf{C}. \tag{2.293}$$

We write $\sqrt{\mathbf{C}}$ *for* $\mathbf{U}$.

Proof. Let

$$\mathbf{C} = \sum_{i=1}^{n} c_i \mathbf{g}_i \otimes \mathbf{g}_i \tag{2.294}$$

be the spectral decomposition of $\mathbf{C}$, with $c_i > 0$. Let us define $\mathbf{U} \in$ Psym in the following way

$$\mathbf{U} = \sum_{i=1}^{n} \sqrt{c_i} \mathbf{g}_i \otimes \mathbf{g}_i, \tag{2.295}$$

(2.293) is trivially verified. To prove the uniqueness of $\mathbf{U}$ let us assume that there exist $\mathbf{U}_1, \mathbf{U}_2 \in$ Psym such that $\mathbf{U}_1^2 = \mathbf{U}_2^2 = \mathbf{C}$. For each $i = 1, ..., n$ we have

$$\mathbf{0} = (\mathbf{U}_1^2 - c_i \mathbf{I})\mathbf{g}_i = (\mathbf{U}_1 + \sqrt{c_i}\mathbf{I})(\mathbf{U}_1 - \sqrt{c_i}\mathbf{I})\mathbf{g}_i, \tag{2.296}$$

putting $\mathbf{v}_i = (\mathbf{U}_1 - \sqrt{c_i}\mathbf{I})\mathbf{g}_i$, from (2.296) we have that $\mathbf{U}_1 \mathbf{v}_i = -\sqrt{c_i}\mathbf{v}_i$, therefore $\mathbf{v}_i = \mathbf{0}$ since the eigenvalues of $\mathbf{U}_1$ are positive. The, we get that $\mathbf{U}_1 \mathbf{g}_i = \sqrt{c_i}\mathbf{g}_i$; analogously we prove that $\mathbf{U}_2 \mathbf{g}_i = \sqrt{c_i}\mathbf{g}_i$, then $\mathbf{U}_1 \mathbf{g}_i = \mathbf{U}_2 \mathbf{g}_i$ for each $i = 1, ..., n$. □

Exercise 13. *Put* $n = 3$. *For each* $\mathbf{E} \in$*Sym, determine the projection* $P_{Sym^-}(\mathbf{E})$ *of* $\mathbf{E}$ *onto Sym*$^-$.

We remark that Sym$^-$ and Sym$^+$ are convex closed cones of Sym [3], then we can apply the minimum norm theorem. Thus, given $\mathbf{E} \in$Sym we have to find $\mathbf{A} \in$Sym$^-$ such that

$$(\mathbf{E} - \mathbf{A}) \cdot (\mathbf{T} - \mathbf{A}) \leq 0 \text{ for each } \mathbf{T} \in \mathrm{Sym}^-. \tag{2.297}$$

Let

$$\mathbf{E} = \sum_{i=1}^{3} e_i \mathbf{g}_i \otimes \mathbf{g}_i \tag{2.298}$$

be the spectral decomposition of $\mathbf{E}$, we have

$$\mathbf{E}^2 = \sum_{i=1}^{3} e_i^2 \mathbf{g}_i \otimes \mathbf{g}_i, \quad \sqrt{\mathbf{E}^2} = \sum_{i=1}^{3} |e_i| \mathbf{g}_i \otimes \mathbf{g}_i. \tag{2.299}$$

[3] Sym$^+$ (Sym$^-$) is closed in Sym since it is the inverse image of the closed set $\{\mathbf{x} \in \mathbb{R}^3 : 0 \leq x_1 \leq x_2 \leq x_3\}$ ($\{\mathbf{x} \in \mathbb{R}^3 : x_1 \leq x_2 \leq x_3 \leq 0\}$) of $\mathbb{R}^3$ under the continuous function that assigns to each tensor its spectrum.

Tensors

$$\mathbf{A} = \frac{\mathbf{E} - \sqrt{\mathbf{E}^2}}{2}, \quad \mathbf{B} = \frac{\mathbf{E} + \sqrt{\mathbf{E}^2}}{2}, \tag{2.300}$$

belong respectively to Sym^- and Sym^+ and are orthogonal,

$$\mathbf{A} \cdot \mathbf{B} = \frac{1}{4} tr(\mathbf{E}^2 - \sqrt{\mathbf{E}^2}\mathbf{E} + \mathbf{E}\sqrt{\mathbf{E}^2} - \mathbf{E}^2) = 0. \tag{2.301}$$

It is an easy matter to prove that $P_{\text{Sym}^-}(\mathbf{E}) = \mathbf{A}$, with $\mathbf{A}$ defined in (2.300), in fact

$$(\mathbf{E} - \mathbf{A}) \cdot (\mathbf{T} - \mathbf{A}) = \mathbf{B} \cdot \mathbf{T} - \mathbf{B} \cdot \mathbf{A} = \mathbf{B} \cdot \mathbf{T} \leq 0 \tag{2.302}$$

for each $\mathbf{T} \in \text{Sym}^-$, because $\mathbf{B} \in \text{Sym}^+$. Finally, we have

$$\mathbf{E} = \mathbf{A} + \mathbf{B} = P_{\text{Sym}^-}(\mathbf{E}) + P_{\text{Sym}^+}(\mathbf{E}). \tag{2.303}$$

The projection P_{Sym^-} is not linear. For $\mathbf{E}_1 = -\mathbf{g}_3 \otimes \mathbf{g}_3 + 2(\mathbf{I} - \mathbf{g}_3 \otimes \mathbf{g}_3)$, $\mathbf{E}_2 = -\mathbf{g}_1 \otimes \mathbf{g}_1 - 3\mathbf{g}_2 \otimes \mathbf{g}_2$, we have $P_{\text{Sym}^-}(\mathbf{E}_1) = -\mathbf{g}_3 \otimes \mathbf{g}_3$, $P_{\text{Sym}^-}(\mathbf{E}_2) = -\mathbf{g}_1 \otimes \mathbf{g}_1 - 3\mathbf{g}_2 \otimes \mathbf{g}_2$ and $P_{\text{Sym}^-}(\mathbf{E}_1 + \mathbf{E}_2) = -(\mathbf{I} - \mathbf{g}_1 \otimes \mathbf{g}_1)$.

Theorem 13. *(Polar decomposition theorem). For each $\mathbf{F} \in Lin^+$, there exist $\mathbf{U}, \mathbf{V} \in Psym$ and $\mathbf{R} \in Orth^+$ such that*

$$\mathbf{F} = \mathbf{R}\mathbf{U} = \mathbf{V}\mathbf{R}. \tag{2.304}$$

Moreover, each of these decomposition is unique; in fact,

$$\mathbf{U} = \sqrt{\mathbf{F}^T\mathbf{F}}, \quad \mathbf{V} = \sqrt{\mathbf{F}\mathbf{F}^T}. \tag{2.305}$$

$\mathbf{F} = \mathbf{R}\mathbf{U}$ *is called right polar decomposition of* $\mathbf{F}$*,* $\mathbf{F} = \mathbf{V}\mathbf{R}$ *is called left polar decomposition of* $\mathbf{F}$*.*

Proof. First of all let us prove that $\mathbf{F}^T\mathbf{F}$ and $\mathbf{F}\mathbf{F}^T$ belong to Psym; we have $\mathbf{F}^T\mathbf{F}, \mathbf{F}\mathbf{F}^T \in \text{Sym}$, moreover

$$\mathbf{v} \cdot \mathbf{F}^T\mathbf{F}\mathbf{v} = \mathbf{F}\mathbf{v} \cdot \mathbf{F}\mathbf{v} \geq 0 \text{ for each } \mathbf{v} \in \mathcal{V},$$

and $\mathbf{F}\mathbf{v} \cdot \mathbf{F}\mathbf{v} = 0$ if and only if $\mathbf{F}\mathbf{v} = \mathbf{0}$ or, if and only if $\mathbf{v} = \mathbf{0}$ since $\mathbf{F}$ is invertible. Analogously we prove that $\mathbf{F}\mathbf{F}^T \in \text{Psym}$. Therefore, $\mathbf{U}$ and $\mathbf{V}$ in (2.305) are well defined. Let us prove the existence of the polar decomposition. For $\mathbf{U} = \sqrt{\mathbf{F}^T\mathbf{F}} \in \text{Psym}$ put $\mathbf{R} = \mathbf{F}\mathbf{U}^{-1}$, we have to prove that $\mathbf{R} \in \text{Orth}^+$. Since $\det \mathbf{F} > 0$ and $\det \mathbf{U} > 0$ we have that $\det \mathbf{R} > 0$, moreover

$$\mathbf{R}^T\mathbf{R} = \mathbf{U}^{-1}\mathbf{F}^T\mathbf{F}\mathbf{U}^{-1} = \mathbf{I}, \tag{2.306}$$

and

$$\mathbf{R}\mathbf{R}^T = \mathbf{F}\mathbf{U}^{-1}\mathbf{U}^{-1}\mathbf{F}^T = \mathbf{F}(\mathbf{U}^2)^{-1}\mathbf{F}^T = \mathbf{F}\mathbf{F}^{-1}\mathbf{F}^{-T}\mathbf{F}^T = \mathbf{I}. \tag{2.307}$$

Finally, let us define $\mathbf{V} = \mathbf{R}\mathbf{U}\mathbf{R}^T$, $\mathbf{V} \in$ Psym; then

$$\mathbf{v} \cdot \mathbf{V}\mathbf{v} = \mathbf{v} \cdot (\mathbf{R}\mathbf{U}\mathbf{R}^T)\mathbf{v} = \mathbf{R}^T\mathbf{v} \cdot \mathbf{U}\mathbf{R}^{\mathbf{T}}\mathbf{v} > 0 \text{ per ogni } \mathbf{v} \neq \mathbf{0},$$

since $\mathbf{U} \in$ Psym. Moreover, we have

$$\mathbf{V}\mathbf{R} = \mathbf{R}\mathbf{U}\mathbf{R}^T\mathbf{R} = \mathbf{R}\mathbf{U} = \mathbf{F}. \tag{2.308}$$

Now, we have to prove the uniqueness of the polar decomposition. Let $\mathbf{F} = \mathbf{R}\mathbf{U}$ the right polar decomposition of $\mathbf{F}$, since $\mathbf{R} \in \mathrm{Orth}^+$ we have

$$\mathbf{F}^T\mathbf{F} = \mathbf{U}^2. \tag{2.309}$$

In virtue of the square root theorem, there is a unique $\mathbf{U} \in$ Psym satisfying (2.309), thus $\mathbf{U} = \sqrt{\mathbf{F}^T\mathbf{F}}$ and $\mathbf{U}$ is unique. Since $\mathbf{R} = \mathbf{F}\mathbf{U}^{-1}$, also $\mathbf{R}$ is unique. Analogously, we prove the uniqueness of the decomposition $\mathbf{F} = \mathbf{V}\mathbf{R}$ and this concludes the proof. □

We point out that in general $\mathbf{U}$ and $\mathbf{V}$ do not coincide, on the other hand, if $\mathbf{F} \in \mathrm{Lin}^+ \cap \mathrm{Sym}$, then from the relations $\mathbf{F}^2 = \mathbf{U}^2 = \mathbf{V}^2$ it follows that $\mathbf{U} = \mathbf{V}$ and then $\mathbf{F} = \mathbf{R}\mathbf{U} = \mathbf{U}\mathbf{R}$.

The tensors $\mathbf{U}$ and $\mathbf{V}$ of the polar decomposition of $\mathbf{F} \in \mathrm{Lin}^+$, are linked by the relation $\mathbf{V} = \mathbf{R}\mathbf{U}\mathbf{R}^T$, and have the same spectrum.

Exercise 14. *Put $n = 3$. Given $\mathbf{e}_1, \mathbf{e}_2$ orthogonal vectors with norm 1, for $\mathbf{e}_3 = \mathbf{e}_1 \wedge \mathbf{e}_2$ let $\mathbf{W} = \mathbf{e}_2 \otimes \mathbf{e}_1 - \mathbf{e}_1 \otimes \mathbf{e}_2$ be the skew-symmetric tensor having $\mathbf{e}_3$ as axial vector. Compute the right polar decomposition of the tensor $\mathbf{F} = \mathbf{I} + \mathbf{W}$.*

Solution. We point out that

$$\mathbf{F}\mathbf{e}_1 = \mathbf{e}_1 + \mathbf{e}_2,$$
$$\mathbf{F}\mathbf{e}_2 = \mathbf{e}_2 - \mathbf{e}_1,$$
$$\mathbf{F}\mathbf{e}_3 = \mathbf{e}_3,$$

and that

$$\det \mathbf{F} = \mathbf{F}\mathbf{e}_1 \cdot (\mathbf{F}\mathbf{e}_2 \wedge \mathbf{F}\mathbf{e}_3) = (\mathbf{e}_1 + \mathbf{e}_2) \cdot (\mathbf{e}_1 + \mathbf{e}_2) = 2,$$

therefore $\mathbf{F} \in \mathrm{Lin}^+$. Now, let us determine the spectral decomposition of $\mathbf{F}^T\mathbf{F} = \mathbf{I} - \mathbf{W}^2$. From relations

$$\mathbf{F}^T\mathbf{F}\mathbf{e}_1 = 2\mathbf{e}_1, \quad \mathbf{F}^T\mathbf{F}\mathbf{e}_2 = 2\mathbf{e}_2, \quad \mathbf{F}^T\mathbf{F}\mathbf{e}_3 = \mathbf{e}_3,$$

we get that

$$\mathbf{F}^T\mathbf{F} = \mathbf{e}_3 \otimes \mathbf{e}_3 + 2(\mathbf{I} - \mathbf{e}_3 \otimes \mathbf{e}_3),$$

from which we get the expression of $\mathbf{U}$,

$$\mathbf{U} = \mathbf{e}_3 \otimes \mathbf{e}_3 + \sqrt{2}(\mathbf{I} - \mathbf{e}_3 \otimes \mathbf{e}_3),$$

and finally

$$\mathbf{R} = \mathbf{F}\mathbf{U}^{-1} = \mathbf{e}_3 \otimes \mathbf{e}_3 + \frac{1}{\sqrt{2}}(\mathbf{I} - \mathbf{e}_3 \otimes \mathbf{e}_3) + \frac{1}{\sqrt{2}}\mathbf{W}.$$

Exercise 15. *For $n = 3$ let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be an orthonormal basis of $\mathcal{V}$. Given the tensor $\mathbf{F} = \mathbf{I} + \gamma \mathbf{e}_2 \otimes \mathbf{e}_1$, $\gamma \in \mathbb{R}$, compute its polar decomposition $\mathbf{F} = \mathbf{VR}$.*

Solution. If $\gamma = 0$, then $\mathbf{F} = \mathbf{I}$ e $\mathbf{V} = \mathbf{R} = \mathbf{I}$. The, assume that $\gamma \neq 0$. The eigenvalues of the tensor $\mathbf{F}\mathbf{F}^T = \mathbf{I} + \gamma(\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1) + \gamma^2 \mathbf{e}_2 \otimes \mathbf{e}_2$ are

$$\varphi_1 = \frac{2 + \gamma^2 - \sqrt{2\gamma^2 + \gamma^4}}{2}, \tag{2.310}$$

$$\varphi_2 = \frac{2 + \gamma^2 + \sqrt{2\gamma^2 + \gamma^4}}{2}, \tag{2.311}$$

$$\varphi_3 = 1, \tag{2.312}$$

and the corresponding eigenvectors are

$$\mathbf{q}_1 = \frac{1}{n_1}\mathbf{e}_1 + \frac{\varphi_1 - 1}{\gamma n_1}\mathbf{e}_2, \tag{2.313}$$

$$\mathbf{q}_2 = \frac{1}{n_2}\mathbf{e}_1 + \frac{\varphi_2 - 1}{\gamma n_2}\mathbf{e}_2, \tag{2.314}$$

and

$$\mathbf{q}_3 = \mathbf{e}_3 \tag{2.315}$$

with

$$n_i = \sqrt{1 + \left(\frac{\varphi_i - 1}{\gamma}\right)^2}, \quad i = 1, 2. \tag{2.316}$$

Then, we have

$$\begin{aligned}\mathbf{V} &= \sqrt{\varphi_1}\mathbf{q}_1 \otimes \mathbf{q}_1 + \sqrt{\varphi_2}\mathbf{q}_2 \otimes \mathbf{q}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3 \\ &= \frac{2}{\sqrt{4+\gamma^2}}\mathbf{e}_1 \otimes \mathbf{e}_1 + \frac{\gamma}{\sqrt{4+\gamma^2}}(\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1) \\ &\quad + \frac{2+\gamma^2}{\sqrt{4+\gamma^2}}\mathbf{e}_2 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3,\end{aligned} \tag{2.317}$$

and

$$\begin{aligned}\mathbf{R} = \mathbf{V}^{-1}\mathbf{F} &= [(\varphi_1)^{-2}\mathbf{q}_1 \otimes \mathbf{q}_1 + (\varphi_2)^{-2}\mathbf{q}_2 \otimes \mathbf{q}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3]\,[\mathbf{I} + \gamma\mathbf{e}_2 \otimes \mathbf{e}_1] \\ &= \frac{2}{\sqrt{4+\gamma^2}}\mathbf{e}_1 \otimes \mathbf{e}_1 + \frac{2}{\sqrt{4+\gamma^2}}\mathbf{e}_2 \otimes \mathbf{e}_2 \\ &\quad + \frac{\gamma}{\sqrt{4+\gamma^2}}(-\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1) + \mathbf{e}_3 \otimes \mathbf{e}_3.\end{aligned} \tag{2.318}$$

Exercise 16. *For $n = 3$ let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be an orthonormal basis of $\mathcal{V}$. Compute the polar decomposition $\mathbf{F} = \mathbf{VR}$ of the tensor*

$$\mathbf{F} = \delta\mathbf{e}_1 \otimes \mathbf{e}_1 + \alpha(\mathbf{I} - \mathbf{e}_1 \otimes \mathbf{e}_1) + \alpha\gamma\mathbf{e}_2 \otimes \mathbf{e}_1, \tag{2.319}$$

con $\alpha, \gamma, \delta \in \mathbb{R}, \alpha > 0, \delta > 0$.

Exercise 17. *For* $\mathbf{F} \in Lin^+$, *let* $\mathbf{F} = \mathbf{R}\mathbf{U} = \mathbf{V}\mathbf{R}$, $\mathbf{U}, \mathbf{V} \in Psym$ *and* $\mathbf{R} \in Orth^+$ *be the polar decompositions of* $\mathbf{F}$. *Prove that* $\mathbf{R}$ *is the rotation closest to* $\mathbf{F}$ *in the sense that*

$$||\mathbf{F} - \mathbf{R}|| < ||\mathbf{F} - \mathbf{Q}||, \quad \textit{for each } \mathbf{Q} \in Orth^+, \ \mathbf{Q} \neq \mathbf{R}. \tag{2.320}$$

Solution. For each $\mathbf{Q} \in$Orth$^+$, $\mathbf{Q} \neq \mathbf{R}$, we have

$$||\mathbf{F} - \mathbf{Q}||^2 = tr(\mathbf{F}^T\mathbf{F} - \mathbf{Q}^T\mathbf{F} - \mathbf{F}^T\mathbf{Q} + \mathbf{I}) = ||\mathbf{F}||^2 + 3 - 2\mathbf{Q} \cdot \mathbf{F}, \tag{2.321}$$

$$||\mathbf{F} - \mathbf{R}||^2 = ||\mathbf{F}||^2 + 3 - 2\mathbf{U} \cdot \mathbf{I}, \tag{2.322}$$

from which we get

$$||\mathbf{F} - \mathbf{Q}||^2 - ||\mathbf{F} - \mathbf{R}||^2 = 2(\mathbf{U} \cdot \mathbf{I} - \mathbf{Q} \cdot \mathbf{F}). \tag{2.323}$$

Since

$$\mathbf{Q} \cdot \mathbf{F} = tr(\mathbf{F}^T\mathbf{Q}) = tr(\mathbf{U}\mathbf{R}^T\mathbf{Q}) = \mathbf{Q}_0 \cdot \mathbf{U}, \tag{2.324}$$

with $\mathbf{Q}_0 = \mathbf{R}^T\mathbf{Q} \in$Orth$^+, \mathbf{Q}_0 \neq \mathbf{I}$, we have

$$||\mathbf{F} - \mathbf{Q}||^2 - ||\mathbf{F} - \mathbf{R}||^2 = 2\mathbf{U} \cdot (\mathbf{I} - \mathbf{Q}_0), \tag{2.325}$$

moreover

$$\begin{aligned} tr[(\mathbf{Q}_0 - \mathbf{I})^T\mathbf{U}(\mathbf{Q}_0 - \mathbf{I})] = \mathbf{U} \cdot [(\mathbf{Q}_0 - \mathbf{I})(\mathbf{Q}_0 - \mathbf{I})^T] = \\ \mathbf{U} \cdot (2\mathbf{I} - \mathbf{Q}_0 - \mathbf{Q}_0^T) = 2\mathbf{U} \cdot (\mathbf{I} - \mathbf{Q}_0). \end{aligned} \tag{2.326}$$

Since $(\mathbf{Q}_0 - \mathbf{I})^T\mathbf{U}(\mathbf{Q}_0 - \mathbf{I}) \in$ Psym, from (2.326) it folloes that $2\mathbf{U} \cdot (\mathbf{I} - \mathbf{Q}_0) > 0$.

2.15 The Cayley-Hamilton theorem

Theorem 14. *(Cayley-Hamilton theorem). Put* $n = 3$. *For* $\mathbf{A} \in Lin$ *we have*

$$\mathbf{A}^3 - I_1(\mathbf{A})\mathbf{A}^2 + I_2(\mathbf{A})\mathbf{A} - I_3(\mathbf{A})\mathbf{I} = \mathbf{0}. \tag{2.327}$$

Proof. We prove the theorem by assuming that $\mathbf{A}$ has three linearly independent eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$

$$\mathbf{A}\mathbf{v}_i = a_i\mathbf{v}_i, \quad i = 1, 2, 3. \tag{2.328}$$

Since from (2.328) it follows that

$$\mathbf{A}^j\mathbf{v}_i = a_i^j\mathbf{v}_i, \quad i = 1, 2, 3, \ j = 1, 2, 3, \tag{2.329}$$

in view of (2.236) we have

$$\begin{aligned} [\mathbf{A}^3 - I_1(\mathbf{A})\mathbf{A}^2 + I_2(\mathbf{A})\mathbf{A} - I_3(\mathbf{A})\mathbf{I}]\mathbf{v}_i = \\ a_i^3\mathbf{v}_i - I_1(\mathbf{A})a_i^2\mathbf{v}_i + I_2(\mathbf{A})a_i\mathbf{v}_i - I_3(\mathbf{A})\mathbf{v}_i = \\ [a_i^3 - I_1(\mathbf{A})a_i^2 + I_2(\mathbf{A})a_i - I_3(\mathbf{A})]\mathbf{v}_i = 0, \quad i = 1, 2, 3. \end{aligned} \tag{2.330}$$

Since $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly independent, (2.327) follows directly from (2.330). □

Exercise 18. *Put $n = 3$. Given $\mathbf{A} \in Lin$, prove that*

$$\det \mathbf{A} = \frac{1}{6}[(tr\mathbf{A})^3 - 3(tr\mathbf{A})tr(\mathbf{A}^2) + 2tr(\mathbf{A}^3)]. \tag{2.331}$$

Solution. Consider that from the Cayley-Hamilton it follows that

$$tr(\mathbf{A}^3) = I_1(\mathbf{A})tr(\mathbf{A}^2) - I_2(\mathbf{A})I_1(\mathbf{A}) + 3I_3(\mathbf{A}).$$

Exercise 19. *Put $n = 3$. Show that $\mathbf{A} \in Lin$ is invertible, then*

$$\mathbf{A}^{-1} = \frac{1}{I_3(\mathbf{A})}[\mathbf{A}^2 - I_1(\mathbf{A})\mathbf{A} + I_2(\mathbf{A})\mathbf{I}], \tag{2.332}$$

and deduce that for each integer k, $\mathbf{A}^k$ can be expressed as linear combination of $\mathbf{I}, \mathbf{A}, \mathbf{A}^2$ with coefficients that depend on the principal invariants of $\mathbf{A}$.

Solution. For the Cayley-Hamilton theorem we have

$$\mathbf{A}^3 - I_1(\mathbf{A})\mathbf{A}^2 + I_2(\mathbf{A})\mathbf{A} = I_3(\mathbf{A})\mathbf{A}\mathbf{A}^{-1},$$

multiplying by $\mathbf{A}^{-1}$ we get the desired expression.

Exercise 20. *Put $n = 3$. Let $\mathbf{A} \in Lin$ be invertible. Show that*

(a) $I_1(\mathbf{A}^{-1}) = \frac{I_2(\mathbf{A})}{I_3(\mathbf{A})}$,

(b) $I_2(\mathbf{A}^{-1}) = \frac{I_1(\mathbf{A})}{I_3(\mathbf{A})}$,

(c) $I_3(\mathbf{A}^{-1}) = \frac{1}{I_3(\mathbf{A})}$.

2.16 The generalized eigenvalue problem

Given the tensors $\mathbf{A} \in$ Sym, $\mathbf{B} \in$ Psym, we say that a (real) number a is a *generalized eigenvalue* of $(\mathbf{A}, \mathbf{B})$ if there exists $\mathbf{u} \in \mathcal{V}$ $\mathbf{u} \neq \mathbf{0}$, such that

$$\mathbf{A}\mathbf{u} = a\mathbf{B}\mathbf{u}; \tag{2.333}$$

$\mathbf{u}$ is called *generalized eigenvector* and problem (2.333) is called *generalized eigenvalue problem*.

Vectors $\mathbf{u}_1, ..., \mathbf{u}_n$ are $\mathbf{B}$-*orthonormal* if

$$\mathbf{u}_i \cdot \mathbf{B}\mathbf{u}_j = \delta_{ij} = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases} \tag{2.334}$$

Proposition 45. *Given the tensors $\mathbf{A} \in Sym$, $\mathbf{B} \in Psym$, there exists a basis of $\mathcal{V}$ constituted by $\mathbf{B}$-orthonormal generalized eigenvectors $\mathbf{u}_1, ..., \mathbf{u}_n$ corresponding to the generalized eigenvalues $a_1, ..., a_n$ of $(\mathbf{A}, \mathbf{B})$,*

$$\mathbf{A}\mathbf{u}_i = a_i\mathbf{B}\mathbf{u}_i, \quad i = 1, ..., n. \tag{2.335}$$

Proof. In virtue of the square root theorem there exists $\mathbf{U} \in$Psym such that $\mathbf{U}^2 = \mathbf{B}$. The problem (2.333) can be rewritten as

$$\widehat{\mathbf{A}}\mathbf{v} = a\mathbf{v}, \tag{2.336}$$

where we put

$$\widehat{\mathbf{A}} = \mathbf{U}^{-1}\mathbf{A}\mathbf{U}^{-1}, \quad \mathbf{v} = \mathbf{U}\mathbf{u}. \tag{2.337}$$

From the spectral theorem, it follows that there exist an orthonormal basis of $\mathcal{V}$ formed by eigenvectors $\mathbf{v}_1, ..., \mathbf{v}_n$ of (2.336) and n real numbers $a_1, ..., a_n$ such that

$$\widehat{\mathbf{A}}\mathbf{v}_i = a_i\mathbf{v}_i, \ i = 1, ..., n. \tag{2.338}$$

It is easy to verify that the vectors $\mathbf{u}_i = \mathbf{U}^{-1}\mathbf{v}_i$, $i = 1, ..., n$ are the eigenvectors of the generalized problem (2.333) corresponding to the eigenvalues $a_1, ..., a_n$ and satisfy (2.336). Moreover, the relations

$$\delta_{ij} = \mathbf{v}_i \cdot \mathbf{v}_j = \mathbf{U}\mathbf{u}_i \cdot \mathbf{U}\mathbf{u}_j = \mathbf{u}_i \cdot \mathbf{U}^2\mathbf{u}_j = \mathbf{u}_i \cdot \mathbf{B}\mathbf{u}_j, \tag{2.339}$$

allow to conclude that vectors $\mathbf{u}_1, ..., \mathbf{u}_n$ are $\mathbf{B}-$orthonormal. □

a is a generalized eigenvalue if and only if $\mathbf{A} - a\mathbf{B}$ is not invertible and then a is a real root of the characteristic polynomial,

$$p(a) = \det(\mathbf{A} - a\mathbf{B}). \tag{2.340}$$

For each vector $\mathbf{u} \in \mathcal{V}$ $\mathbf{u} \neq \mathbf{0}$, the ratio

$$q(\mathbf{u}) = \frac{\mathbf{u} \cdot \mathbf{A}\mathbf{u}}{\mathbf{u} \cdot \mathbf{B}\mathbf{u}}, \tag{2.341}$$

is the *Rayleigh quotient* of the generalized problem (2.333).

Proposition 46. *Let*

$$a_1 \leq a_2 \leq ... \leq a_n \tag{2.342}$$

be the generalized eigenvalues of the problem (2.333). The Rayleigh quotient (2.341) satisfies the inequalities

$$a_1 \leq q(\mathbf{u}) \leq a_n, \quad \textit{per ogni}\ \mathbf{u} \in \mathcal{V} \quad \mathbf{u} \neq \mathbf{0}. \tag{2.343}$$

Proof. Let $\mathbf{u}_1, ..., \mathbf{u}_n$ be the $\mathbf{B}-$orthonormal eigenvectors of (2.333). For each $\mathbf{u} \in \mathcal{V}$ $\mathbf{u} \neq \mathbf{0}$ we have

$$\mathbf{u} = \alpha_1\mathbf{u}_1 + ... + \alpha_n\mathbf{u}_n, \tag{2.344}$$

and

$$q(\mathbf{u}) = \frac{(\alpha_1\mathbf{u}_1 + ... + \alpha_n\mathbf{u}_n) \cdot (\alpha_1 a_1\mathbf{B}\mathbf{u}_1 + ... + \alpha_n a_n\mathbf{B}\mathbf{u}_n)}{(\alpha_1\mathbf{u}_1 + ... + \alpha_n\mathbf{u}_n) \cdot (\alpha_1\mathbf{B}\mathbf{u}_1 + ... + \alpha_n\mathbf{B}\mathbf{u}_n)}$$

$$= \frac{\alpha_1^2 a_1 + ... + \alpha_n^2 a_n}{\alpha_1^2 + ... + \alpha_n^2}. \tag{2.345}$$

From (2.345), taking (2.342) into account, we get

$$a_1 = \frac{(\alpha_1^2 + ... + \alpha_\cdot^2)a_1}{\alpha_1^2 + ... + \alpha_\cdot^2} \leq q(\mathbf{u}) \leq \frac{(\alpha_1^2 + ... + \alpha_\cdot^2)a_n}{\alpha_1^2 + ... + \alpha_\cdot^2} = a_n. \tag{2.346}$$

□

2.17 Third and fourth-order tensors

A *third-order tensor* $\boldsymbol{F}$ can be considered a linear mapping from Lin to $\mathcal{V}$ or a linear mapping from $\mathcal{V}$ to Lin. In particular, given $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V}$, $\mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w}$ denotes the third-order tensor defined by

$$\mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w}[\mathbf{H}] = (\mathbf{v} \otimes \mathbf{w} \cdot \mathbf{H})\mathbf{u}, \quad \mathbf{H} \in \text{Lin}, \tag{2.347}$$

$$\mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w}[\mathbf{h}] = (\mathbf{w} \cdot \mathbf{h})\mathbf{u} \otimes \mathbf{v}, \quad \mathbf{h} \in \mathcal{V}. \tag{2.348}$$

Example 11. *Put $n = 3$. The mapping $\boldsymbol{E}$ from $\mathcal{V}$ to Skw that assigns to each vector $\mathbf{w}$ the skew-symmetric tensor $\mathbf{W}$ having $\mathbf{w}$ as axial vector*

$$\boldsymbol{E}(\mathbf{w}) = \mathbf{W} \quad \textit{with} \quad \mathbf{W}\mathbf{v} = \mathbf{w} \wedge \mathbf{v}, \quad \textit{for each } \mathbf{v} \in \mathcal{V}, \tag{2.349}$$

is a third-order tensor.

Let us put

$$\varepsilon_{ijk} = \begin{cases} 1 & \textit{se } ijk \textit{ is an even permutation of } 1,2,3, \\ -1 & \textit{se } ijk \textit{ is an odd permutation of } 1,2,3, \\ 0 & \textit{otherwise} \end{cases}, \quad i,j,k = 1,2,3. \tag{2.350}$$

Given an orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ of $\mathcal{V}$ and denoted by w_1, w_2, w_3 the components of $\mathbf{w}$, in view of (2.155) the components of $\mathbf{W}$ are

$$W_{ij} = -\sum_{k=1}^{3} \varepsilon_{ijk} w_k, \tag{2.351}$$

therefore

$$\mathbf{W} = \sum_{i,j=1}^{3} W_{ij} \mathbf{e}_i \otimes \mathbf{e}_j = -\sum_{i,j,k=1}^{3} \varepsilon_{ijk} (\mathbf{w} \cdot \mathbf{e}_k) \mathbf{e}_i \otimes \mathbf{e}_j$$

$$= -\sum_{i,j,k=1}^{3} \varepsilon_{ijk} (\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k)\mathbf{w} \tag{2.352}$$

and, finally,

$$\boldsymbol{E} = -\sum_{i,j,k=1}^{3} \varepsilon_{ijk}(\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k). \tag{2.353}$$

Analogously, the function $\boldsymbol{F}$ *from Skw to* $\mathcal{V}$ *that assigns to each skew-symmetric tensor* $\mathbf{W}$ *the corresponding axial vector* $\mathbf{w}$*,*

$$\boldsymbol{F}(\mathbf{W}) = \mathbf{w} \quad \text{with } \mathbf{W}\mathbf{v} = \mathbf{w} \wedge \mathbf{v}, \quad \text{for each } \mathbf{v} \in \mathcal{V}, \tag{2.354}$$

is a third-order tensor. Given the components of $\mathbf{W}$*, the components of* $\mathbf{w}$ *are*

$$w_i = -\frac{1}{2}\sum_{j,k=1}^{3} \varepsilon_{ijk} W_{jk}, \tag{2.355}$$

then

$$\mathbf{w} = \sum_{i=1}^{3} w_i \mathbf{e}_i = -\frac{1}{2}\sum_{i,j,k=1}^{3} \varepsilon_{ijk} W_{jk}\mathbf{e}_i = -\frac{1}{2}\sum_{i,j,k=1}^{3} \varepsilon_{ijk}(\mathbf{e}_j \cdot \mathbf{W}\mathbf{e}_k)\mathbf{e}_i$$

$$= -\frac{1}{2}\sum_{i,j,k=1}^{3} \varepsilon_{ijk}[(\mathbf{e}_j \otimes \mathbf{e}_k)\cdot \mathbf{W}]\mathbf{e}_i = -\frac{1}{2}\sum_{i,j,k=1}^{3} \varepsilon_{ijk}(\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k)\mathbf{W}, \tag{2.356}$$

and, finally,

$$\boldsymbol{F} = -\frac{1}{2}\sum_{i,j,k=1}^{3} \varepsilon_{ijk}(\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k). \tag{2.357}$$

Also the function $\boldsymbol{E}_2$ *from Lin to* $\mathcal{V}$ *that assigns to each tensor* $\mathbf{A}$ *the axial vector of the skew-symmetric part* $(\mathbf{A} - \mathbf{A}^T)/2$ *of* $\mathbf{A}$ *is a third-order tensor.*

A *fourth-order tensor* $\mathbb{A}$ is a linear mapping from Lin to Lin. Let us denote by $\mathbb{I}$ the fourth-order identity defined by $\mathbb{I}[\mathbf{H}] = \mathbf{H}$ for each $\mathbf{H} \in \text{Lin}$. The tensor product $\mathbf{A} \otimes \mathbf{B}$ of the second-order tensors $\mathbf{A}$ and $\mathbf{B}$ is the fourth order tensor defined by

$$\mathbf{A} \otimes \mathbf{B}[\mathbf{H}] = (\mathbf{B} \cdot \mathbf{H})\mathbf{A}, \quad \mathbf{H} \in \text{Lin}. \tag{2.358}$$

From tensors $\mathbf{A}$ and $\mathbf{B}$ it is possible to define the fourth-order tensor $\mathbf{A} \boxtimes \mathbf{B}$,

$$\mathbf{A} \boxtimes \mathbf{B}[\mathbf{H}] = \mathbf{A}\mathbf{H}\mathbf{B}^T, \quad \mathbf{H} \in \text{Lin}.$$

Let us denote by $\mathbb{L}\text{in}$ the vector space of all fourth-order tensors. Let us consider the orthonormal basis $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ of $\mathcal{V}$, the *components* of the fourth-order tensor $\mathbb{A}$ are

$$\mathbb{A}_{ijkl} = (\mathbf{e}_i \otimes \mathbf{e}_j) \cdot \mathbb{A}[\mathbf{e}_k \otimes \mathbf{e}_l], \quad i,j,k,l = 1, ..., n. \tag{2.359}$$

Putting $\mathbf{K} = \mathbb{A}[\mathbf{H}]$, from (2.46) we get that

$$K_{ij} = \sum_{k,l=}^{n} \mathbb{A}_{ijkl} H_{kl}.$$

From the linear independence of vectors $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ in $\mathcal{V}$ it follows that the elements $(\mathbf{e}_i \otimes \mathbf{e}_j) \otimes (\mathbf{e}_k \otimes \mathbf{e}_l)$, $i, j, k, l = 1, ..., n$ in $\mathbb{L}$in are linearly independent, moreover, for each $\mathbb{A} \in \mathbb{L}$in the following representation holds,

$$\mathbb{A} = \sum_{i,j,k,k=1}^{n} \mathbb{A}_{ijkl} (\mathbf{e}_i \otimes \mathbf{e}_j) \otimes (\mathbf{e}_k \otimes \mathbf{e}_l). \tag{2.360}$$

In fact,

$$\mathbb{A}[\mathbf{U}] = \sum_{i,j=1}^{n} (\mathbb{A}[\mathbf{U}])_{ij} (\mathbf{e}_i \otimes \mathbf{e}_j) = \sum_{i,j,k,k=1}^{n} \mathbb{A}_{ijkl} U_{kl} \mathbf{e}_i \otimes \mathbf{e}_j$$

$$= \sum_{i,j,k,k=1}^{n} \mathbb{A}_{ijkl} (\mathbf{e}_k \cdot \mathbf{U}\mathbf{e}_l) \mathbf{e}_i \otimes \mathbf{e}_j = \sum_{i,j,k,k=1}^{n} \mathbb{A}_{ijkl} ((\mathbf{e}_k \otimes \mathbf{e}_l) \cdot \mathbf{U}) \mathbf{e}_i \otimes \mathbf{e}_j$$

$$= \sum_{i,j,k,k=1}^{n} \mathbb{A}_{ijkl} (\mathbf{e}_i \otimes \mathbf{e}_j) \otimes (\mathbf{e}_k \otimes \mathbf{e}_l)[\mathbf{U}]. \tag{2.361}$$

Thus, the fourth-order tensors $\{(\mathbf{e}_i \otimes \mathbf{e}_j) \otimes (\mathbf{e}_k \otimes \mathbf{e}_l)\}_{i,j,k,l=1,...,n}$ are a basis of the vector space $\mathbb{L}$in, which has dimension n^4 .

$\mathbb{L}$in is a normed space, with the natural norm

$$||\mathbb{A}||_N = \sup_{\mathbf{H} \in \mathrm{Lin},\ \mathbf{H} \neq \mathbf{0}} \frac{||\mathbb{A}[\mathbf{H}]||}{||\mathbf{H}||}. \tag{2.362}$$

The *transpose* of $\mathbb{A}$ is the unique fourth-order tensor $\mathbb{A}^T$ such that

$$\mathbb{A}^T[\mathbf{H}] \cdot \mathbf{K} = \mathbb{A}[\mathbf{K}] \cdot \mathbf{H}, \quad \text{for each } \mathbf{H}, \mathbf{K} \in \mathrm{Lin}. \tag{2.363}$$

Exercise 21. *Compute the transpose of the fourt order tensors* $\mathbf{A} \otimes \mathbf{B}$ *and* $\mathbf{A} \boxtimes \mathbf{B}$.

Solution. Given $\mathbf{H}, \in \mathbf{K} \in$Lin, we have

$$\mathbf{H} \cdot \mathbf{A} \otimes \mathbf{B}[\mathbf{K}] = \mathbf{H} \cdot (\mathbf{B} \cdot \mathbf{K})\mathbf{A} = (\mathbf{B} \cdot \mathbf{K})(\mathbf{A} \cdot \mathbf{H}) =$$

$$\mathbf{K} \cdot \mathbf{B} \otimes \mathbf{A}[\mathbf{H}],$$

then $(\mathbf{A} \otimes \mathbf{B})^T = \mathbf{B} \otimes \mathbf{A}$. Moreover, we have that

$$\mathbf{H} \cdot \mathbf{A} \boxtimes \mathbf{B}[\mathbf{K}] = \mathbf{H} \cdot \mathbf{A}\mathbf{K}\mathbf{B}^T = tr(\mathbf{B}\mathbf{K}^T\mathbf{A}^T\mathbf{H}) =$$

$$tr(\mathbf{K}^T\mathbf{A}^T\mathbf{H}\mathbf{B}) = \mathbf{K} \cdot \mathbf{A}^T \boxtimes \mathbf{B}^T[\mathbf{H}],$$

and then, $(\mathbf{A} \boxtimes \mathbf{B})^T = \mathbf{A}^T \boxtimes \mathbf{B}^T$.

The fourth-order tensor $\mathbb{A}$ has the *major symmetry* (or is symmetric) if $\mathbb{A}^T = \mathbb{A}$. In terms of indices, this means that $\mathbb{A}_{ijkl} = \mathbb{A}_{klij}$.

The symmetry in the first couple of indices ($\mathbb{A}_{ijkl} = \mathbb{A}_{jikl}$) means that $\mathbb{A}$ has values in Sym

$$\mathbb{A}[\mathbf{H}]^T = \mathbb{A}[\mathbf{H}], \quad \mathbf{H} \in \text{Lin}, \tag{2.364}$$

and the symmetry in the second couple of indices ($\mathbb{A}_{ijkl} = \mathbb{A}_{ijlk}$) means that

$$\mathbb{A}[\mathbf{H}^T] = \mathbb{A}[\mathbf{H}], \quad \mathbf{H} \in \text{Lin}, \tag{2.365}$$

or, equivalently, that $\mathbb{A}$ is zero on Skw,

$$\mathbb{A}[\mathbf{W}] = \mathbf{0}, \quad \mathbf{W} \in \text{Skw}.$$

We say that $\mathbb{A}$ has the *minor symmetry* if has the symmetries in both first and second couples of indices.

Exercise 22. *Prove that the components of the identity tensor $\mathbb{I}$, defined by $\mathbb{I}[\mathbf{A}] = \mathbf{A}$, for each $\mathbf{A} \in$Lin, are*

$$\mathbb{I}_{ijhk} = \begin{cases} 1 & se\ i = h\ \ e\ j = k, \\ 0 & \quad otherwise, \end{cases} \quad , \tag{2.366}$$

and that $\mathbf{I} \boxtimes \mathbf{I} = \mathbb{I}$, where $\mathbf{I}$ is the identity of Lin.

The mapping $\mathbb{T}$: Lin→ Lin such that $\mathbb{T}[\mathbf{A}] = \mathbf{A}^T$, for each $\mathbf{A}$ ∈Lin is a fourth-order tensor and the fourth-order tensors $\mathbb{S}$ and $\mathbb{W}$ defined by

$$\mathbb{S}[\mathbf{A}] = \frac{\mathbf{A} + \mathbf{A}^T}{2}, \quad \mathbb{W}[\mathbf{A}] = \frac{\mathbf{A} - \mathbf{A}^T}{2}, \text{ for each } \mathbf{A} \in \text{Lin}, \tag{2.367}$$

are called *symmetrizer* and *skew-symmetrizer.*

Exercise 23. *Compute the components of the fourth-order tensors $\mathbf{A} \otimes \mathbf{B}$, $\mathbf{A} \boxtimes \mathbf{B}$, $\mathbb{T}$, $\mathbb{S}$ e $\mathbb{W}$.*

Exercise 24. *Prove that $\mathbf{A} \boxtimes \mathbf{B}$ is symmetric if and only if $\mathbf{A}$ and $\mathbf{B}$ are symmetric and that $\mathbf{A} \otimes \mathbf{B}$ is symmetric if and only if $\mathbf{A} = \alpha\mathbf{B}$, $\alpha \in \mathbb{R}$.*

Exercise 25. *Given $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D} \in$Lin and $\mathbb{A} \in \mathbb{L}in$, prove the following composition rules*

$$(\mathbf{A} \boxtimes \mathbf{B})(\mathbf{C} \boxtimes \mathbf{D}) = \mathbf{AC} \boxtimes \mathbf{BD}, \tag{2.368}$$

$$(\mathbf{A} \otimes \mathbf{B})\mathbb{A} = \mathbf{A} \otimes \mathbb{A}^T[\mathbf{B}], \tag{2.369}$$

$$\mathbb{A}(\mathbf{A} \otimes \mathbf{B}) = \mathbb{A}[\mathbf{A}] \otimes \mathbf{B}. \tag{2.370}$$

Given $\mathbf{Q}$ $\in$Orth, let us consider the fourth-order tensor $\mathbb{Q}$ =$\mathbf{Q} \boxtimes \mathbf{Q}$, we have

$$\mathbb{Q}[\mathbf{A}] \cdot \mathbb{Q}[\mathbf{B}] = \mathbf{Q}\mathbf{A}\mathbf{Q}^T \cdot \mathbf{Q}\mathbf{B}\mathbf{Q}^T$$

$$= tr(\mathbf{Q}\mathbf{A}^T\mathbf{Q}^T\mathbf{Q}\mathbf{B}\mathbf{Q}^T) = tr(\mathbf{A}^T\mathbf{B}) = \mathbf{A} \cdot \mathbf{B}, \quad \text{per ogni } \mathbf{A}, \mathbf{B} \in \text{Lin}, \quad (2.371)$$

then $\mathbb{Q} \in \mathbb{L}$in is an isometry (cfr. (1.67)).

Let us denote by $\mathbb{S}$ym the vector space of all fourth-order tensors defined from Sym into Sym and denote by $\mathbb{I}_{Sym}$ the restriction of $\mathbb{I}$ to Sym. Herein after we shall limit ourselves to consider tensors $\mathbb{A} \in \mathbb{S}$ym

A symmetric tensor $\mathbb{A} \in \mathbb{S}$ym is called *positive definite* if

$$\mathbf{A} \cdot \mathbb{A}[\mathbf{A}] > 0, \quad \text{for each } \mathbf{A} \in \text{Sym}, \ \mathbf{A} \neq \mathbf{0}. \quad (2.372)$$

$\mathbb{A}$ is called *invertible* if it is bijective. Tensor $\mathbb{A}^{-1}$ such that $\mathbb{A}^{-1}\mathbb{A} = \mathbb{A}\mathbb{A}^{-1} = \mathbb{I}_{Sym}$ is the *inverse* of $\mathbb{A}$.

Let $\mathbb{C} \in \mathbb{S}$ym be a symmetric fourth-order tensor. The *spectral problem* relative to $\mathbb{C}$ consists in determining the pairs $(\gamma, \mathbf{C})$ with $\gamma \in \mathbb{R}$, $\mathbf{C}$ $\in$Sym, $||\mathbf{C}|| = 1$ and $\mathbb{C}[\mathbf{C}] = \gamma\mathbf{C}$; γ is an *eigenvalue* of $\mathbb{C}$ and $\mathbf{C}$ the corresponding *eigentensor*. As for the symmetric second-order tensors, for symmetric fourth-order tensors the following spectral theorem [9] holds.

Theorem 15. *Let $\mathbb{C}$:Sym$\rightarrow$Sym be a symmetric fourth-order tensor. There exist $\gamma_i \in \mathbb{R}$ and $\mathbf{C}_i$ $\in$Sym , $i = 1, ..., \frac{n(n+1)}{2}$, such that*

$$\mathbf{C}_i \cdot \mathbf{C}_j = \delta_{ij}, \quad , \sum_{i=1}^{\frac{n(n+1)}{2}} \mathbf{C}_i \otimes \mathbf{C}_i = \mathbb{I}_{Sym}, \quad (2.373)$$

and

$$\mathbb{C}[\mathbf{C}_i] = \gamma_i \mathbf{C}_i, \quad \mathbb{C} = \sum_{i=1}^{\frac{n(n+1)}{2}} \gamma_i \mathbf{C}_i \otimes \mathbf{C}_i, \quad (2.374)$$

Exercise 26. *For $n = 3$, let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be an orthonormal basis of $\mathcal{V}$, and let us consider the symmetric tensors*

$$\mathbf{O}_1 = \mathbf{e}_1 \otimes \mathbf{e}_1, \quad \mathbf{O}_2 = \mathbf{e}_2 \otimes \mathbf{e}_2, \quad \mathbf{O}_3 = \mathbf{e}_3 \otimes \mathbf{e}_3, \quad (2.375)$$

$$\mathbf{O}_4 = \frac{1}{\sqrt{2}}(\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1), \ \mathbf{O}_5 = \frac{1}{\sqrt{2}}(\mathbf{e}_1 \otimes \mathbf{e}_3 + \mathbf{e}_3 \otimes \mathbf{e}_1), \quad (2.376)$$

$$\mathbf{O}_6 = \frac{1}{\sqrt{2}}(\mathbf{e}_2 \otimes \mathbf{e}_3 + \mathbf{e}_3 \otimes \mathbf{e}_1). \quad (2.377)$$

Compute eigenvalues and eigentensors of the symmetric fourth-order tensor $\mathbb{A} \in \mathbb{S}$ym

$$\mathbb{A} = \mathbf{O}_1 \otimes \mathbf{O}_1 + \mathbf{O}_2 \otimes \mathbf{O}_2 + \mathbf{O}_1 \otimes \mathbf{O}_2 + \mathbf{O}_2 \otimes \mathbf{O}_1. \quad (2.378)$$

Solution. We have

$$\mathbb{A}[\mathbf{O}_1] = \mathbf{O}_1 + \mathbf{O}_2,\ \mathbb{A}[\mathbf{O}_2] = \mathbf{O}_1 + \mathbf{O}_2,\ \mathbb{A}[\mathbf{O}_i] = \mathbf{0},\ i = 3, ..., 6, \tag{2.379}$$

then, the eigenvalues of $\mathbb{A}$ are $\gamma_1 = 2$ with eigentensor $\frac{1}{\sqrt{2}}(\mathbf{O}_1 + \mathbf{O}_2)$ and $\gamma_2 = 0$ with eigentensors $\frac{1}{\sqrt{2}}(\mathbf{O}_1 - \mathbf{O}_2)$, $\mathbf{O}_3, \mathbf{O}_4, \mathbf{O}_5$ e $\mathbf{O}_6$ and the spectral decomposition of $\mathbb{A}$ is

$$\mathbb{A} = 2\frac{\mathbf{O}_1 + \mathbf{O}_2}{\sqrt{2}} \otimes \frac{\mathbf{O}_1 + \mathbf{O}_2}{\sqrt{2}}. \tag{2.380}$$

Exercise 27. *For $n = 3$, compute eigenvalues and eigentensors of the fourth-order tensor*

$$\mathbb{C} = 2\mu\mathbb{I}_{Sym} + \lambda\mathbf{I} \otimes \mathbf{I}, \quad \lambda, \mu \in \mathbb{R}. \tag{2.381}$$

Solution. We have $\mathbb{C}[\mathbf{I}] = (2\mu + 3\lambda)\mathbf{I}$ and $\mathbb{C}[\mathbf{A}] = 2\mu\mathbf{A}$ for each $A \in$Dev, then the eigenvalues of $\mathbb{C}$ are $\gamma_1 = 2\mu + 3\lambda$ with eigentensor $\mathbf{C}_1 = \frac{1}{\sqrt{3}}\mathbf{I}$ and $\gamma_2 = 2\mu$ with eigentensors orthogonal to $\mathbf{I}$. The spectral decomposition of $\mathbb{C}$ is

$$\mathbb{C} = (2\mu + 3\lambda)\frac{\mathbf{I}}{\sqrt{3}} \otimes \frac{\mathbf{I}}{\sqrt{3}} + 2\mu(\mathbb{I}_{Sym} - \frac{\mathbf{I}}{\sqrt{3}} \otimes \frac{\mathbf{I}}{\sqrt{3}}). \tag{2.382}$$

Exercise 28. *Prove that a symmetric fourth-order tensor $\mathbb{C}$ is definite positive if and only if its eigenvalues are positive.*

Exercise 29. *Prove that the tensor $\mathbb{C}$ defined in (2.381) is positive definite if and only if*

$$\mu > 0, \quad 2\mu + 3\lambda > 0. \tag{2.383}$$

Prove that if μ and λ satisfy (2.383), the inverse of $\mathbb{C}$ is

$$\mathbb{C}^{-1} = \frac{1}{2\mu + 3\lambda}\frac{\mathbf{I}}{\sqrt{3}} \otimes \frac{\mathbf{I}}{\sqrt{3}} + \frac{1}{2\mu}(\mathbb{I}_{Sym} - \frac{\mathbf{I}}{\sqrt{3}} \otimes \frac{\mathbf{I}}{\sqrt{3}}). \tag{2.384}$$

Exercise 30. *For $n = 3$, and $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ an orthonormal basis of $\mathcal{V}$, compute eigenvalues and eigenvector of the fourth-order tensor*

$$\mathbb{A} = \alpha\mathbf{O}_1 \otimes \mathbf{O}_1 + \beta\mathbf{O}_2 \otimes \mathbf{O}_2 + \gamma(\mathbf{O}_1 \otimes \mathbf{O}_2 + \mathbf{O}_2 \otimes \mathbf{O}_1), \tag{2.385}$$

with $\mathbf{O}_1 = \mathbf{e}_1 \otimes \mathbf{e}_1, \mathbf{O}_2 = \mathbf{e}_2 \otimes \mathbf{e}_2,\ \alpha, \beta, \gamma \in \mathbb{R}$.

2.18 Isotropic functions

In this section we limit ourselves to consider the case $n = 3$. Given $\mathfrak{I} \subset$Orth, a subset $\mathcal{A} \subset$Lin is *invariant* with respect to $\mathfrak{I}$ if $\mathbf{QAQ}^T \in \mathcal{A}$ for each $\mathbf{A} \in \mathcal{A}$, $\mathbf{Q} \in \mathfrak{I}$. $\mathbf{QAQ}^T$ is called the *orthogonal conjugate* of $\mathbf{A}$ with respect to $\mathbf{Q}$.

The sets Lin, Lin$^+$, Orth, Orth$^+$, Sym, Skw, Sym$^-$, Sym$^+$, Psym e Nsym are invariant with respect to Orth. In fact, by limiting ourselves to the case Lin$^+$ we have

$$\det(\mathbf{QAQ}^T) = \det\mathbf{A}(\det\mathbf{Q})^2 = \det\mathbf{A}.$$

For $\mathcal{A}$ $\subset$Lin, a functional $\varphi : \mathcal{A} \to \mathbb{R}$ is *invariant* with respect to $\mathfrak{I}$ if $\mathcal{A}$ is invariant with respect to $\mathfrak{I}$ and

$$\varphi(\mathbf{A}) = \varphi(\mathbf{QAQ}^T), \quad \text{for each } \mathbf{A} \in \mathcal{A}, \mathbf{Q} \in \mathfrak{I}. \tag{2.386}$$

A function $G : \mathcal{A} \to$Lin is *invariant* with respect to $\mathfrak{I}$ if $\mathcal{A}$ is invariant with respect to $\mathfrak{I}$ and if

$$\mathbf{Q}G(\mathbf{A})\mathbf{Q}^T = G(\mathbf{QAQ}^T), \quad \text{for each } \mathbf{A} \in \mathcal{A}, \mathbf{Q} \in \mathfrak{I}. \tag{2.387}$$

A functional (a function) is called *isotropic* if it is invariant with respect to Orth.

Proposition 47. *Let ϕ be a function on Lin with scalar or tensorial values, then ϕ is isotropic if and only if ϕ is isotropic with respect to Orth$^+$.*

Example 12. *The functionals I_1, I_2 e I_3 on Lin are isotropic. In particular,*

$$\eta(\mathbf{A}) = \eta(\mathbf{QAQ}^T), \quad \textit{for each } \mathbf{Q} \in \textit{Orth}. \tag{2.388}$$

Let us denote by $\mathfrak{P}(\mathcal{A}) = \{\eta(\mathbf{A}) : \mathbf{A} \in \mathcal{A}\}$ the set of all possible lists $\eta(\mathbf{A})$ of principal invariants, with $\mathbf{A} \in \mathcal{A}$.

We shall prove some important representation theorems for functions on $\mathcal{A}$ $\subset$Sym. Herein after, we shall assume that $\mathcal{A}$ is invariant with respect to Orth.

Theorem 16. *(Representation theorem for isotropic functionals). A functional $\varphi : \mathcal{A} \to \mathbb{R}$ is isotropic if and only if there exists a function $\widetilde{\varphi} : \mathfrak{P}(\mathcal{A}) \to \mathbb{R}$ such that*

$$\varphi(\mathbf{A}) = \widetilde{\varphi}(\eta(\mathbf{A})), \quad \textit{for each } \mathbf{A} \in \mathcal{A}. \tag{2.389}$$

Proof. Assume that φ is isotropic, to show (2.389) it is sufficient to show that

$$\varphi(\mathbf{A}) = \varphi(\mathbf{B}) \tag{2.390}$$

whenever

$$\eta(\mathbf{A}) = \eta(\mathbf{B}). \tag{2.391}$$

Let $\mathbf{A}, \mathbf{B} \in \mathcal{A}$ satisfy (2.391), then $\mathbf{A}$ and $\mathbf{B}$ have the same spectrum and in virtue of the spectral theorem there exist two orthonormal bases $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and$\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ such that

$$\mathbf{A} = \sum_{i=1}^{3} \omega_i \mathbf{e}_i \otimes \mathbf{e}_i, \quad \mathbf{B} = \sum_{i=1}^{3} \omega_i \mathbf{f}_i \otimes \mathbf{f}_i. \tag{2.392}$$

Let $\mathbf{Q}$ be the orthogonal tensor such that

$$\mathbf{Q}\mathbf{f}_i = \mathbf{e}_i; \tag{2.393}$$

since $\mathbf{Q}(\mathbf{f}_i \otimes \mathbf{f}_i)\mathbf{Q}^T = (\mathbf{Q}\mathbf{f}_i) \otimes (\mathbf{Q}\mathbf{f}_i)$, we have that $\mathbf{QBQ}^T = \mathbf{A}$. But, since φ is isotropic we have , $\varphi(\mathbf{A}) = \varphi(\mathbf{QBQ}^T) = \varphi(\mathbf{B})$. The inverse implication is a trivial consequence of the fact that $\eta(\mathbf{A}) = \eta(\mathbf{QAQ}^T)$ for each $\mathbf{Q}$ $\in$Orth. $\square$

Theorem 17. *(Transfer theorem) Let $G : \mathcal{A} \to$Lin be an isotropic function. Then, each eigenvector of $\mathbf{A} \in \mathcal{A}$ is an eigenvector of $G(\mathbf{A})$.*

Proof. Let $\mathbf{e}$ be an eigenvector of $\mathbf{A} \in \mathcal{A}$ and $\mathbf{Q} \in$Orth the reflection with respect to the plane orthogonal to $\mathbf{e}$,

$$\mathbf{Qe} = -\mathbf{e}, \quad \mathbf{Qf} = \mathbf{f}, \quad \text{for each } \mathbf{f} \in \text{Span}(\mathbf{e})^{\perp}, \tag{2.394}$$

It is easy to prove that $\mathbf{QAQ}^T = \mathbf{A}$. Now, since G is isotropic

$$\mathbf{Q}G(\mathbf{A})\mathbf{Q}^T = G(\mathbf{QAQ}^T) = G(\mathbf{A}), \tag{2.395}$$

then, $\mathbf{Q}$ commutes with $G(\mathbf{A})$. Moreover,

$$\mathbf{Q}G(\mathbf{A})\mathbf{e} = G(\mathbf{A})\mathbf{Qe} = -G(\mathbf{A})\mathbf{e} \tag{2.396}$$

which, along with (2.394) implies that $G(\mathbf{A})\mathbf{e} \in$Span$(\mathbf{e})$,

$$G(\mathbf{A})\mathbf{e} = \omega\mathbf{e}, \tag{2.397}$$

and then $\mathbf{e}$ is an eigenvector of $G(\mathbf{A})$. □

Proposition 48. *(Wang's lemma). Consider $\mathbf{A} \in$Sym.*

(a) *If the eigenvalues of $\mathbf{A}$ are distinct,*

$$\mathbf{A} = \sum_{i=1}^{3} \omega_i \mathbf{e}_i \otimes \mathbf{e}_i, \tag{2.398}$$

then $\mathbf{I}$, $\mathbf{A}$ e $\mathbf{A}^2$ are linearly independent and

$$\mathit{Span}(\mathbf{I}, \mathbf{A}, \mathbf{A}^2) = \mathit{Span}(\mathbf{e}_1 \otimes \mathbf{e}_1, \mathbf{e}_2 \otimes \mathbf{e}_2, \mathbf{e}_3 \otimes \mathbf{e}_3). \tag{2.399}$$

(b) *If $\mathbf{A}$ has two distinct eigenvalues,*

$$\mathbf{A} = \omega_1 \mathbf{e} \otimes \mathbf{e} + \omega_2(\mathbf{I} - \mathbf{e} \otimes \mathbf{e}), \quad ||\mathbf{e}|| = 1, \tag{2.400}$$

then $\mathbf{I}$ and $\mathbf{A}$ are linearly independent and

$$\mathit{Span}(\mathbf{I}, \mathbf{A}) = \mathit{Span}(\mathbf{e} \otimes \mathbf{e}, \mathbf{I} - \mathbf{e} \otimes \mathbf{e}). \tag{2.401}$$

Proof. As far as (a) is concerned, to prove that $\mathbf{I}$, $\mathbf{A}$ and $\mathbf{A}^2$ are linearly independent we have to prove that if

$$\alpha\mathbf{A}^2 + \beta\mathbf{A} + \gamma\mathbf{I} = \mathbf{0}, \tag{2.402}$$

then

$$\alpha = \beta = \gamma = 0. \tag{2.403}$$

From (2.398) we get

$$\begin{cases} \alpha\omega_1^2 + \beta\omega_1 + \gamma = 0, \\ \alpha\omega_2^2 + \beta\omega_2 + \gamma = 0, \\ \alpha\omega_3^2 + \beta\omega_3 + \gamma = 0. \end{cases} \tag{2.404}$$

The matrix of the system (2.404) is the Vandermonde matrix, whose determinant is given by $\prod_{1\le i<j\le 3} (\omega_i - \omega_j)$. Since ω_i are distinct, the solution to system (2.404) is given by (2.403). The subspace $\mathcal{H}$ =Span($\mathbf{e}_1 \otimes \mathbf{e}_1, \mathbf{e}_2 \otimes \mathbf{e}_2, \mathbf{e}_3 \otimes \mathbf{e}_3$) has dimension 3, and since $\mathbf{A}^2 = \sum_{i=1}^3 \omega_i^2 \mathbf{e}_i \otimes \mathbf{e}_i$, we have that $\mathbf{I}, \mathbf{A}, \mathbf{A}^2$ belong to $\mathcal{H}$, then $\mathcal{H}$ =Span($\mathbf{I}, \mathbf{A}, \mathbf{A}^2$). Point (b) can be proved analogously. □

Theorem 18. *(First representation theorem for isotropic functions) A function $G : \mathcal{A} \to$Sym ($\mathcal{A} \subset$Sym) is isotropic if and only if there exist functionals $\varphi_0, \varphi_1, \varphi_2 : \mathfrak{P}(\mathcal{A}) \to \mathbb{R}$ such that*

$$G(\mathbf{A}) = \varphi_0(\eta(\mathbf{A}))\mathbf{I} + \varphi_1(\eta(\mathbf{A}))\mathbf{A} + \varphi_2(\eta(\mathbf{A}))\mathbf{A}^2 \quad \textit{for each } \mathbf{A} \in \mathcal{A}. \tag{2.405}$$

Proof. Assume that G has the representation (2.405). Given $\mathbf{A} \in \mathcal{A}$ and $\mathbf{Q} \in$Orth, in view of (2.388) we have

$$\begin{gathered} G(\mathbf{QAQ}^T) = \varphi_0(\eta(\mathbf{QAQ}^T))\mathbf{I} + \varphi_1(\eta(\mathbf{QAQ}^T))\mathbf{QAQ}^T + \\ \varphi_2(\eta(\mathbf{QAQ}^T))\mathbf{QAQ}^T\mathbf{QAQ}^T = \\ \varphi_0(\eta(\mathbf{A}))\mathbf{QQ}^T + \varphi_1(\eta(\mathbf{A}))\mathbf{QAQ}^T + \\ \varphi_2(\eta(\mathbf{A}))\mathbf{QA}^2\mathbf{Q}^T = \mathbf{Q}G(\mathbf{A})\mathbf{Q}^T, \end{gathered} \tag{2.406}$$

then G is isotropic. Vice versa assume that G is isotropic and take $\mathbf{A} \in \mathcal{A}$. The following cases occur.

Case 1. $\mathbf{A}$ has three distinct eigenvalues. Let (2.398) be the spectral decomposition of $\mathbf{A}$, by virtue of theorem 17

$$G(\mathbf{A}) = \sum_{i=1}^{3} \beta_i \mathbf{e}_i \otimes \mathbf{e}_i, \tag{2.407}$$

from (2.399) we conclude that there exist three scalar functions $\alpha_0(\mathbf{A})$, $\alpha_1(\mathbf{A})$, $\alpha_2(\mathbf{A})$ such that

$$G(\mathbf{A}) = \alpha_0(\mathbf{A})\mathbf{I} + \alpha_1(\mathbf{A})\mathbf{A} + \alpha_2(\mathbf{A})\mathbf{A}^2. \tag{2.408}$$

Case 2. The proof is similar to the proof of case 1.

Case 3. $\mathbf{A}$ has exactly one distinct eigenvalue , $\mathbf{A} = \omega\mathbf{I}$. In particular, $\mathbf{QAQ}^T = \mathbf{A}$ for each $\mathbf{Q} \in$Orth, then from the isotropy of $G(\mathbf{A})$ it follows that $\mathbf{QG(A)Q}^T = G(\mathbf{A})$ for each $\mathbf{Q} \in$Orth, then, in view of proposition 41 $G(\mathbf{A}) = \beta\mathbf{I}$ e $G(\mathbf{A})$ has the representation (2.408) con $\alpha_0(\mathbf{A}) = \beta$ e $\alpha_1(\mathbf{A}) = \alpha_2(\mathbf{A}) = 0$.

Now we have proved that if G is isotropic then has the representation (2.408). By virtue of the representation theorem for isotropic functionals, to complete the proof we have to prove that $\alpha_0, \alpha_1, \alpha_2$ are isotropic functionals

$$\alpha_k(\mathbf{QAQ}^T) = \alpha_k(\mathbf{A}), \quad k = 1, 2, 3 \tag{2.409}$$

for each $\mathbf{A} \in \mathcal{A}$, $\mathbf{Q} \in$Orth. Then, consider $\mathbf{A} \in \mathcal{A}$, $\mathbf{Q} \in$Orth, from the isotropy of G it follows that

$$G(\mathbf{A}) - \mathbf{Q}^T\mathbf{G}(\mathbf{QAQ^T})\mathbf{Q} = \mathbf{0},$$

from which, taking both (2.408) and

$$\mathbf{Q}^T(\mathbf{QAQ^T})^2\mathbf{Q} = \mathbf{A}^2, \tag{2.410}$$

into account, we get

$$[\alpha_0(\mathbf{A}) - \alpha_0(\mathbf{QAQ^T})]\mathbf{I} + [\alpha_1(\mathbf{A}) - \alpha_1(\mathbf{QAQ^T})]\mathbf{A} +$$
$$[\alpha_2(\mathbf{A}) - \alpha_2(\mathbf{QAQ^T})]\mathbf{A}^2 = \mathbf{0}. \tag{2.411}$$

Now, we have to consider the three cases previously analyzed.

Case 1. By virtue of the Wang's lemma, $\mathbf{I}$, $\mathbf{A}$ e $\mathbf{A}^2$ are linearly independent and (2.411) implies (2.409).

Case 2. In view of (2.388) and proposition 39 $\mathbf{A}$ and $\mathbf{QAQ^T}$ have the same spectrum, then, $\mathbf{QAQ^T}$ as $\mathbf{A}$ has two distinct eigenvalues and from (**??**) we get $\alpha_2(\mathbf{A}) = \alpha_2(\mathbf{QAQ^T}) = 0$. Moreover, due to the Wang's lemma $\mathbf{I}$ and $\mathbf{A}$ are linearly independent and once again from (2.411) we get (2.409).

Case 3. In this case $\mathbf{A} = \omega\mathbf{I}$ and $\mathbf{QAQ^T} = \mathbf{A}$, then (2.409) is trivially verified.

□

If $\mathbf{A}$ is invertible, from the Cayley-Hamilton theorem it follows that

$$\mathbf{A}^2 = I_1(\mathbf{A})\mathbf{A} - I_2(\mathbf{A})\mathbf{I} + I_3(\mathbf{A})\mathbf{A}^{-1}, \tag{2.412}$$

thus, theorem 18 has the following corollary.

Theorem 19. *(Second representation theorem for isotropic functions) Let $\mathcal{A}$ be the set of all invertible symmetric tensors. A function $G : \mathcal{A} \to$Sym is isotropic if and only if there exist functionals $\psi_0, \psi_1, \psi_2 : \mathfrak{P}(\mathcal{A}) \to \mathbb{R}$ such that*

$$G(\mathbf{A}) = \psi_0(\eta(\mathbf{A}))\mathbf{I} + \psi_1(\eta(\mathbf{A}))\mathbf{A} + \psi_2(\eta(\mathbf{A}))\mathbf{A}^{-1} \quad \textit{per ogni } \mathbf{A} \in \mathcal{A}. \tag{2.413}$$

For the linear applications (fourth-order tensors) the following result holds.

Theorem 20. *(Representation theorem for isotropic fourth-order tensors) A fourth order tensor $\mathbb{A}$:Sym→Sym is isotropic if and only if these exist two scalars μ and λ such that*

$$\mathbb{A}[\mathbf{A}] = 2\mu\mathbf{A} + \lambda\ tr(\mathbf{A})\mathbf{I}, \quad \textit{for each } \mathbf{A} \in \ \textit{Sym}. \tag{2.414}$$

Proof. Clearly (2.414) defines an isotropic function. Let $\mathcal{N}$ be the set of all vectors with 1 norm. For each $\mathbf{e} \in \mathcal{N}$, tensor $\mathbf{e} \otimes \mathbf{e}$ has spectrum $\{0, 0, 1\}$ and characteristic spaces $\text{Span}(\mathbf{e})^{\perp}$ and $\text{Span}(\mathbf{e})$. Then the same procedure used to prove (**??**) implies the existence of two functionals $\mu,\ \lambda : \mathcal{N} \to \mathbb{R}$ such that

$$\mathbb{A}[\mathbf{e} \otimes \mathbf{e}] = 2\mu(\mathbf{e})\mathbf{e} \otimes \mathbf{e} + \lambda(\mathbf{e})\mathbf{I}, \quad \text{for each } \mathbf{e} \in \mathcal{N}. \tag{2.415}$$

Now, consider $\mathbf{e}, \mathbf{f} \in \mathcal{N}$ and let $\mathbf{Q}$ be the orthogonal tensor such that $\mathbf{Qe} = \mathbf{f}$. Since

$$\mathbf{Qe} \otimes \mathbf{eQ}^T = \mathbf{f} \otimes \mathbf{f}, \tag{2.416}$$

and $\mathbb{A}$ is isotropic, we have

$$\mathbf{0} = \mathbf{Q}\mathbb{A}[\mathbf{e} \otimes \mathbf{e}]\mathbf{Q}^T - \mathbb{A}[\mathbf{f} \otimes \mathbf{f}] =$$

$$2[\mu(\mathbf{e}) - \mu(\mathbf{f})]\mathbf{f} \otimes \mathbf{f} + [\lambda(\mathbf{e}) - \lambda(\mathbf{f})]\mathbf{I}. \tag{2.417}$$

Since $\mathbf{f} \otimes \mathbf{f}$ and $\mathbf{I}$ are linearly independent, from (2.417) we get

$$\mu(\mathbf{e}) = \mu(\mathbf{f}), \quad \lambda(\mathbf{e}) = \lambda(\mathbf{f}), \tag{2.418}$$

then μ and λ are constant scalar quantities and from (2.415) we conclude that

$$\mathbb{A}[\mathbf{e} \otimes \mathbf{e}] = 2\mu\mathbf{e} \otimes \mathbf{e} + \lambda\mathbf{I}, \quad \text{for each } \mathbf{e} \in \mathcal{N}. \tag{2.419}$$

Now, let us consider $\mathbf{A}$ ∈Sym, in view of the spectral theorem $\mathbf{A}$ has the representation (2.398), and by virtue of (2.419) and the linearity of $\mathbb{A}$, we have

$$\mathbb{A}[\mathbf{A}] = 2\mu\mathbf{A} + \lambda(\omega_1 + \omega_2 + \omega_3)\mathbf{I}. \tag{2.420}$$

□

Corollary 1. *Let $\mathbb{A}$:Sym_0 →Sym be a fourth-order tensor, with $Sym_0 = \{\mathbf{A} \in Sym : tr\mathbf{A} = 0\}$. $\mathbb{A}$ is isotropic if and only if it exists a scalar μ such that*

$$\mathbb{A}[\mathbf{A}] = 2\mu\mathbf{A}, \quad \textit{for each } \mathbf{A} \in Sym_0. \tag{2.421}$$

In particular, if $\mathbb{A}$ is isotropic, then $\mathbf{A}$ and $\mathbb{A}[\mathbf{A}]$ commute, $\mathbf{A}\mathbb{A}[\mathbf{A}] = \mathbb{A}[\mathbf{A}]\mathbf{A}$, and are coaxial.

Exercise 31. *Prove that the function R :Psym→Psym which associates to each $\mathbf{C}$ the tensor $\sqrt{\mathbf{C}}$ is isotropic.*

Solution. For each $\mathbf{C}$ ∈Psym, $\mathbf{Q}$ ∈Orth, we have

$$(\mathbf{Q}\sqrt{\mathbf{C}}\mathbf{Q}^T)^2 = \mathbf{Q}\sqrt{\mathbf{C}}\mathbf{Q}^T\mathbf{Q}\sqrt{\mathbf{C}}\mathbf{Q}^T = \mathbf{Q}\mathbf{C}\mathbf{Q}^T. \tag{2.422}$$

Exercise 32. *Let G :Sym→Sym be an invertible isotropic function, prove that G^{-1} is isotropic.*

Solution. For each $\mathbf{A}$ ∈Sym, $\mathbf{Q}$ ∈Orth, we have

$$G(\mathbf{Q}G^{-1}(\mathbf{A})\mathbf{Q}^T) = \mathbf{Q}\mathbf{A}\mathbf{Q}^T, \tag{2.423}$$

applying G^{-1} we get

$$\mathbf{Q}G^{-1}(\mathbf{A})\mathbf{Q}^T = G^{-1}(\mathbf{Q}\mathbf{A}\mathbf{Q}^T). \tag{2.424}$$

Exercise 33. *The function G :Lin→Lin defined by $G(\mathbf{A}) = \mathbf{A}^k$, $k \in \mathbb{N}$ is isotropic.*

Exercise 34. *The function G :Lin$^+$ →Lin$^+$ defined by $G(\mathbf{A}) = \mathbf{A}^{-1}$, is isotropic.*

From exercises 31 and 34 it follows that the function S :Psym→Psym defined by $S(\mathbf{C}) = (R(\mathbf{C}))^{-1} = (\sqrt{\mathbf{C}})^{-1}$ is isotropic. Moreover, the function that associates to each $\mathbf{F}$ ∈Lin$^+$ the tensor $\mathbf{F}\mathbf{F}^T$ ∈Psym is isotropic; since the composition of two isotropic functions is isotropic, the functions defined from Lin$^+$ into in Psym that assigns to each tensor $\mathbf{F}$ the tensor $\mathbf{U}$ of the right polar decomposition of $\mathbf{F}$ and $\mathbf{U}^{-1}$ are isotropic. The function from Lin$^+$ into Orth$^+$ that assigns to each tensor $\mathbf{F}$ the tensor $\mathbf{R}$ of the right polar decomposition of $\mathbf{F}$ is isotropic.

2.19 Convergence of tensors

Let us consider the natural norm on Lin (cfr. (1.104)),

$$||\mathbf{A}||_N = \sup_{\mathbf{u}\neq\mathbf{0}} \frac{||\mathbf{A}\mathbf{u}||}{||\mathbf{u}||}, \quad \mathbf{A} \in \text{Lin}. \tag{2.425}$$

The natural norm is submultiplicative

$$||\mathbf{A}\mathbf{B}||_N \leq ||\mathbf{A}||_N||\mathbf{B}||_N, \quad \mathbf{A},\mathbf{B} \in \text{Lin}, \tag{2.426}$$

and $||\mathbf{I}||_N = 1$.

Given $\mathbf{A}$ ∈Lin, from (2.80) it follows that

$$||\mathbf{A}|| \geq ||\mathbf{A}||_N \tag{2.427}$$

where $||\mathbf{A}||$ is given in (2.71). In particular, if $n = 3$, we have

$$||\mathbf{A}||^2 = tr(\mathbf{A}^T\mathbf{A}) = \lambda_1 + \lambda_2 + \lambda_3, \tag{2.428}$$

where $0 \leq \lambda_1 \leq \lambda_2 \leq \lambda_3$ are the eigenvalues of the tensor $\mathbf{A}^T\mathbf{A} \in$Sym$^+$ and

$$\frac{||\mathbf{Au}||^2}{||\mathbf{u}||^2} = \frac{\mathbf{A}^T\mathbf{Au}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}} \tag{2.429}$$

is the Rayleigh quotient of the tensor $\mathbf{A}^T\mathbf{A}$ which satisfies the inequalities

$$\lambda_1 \leq \frac{\mathbf{A}^T\mathbf{Au}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}} \leq \lambda_3, \tag{2.430}$$

from which it follows that

$$||\mathbf{A}||_N^2 = \lambda_3. \tag{2.431}$$

In particular,

$$||\mathbf{A}||_N^2 \leq ||\mathbf{A}||^2 \leq 3||\mathbf{A}||_N^2. \tag{2.432}$$

Now we want to define the convergence of a sequence of tensors $\{\mathbf{A}_k\}_{k\in\mathbb{N}}$ to a tensor $\mathbf{A} \in$Lin. We say that $\{\mathbf{A}_k\}_{k\in\mathbb{N}}$ converges to $\mathbf{A}$ if if for each $\varepsilon > 0$ there exists $\overline{k} > 0$ such that

$$\|\mathbf{A}_k - \mathbf{A}\| < \varepsilon \quad \text{for each} \quad k \geq \overline{k}. \tag{2.433}$$

Proposition 49. *The following conditions are equivalent.*

(i) $||\mathbf{A}_k - \mathbf{A}|| \to 0$ *for* $k \to \infty$.

(ii) $||\mathbf{A}_k\mathbf{u} - \mathbf{Au}|| \to 0$ *when* $k \to \infty$, *for each fixed* $\mathbf{u} \in \mathcal{V}$.

(iii) $|\mathbf{A}_k\mathbf{u}\cdot\mathbf{v} - \mathbf{Au}\cdot\mathbf{v}| \to 0$ *quando* $k \to \infty$, *for each* $\mathbf{u}, \mathbf{v} \in \mathcal{V}$.

Proof. If (i) holds, then, for each $\mathbf{u} \in \mathcal{V}$ we have

$$||\mathbf{A}_k\mathbf{u} - \mathbf{Au}|| = ||(\mathbf{A}_k - \mathbf{A})\mathbf{u}|| \leq ||\mathbf{A}_k - \mathbf{A}||\ ||\mathbf{u}|| \to 0, \tag{2.434}$$

then (i)⇒(ii). In section 1.8 we have proved that (ii)⇒(iii) and that in finite-dimensional vector spaces (iii)⇒(ii). Thus, we have to prove that in a finite-dimensional vector space (ii)⇒(i). Let $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ be an orthonormal basis of $\mathcal{V}$, if (ii) holds, then for each $\varepsilon > 0$ there is $k_0 = k(\varepsilon)$ such that $||\mathbf{A}_k\mathbf{u}_i - \mathbf{Au}_i|| < \varepsilon$ for $k \geq k_0$ and for $i = 1, 2, 3$. Given $\mathbf{u} \in \mathcal{V}$, we have $\mathbf{u} = \sum_{i=1}^3(\mathbf{u}\cdot\mathbf{u}_i)\mathbf{u}_i$ and then

$$||(\mathbf{A}_k - \mathbf{A})\mathbf{u}|| = ||\sum_{i=1}^3(\mathbf{u}\cdot\mathbf{u}_i)(\mathbf{A}_k - \mathbf{A})\mathbf{u}_i|| \leq$$

$$\sum_{i=1}^3 ||\mathbf{u}||\ ||(\mathbf{A}_k - \mathbf{A})\mathbf{u}_i|| \leq 3\varepsilon||\mathbf{u}||, \tag{2.435}$$

therefore $||\mathbf{A}_k - \mathbf{A}||_N \to 0$, which, along with (2.432) gives (i). □

The inner product on $\mathcal{V}$ and the vector product, as bilinear functions, are continuous. In fact, given $\mathbf{u}_k \to \mathbf{u}$, $\mathbf{v}_k \to \mathbf{v}$, we have

$$|\mathbf{u}_k \cdot \mathbf{v}_k - \mathbf{u} \cdot \mathbf{v}| \leq |\mathbf{u}_k \cdot \mathbf{v}_k - \mathbf{u} \cdot \mathbf{v}_k| + |\mathbf{u} \cdot \mathbf{v}_k - \mathbf{u} \cdot \mathbf{v}|,$$

$$||\mathbf{u}_k \wedge \mathbf{v}_k - \mathbf{u} \wedge \mathbf{v}|| \leq ||\mathbf{u}_k \wedge \mathbf{v}_k - \mathbf{u} \wedge \mathbf{v}_k|| + ||\mathbf{u} \wedge \mathbf{v}_k - \mathbf{u} \wedge \mathbf{v}|| \leq$$
$$||(\mathbf{u}_k - \mathbf{u}) \wedge \mathbf{v}_k|| + ||\mathbf{u} \wedge (\mathbf{v}_k - \mathbf{v})|| \leq$$
$$||\mathbf{u}_k - \mathbf{u}||\ ||\mathbf{v}_k|| + ||\mathbf{v}_k - \mathbf{v}||\ ||\mathbf{u}||.$$

Exercise 35. *Prove that*

(a) $\varphi_1 : Lin \to \mathbb{R}$, $\varphi_1(\mathbf{A}) = ||\mathbf{A}||$,

(b) $T_1 : Lin \times \mathcal{V} \to \mathcal{V}$, $T_1(\mathbf{A}, \mathbf{u}) = \mathbf{A}\mathbf{u}$,

(c) $\varphi_3 : Lin \times \mathcal{V} \times \mathcal{V} \to \mathbb{R}$, $\varphi_3(\mathbf{A}, \mathbf{u}, \mathbf{v}) = \mathbf{A}\mathbf{u} \cdot \mathbf{v}$,

(d) $\varphi_2 : Lin \times \mathcal{V} \to \mathbb{R}$, $\varphi_2(\mathbf{A}, \mathbf{u}) = ||\mathbf{A}\mathbf{u}||$,

(e) $T_2 : Lin \times Lin \to Lin$, $T_2(\mathbf{A}, \mathbf{B}) = \mathbf{A} + \mathbf{B}$,

(f) $T_3 : Lin \times \mathbb{R} \to Lin$, $T_3(\mathbf{A}, \alpha) = \alpha\mathbf{A}$,

(g) $T_4 : Lin \times Lin \to Lin$, $T_4(\mathbf{A}, \mathbf{B}) = \mathbf{A}\mathbf{B}$,

(h) $T_5 : Lin \to Lin$, $T_5(\mathbf{A}) = \mathbf{A}^T$,

(i) $\varphi_4 : Lin \to \mathbb{R}$, $\varphi_4(\mathbf{A}) = tr\mathbf{A}$,

(j) $\varphi_5 : Lin \to \mathbb{R}$, $\varphi_5(\mathbf{A}) = \det \mathbf{A}$,

(l) $T_6 : Inv \to Inv$, $T_5(\mathbf{A}) = \mathbf{A}^{-1}$, *with Inv the set of all invertible tensors,*

are continuous functions.

Solution.

(a) Given $\mathbf{A}_k \to \mathbf{A}$, from the second triangle inequality (1.35) it follows that $|||\mathbf{A}_k|| - ||\mathbf{A}||| \leq ||\mathbf{A}_k - \mathbf{A}||$.

(b) Given $\mathbf{A}_k \to \mathbf{A}$ and $\mathbf{u}_k \to \mathbf{u}$,

$$||\mathbf{A}_k\mathbf{u}_k - \mathbf{A}\mathbf{u}|| \leq ||\mathbf{A}_k\mathbf{u}_k - \mathbf{A}_k\mathbf{u}|| + ||\mathbf{A}_k\mathbf{u} - \mathbf{A}\mathbf{u}|| \leq$$

$$||\mathbf{A}_k||\ ||\mathbf{u}_k - \mathbf{u}|| + ||\mathbf{A}_k\mathbf{u} - \mathbf{A}\mathbf{u}||.$$

(c) Given $\mathbf{A}_k \to \mathbf{A}$, $\mathbf{u}_k \to \mathbf{u}$, $\mathbf{v}_k \to \mathbf{v}$, we have

$$|\mathbf{A}_k\mathbf{u}_k \cdot \mathbf{v}_k - \mathbf{A}\mathbf{u} \cdot \mathbf{v}| \leq |\mathbf{A}_k\mathbf{u}_k \cdot \mathbf{v}_k - \mathbf{A}\mathbf{u}_k \cdot \mathbf{v}_k| +$$

$$|\mathbf{A}\mathbf{u}_k \cdot \mathbf{v}_k - \mathbf{A}\mathbf{u} \cdot \mathbf{v}_k| + |\mathbf{A}\mathbf{u} \cdot \mathbf{v}_k - \mathbf{A}\mathbf{u} \cdot \mathbf{v}|.$$

(j) Let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be an orthonormal basis of $\mathcal{V}$ with $\mathbf{e}_3 = \mathbf{e}_1 \wedge \mathbf{e}_2$ and $\mathbf{A}_k \to \mathbf{A}$, in view of (2.175) we have

$$|\det \mathbf{A}_k - \det \mathbf{A}| = |\mathbf{A}_k\mathbf{e}_1 \cdot (\mathbf{A}_k\mathbf{e}_2 \wedge \mathbf{A}_k\mathbf{e}_3) - \mathbf{A}\mathbf{e}_1 \cdot (\mathbf{A}\mathbf{e}_2 \wedge \mathbf{A}\mathbf{e}_3)|.$$

The thesis follows from (d) and from the continuity of the vector product and inner product on $\mathcal{V}$.

(l) By virtue of the Cayley-Hamilton theorem, $T_6(\mathbf{A}) = \frac{1}{\det(\mathbf{A})}[\mathbf{A}^2 - I_1(\mathbf{A})\mathbf{A} + I_2(\mathbf{A})\mathbf{I}]$, therefore T_6 is continuous because it is sum, product and quotient of continuous functions.

2.20 Derivatives of functionals and vector and tensor-valued functions

Put $n = 3$.

Exercise 36. *Compute the derivative of the following functions.*

(a) $\varphi : \mathcal{V} \to \mathbb{R}$ *defined by*

$$\varphi(\mathbf{v}) = \mathbf{v} \cdot \mathbf{v}, \quad \mathbf{v} \in \mathcal{V}.$$

(b) F *:Lin→Lin defined by*

$$G(\mathbf{A}) = \mathbf{A}^2, \quad \mathbf{A} \in Lin. \tag{2.436}$$

(c) F *:Lin→Lin defined by*

$$F(\mathbf{A}) = \mathbf{A}^3, \quad \mathbf{A} \in Lin. \tag{2.437}$$

Solution.

(a) For $\mathbf{v} \in \mathcal{V}$ we have

$$\varphi(\mathbf{v} + \mathbf{u}) = (\mathbf{v} + \mathbf{u}) \cdot (\mathbf{v} + \mathbf{u}) = \mathbf{v} \cdot \mathbf{v} + 2\mathbf{v} \cdot \mathbf{u} + \mathbf{u} \cdot \mathbf{u} =$$

$$\varphi(\mathbf{v}) + 2\mathbf{v} \cdot \mathbf{u} + o(\mathbf{u}) \quad \mathbf{u} \to \mathbf{0}, \tag{2.438}$$

from which

$$D\varphi(\mathbf{v})[\mathbf{u}] = 2\mathbf{v} \cdot \mathbf{u}, \quad \mathbf{u} \in \mathcal{V}.$$

(b) For $\mathbf{A} \in$Lin we have

$$G(\mathbf{A} + \mathbf{U}) = (\mathbf{A} + \mathbf{U})(\mathbf{A} + \mathbf{U}) =$$

$$\mathbf{A}^2 + \mathbf{A}\mathbf{U} + \mathbf{U}\mathbf{A} + \mathbf{U}^2 = G(\mathbf{A}) + \mathbf{A}\mathbf{U} + \mathbf{U}\mathbf{A} + o(\mathbf{U}), \quad \mathbf{U} \to \mathbf{0}, \tag{2.439}$$

where the last equality follows from the fact that the norm is submultiplicative,

$$||\mathbf{U}^2|| \leq ||\mathbf{U}||^2.$$

From (2.439) we obtain

$$DG(\mathbf{A})[\mathbf{U}] = \mathbf{AU} + \mathbf{UA}, \quad \mathbf{U} \in \text{Lin}, \tag{2.440}$$

and then $DG(\mathbf{A}) = \mathbf{A} \boxtimes \mathbf{I} + \mathbf{I} \boxtimes \mathbf{A}^T$. The function DG :Lin$\rightarrow \mathcal{L}$(Lin,Lin) that assigns to each $\mathbf{A}$ ∈Lin the fourth-order tensor $DG(\mathbf{A})$ is continuous. To prove the continuity, given $\mathbf{A}_k \rightarrow \mathbf{A}$, we have to prove that $DG(\mathbf{A}_k)$ converges to $DG(\mathbf{A})$. In view of (2.440) we have

$$||DG(\mathbf{A}_k) - DG(\mathbf{A})|| = \sup_{\substack{\mathbf{H}\in\text{Lin}\\ \mathbf{H}\neq\mathbf{0}}} \frac{||DG(\mathbf{A}_k)[\mathbf{H}] - DG(\mathbf{A})[\mathbf{H}]||}{||\mathbf{H}||} =$$

$$\sup_{\substack{\mathbf{H}\in\text{Lin}\\ \mathbf{H}\neq\mathbf{0}}} \frac{||\mathbf{A}_k\mathbf{H} + \mathbf{HA}_k - \mathbf{AH} - \mathbf{HA}||}{||\mathbf{H}||} \leq \sup_{\substack{\mathbf{H}\in\text{Lin}\\ \mathbf{H}\neq\mathbf{0}}} \frac{||\mathbf{A}_k\mathbf{H} - \mathbf{AH}|| + ||\mathbf{HA}_k - \mathbf{HA}||}{||\mathbf{H}||} \leq$$

$$\sup_{\substack{\mathbf{H}\in\text{Lin}\\ \mathbf{H}\neq\mathbf{0}}} \frac{||\mathbf{A}_k\mathbf{H} - \mathbf{AH}||}{||\mathbf{H}||} + \sup_{\substack{\mathbf{H}\in\text{Lin}\\ \mathbf{H}\neq\mathbf{0}}} \frac{||\mathbf{HA}_k - \mathbf{HA}||}{||\mathbf{H}||} \leq 2||\mathbf{A}_k - \mathbf{A}||. \tag{2.441}$$

This allows to conclude that G is of class C^1.

(c)For $\mathbf{A}$ ∈Lin we have

$$F(\mathbf{A} + \mathbf{U}) = \mathbf{A}^3 + \mathbf{A}^2\mathbf{U} + \mathbf{UA}^2 + \mathbf{AUA} +$$

$$\mathbf{AU}^2 + \mathbf{U}^2\mathbf{A} + \mathbf{UAU} + \mathbf{U}^3 =$$

$$F(\mathbf{A}) + \mathbf{A}^2\mathbf{U} + \mathbf{UA}^2 + \mathbf{AUA} + o(\mathbf{U}), \quad \mathbf{U} \rightarrow \mathbf{0}, \tag{2.442}$$

where we have taken into account the fact that

$$||\mathbf{AU}^2 + \mathbf{U}^2\mathbf{A} + \mathbf{UAU} + \mathbf{U}^3|| \leq 3||\mathbf{A}||\ ||\mathbf{U}||^2 + ||\mathbf{U}||^3,$$

and then $\mathbf{AU}^2 + \mathbf{U}^2\mathbf{A} + \mathbf{UAU} + \mathbf{U}^3 = o(\mathbf{U}), \quad \mathbf{U} \rightarrow \mathbf{0}$.

From (2.442) it follows that $DF(\mathbf{A})$ is the fourth order tensor defined by

$$DF(\mathbf{A})[\mathbf{U}] = \mathbf{A}^2\mathbf{U} + \mathbf{UA}^2 + \mathbf{AUA}, \quad \mathbf{U} \in \text{Lin}, \tag{2.443}$$

with $DF(\mathbf{A}) = \mathbf{A}^2 \boxtimes \mathbf{I} + \mathbf{I} \boxtimes (\mathbf{A}^2)^T + \mathbf{A} \boxtimes \mathbf{A}^T$.

Exercise 37. *Compute the derivative of the following functions from Lin to Lin:*

(a) $G(\mathbf{A}) = (tr\mathbf{A})\mathbf{A}$, *for each* $\mathbf{A}$ *∈Lin.*

(b) $G(\mathbf{A}) = \mathbf{ABA}$, *for each* $\mathbf{A}$ *∈Lin,* $\mathbf{B}$ *∈Lin fixed.*

(c) $G(\mathbf{A}) = \mathbf{A}^T\mathbf{A}$, *for each* $\mathbf{A}$ *∈Lin.*

(d) $G(\mathbf{A}) = (\mathbf{u} \cdot \mathbf{Au})\mathbf{A}$, *for each* $\mathbf{A}$ *∈Lin,* $\mathbf{u} \in \mathcal{V}$ *fixed.*

Solution.

(a) For every $\mathbf{U}$ $\in$Lin we have

$$G(\mathbf{A}+\mathbf{U}) = tr(\mathbf{A}+\mathbf{U})(\mathbf{A}+\mathbf{U}) =$$

$$tr(\mathbf{A})\mathbf{A} + tr(\mathbf{U})\mathbf{A} + tr(\mathbf{A})\mathbf{U} + tr(\mathbf{U})\mathbf{U},$$

since $||tr(\mathbf{U})\mathbf{U}|| = |\mathbf{I}\cdot\mathbf{U}|\ ||\mathbf{U}|| \leq \sqrt{3}||\mathbf{U}||^2$, we have that $tr(\mathbf{U})\mathbf{U} = o(\mathbf{U})$ per $\mathbf{U} \to \mathbf{0}$, then,

$$DG(\mathbf{A})[\mathbf{U}] = tr(\mathbf{U})\mathbf{A} + tr(\mathbf{A})\mathbf{U}, \quad \text{for each } \mathbf{U} \in \text{Lin}$$

and

$$DG(\mathbf{A}) = \mathbf{A}\otimes\mathbf{I} + (\mathbf{A}\cdot\mathbf{I})\mathbb{I}.$$

(b) For every $\mathbf{U}$ $\in$Lin we have

$$G(\mathbf{A}+\mathbf{U}) = (\mathbf{A}+\mathbf{U})\mathbf{B}(\mathbf{A}+\mathbf{U}) =$$

$$G(\mathbf{A}) + \mathbf{UBA} + \mathbf{ABU} + \mathbf{UBU},$$

Since $\mathbf{UBU} = o(\mathbf{U})$, $\mathbf{U}\to\mathbf{0}$, we have

$$DG(\mathbf{A})[\mathbf{U}] = \mathbf{UBA} + \mathbf{ABU}, \quad \text{for each } \mathbf{U} \in \text{Lin},$$

and

$$DG(\mathbf{A}) = \mathbf{AB}\boxtimes\mathbf{I} + \mathbf{I}\boxtimes(\mathbf{BA})^T.$$

(c) For every $\mathbf{U}$ $\in$Lin we have

$$G(\mathbf{A}+\mathbf{U}) = (\mathbf{A}^T+\mathbf{U}^T)(\mathbf{A}+\mathbf{U}) =$$

$$G(\mathbf{A}) + \mathbf{A}^T\mathbf{U} + \mathbf{U}^T\mathbf{A} + \mathbf{U}^T\mathbf{U}.$$

Since $\mathbf{U}^T\mathbf{U} = o(\mathbf{U})$, $\mathbf{U}\to\mathbf{0}$, we have

$$DG(\mathbf{A})[\mathbf{U}] = \mathbf{A}^T\mathbf{U} + \mathbf{U}^T\mathbf{A}, \quad \text{for each } \mathbf{U} \in \text{Lin}.$$

Theorem 21. *Let φ be the functional defined on the subset Inv of Lin constituted by all invertible tensors*

$$\varphi(\mathbf{A}) = \det\mathbf{A}. \tag{2.444}$$

φ is of class C^1 and

$$D\varphi(\mathbf{A})[\mathbf{U}] = (\det\mathbf{A})tr(\mathbf{U}\mathbf{A}^{-1}), \quad \textit{for each } \mathbf{U} \in \textit{Lin}. \tag{2.445}$$

Proof. Let us start by remarking that the set Inv= $\{\mathbf{A}$ $\in$Lin : $\det\mathbf{A} \neq 0\}$ is open in Lin because it is the complement of the set Ninv= $\{\mathbf{A}$ $\in$Lin : $\det\mathbf{A} = 0\}$ which is closed as it is the inverse image of the closed set $\{0\}$ in $\mathbb{R}$ under the continuous function det. Given $\mathbf{B}$ $\in$Lin, from (2.235) and (2.236) with $a = -1$ we obtain

$$\det(\mathbf{B}+\mathbf{I}) = 1 + I_1(\mathbf{B}) + I_2(\mathbf{B}) + I_3(\mathbf{B}). \tag{2.446}$$

From the relation $\det \mathbf{B} = \frac{1}{6}[(tr\mathbf{B})^3 - 3(tr\mathbf{B})tr(\mathbf{B}^2) + 2tr(\mathbf{B}^3)]$, it follows that

$$|\det \mathbf{B}| \leq \frac{1}{6}[|tr\mathbf{B}|^3 + 3|tr\mathbf{B}|\ |tr(\mathbf{B}^2)| + 2|tr(\mathbf{B}^3)|] \leq$$

$$\frac{\sqrt{3}}{6}[3||\mathbf{B}||^3 + 9|\mathbf{B}||^2 + 2||\mathbf{B}||^3], \tag{2.447}$$

then $\det \mathbf{B} = o(\mathbf{B})$, $\mathbf{B} \to \mathbf{0}$ e

$$\det(\mathbf{B} + \mathbf{I}) = 1 + I_1(\mathbf{B}) + o(\mathbf{B}), \quad \mathbf{B} \to \mathbf{0}. \tag{2.448}$$

Thus, for $\mathbf{A}$ ∈Inv fixed, for each $\mathbf{U}$ ∈Lin, we have

$$\det(\mathbf{A} + \mathbf{U}) = \det[(\mathbf{I} + \mathbf{U}\mathbf{A}^{-1})\mathbf{A}] =$$

$$(\det \mathbf{A})\det(\mathbf{I} + \mathbf{U}\mathbf{A}^{-1}) = (\det \mathbf{A})[1 + tr(\mathbf{U}\mathbf{A}^{-1}) + o(\mathbf{U})], \quad \mathbf{U} \to \mathbf{0}. \tag{2.449}$$

Since the function $\mathbf{U} \mapsto tr(\mathbf{U}\mathbf{A}^{-1})$ is linear, from (2.449), (2.445) follows. Moreover, the continuity of the function $D\varphi$ from Inv to $\mathcal{L}(\text{Lin}, \mathbb{R})$, follows from the continuity of the determinant and of the inverse. In particular, we have to prove that if $\mathbf{A}_k \to \mathbf{A}$, in Inv, then $D\varphi(\mathbf{A}_k) \to D\varphi(\mathbf{A})$ in $\mathcal{L}(\text{Lin}, \mathbb{R})$,

$$||D\varphi(\mathbf{A}_k) - D\varphi(\mathbf{A})||_N = \sup_{\substack{\mathbf{H}\in\text{Lin} \\ \mathbf{H}\neq\mathbf{0}}} \frac{|D\varphi(\mathbf{A}_k)[\mathbf{H}] - D\varphi(\mathbf{A})[\mathbf{H}]|}{||\mathbf{H}||} =$$

$$\sup_{\substack{\mathbf{H}\in\text{Lin} \\ \mathbf{H}\neq\mathbf{0}}} \frac{|(\det \mathbf{A}_k)tr(\mathbf{H}\mathbf{A}_k^{-1}) - (\det \mathbf{A})tr(\mathbf{H}\mathbf{A}^{-1})|}{||\mathbf{H}||} =$$

$$\sup_{\substack{\mathbf{H}\in\text{Lin} \\ \mathbf{H}\neq\mathbf{0}}} \frac{|(\det \mathbf{A}_k)\mathbf{H}^T \cdot \mathbf{A}_k^{-1} - (\det \mathbf{A})\mathbf{H}^T \cdot \mathbf{A}^{-1}|}{||\mathbf{H}||} \leq$$

$$\sup_{\substack{\mathbf{H}\in\text{Lin} \\ \mathbf{H}\neq\mathbf{0}}} \frac{|\mathbf{H}^T \cdot [(\det \mathbf{A}_k)\mathbf{A}_k^{-1} - (\det \mathbf{A})\mathbf{A}^{-1}]|}{||\mathbf{H}||} \leq$$

$$||(\det \mathbf{A}_k)\mathbf{A}_k^{-1} - (\det \mathbf{A}_k)\mathbf{A}^{-1}|| +$$

$$||(\det \mathbf{A}_k)\mathbf{A}^{-1} - (\det \mathbf{A})\mathbf{A}^{-1}|| \leq$$

$$|\det \mathbf{A}_k|\ ||\mathbf{A}_k^{-1} - \mathbf{A}^{-1}|| + ||\mathbf{A}^{-1}||\ |\det \mathbf{A}_k - \det \mathbf{A}|,$$

and the thesis follows from the continuity of the determinant (exercise 35 (j)) and the function T_6 (exercise 35 (l)). □

From (2.445) and (2.192) it follows that the derivative of the determinant of a tensor $\mathbf{A}$ ∈Inv coincides with its cofactor $\mathbf{A}^*$.

Exercise 38. *Consider G :Inv→Lin such that $G(\mathbf{A}) = \mathbf{A}^{-1}$. Assuming that G is differentiable, prove that*

$$DG(\mathbf{A})[\mathbf{H}] = -\mathbf{A}^{-1}\mathbf{H}\mathbf{A}^{-1}, \quad \mathbf{A} \in Lin. \tag{2.450}$$

Solution. Let us consider the linear function F :Inv$\to$Inv such that $F(\mathbf{A}) = \mathbf{A}$. Let us consider the function product

$$F(\mathbf{A})G(\mathbf{A}) = \mathbf{I}, \quad \mathbf{A} \in \text{Inv}. \tag{2.451}$$

From the product rule it follows that

$$DF(\mathbf{A})[\mathbf{H}]G(\mathbf{A}) + F(\mathbf{A})DG(\mathbf{A})[\mathbf{H}] = \mathbf{0}, \quad \mathbf{H} \in \text{Lin}, \tag{2.452}$$

from which

$$\mathbf{H}\mathbf{A}^{-1} + \mathbf{A}DG(\mathbf{A})[\mathbf{H}] = \mathbf{0}, \tag{2.453}$$

and then (2.450) is satisfied.

Exercise 39. *Given ψ :Inv$\to \mathbb{R}$ such that $\psi(\mathbf{A}) = \det(\mathbf{A}^2)$, compute $D\psi(\mathbf{A})$ for each $\mathbf{A}$ ∈Inv.*

Solution. Consider φ :Inv$\to \mathbb{R}$ such that $\varphi(\mathbf{A}) = \det(\mathbf{A})$ and G :Inv$\to$Inv such that $G(\mathbf{A}) = \mathbf{A}^2$. Taking into account that $\psi = \varphi \circ G$, for each $\mathbf{H}$ ∈Lin we have

$$\begin{gathered} D\psi(\mathbf{A})[\mathbf{H}] = D\varphi(G(\mathbf{A}))[DG(\mathbf{A})[\mathbf{H}]] = \\ \det(\mathbf{A}^2)tr(DG(\mathbf{A})[\mathbf{H}]\mathbf{A}^{-2}) = \\ \det(\mathbf{A}^2)tr((\mathbf{A}\mathbf{H} + \mathbf{H}\mathbf{A})\mathbf{A}^{-2}) = 2\det(\mathbf{A}^2)tr(\mathbf{H}\mathbf{A}^{-1}). \end{gathered} \tag{2.454}$$

Exercise 40. *Consider ψ :Inv$\to \mathbb{R}$ such that $\psi(\mathbf{A}) = (\det \mathbf{A})tr(\mathbf{A}^{-1})$, $\mathbf{A}$ ∈Inv. Compute $D\psi(\mathbf{A})$.*

Solution. Consider φ :Inv$\to \mathbb{R}$ such that $\varphi(\mathbf{A}) = \det(\mathbf{A})$ and G :Inv$\to$Inv such that $G(\mathbf{A}) = \mathbf{A}^{-1}$, then, we have $\psi(\mathbf{A}) = \varphi(\mathbf{A})tr(G(\mathbf{A}))$. Therefore

$$\begin{gathered} D\psi(\mathbf{A})[\mathbf{H}] = D\varphi(\mathbf{A})[\mathbf{H}]tr(G(\mathbf{A})) + \varphi(\mathbf{A})Dtr(G(\mathbf{A}))[DG(\mathbf{A})[\mathbf{H}]] = \\ (\det \mathbf{A})tr(\mathbf{H}\mathbf{A}^{-1})tr(\mathbf{A}^{-1}) + \\ (\det \mathbf{A})tr(-\mathbf{A}^{-1}\mathbf{H}\mathbf{A}^{-1}) = \\ (\det \mathbf{A})\{tr(\mathbf{A}^{-1})tr(\mathbf{H}\mathbf{A}^{-1}) - tr(\mathbf{H}\mathbf{A}^{-2})\} = \\ (\det \mathbf{A})\{tr(\mathbf{A}^{-1})\mathbf{A}^{-T} - \mathbf{A}^{-2T}\} \cdot \mathbf{H}, \quad \mathbf{H} \in \text{Lin}. \end{gathered} \tag{2.455}$$

Exercise 41. *Let I_2 :Lin$\to \mathbb{R}$ be the functional defined by $I_2(\mathbf{A}) = \frac{1}{2}[(tr\mathbf{A})^2 - tr(\mathbf{A}^2)]$. Compute $DI_2(\mathbf{A})$.*

Solution.

$$\begin{gathered} I_2(\mathbf{A} + \mathbf{H}) = \frac{1}{2}[(tr(\mathbf{A} + \mathbf{H}))^2 - tr((\mathbf{A} + \mathbf{H})^2)] = \\ I_2(\mathbf{A}) + tr(\mathbf{A})tr(\mathbf{H}) - tr(\mathbf{A}\mathbf{H}) + o(\mathbf{H}), \quad \mathbf{H} \to 0, \end{gathered}$$

then

$$DI_2(\mathbf{A})[\mathbf{H}] = \{tr(\mathbf{A})\mathbf{I} - \mathbf{A}^T\} \cdot \mathbf{H}, \quad \mathbf{H} \in \text{Lin}$$

and

$$DI_2(\mathbf{A}) = tr(\mathbf{A})\mathbf{I} - \mathbf{A}^T.$$

Exercise 42. *Consider f :Lin→ $\mathbb{R}$ defined by*

$$f(\mathbf{K}) = ||\mathbf{K}_0||^2 - \rho^2(tr\mathbf{K}), \tag{2.456}$$

with $\mathbf{K}_0 = \mathbf{K} - (tr\mathbf{K})\mathbf{I}/3$ the deviator of $\mathbf{K}$ and $\rho : \mathbb{R} \to \mathbb{R}$ a continuous and differentiable function. Compute $Df(\mathbf{K})$.

Solution. Let us remark that $D\mathbf{K}_0(\mathbf{K})[\mathbf{U}] = \mathbf{U}_0$; having in mind that $||\mathbf{K}_0||^2 = \mathbf{K}_0 \cdot \mathbf{K}_0$, exploiting the chain rule, we obtain

$$Df(\mathbf{K})[\mathbf{U}] = 2(\mathbf{K}_0 - 2\rho(tr\mathbf{K})\rho'(tr\mathbf{K})\mathbf{I}) \cdot \mathbf{U}, \quad \text{for each } \mathbf{U} \in \text{Lin}, \tag{2.457}$$

where ρ' denotes the derivative of ρ with respect to the independent variable.

Exercise 43. *For each integer $k \geq 1$ let us consider the functional τ_k :Lin→ $\mathbb{R}$ defined by $\tau_k(\mathbf{A}) = tr(\mathbf{A}^k)$, with $\mathbf{A}^k$ given in (2.9). Prove that*

$$D\tau_k(\mathbf{A}) = k(\mathbf{A}^{k-1})^T. \tag{2.458}$$

Solution. We can prove the following relation by induction

$$(\mathbf{A} + \mathbf{H})^k = \mathbf{A}^k + \sum_{i=0}^{k-1} \mathbf{A}^i\mathbf{H}\mathbf{A}^{k-1-i} + o(\mathbf{H}), \quad \mathbf{H} \to 0. \tag{2.459}$$

Calculating the trace of both sides of (2.459) we get that

$$D\tau_k(\mathbf{A})[\mathbf{H}] = k\ tr(\mathbf{H}\mathbf{A}^{k-1}), \quad \mathbf{H} \in \text{Lin}, \tag{2.460}$$

from which the thesis follows.

Exercise 44. *Given a second-order tensor $\mathbf{L}$, for each integer $k \geq 1$ consider the functional ψ_k :Lin→ $\mathbb{R}$ defined by $\psi_k(\mathbf{A}) = tr(\mathbf{A}^k\mathbf{L})$. Compute $D\psi_k(\mathbf{A})$.*

Solution. In view of (2.459) we have

$$\psi_k(\mathbf{A} + \mathbf{H}) = \psi_k(\mathbf{A}) + tr(\sum_{i=0}^{k-1} \mathbf{A}^i\mathbf{H}\mathbf{A}^{k-1-i}\mathbf{L}) + o(\mathbf{H}), \quad \mathbf{H} \to 0, \tag{2.461}$$

from which it follows that

$$D\psi_k(\mathbf{A})[\mathbf{H}] = \sum_{i=0}^{k-1} (\mathbf{A}^i\mathbf{L}\mathbf{A}^{k-1-i})^T \cdot \mathbf{H}, \quad \mathbf{H} \in \text{Lin}. \tag{2.462}$$

Now, we can prove the following proposition, which expresses the invariance of the derivative of a function invariant with respect to a subset $\mathfrak{I} \subset$Orth.

Proposition 50. *Let $\mathfrak{I}$ be a subset of Orth, $\mathcal{A}$ an open set contained in a subspace $\mathcal{U}$ of Lin, with $\mathcal{A}$ invariant with respect to $\mathfrak{I}$. Assume that $G : \mathcal{A} \to$Lin is invariant with respect to $\mathfrak{I}$ and of class C^1. Then,*

$$\mathbf{Q}DG(\mathbf{A})[\mathbf{U}]\mathbf{Q}^T = DG(\mathbf{Q}\mathbf{A}\mathbf{Q}^T)[\mathbf{Q}\mathbf{U}\mathbf{Q}^T], \tag{2.463}$$

for each $\mathbf{A} \in \mathcal{A}, \mathbf{U} \in \mathcal{U}, \mathbf{Q} \in \mathfrak{I}$.

Proof. Let us start by proving that $\mathcal{U}$ is invariant with respect to $\mathfrak{I}$. Take $\mathbf{U} \in \mathcal{U}, \mathbf{A} \in \mathcal{A}, \mathbf{Q} \in \mathfrak{I}$, since $\mathcal{A}$ is open, there exists $\alpha > 0$ such that $\mathbf{A}+\alpha\mathbf{U} \in \mathcal{A}$. From

$$\mathbf{Q}(\mathbf{A} + \alpha\mathbf{U})\mathbf{Q}^T = \mathbf{Q}\mathbf{A}\mathbf{Q}^T + \alpha\mathbf{Q}\mathbf{U}\mathbf{Q}^T \in \mathcal{A} \subset \mathcal{U} \tag{2.464}$$

taking into account that $\mathcal{U}$ is a subspace, we get that $\mathbf{Q}\mathbf{U}\mathbf{Q}^T \in \mathcal{U}$.

Given $\mathbf{A} \in \mathcal{A}, \mathbf{U} \in \mathcal{U}, \mathbf{Q} \in \mathfrak{I}$, we ha

$$G(\mathbf{Q}(\mathbf{A} + \mathbf{U})\mathbf{Q}^T) = G(\mathbf{Q}\mathbf{A}\mathbf{Q}^T + \mathbf{Q}\mathbf{U}\mathbf{Q}^T)$$

$$= G(\mathbf{Q}\mathbf{A}\mathbf{Q}^T) + DG(\mathbf{Q}\mathbf{A}\mathbf{Q}^T)[\mathbf{Q}\mathbf{U}\mathbf{Q}^T] + o(\mathbf{U}), \quad \mathbf{U} \to \mathbf{0}; \tag{2.465}$$

on the other hand, since G is invariant with respect to $\mathfrak{I}$ we have

$$G(\mathbf{Q}(\mathbf{A} + \mathbf{U})\mathbf{Q}^T) = \mathbf{Q}G(\mathbf{A} + \mathbf{U})\mathbf{Q}^T, \tag{2.466}$$

and

$$\mathbf{Q}G(\mathbf{A} + \mathbf{U})\mathbf{Q}^T = \mathbf{Q}G(\mathbf{A})\mathbf{Q}^T + \mathbf{Q}\mathbf{D}G(\mathbf{A})[\mathbf{U}]\mathbf{Q}^T + o(\mathbf{U}), \quad \mathbf{U} \to \mathbf{0}. \tag{2.467}$$

Comparing the relations (2.465) and (2.467) we finally get (2.463). □

Exercise 45. *Let T :Inv→Inv be the function defined by*

$$T(\mathbf{V}) = \mu(\mathbf{V}\mathbf{V}^T - \mathbf{I}) + \lambda[(\det \mathbf{V})^2 - 1]\mathbf{I}, \quad \textit{con } \mu, \lambda \in \mathbb{R}; \tag{2.468}$$

compute $DT(\mathbf{V})$.

Solution. We have

$$DT(\mathbf{V})[\mathbf{H}] = \mu(\mathbf{V}\mathbf{H}^T + \mathbf{H}\mathbf{V}^T) + 2\lambda(\det \mathbf{V})^2\mathbf{I} \otimes \mathbf{V}^{-T}[\mathbf{H}], \quad \mathbf{H} \in \text{Lin}. \tag{2.469}$$

Let T :Lin→Lin be a differentiable function. For $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ an orthonormal basis of $\mathcal{V}$ and $\mathbf{A} \in$ Lin, we want to calculate the components of the fourth-order tensor $DT[\mathbf{A}]$. Taking into account that

$$\mathbf{A} = \sum_{i,j=1}^{n} A_{ij}\mathbf{e}_i \otimes \mathbf{e}_j \tag{2.470}$$

from the differentiability of T, we get

$$T(\mathbf{A} + \alpha\mathbf{e}_k \otimes \mathbf{e}_l) = T(\mathbf{A}) + \alpha DT(\mathbf{A})[\mathbf{e}_k \otimes \mathbf{e}_l] + o(\alpha), \quad \alpha \to 0, \tag{2.471}$$

and,

$$(\mathbf{e}_i \otimes \mathbf{e}_j) \cdot T(\mathbf{A} + \alpha \mathbf{e}_k \otimes \mathbf{e}_l) = (\mathbf{e}_i \otimes \mathbf{e}_j) \cdot T(\mathbf{A})$$
$$+\alpha(\mathbf{e}_i \otimes \mathbf{e}_j) \cdot DT(\mathbf{A})[\mathbf{e}_k \otimes \mathbf{e}_l] + o(\alpha), \quad \alpha \to 0, \tag{2.472}$$

from which we get

$$\mathbf{e}_i \otimes \mathbf{e}_j \cdot DT(\mathbf{A})[\mathbf{e}_k \otimes \mathbf{e}_l] = \lim_{\alpha \to 0} \frac{T(\mathbf{A} + \alpha \mathbf{e}_k \otimes \mathbf{e}_l)_{ij} - T(\mathbf{A})_{ij}}{\alpha}, \tag{2.473}$$

and finally,

$$DT(\mathbf{A})_{ijkl} = \frac{\partial T(\mathbf{A})_{ij}}{\partial A_{kl}}. \tag{2.474}$$

2.21 Derivatives of functions defined over an open set of $\mathbb{R}$

The following proposition follows directly from (1.159).

Proposition 51. *Given an open set $\mathcal{D}$ of $\mathbb{R}$ let*

$$\varphi : \mathcal{D} \to \mathbb{R},$$

$$\mathbf{v}, \mathbf{w} : \mathcal{D} \to \mathcal{V},$$

$$\mathbf{A}, \mathbf{B} : \mathcal{D} \to Lin,$$

be functions of class C^1. Then

$$(\varphi \mathbf{v})^{\cdot} = \varphi \dot{\mathbf{v}} + \dot{\varphi} \mathbf{v}, \tag{2.475}$$

$$(\mathbf{v} \cdot \mathbf{w})^{\cdot} = \mathbf{v} \cdot \dot{\mathbf{w}} + \dot{\mathbf{v}} \cdot \mathbf{w}, \tag{2.476}$$

$$(\mathbf{A}\mathbf{B})^{\cdot} = \mathbf{A} \cdot \dot{\mathbf{B}} + \dot{\mathbf{A}} \cdot \mathbf{B}, \tag{2.477}$$

$$(\mathbf{A} \cdot \mathbf{B})^{\cdot} = \mathbf{A} \cdot \dot{\mathbf{B}} + \dot{\mathbf{A}} \cdot \mathbf{B}, \tag{2.478}$$

$$(\mathbf{A}\mathbf{v})^{\cdot} = \mathbf{A}\dot{\mathbf{v}} + \dot{\mathbf{A}}\mathbf{v}, \tag{2.479}$$

$$(\mathbf{v} \wedge \mathbf{w})^{\cdot} = \mathbf{v} \wedge \dot{\mathbf{w}} + \dot{\mathbf{v}} \wedge \mathbf{w}, \tag{2.480}$$

$$(\mathbf{v} \otimes \mathbf{w})^{\cdot} = \mathbf{v} \otimes \dot{\mathbf{w}} + \dot{\mathbf{v}} \otimes \mathbf{w}, \tag{2.481}$$

$$(\varphi \mathbf{A})^{\cdot} = \varphi \dot{\mathbf{A}} + \dot{\varphi} \mathbf{A}. \tag{2.482}$$

For $n = 3$, let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be an orthonormal basis of $\mathcal{V}$, given the vector function $\mathbf{v}(t)$ and the tensor function $\mathbf{A}(t)$, we have

$$\dot{\mathbf{v}}(t) = \sum_{i=1}^{3} \dot{v}_i(t)\mathbf{e}_i, \quad \dot{\mathbf{A}}(t) = \sum_{i,j=1}^{3} \dot{A}_{ij}(t)\mathbf{e}_i \otimes \mathbf{e}_j. \tag{2.483}$$

If $\mathbf{A}(t)$ is a non-null tensor function, for $\varphi(t) = ||\mathbf{A}(t)||$, we have

$$\dot{\varphi}(t) = \frac{\mathbf{A}(t)}{||\mathbf{A}(t)||} \cdot \dot{\mathbf{A}}(t), \tag{2.484}$$

from which we get the identity

$$(||\mathbf{A}(t)||)^{\cdot} = \frac{\mathbf{A}(t)}{||\mathbf{A}(t)||} \cdot \dot{\mathbf{A}}(t). \tag{2.485}$$

Proposition 52. *Let $\mathcal{D}$ be on open set of $\mathbb{R}$ and $\mathbf{B} : \mathcal{D} \to Lin$ be a function of class C^1. We have*

$$(\mathbf{B}^T)^{\cdot} = (\dot{\mathbf{B}})^T. \tag{2.486}$$

Moreover, if $\mathbf{B}(t)$ is invertible for each $t \in \mathcal{D}$, we have

$$(\det \mathbf{B})^{\cdot} = (\det \mathbf{B}) tr(\dot{\mathbf{B}}\mathbf{B}^{-1}), \tag{2.487}$$

and

$$(\mathbf{B}^{-1})^{\cdot} = -\mathbf{B}^{-1}\dot{\mathbf{B}}\mathbf{B}^{-1}. \tag{2.488}$$

Proof. Let L :Lin→Lin be the linear function defined by $L(\mathbf{A}) = \mathbf{A}^T$, $\mathbf{A} \in$Lin. Since $DL(\mathbf{A}) = L$, from the chain rule ir follows that

$$(\mathbf{B}^T)^{\cdot} = (L(\mathbf{B}))^{\cdot} = L(\dot{\mathbf{B}}) = (\dot{\mathbf{B}})^T.$$

From the relations (1.161) and (2.445), for $\varphi(\mathbf{B}) = \det \mathbf{B}$ we have

$$(\varphi(\mathbf{B}(t)))^{\cdot} = D\varphi(\mathbf{B}(t))[\dot{\mathbf{B}}(t)] = (\det \mathbf{B}(t)) tr\left(\dot{\mathbf{B}}(t)\mathbf{B}(t)^{-1}\right).$$

□

Exercise 46. *Assume that $\mathbf{Q} : \mathbb{R} \to Orth$ is differentiable. Show that*

$$\mathbf{Q}(t)\dot{\mathbf{Q}}(t)^T \in \; Skw \quad \textit{for each } t \in \mathbb{R}. \tag{2.489}$$

Solution. Consider L :Lin→Lin such that $L(\mathbf{A}) = \mathbf{A}^T$, since $\mathbf{Q}(t)\mathbf{Q}(t)^T = \mathbf{I}$ for each $t \in \mathbb{R}$, differentiating with respect to t we get

$$\mathbf{0} = \dot{\mathbf{Q}}\mathbf{Q}^T + \mathbf{Q}DL(\mathbf{Q})[\dot{\mathbf{Q}}] = \dot{\mathbf{Q}}\mathbf{Q}^T + \mathbf{Q}\dot{\mathbf{Q}}^T, \tag{2.490}$$

from which the thesis follows.

Bibliography

[1] R. M. Bowen, C. C. Wang, *Introduction to Vectors and Tensors*, Volumes 1 and 2. Plenum Press, New York 1976.

[2] H. Brezis, *Analyse Fonctionnelle. Theorie et Applications*, Masson Editeur 1983.

[3] V. Checcucci, A. Tognoli, E. Vesentini. *Lezioni di Topologia Generale*. Feltrinelli, 1977.

[4] C. L. DeVito, *Functional Analysis and Linear Operator Theory*. Addison-Wesley Publishing Company, 1990.

[5] J. Dieudonné, *Foundations of Modern Analysis*. Academic Press, New York and London, 1969.

[6] M. E. Gurtin, *An Introduction to Continuum Mechanics*. Academic Press, 1981.

[7] M. E. Gurtin, *The Linear Theory of Elasticity*. Encyclopedia of Physics, vol. VIa/2 Mechanics of Solids II, Springer-Verlag, 1972.

[8] M. E. Gurtin, E. Fried, L. Anand, *The Mechanics and Thermodynamics of Continua*. Cambridge University Press, 2010.

[9] P. R. Halmos, *Finite-Dimensional Vector Spaces*. Springer-Verlag, 1987.

[10] M. Itskov, *Tensor Algebra and Tensor Analysis for Engineers*. Springer-Verlag, 2007.

[11] J. L. Kelley, *General Topology*. Van Nostrand, 1955.

[12] P. Podio-Guidugli, A primer in elasticity. *Journal of Elasticity* **58**(1), 2000.

[13] M. Silhavy, *The Mechanics and Thermodynamics of Continuous Media*. Text and Monographs in Physics, Springer-Verlag Berlin Heidelberg 1997.