MATHEMATICAL FOUNDATIONS OF ELEMENTARY CONTINUUM MECHANICS

Roberto Dias Algarte

2023

Springer-Verlag Berlin Heidelberg GmbH

Synopsis of “Mathematical Foundations of Elementary Continuum Mechanics”

The main purpose of the book is to present the fundamental theoretical aspects of Continuum Mechanics in a comprehensive mathematical approach. In order to achieve this goal, the study is divided in two parts. The first one presents the mathematical elements, starting from the basics of Algebra – Abstract and Linear – and then advancing to Tensor Algebra. After that, geometrical concepts are included in this algebraic context by presenting topics of Affine Geometry. Then, fundamental and advanced topics on the Calculus of Tensor Functions are studied. The second part deals with the mechanical introductory aspects of continua, using all the strong mathematical background presented in the first part. The study on this second part starts with a brief biobibliographical presentation of some important authors of the early years of Continuum Mechanics, and then develops the Kinematics and Dynamics of continua. Finally, the subject of Constitutive Equations is exemplified by introducing Elasticity Theory. The book strives to be mathematically rigorous and as self-contained as possible, requiring the reader to be skilled on the basics of Linear Algebra and Calculus. There are many figures and a few tables.

Brasília, February 2023
Roberto Dias Algarte

Contents

List of Theorems and Corollaries

List of Figures

Part I

Mathematical Foundations

Chapter

1 Collections and Relationships

It is at least innocuous to study a fundamental object which is completely isolated, without getting it together with other objects. Resulting or not from some selection criteria, this gathering of objects that defines a scope of study, a comprehensiveness of analysis, we shall call it a collection. Thereby, between these collected objects, called **elements**, it is possible to establish relationships; even when these elements belong to different collections. A **rule** expresses generically how such relationship must occur and, because of this generic character, we describe it, together with collections and elements, in an algebraic approach. Thus, before developing the main subject of this chapter, let's start with some basic remarks about Algebra.

1.1 What is Algebra?

At some moment during the development of our vital urge for communication and our capacity to apprehend the world around, we humans felt the need to transmit less subjectively, *without dubieties*, some of the sense impressions captured from the physical environment. Chronologically speaking, the first two human actions that emerged out of this necessity were to count and then to measure. In these ancient times, the beautiful art of Mathematics arose as a result of the effort to codify quantities and sizes in a particular symbology, where subjective interpretations were minimized or simply suppressed.

The ever growing sophistication of life demands brought complexity to the problems of Mathematics and the need for generalization emerged naturally, as a strategy to broaden its applicability, to simplify its description and solution methods. If there are conceptual structures that occur repeatedly on the treatment of different problems, then the **generalization** of theses structures can make them applicable to new problems, to creating new structures or simplifying others. In this context, the act of generalizing encompasses three distinct actions: a) to abstract, when we mentally extract from the global structure some part or substructure of interest, in order to concentrate only on it; b) to analyze, when we decompose the abstraction to understand it better; c) to conceptualize, when we create new concepts from analyzing the abstraction. The product or result of the generalization process is called the **abstract**, a noun; and it is precisely the abstract – as a rigorously conceived mental construct – that allows Mathematics to describe the concrete with fewer, less complex, structures. The human endless quest for knowledge and the unending practice of generalization over generalization, resulting in an increasingly stable and generic abstract, bestow upon Mathematics a strong psychological character if we consider it as an observable expression of the deepest manifestations of the human mind. From this point of view, Mathematics, such as Painting or Music, is an inherent part of human nature; in other words, it is undoubtedly *art*.

From the ideas already exposed, we state that **Algebra** is the branch if Mathematics that deals with *generalizations of structures formed by symbols and their collections, by relationships between these symbols and by restrictions governing these relationships*. As an example, letters representing real numbers, sets of these letters, functions that have these letters as arguments and the rules expressing these functions are respectively the symbols, collections, relationships and restrictions that constitute the fundamental objects of Algebra. It is surely a wide field of study, so wide that Algebra, with its deep generic structures, approximated other branches of Mathematics seemingly distant from each other. This aggregative aspect that pervades different areas of study allows us to regard Algebra as a fundamental mathematical branch, one of its pillars, both because of its theoretical significance and also for enabling mathematical knowledge as a whole.

As a consequence of its influence in other branches and also of didactic particularizations, Algebra has many divisions. Among them, the most important are **elementary algebra** and **abstract algebra**: the former deals with the lowest generalization level of the **Arithmetics** and the latter reaches deeper and wider generalizations. As Mathematics is the art of abstraction, then the adjective in "abstract algebra" is indeed a pleonasm, in order to emphasize its non-numeric, non-specific symbolic character. In this work, we shall mostly study **linear algebra**, a subdivision of abstract algebra that deals with vectors (symbols), vector spaces (collections) and linear functions (relationships and rules).

Now, let's talk a little about historical matters. The first known record closest to the current algebraic thought was written by the greek mathematician Diophantus of Alexandria on the third century A.D. From this work, entitled *Arithmetica*, composed originally of many books, only 189 problems remain, all expressed in a specific notation, very similar to the current practice of writing equations: the unknowns represented by non numeric symbols with an equality separating the operations. The equation currently expressed by

$$x^3 - 2x^2 + 10x - 1 = 5$$

Diophantus wrote it the following way[1]:

$$K^{Y}\overline{\alpha}\sigma\overline{\iota}\pitchfork\Delta^{Y}\overline{\beta}\mathbf{M}\overline{\alpha}{'}\acute{\iota}\sigma\mathbf{M}\overline{\varepsilon},$$

where the symbols with overbars are numerical constants, $'\acute{\iota}\sigma$ means "equals to", $\pitchfork$ represents difference an the other symbols are related to the unknown σ. Because of this symbolic approach – unprecedented until that time, according to known records – in handling problems, some historians consider Diophantus the father of Algebra. Many others argue that the work of Diophantus did not bring methodological evolution to solving the purposed problems: every solution is applicable specifically, valid only for each particular case. There is no effort for generalization, for creating solution procedures extensible to different problems.

Six hundred years after the *Arithmetica* of Diophantus, around 820 A.D., the persian polymath Abū 'Abd Muhammad Ibn Mūsā al-Khwārizmī[2] (780–850), who was the head of the famous House of Wisdom in Baghdad, wrote *Al-kitāb al-mukhtasar fī hisāb al-ǧabr wa'l-muqābala*, or *Handbook of "al-jabr" and of "al-muqabala"* in a literal translation. There are no words in english language that express accurately the two transliterated arab words in quotation marks, whose meaning can be understood from the methodology proposed by the author. The book has three parts, with the first one devoted to solving quadratic and linear equations, reducible to one of the six following types.

1. Squares equal roots: $ax^2 = bx$;
2. Squares equal numbers: $ax^2 = c$;
3. Roots equal numbers: $bx = c$;
4. Squares and roots equal numbers: $ax^2 + bx = c$;
5. Squares and numbers equal roots: $ax^2 + c = bx$;
6. Roots and numbers equal squares: $bx + c = ax^2$.

These generic problems, al-Khwārizmī did not solve them using a symbolic notation, as Diophantus did, but in literal terms, just like the descriptions in the six items. To each of these problems, the author created literal solution methods, applicable to any specific problem reducible to one of the six types. In order to do this, he proposed two procedures, presented as follows.

[1] See Derbyshire[15].

[2] According to Knuth[31], the name means "*Father of Abdullah, Mohammad, son of Moses, native from Khwārizmī*", southern region of the Sea of Aral.

a) "al-jabr" involves the acts of adding to the side where there is a subtraction a value that "restores" the subtracted term and of balancing the equation by adding this same value to the other side. The word "al-jabr" is the etymological ancestor of the word "algebra"[3], the latter being constructed from pronouncing the former. Using a symbolic notation, this is the example presented by al-Khwārizmī in his handbook:

$$\begin{aligned} x^2 &= 40x - 4x^2 \\ 5x^2 &= 40x. \end{aligned}$$

b) "al-muqabala" means subtracting both sides by a value that eliminates one of the terms. Here is an example of this procedure from the handbook:

$$\begin{aligned} 50 + x^2 &= 29 + 10x \\ 21 + x^2 &= 10x. \end{aligned}$$

Translations to the latin language of the al-Khwārizmī work in 1145 helped to incorporate the arab mathematics in the western thought. The works of Diophantus and al-Khwārizmī, dealing essentially with the solution of equations, are the two most relevant origins of what today we call Algebra.

1.2 Sets

In conceptual terms, the least restricted algebraic collection is called set, which can have a finite or infinite number of distinct elements, or none[4]. Therefore, from the finite collection $\{a, b, b, a, c\}$, we can define a finite set $\{a, b, c\}$ of three distinct[5] elements. In concrete examples, this distinction – that can be done abstractly by labeling objects with letters – depends on the element characteristics elected to distinguish each other, as in figure 1.1. Sometimes, it is useful to write $\{a, b, b, a, c\}_{\neq} = \{a, b, c\}$ to denote the set defined from the distinct elements of collection $\{a, b, b, a, c\}$. It is important to say that the descriptive sequence of its elements does not alter the definition of a set: for example, definitions $A := \{a, b\}$ and $A := \{b, a\}$ are identical. Speaking of comparisons, the sets A and B are said to be equal, $A = B$, if they have the same elements; otherwise, they are different: $A \neq B$.

[3]In CERVANTES[9], there's an interesting passage at p. 476: *"En esto fueron razonando los dos, hasta que llegaon a un pueblo donde fue ventura hallar un algebrista, con quien se curó...'.* The quote says that Don Quixote and his faithful esquire are lucky to find an algebraist, who was a healer that restored displaced bones. It should be said that the spanish language was strongly influenced by the successive arab invasions coming from the south. In this same novel, Cervantes also states that every spanish word started by "al" has an arabian source.

[4]This is not the approach of the Axiomatic Set Theory. See CAMERON[8].

[5]A set does not admit repeated elements. See SHEN & VERESHCHAGIN[46].

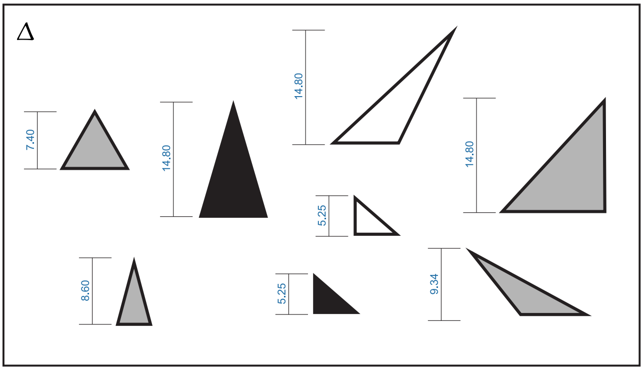

Figure 1.1 – If the selection criteria is "`triangular plates with heights between 5cm and 15cm`", then Δ is a mathematical collection if equality is based only on the feature "number of sides"; additionally, if it is based on "height", Δ is still a collection; but, if it is based also on "fill color", then Δ is a mathematical set.

The **empty set** $\emptyset$ has no elements and it enforces the idea of set as a restricted collection that can be build from some selection criteria: the empty set may be the result of a criteria that no object obeyed. For example, if the selection criteria is "`prime even numbers different from two`", the result will be an empty set. In set theory, this "selection criteria" is called **specification**, whose mathematical syntax is the following:

$$\mathtt{set} := \{ \ \mathtt{selection} \ : \ \mathtt{criteria} \ \}. \tag{1.1}$$

From this syntax pattern and considering x a representation for an arbitrary integer,

$$E := \{x \in \mathbb{Z} : x \bmod 2 = 0\} \tag{1.2}$$

is the set of even numbers. This specification reads "the set E defined by every element of the set of integers whose division by two has a zero remainder´´.

By the intuitive sense of belonging, the most basic relationship between an object and a set determines whether the former is element of the latter or not. This idea has a fundamental importance in the so called Naive Set Theory, which we adopt here, following HALMOS[26]. In mathematical terms, if a is element of the set A, we say that it **belongs** to the set or that $a \in A$; otherwise, it doesn't belong to the set: $a \notin A$. When all the elements of a set A_1 belong to the set A, we say that A_1 is a **subset** of A. If this is the case, when the equality $A_1 = A$ is not admissible, A_1 is called a **proper subset** of A, written also as $A_1 \subset A$; when it is admissible, A_1 is an **improper subset** of A, or $A_1 \subseteq A$, from which we can state that every set is an improper subset of itself. If set A_1 is not subset or proper subset of A, we represent $A_1 \not\subset A$ or $A_1 \not\subseteq A$ respectively. The set $\emptyset$ would not be a subset of an arbitrary set A if it had some element not belonging to A; but this is impossible because $\emptyset$ has no elements, and then we can state that every set has an empty subset.

The concept of belonging just presented can also be used to create sets. A **union** of the sets A_1 and A_2 is the set $A_1 \cup A_2$ to which all the elements of A_1 and A_2 belong. The compact representation $\bigcup_{i=1}^{n} A_i$ is the union of the n sets A_i. A set of elements that belong both to A_1 and to A_2 is the **intersection** $A_1 \cap A_2$. In other words, if the element $x \in A_1 \cap A_2$ then $(x \in A_1) \wedge (x \in A_2)$, where $\wedge$ means "AND" in english. Similarly to union, $\bigcap_{i=1}^{n} A_i$ is how we write the intersection of n sets A_i. By using the so called Venn diagrams (figure 1.2), it can be verified that intersection is distributive in union, that is,

$$A_1 \cap (A_2 \cup A_3) = (A_1 \cap A_2) \cup (A_1 \cap A_3)\,. \tag{1.3}$$

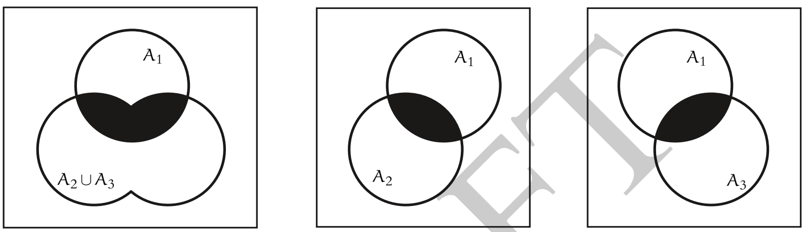

Figure 1.2 – Intersection is distributive in union.

When a set $A_1 \cap A_2$ is empty, we say that A_1 and A_2 are **disjoint**. In this case, if $x \in A_1 \cup A_2$ then $(x \in A_1) \vee (x \in A_2)$, where the symbol $\vee$ means "OR" literally. Concerning disjoint sets, a trivial but important property is the following: if $B \subset \bigcup_{i=1}^{n} A_i$ where A_i and A_j are disjoint whenever $i \neq j$, then it is clear that $B = \bigcup_{i=1}^{n} B \cap A_i$.

The **difference set** $A_1 \setminus A_2$ is the set of elements of A_1 not belonging to A_2. From this definition, if A_1 is subset of A, the set $A'_1 := A \setminus A_1$ is called the **complement** of A_1 in A. Thereby, if A_1 and A_2 are subsets of A, it is possible to verify, also by Veen diagrams, that the union complement $(A_1 \cup A_2)' = A'_1 \cap A'_2$, that the intersection complement $(A_1 \cap A_2)' = A'_1 \cup A'_2$ and that the difference $A_1 \setminus A_2 = A_1 \cap A'_2$. From these three equalities, we can say that in union the difference is distributed according to

$$A_1 \setminus (A_2 \cup A_3) = (A_1 \setminus A_2) \cap (A_1 \setminus A_3)\,, \tag{1.4}$$

and in intersection according to

$$A_1 \setminus (A_2 \cap A_3) = (A_1 \setminus A_2) \cup (A_1 \setminus A_3)\,. \tag{1.5}$$

Proof. Let the sets A_1, A_2, A_3 be propers subsets of A. Therefore, we have equalities $A_1 \setminus (A_2 \cup A_3) = A_1 \cap (A_2 \cup A_3)' = A_1 \cap A'_2 \cap A'_3$. We also have equalities $(A_1 \setminus A_2) \cap (A_1 \setminus A_3) = A_1 \cap A'_2 \cap A_1 \cap A'_3 = A_1 \cap A'_2 \cap A'_3$. The equality (1.4) is thus verified. The following development proves (1.5):

$$\begin{aligned} A_1 \setminus (A_2 \cap A_3) &= A_1 \cap (A_2 \cap A_3)' \\ &= A_1 \cap (A_2' \cup A_3') \\ &= (A_1 \cap A_2') \cup (A_1 \cap A_3') \\ &= (A_1 \setminus A_2) \cup (A_1 \setminus A_3). \end{aligned}$$

□

A set C of sets is called a **class**, here represented by $\mathfrak{C}$. When all the elements of $\mathfrak{C}$ are disjoint then $\mathfrak{C}$ is also called disjoint. If $\mathfrak{C}$ is constituted by subsets of a set M, then $\mathfrak{C}$ is said to be a class of M. A class $\mathfrak{C}$ is **countable** if each of its elements can be uniquely identified by natural numbers, that is, $\mathfrak{C} = \{C_1, C_2, \cdots\}$. Given a non empty class $\mathfrak{M}$, an arbitrary countable class $\{A_1, \cdots, A_n\} \subseteq \mathfrak{M}$ and arbitrary elements $A, B \in \mathfrak{M}$, if both sets $\bigcup_{i=1}^{n} A_i$ and $A \setminus B$ also belong to $\mathfrak{M}$, then this class is called a **ring**. From this definition, if $\mathfrak{M}$ is a ring and we choose $A = B$, it is clear that every ring has the empty set because $A \setminus B = \emptyset$. Considering the arbitrary countable class $\{A_1, A_2, \cdots\}$ an improper subset of ring $\mathfrak{M}$, if $\bigcup_{i=1}^{\infty} A_i \in \mathfrak{M}$, then $\mathfrak{M}$ is called a σ-**ring**.

1.3 Sequences

In Algebra, a collection is called sequence when its elements need to be ordered. For this ordering, each element of the sequence has a position identified by an ordinal number[6], called **index**, which grows from left to right on the sequence notation (a, b, c). Thereby, every element of a sequence has a unique position, labeled by an index; and from this we conclude that two sequences are equal if and only if they have the same elements equally indexed. For example, the sequence $(a, b, c) \neq (a, c, b)$ because the elements b and c have different indexes in each of the sequences. Differently from sets, the positional restriction of sequences does not forbid indistinct elements: the collection a, b, c, a is valid as a sequence (a, b, c, a) since the two a elements have different indexes. When a sequence is finite, it is called a **tuple**. The tuple that has one element is a **monad**; two elements, a **double**; three elements, a **triple**; four elements, a **quadruple**; n elements, a n-**tuple**, where $n \in \mathbb{N}$. There is also an **empty sequence**, called 0-tuple. When two arbitrary elements in a n-tuple, $n > 1$, interchange positions, we called it a **transposition**.

Elements of sets can be used to build sequences. For instance, we can build doubles of the type $(x, x/2)$, where $x \in \mathbb{Z}$ and $x/2 \in \mathbb{Q}$. If there is a collection of sets $A_1, A_2, \cdots, A_n$, not necessarily distinct, n-tuples of the type $(a_1, \cdots, a_n)$, where each element $a_i \in A_i$, can also be built. Thereby, the set of all these constructed n-tuples is called the **cartesian product** of the sets A_i. In mathematical terms, if a collection

[6]Ordinals are integer numbers, elements of $\mathbb{N}$, used to label positions sequentially.

$A_1, A_2, \cdots, A_n$ of sets is given, the set

$$A_1 \times A_2 \times \cdots \times A_n := \{(a_1, a_2, \cdots, a_n) : a_i \in A_i\,, i = 1, \cdots, n\}\,, n > 1\,, \tag{1.6}$$

is their cartesian product. In order to simplify notation, $A_1 \times A_2 \times \cdots \times A_n$ is compacted to $A^{\times n}$. When all the n sets are equal to A, we adopt the format A^n, called **cartesian power**. If one of the terms in a cartesian product is the empty set, the result is also the empty set: $A_1 \times \emptyset = \emptyset \times A_1 = \emptyset$. Since element ordering distinguishes sequences, the cartesian product $A_1 \times A_2$ is commutative only when one of the sets is empty or when they are equal; in other words, if a set $A_1 \neq A_2 \neq \emptyset$, then $A_1 \times A_2 \neq A_2 \times A_1$. Element ordering in sequences also makes the cartesian product non-associative when sets involved are not empty. Thereby, the set $(A \times B) \times C \neq A \times (B \times C)$ because double of double and element differs from double of element and double; that is, the double $((a, b), c) \neq (a, (b, c))$, where a, b, c are arbitrary elements of A, B, C respectively, even when such sets are not disjoint. The cartesian product is distributive in union, intersection and difference of sets. Therefore, given arbitrary sets A, B, C, we can write the following:

$$A \times (B \cup C) = (A \times B) \cup (A \times C)\,; \tag{1.7}$$

$$A \times (B \cap C) = (A \times B) \cap (A \times C)\,; \tag{1.8}$$

$$A \times (B \setminus C) = (A \times B) \setminus (A \times C)\,; \tag{1.9}$$

Proof. Let the sets $A = \{a_1, a_2, \cdots\}$, $B = \{b_1, b_2, \cdots\}$ and $C = \{c_1, c_2, \cdots\}$. In order to verify (1.7), here's the following development:

$$\begin{aligned}\{a_1, a_2, \cdots\} \times (\{b_1, b_2, \cdots\} \cup \{c_1, c_2, \cdots\}) &= \{a_1, a_2, \cdots\} \times (\{b_1, b_2, \cdots, c_1, c_2, \cdots\}) = \\ = \{(a_1, b_1), (a_1, b_2), (a_2, b_1), (a_2, b_2), &\cdots, (a_1, c_1), (a_1, c_2), (a_2, c_1), (a_2, c_2), \cdots\} = \\ &= (\{a_1, a_2, \cdots\} \times \{b_1, b_2, \cdots\}) \cup (\{a_1, a_2, \cdots\} \times \{c_1, c_2, \cdots\})\,.\end{aligned}$$

Equality (1.8) can be demonstrated through the following reasoning: if $(a_1, x) \in A \times (B \cap C)$ then, by the concepts of intersection and cartesian product,

$$(a_1 \in A) \wedge (x \in (B \cap C)) = (a_1 \in A) \wedge (x \in B) \wedge (x \in C)\,.$$

And then the double $(a_1, x) \in (A \times B) \wedge (a_1, x) \in (A \times C)$. This same strategy can be used to prove the last equality: if $(a_1, x) \in A \times (B \setminus C)$ then

$$(a_1 \in A) \wedge (x \in (B \setminus C)) = (a_1 \in A) \wedge (x \in B) \wedge (x \notin C)\,.$$

Therefore, double $(a_1, x) \in (A \times B) \wedge (a_1, x) \notin (A \times C)$. □

1.4 Functions

The act of thinking is fundamentally based on the capacity of making relationships between entities in order to clarify something obscure or, more pretentiously, to disclose the unknown. To correlate entities in algebraic thinking means to describe or

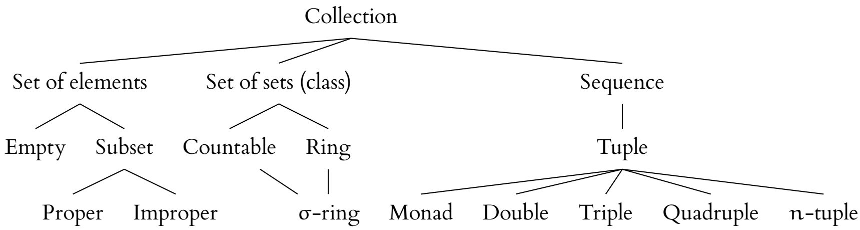

Figure 1.3 – Conceptual hierarchy of sets and sequences.

establish a mathematical link between them. Thereby, among the many types of interactions that justify pairing these entities or objects, there can be links of cause and effect, transformations, dependencies, attributions, associations and so on. These relationships are described by rules that specify, through mathematical expressions, how a certain pairing of objects is done.

The fundamental concept that establishes algebraic relationships we call it function, here understood as *a systematic assignment of one and only one object to each element of a given set*. More precisely, we define function as a double (D,f), where D is this given set, called the **domain** of the function, and f is the rule that implements the so called systematic assignment. In order to avoid confusions, we adopt here the usual notation that considers the rule also a function, that is, the context will define if f is a rule or a function (D,f). The object related to an element $d \in D$ is represented by $f(d)$, called the value of function f in d, which allows us to write the fundamental characteristic of functions, namely,

$$f(d_1) \neq f(d_2) \implies d_1 \neq d_2,\ \forall d_1, d_2 \in D. \tag{1.10}$$

When we want to emphasize the domain D of function f, we use the combined notation D_f. There is also an alternative notation for the function f that makes its domain explicit: $d \mapsto f(d)$, where each element $d \in D$ is related to a value $f(d)$ by f in terms already described. Additionally, it is important to say that just like any other set, classes can also be the domain of functions, which are usually called **set functions**.

The description of f, as a rule, is done through an algebraic expression, where the symbol that represents an arbitrary element of the domain is called a **variable**. For example, let $(\mathbb{R}, f)$ be a function and

$$f(x) = x^2 + 2, \tag{1.11}$$

where variable x represents an arbitrary real value. This sentence says that the value of function f, on the left side, equals the value of the algebraic expression on the right. We also say that variable x is the **argument** of f. Concerning the number of arguments, if the domain $D_f = A^{\times n}$, then f is said to be a **univariate** function when $n = 1$, **bivariate** when $n = 2$, **trivariate** when $n = 3$ or **multivariate** when $n > 3$.

Besides domain, there are at least two additional special sets when we study functions. The first one, represented by R_f, is the **image** of the function f, defined by all the values of f; in other words,

$$R_f := \{f(d) : d \in D_f\}. \tag{1.12}$$

The second one arises when we want to study the part of the function domain, called **preimage**, which is related to a certain subset of the image. In other words, given a subset $B \subseteq R_f$, the preimage of B is the set $R_B^{-1} \subseteq D_f$ such that

$$R_B^{-1} := \{d \in D_f : f(d) \in B\}. \tag{1.13}$$

A function f is said to be **invertible** if it assigns distinct values to its domain elements, resulting in an element-value correlation of one-to-one, called **biunivocal correlation**. In more rigorous terms, f is invertible when

$$d_1 \neq d_2 \Leftrightarrow f(d_1) \neq f(d_2), \forall d_1, d_2 \in D_f. \tag{1.14}$$

When all the values of the invertible function f define an image R_f, the function f^{-1} is called the **inverse** of f if

$$D_{f^{-1}} = R_f \quad \text{and} \quad f^{-1}(f(d)) = d, \forall d \in D_f. \tag{1.15}$$

As a consequence, we can state that f^{-1} is also invertible. Therefore, considering function $g := f^{-1}$ and its image R_g, there exists a function g^{-1} where, according to the previous definition,

$$f(d) = g^{-1}(g(f(d))) = g^{-1}(f^{-1}(f(d))) = g^{-1}(d), \tag{1.16}$$

for an arbitrary element $d \in D_f$. Thereby, we state that f is the inverse function of f^{-1} and then both are the inverse of each other. The superposition of a function and its inverse results in a special function, whose values equals the arguments, called the **identity** function, represented by i. Thereby, since $i(d) = d, \forall d \in D_i$, image $R_i = D_i$, and then we conclude that $i = i^{-1}$ because i is obviously invertible.

1.5 Mappings

There is another algebraic relationship whose main purpose is to relate sets by using functions. In order to do this, it is considered that any function value is also a set element. Thereby, we have a source-set U, on which a function f "acts" and a target-set V, to which all the values of f belong. This relationship is called mapping when the domain $D_f = U$ and the image $R_f \subseteq V$, where V is called the **codomain** of f. In more rigorous terms, a mapping is a triple (U, V, f) where f maps U to V, that is, $u \mapsto f(u) \in V$

for every $u \in U$. Instead of representing a mapping by a tuple, we prefer to notate it as $f: U \mapsto V$, where the arrow makes source–target relationship explicit.

If the image R_f equals codomain V, the mapping $f: U \mapsto V$ is called **surjective** and the function f is a **surjection**. Thereby, we can say that in a surjective mapping notation, the function image is always explicit. Now, when f is an invertible function, the mapping is said to be **injective** and its function to be an **injection**. In this context, considering the image of the injection f, it is possible to define a mapping $f^{-1}: R_f \mapsto U$, where R_f, domain of f^{-1}, is an improper subset of V. The function f can cumulatively be an injection and a surjection, when it is called a **bijection** and its respective mapping a **bijective mapping**. This bijection f invariably implies the existence of the bijective mapping $f^{-1}: V \mapsto U$.

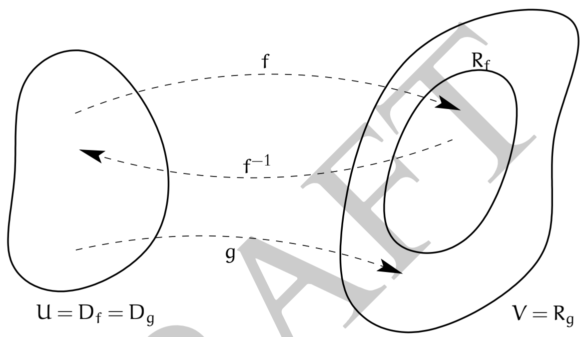

Figure 1.4 – Functions f and g map U to V, where f is an injection and g a surjection. If image R_f were equal to the codomain V, f would be a bijection.

The mapping $f: U \mapsto V$ is said to be an **operation** and its function an **operator** if the domain $U = V^n$. In this case, when n is 1, 2, 3 or 4, operation and operator are classified as **unary**, **binary**, **ternary** and **quaternary** respectively; when $n > 4$, they are called n-**ary**. The arguments of an operator are called **operands** and integer n defines their quantities. It is interesting to note that the unary injective operation $f: V \mapsto V$ is always surjective since the condition of invertibility (1.14) assures that any pair of distinct elements of V is related to a pair of distinct elements of V through f; thereby, image $R_f = V$ and then we can state that any injective unary operator is a bijection. In this case, since unary operator f is a bijection, we can say that if v is an arbitrary element of set V, so is $f(v)$.

For future purposes, it is important to define here the **function graph** of a mapping $f: U^{\times n} \mapsto V$, which is the set of ordered $(n+1)$-tuples

$$\{(x_1, \cdots, x_n, f(x_1, \cdots, x_n)) : x_i \in U_i\} \subseteq U^{\times n} \times V.$$

If U_i and V are sets of real numbers, it is possible to depict pictorially the classical cartesian representation of each tuple $(x_1, \cdots, x_n, f(x_1, \cdots, x_n))$. For example, the set $\{(x_1, x_2, x_1^2 + x_2^2) : x_1, x_2 \in \mathbb{R}\}$ is the function graph of a circular paraboloid.

At the end of the last section, we talked superficially about "superposition" of functions, namely, when a function has a function value as argument. This important concept, in more precise terms, can be presented as follows. Let's say that $g : U \mapsto V$, $f : V \mapsto W$ and $h : U \mapsto W$ are mappings where $h(u) = f(g(u))$, for all $u \in U$. In this context, it is said that h is a **composite function** of f and g, usually represented by $f \circ g$. Note that the composition of functions is not generally commutative, except in particular mappings whose functions and domains permit. Moreover, the following properties are valid:

i. Given $k : W \mapsto L$, we have $k \circ (f \circ g) = (k \circ f) \circ g$;

ii. If f and g are bijections, $f \circ g$ is also a bijection and

$$\begin{aligned} (f \circ g)^{-1} &= g^{-1} \circ f^{-1}, \\ f \circ f^{-1} &= i_W, \\ f^{-1} \circ f &= i_V ; \end{aligned}$$

iii. $f \circ i_V = i_W \circ f = f$.

Proof. For the second item, if the relationship between v and u is biunivocal in $v = g(u)$ for all $v \in V$, then the relationship between u and w is also biunivocal in $f^{-1}(w) = g(u)$ or $w = f(g(u))$ for all $w \in W$; from where $f \circ g$ results a bijection. Now, considering u, v and w arbitrary elements of U, V and W respectively, the first equality in ii is verified as follows:

$$(f \circ g)^{-1}(w) = u = g^{-1}(v) = g^{-1}(f^{-1}(w)) = g^{-1} \circ f^{-1}(w).$$

The other equalities on the list can be easily proved from the definition of composite functions. □

1.6 Groups

The most fundamental algebraic entity that gathers the concepts of collection and relationship is called group, defined by a set and a mapping: a pair of set elements is related to an element of the same set by a mapping, namely, a binary operation that must obey certain restrictions. In other words, when a set defines a group, a double of set elements is functionally related to an element of the same set.

Addition and multiplication of real numbers, composition of invertible functions, subtraction of integers are all examples of mathematical combinations that the concept of group generalizes: they are all associative, they admit identity and inverse elements. Thereby, we can now define in more rigorous terms these intuitive concept. Let G be a non-empty set and $* : G^2 \mapsto G$ a binary operation from which the notation $*(g_1, g_2)$ is shortened to $g_1 * g_2$, where $g_1, g_2 \in G$. The double $(G, *)$ is called a group when the following axioms are valid:

i. Associativity, where $g_1 * (g_2 * g_3) = (g_1 * g_2) * g_3 , \forall g_1, g_2, g_3 \in G$;

ii. Identity element, if $\exists!\, e \in G$ such that $g_1 * e = e * g_1 = g_1 , \forall g_1 \in G$;

iii. Inverse element, if $\exists! b \in G$ such that $g_1 * b = b * g_1 = e$, $\forall g_1 \neq e$.

As an example, the set P of all unary invertible operators whose (co)domain is V, defines a group $(P, \circ)$, according to the properties of function composition.

When it is convenient to make the operator explicit, we shall represent a group by $(G, *)$; otherwise, we'll refer to a group G, indistinct from a set, in order to avoid abuse of notation. Thereby, considering this group G, there are operations $*$, like addition and multiplication of real numbers, that also obey the axiom of

i. Commutativity: $g_1 * g_2 = g_2 * g_1, \forall g_1, g_2 \in G$,

from which the group G becomes **commutative** or **abelian**. In contrast, groups in which the order of operands affects the operator value are called **non-abelian** or **non-commutative**. The abelian group that implements the generalized concept of addition is called **additive** and of multiplication, **multiplicative**, both represented respectively by $(G, +)$ and $(G, \cdot)$. We adopt the notations g_1^{-1} and $-g_1$ as the inverse elements of $g_1 \in G$ in multiplication and addition respectively.

Groups can also be used to define mappings. When this happens, the function must admit as argument any value of the operation involved, since the set of all these values is the domain itself because the inverse and identity element properties force operator $*$ to be surjective. In other words, given a mapping $h : G \mapsto W$, where the sets involved define groups $(G, *)$ and $(W, \rtimes)$, we have each element $g \in G$ as a value of some $g_1 * g_2$, where $g_1, g_2 \in G$. Therefore, it is evident that the value $h(g) = h(g_1 * g_2)$. The operation in group W can take the elements $h(g_1)$ and $h(g_2)$ of W as operands, that is, $h(g_1) \rtimes h(g_2)$, and be defined to present a resulting value of $h(g_1 * g_2) \in W$. Moreover, if h maps the identity element of G to the identity element of W, we say that these two groups, in an operational context, are structurally similar or **homomorphic** in h. In mathematical terms, the function in $h : G \mapsto W$ is said to be a **group homomorphism** if it makes G and W homomorphic, that is, if

i. $h(g_1 * g_2) = h(g_1) \rtimes h(g_2), \forall g_1, g_2 \in G$ and

ii. $h(e_G) = e_W$, where $e_G \in G$ and $e_W \in W$ are identity elements.

A bijection-homomorphism is named **isomorphism** and the groups involved are **isomorphic** in h. If the function in mapping $f : G \mapsto G$ is an isomorphism, then f is called an **automorphism**.

Now, let's consider the mapping $k : G^{\times n} \mapsto W$ whose domain is the cartesian product of n sets, each of them defining a group. In this case, the function k can be called a group homomorphism if, for a group G_i and arbitrary elements $g_{i_1}, g_{i_2} \in G_i$,

$$k(g_1, \cdots, g_{i_1} * g_{i_2}, \cdots, g_n) = k(g_1, \cdots, g_{i_1}, \cdots, g_n) \rtimes k(g_1, \cdots, g_{i_2}, \cdots, g_n) \tag{1.17}$$

and also

$$k(e_{G_1},\cdots,e_{G_n}) = e_W. \tag{1.18}$$

Similarly, k is considered an isomorphism if it is a bijection-homomorphism.

In our study, we shall need to establish a relation between a group, which is a set with an algebraic character, whose elements can be operated, and a set with a geometric character, where sizes, shapes and positions can be observed. A strategy to consistently accomplish this algebraic-geometric relation is to use a function called **group action**. Thereby, let $\varphi : G \times B \mapsto B$ be a mapping where G is a group and B is any non empty set. Function φ is said to be a group action[7] of the set G on B if the following axioms are valid:

i. Identity element, $\varphi(e,b) = b, \forall b \in B$, and

ii. Associativity, $\varphi(g_1,\varphi(g_2,b)) = \varphi(g_1 * g_2, b), \forall b \in B, \forall g_1, g_2 \in G$.

When φ observe these axioms, the structure of B is preserved – now called the G-**set** of G – since its fundamental mathematical attributes remain unaltered. Thereby, given arbitrary elements $b_1, b_2 \in B$, the group action φ is classified as **simply transitive** if

$$\exists!\, g \in G \text{ such that } \varphi(g,b_1) = b_2. \tag{1.19}$$

From these definition, *in a simply transitive group action of* G *on* B*, there is a biunivocal relationship between sets* G *and* B^2*. Moreover, when an element of* B *is fixed in the domain, there results a biunivocal correspondence between the elements of* G *and* B*.*

There can be an abelian group F that is simultaneously additive and multiplicative, on which the multiplication of additions results the addition of multiplications, or in other words, on which the following distributivity is valid:

$$\alpha \cdot \sum_{i=1}^{n} \alpha_i = \sum_{i=1}^{n} \alpha \cdot \alpha_i \,, \forall \alpha, \alpha_i \in F \text{ and } n \in \mathbb{N}^+. \tag{1.20}$$

In these circumstances, the triple $(F,+,\cdot)$ is called a **field**, abbreviated by $\mathbb{F}$. When it is convenient, in order to simplify notation, $\mathbb{F}$ will represent the definer set F. An element of $\mathbb{F}$ is named a **scalar**, of which $\beta \in \mathbb{R}$ and $\gamma \in \mathbb{C}$ are examples. Henceforth, for the purposes of our study, an arbitrary field $\mathbb{F}$ will always refer to either a complex field $\mathbb{C}$ or a real field $\mathbb{R}$, that is, $\mathbb{F} \in \{\mathbb{R},\mathbb{C}\}$. In this sense, where the real field will not be considered a subset of the complex field[8], when $\mathbb{F}$ is $\mathbb{R}$, we must define that the conjugate $\overline{\alpha} = \alpha$, the real part $\Re(\alpha) = \alpha$ and imaginary part $\Im(\alpha) = 0$, for all $\alpha \in \mathbb{F}$. Moreover, considering f an arbitrary scalar valued function and $\overline{f}(x) := \overline{f(x)}$, we have $\overline{f} = f$ in the case of $\mathbb{F} = \mathbb{R}$.

[7] In more precise terms, φ in this case is called a **left group action** in contrast to a **right group action** $\tilde{\varphi}$ where $\tilde{\varphi} : B \times G \mapsto B$. In our study, we shall use the left action.

[8] Since the complex field does not allow the concepts of "greater than" and "less than".

1.7 Arrays

We already learned that sequences are collections of ordered elements; or, more precisely, elements arranged in a queue, where each one has a position labeled by an index. Generalizing this idea, we shall build collections whose elements are arranged in a higher number of perspectives, like a table arrangement, for instance. Thereby, let H be a collection of scalars, arranged by the mapping $\mathsf{h} : \mathsf{N}^{\times \mathsf{q}} \mapsto \mathbb{F}$, where the function h is called **addressing** and each $\mathsf{N}_i = \{1, 2, \cdots, \mathsf{n}_i\}$ is a finite subset of $\mathbb{N}^+$ whose elements are ordinals. In this context, we call the collection H an array, represented by H, whose scalars $\mathsf{h}((\mathsf{i}_1, \cdots, \mathsf{i}_q))$ are notated by $H_{\mathsf{i}_1 \cdots \mathsf{i}_q}$, where the subscript show explicitly the element position. The description of perspectives or **dimension** of an array is given by the n-tuple $(\mathsf{n}_1, \cdots, \mathsf{n}_q)$ or, more usually, $\mathsf{n}_1 \times \cdots \times \mathsf{n}_q$, where the number of perspectives q expresses the **order** of the array and n_i the **size** of each perspective.

Arrays can be added and multiplied. If an array $A + B$ is the sum of A and B, these three arrays have equal dimension $\mathsf{n}_1 \times \cdots \times \mathsf{n}_q$ and each element

$$(A+B)_{\mathsf{i}_1 \cdots \mathsf{i}_q} = A_{\mathsf{i}_1 \cdots \mathsf{i}_q} + B_{\mathsf{i}_1 \cdots \mathsf{i}_q} . \tag{1.21}$$

On the other hand, the multiplication $A *_q B$ requires that the last q elements of the dimension of A and the first q elements of the dimension of B are equal; in other words, if array A has dimension $\mathsf{m}_1 \times \cdots \times \mathsf{m}_p \times \mathsf{n}_1 \times \cdots \times \mathsf{n}_q$, then B must have dimension $\mathsf{n}_1 \times \cdots \times \mathsf{n}_q \times \mathsf{l}_1 \times \cdots \times \mathsf{l}_s$. Thereby, $\mathsf{m}_1 \times \cdots \times \mathsf{m}_p \times \mathsf{l}_1 \times \cdots \times \mathsf{l}_s$ is the resulting dimension of array $A *_q B$ and each element

$$(A *_q B)_{\mathsf{i}_1 \cdots \mathsf{i}_p \mathsf{j}_1 \cdots \mathsf{j}_s} = \sum_{\mathsf{k}_1 = 1}^{\mathsf{n}_1} \cdots \sum_{\mathsf{k}_q = 1}^{\mathsf{n}_q} A_{\mathsf{i}_1 \cdots \mathsf{i}_p \mathsf{k}_1 \cdots \mathsf{k}_q} B_{\mathsf{k}_1 \cdots \mathsf{k}_q \mathsf{j}_1 \cdots \mathsf{j}_s} . \tag{1.22}$$

From now on, we shall adopt that $AB := A *_1 B$. It is also very important for our study to present the so called the **scalar product** or **Frobenius product** of arrays: given two arrays A and B with the same dimension $\mathsf{m}_1 \times \cdots \times \mathsf{m}_p$, the scalar product

$$A : B := \sum_{\mathsf{i}_1}^{\mathsf{m}_1} \cdots \sum_{\mathsf{i}_p}^{\mathsf{m}_p} A_{\mathsf{i}_1 \cdots \mathsf{i}_p} \overline{B_{\mathsf{i}_1 \cdots \mathsf{i}_p}} . \tag{1.23}$$

Speaking of scalars, it is also possible to multiply a scalar α and an array C with dimension $\mathsf{n}_1 \times \cdots \times \mathsf{n}_q$ according to

$$(\alpha C)_{\mathsf{i}_1 \cdots \mathsf{i}_q} := \alpha \cdot C_{\mathsf{i}_1 \cdots \mathsf{i}_q} . \tag{1.24}$$

This definition of multiplication by scalars in the case of $\alpha = -1$, from which we can establish the additive inverse $-C$, together with the addition described in (1.21), allow us to state that the set Y of all $\mathsf{n}_1 \times \cdots \times \mathsf{n}_q$ arrays defines an additive group considering the existence of the null array $0 \in Y$, whose elements are all zero. Conversely, since it is not possible to obtain a multiplicative inverse for every element of Y, this set can not define a multiplicative group. Additionally, it is of fundamental importance for upcoming concepts to define $l_1 \times \cdots \times l_q \times n_1 \times \cdots \times n_q$ arrays A^{T} and $A^{\dagger}$, which we call here the **transpose** and the **conjugate transpose** of array A respectively, whose elements

$$A^{\mathrm{T}}_{i_1\cdots i_q j_1\cdots j_q} := A_{j_1\cdots j_q i_1\cdots i_q} \qquad \text{and} \qquad A^{\dagger}_{i_1\cdots i_q j_1\cdots j_q} := \overline{A_{j_1\cdots j_q i_1\cdots i_q}}\,. \tag{1.25}$$

An example of array that is widely used to accomplish indicial notation of sophisticated quantities and operations is called **Levi-Civita Symbol** or **Permutation Symbol**, notated by the letter ϵ, whose scalars are defined the following way:

$$\epsilon_{i_1\cdots i_n} = \begin{cases} (-1)^{\alpha_p(1,\cdots,n)} & \text{if } \exists\, p(1,\cdots,n) = (i_1,\cdots,i_n) \\ 0 & \text{if } \nexists\, p(1,\cdots,n) = (i_1,\cdots,i_n) \end{cases}, \tag{1.26}$$

where function p permutes the n-tuple elements $(1,\cdots,n)$. In this array, the addressing domain is N^n, where the set of ordinals $N = \{1,2,\cdots,n\}$. The term $\alpha_p(1,\cdots,n)$ means the number of transpositions made on $(1,\cdots,n)$, after which the resulting n-tuple is $p(1,\cdots,n)$. Because of its definition, the Levi-Civita Symbol is an array of order n whose dimension is $n \times \cdots \times n$. Arrays having the same size n in each perspective, like the Levi-Civita Symbol, are usually called **hypercubic** of size n. Therefore, ϵ is a particular hypercubic array of size n that also has order n. For instance, when $n = 2$, ϵ can be represented in a tabular arrangement; more precisely,

$$\epsilon = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \tag{1.27}$$

The Levi-Civita symbol enables us to define a notable mapping $\mathrm{Det}: \bar{A} \mapsto \mathbb{F}$, where $\bar{A}$ is the set of all q-th order hypercubic arrays of size n, $\mathbb{F}$ is a real or complex field and function Det is called **hyperdeterminant**[9], whose rule is

$$\mathrm{Det}(X) = \frac{1}{n!} \sum_{i_1^{(1)}=1}^{n} \cdots \sum_{i_n^{(1)}=1}^{n} \cdots \sum_{i_1^{(q)}=1}^{n} \cdots \sum_{i_n^{(q)}=1}^{n} \epsilon_{i_1^{(1)}\cdots i_n^{(1)}} \cdots \epsilon_{i_1^{(q)}\cdots i_n^{(q)}} \prod_{k=1}^{n} X_{i_k^{(1)}\cdots i_k^{(q)}} \tag{1.28}$$

when the order q is even and $\mathrm{Det}(X) = 0$ when it is odd. Although this definition lacks an adequate justification at this point, it will support future concepts that will have their own intuitive meaning. The term $1/n!$ is justified by the equality

$$\sum_{i_1^{(1)}=1}^{n} \cdots \sum_{i_n^{(1)}=1}^{n} \cdots \sum_{i_1^{(q)}=1}^{n} \cdots \sum_{i_n^{(q)}=1}^{n} \epsilon_{i_1^{(1)}\cdots i_n^{(1)}} \cdots \epsilon_{i_1^{(q)}\cdots i_n^{(q)}} = n!, \tag{1.29}$$

[9] See LUQUE & THIBON[38].

since $n!$ must be eliminated from the calculation of the hyperdeterminant.

There is another hypercubic array δ, also widely used in indicial notations, called **Kronecker Delta**, whose elements

$$\delta_{i_1\cdots i_r j_1\cdots j_r} := \begin{cases} (-1)^{\alpha p(j_1,\cdots,j_r)} & \text{if } \exists\, p(j_1,\cdots,j_r) = (i_1,\cdots,i_r) \\ 0 & \text{se } \nexists\, p(j_1,\cdots,j_r) = (i_1,\cdots,i_r) \end{cases}, \tag{1.30}$$

where the addressing domain is N^{2r} and the set $\mathsf{N} = \{1, 2, \cdots, n\}$. Differently from the Levi-Civita Symbol, in order to build δ it is necessary to specify an order of $2r$ and all dimensions equal to n. In the particular case of $r = 1$, the Kronecker Delta results a $n \times n$ array, that is, a tabular array, as can be observed in

$$\delta = \begin{bmatrix} \delta_{11} & \delta_{12} & \cdots & \delta_{1n} \\ \delta_{21} & \delta_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \delta_{n-1\,n} \\ \delta_{n1} & \cdots & \delta_{n\,n-1} & \delta_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix}. \tag{1.31}$$

Along this text, the following important property of the Kronecker Delta with $r = 1$ will be used to handle indices in various proofs and expression developments:

$$\sum_{i_q=1}^{n} A_{i_1\cdots i_q}\delta_{i_q j} = A_{i_1\cdots i_{q-1} j}, \tag{1.32}$$

where A is a $n \times \cdots \times n$ array, with n repeated q times. In this equality, the index i_q of the array, on which the sum is made, results changed by the index j of the Kronecker Delta. Similarly,

$$\sum_{i_1=1}^{n} \delta_{j i_1} A_{i_1\cdots i_q} = A_{j i_2\cdots i_q}. \tag{1.33}$$

At this point, it is convenient to say that every $n_1 \times n_2$ array, whose elements are arranged in a tabular format, is called a **matrix**. The elements of a matrix are addressed by lines and columns: the escalar with ij position is located on the line i and column j. If $n_1 = n_2 = n$, we have a hypercubic matrix, called **square matrix**, whose size is n. Henceforth, recalling the additive group Y of arrays, we will consider $\bar{\mathsf{Y}} \subset \mathsf{Y}$ the set of hypercubic arrays of size n, $\mathsf{M} \subset \mathsf{Y}$ the set of matrices and $\bar{\mathsf{M}} \subset \bar{\mathsf{Y}}$ the set of square matrices with size n. For example, the set $\bar{\mathsf{M}}$ has the $n \times n$ Kronecker Delta, represented in (1.31), as an element, which we usually call **identity matrix** I because $AI = IA = A$, for an arbitrary $A \in \bar{\mathsf{M}}$.

If there exists a matrix $B \in \bar{\mathsf{M}}$ such that $AB = I$, we call it the **inverse** of A, notated by A^{-1}. The condition of invertibility of a matrix depends on the value of its hyperdeterminant. In the context of matrices, the hyperdeterminant function is called

determinant, which defines a mapping $\det:\bar{\mathbb{M}}\mapsto\mathbb{F}$ where det is described by the rule called **Leibniz formula**, namely

$$\det(X)=\frac{1}{n!}\sum_{i_1=1}^{n}\cdots\sum_{i_n=1}^{n}\sum_{j_1=1}^{n}\cdots\sum_{j_n=1}^{n}\epsilon_{i_1\cdots i_n}\epsilon_{j_1\cdots j_n}\prod_{k=1}^{n}X_{i_kj_k}\,, \tag{1.34}$$

which is actually equality (1.28) applied to square matrices of size n. This equality can be simplified by considering $j_k=k$, from which it can be obtained that

$$\det(X)=\sum_{i_1=1}^{n}\cdots\sum_{i_n=1}^{n}\epsilon_{i_1\cdots i_n}\prod_{k=1}^{n}X_{i_kk}\,, \tag{1.35}$$

where $1/n!$ is no longer necessary because equality (1.29) does not occur. When $\det(X)=0$, matrix X is said to be **singular** or **non-invertible**; otherwise, it is called **non-singular** or **invertible**. From the above definition, we can obtain a zero value determinant when its argument is the zero matrix and a unitary value for the identity matrix. Moreover, since this definition involves sums of products of n elements, it is straightforward to conclude that $\det(\alpha A)=\alpha^n\det(A)$, where α is a scalar. But the main property of determinants of square matrices is to preserve multiplications, that is, the determinant of the product is the product of determinants; in other words $\det(AB)=\det(A)\cdot\det(B)$. Recalling our classification of groups, we can say that a subset of $\bar{\mathbb{M}}$ whose matrices are all invertible defines, together with the operation of multiplication, a non-abelian multiplicative group. From the above definition of determinant of matrices, it is possible to obtain an important equality, called **contracted epsilon identity**, which correlates the Kronecker Delta with the Levi-Civita Symbol when size $n=3$, namely

$$\sum_{i=1}^{3}\epsilon_{ijk}\epsilon_{ilm}=\delta_{jl}\delta_{km}-\delta_{jm}\delta_{kl}\,. \tag{1.36}$$

Proof. First, we want to verify the above multiplicative property of the determinant; and for that, we need a preliminar definition, presented as following. An invertible matrix is called **elementary** when it differs from the identity matrix by one of the following three line or column actions: exchange, scalar multiplication and addition with the multiple of another line. The pre-multiplication of an elementary matrix E by an arbitrary $B\in\bar{\mathbb{M}}$ means to accomplish in B the same line action accomplished on I to arrive at E; conversely, post-multiplication of B by an elementary matrix E means to accomplish in B the same column action accomplished on I to arrive at E. On an invertible matrix A, it is possible to accomplish successive line actions resulting the identity matrix. It means that

$$E_rE_{r-1}\cdots E_2E_1A=I,$$

where matrices E_i are elementary. Thus, the equality

$$A=E_1^{-1}E_2^{-1}\cdots E_{r-1}^{-1}E_r^{-1}$$

is a decomposition of A in elementary matrices, since the inverse of an elementary matrix is also elementary. The determinant of a matrix B' that results from a line action on the matrix B is given by

$\det(B') = \alpha\det(B)$, where $\alpha \in \mathbb{R}$. Therefore, we can state that $\det(E) = \beta\det(I) = \beta$, where $\beta \in \mathbb{R}$. Considering what has been defined so far, we can say that

$$\det(AB) = \det(E_1^{-1}E_2^{-1}\cdots E_{r-1}^{-1}E_r^{-1}B) = \kappa\det(B),$$

where $\kappa \in \mathbb{R}$ is obtained from the r line actions on B. We can also state that

$$\det(AI) = \det(E_1^{-1}E_2^{-1}\cdots E_{r-1}^{-1}E_r^{-1}I) = \kappa\det(I) = \kappa.$$

Therefore, we conclude that $\det(AB) = \det(A)\det(B)$. Now, considering A singular, it is clear that $\det(A)\det(B) = 0$. If the matrix AB is invertible, then there exists a square matrix C where $ABC = I$. But, saying this means saying also that matrix BC is an inverse of A, which is inconsistent since A is singular. Therefore, AB is also singular, and then $\det(AB) = 0$. In this case, we have also $\det(AB) = \det(A)\det(B)$. Now, we prove equality (1.36) by considering that $\det(E_1E_2\cdots E_kBX_1X_2\cdots X_r) = (-1)^{k+r}\det(B)$, where elementary matrices E_i and X_i promote a line and column exchange respectively, and that B is singular if a pair of its rows or columns are equal. Therefore, since $\det I = 1$, we can write that

$$\det\begin{bmatrix}\delta_{11} & \delta_{12} & \delta_{13}\\ \delta_{21} & \delta_{22} & \delta_{23}\\ \delta_{31} & \delta_{32} & \delta_{33}\end{bmatrix} = 1 \Longrightarrow \det\begin{bmatrix}\delta_{in} & \delta_{il} & \delta_{im}\\ \delta_{jn} & \delta_{jl} & \delta_{jm}\\ \delta_{kn} & \delta_{kl} & \delta_{km}\end{bmatrix} = \underbrace{\epsilon_{ijk}}_{\text{line exchange}} \overbrace{\epsilon_{nlm}}^{\text{column exchange}}.$$

for a generic line and column exchanges. From equalities (1.32) and (1.33), summing $\epsilon_{ijk}\epsilon_{nlm}$ on i when $n = i$, we have

$$\begin{aligned}\sum_{i=1}^{3}\epsilon_{ijk}\epsilon_{ilm} &= \sum_{i=1}^{3}(\delta_{ii}\delta_{jl}\delta_{km} + \delta_{il}\delta_{jm}\delta_{ki} + \delta_{im}\delta_{ji}\delta_{kl} - \delta_{im}\delta_{jl}\delta_{ki} - \delta_{ii}\delta_{jm}\delta_{kl} - \delta_{il}\delta_{ji}\delta_{km})\\ &= \sum_{i=1}^{3}(\delta_{ii}\delta_{jl}\delta_{km}) + \delta_{jm}\delta_{kl} + \delta_{jm}\delta_{kl} - \delta_{jl}\delta_{km} - \sum_{i=1}^{3}(\delta_{ii}\delta_{jm}\delta_{kl}) - \delta_{jl}\delta_{km}\\ &= \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}.\end{aligned}$$

□

If $A, B \in \mathrm{M}$ are $n_1 \times n_2$ and $n_2 \times n_1$ matrices respectively in such a way that $A_{ij} = B_{ji}$, we call them **transpose** or that one is the transpose of the other, from which we adopt the representation A^{T} to B and B^{T} to A. In particular, the matrix $S \in \bar{\mathrm{M}}$ is said to be **symmetric** when it is identical to its transpose and **antisymmetric** if $S = -S^{\mathrm{T}}$. From the definition of transpose, we have the following:

i. $(A^{\mathrm{T}})^{\mathrm{T}} = A$;

ii. $(A+B)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}}$;

iii. $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$;

iv. $\left(A^{-1}\right)^{\mathrm{T}} = \left(A^{\mathrm{T}}\right)^{-1}$, if A is invertible.

Additionally, if A is invertible and $A^{-1} = A^{\mathrm{T}}$, it is classified as **orthogonal**. In this context, the property $\det(A)^{\mathrm{T}} = \det(A)$ enables us to write

$$1 = \det(AA^{-1}) = \det(AA^{\mathrm{T}}) = \det(A)\det(A) = [\det(A)]^2, \tag{1.37}$$

which means that $\det(A) = \pm 1$. Considering this result, matrix A is called a **proper orthogonal matrix** when it has a positive determinant or an **improper orthogonal matrix** when its determinant is negative.

Proof. The verification of the first item is trivial. For the second, considering the main property of Kronecker Delta and the representation $(A+B)_{ji}$ as an element of $(A+B)^{\mathrm{T}}$, we can state that

$$\begin{aligned}
(A+B)_{ji} &= \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} \delta_{ji}\,(A+B)_{ij}\,\delta_{ji} \\
&= \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} \delta_{ji}\left(A_{ij}+B_{ij}\right)\delta_{ji} \\
&= \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} \delta_{ji}A_{ij}\delta_{ji} + \delta_{ji}B_{ij}\delta_{ji} \\
&= \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} A_{ji}+B_{ji} = A_{ji}+B_{ji}.
\end{aligned}$$

The third item can be verified similarly:

$$\begin{aligned}
(AB)_{ji} &= \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} \delta_{ji}\,(AB)_{ij}\,\delta_{ji} \\
&= \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} \delta_{ji}\left(\sum_{k=1}^{n_2} A_{ik}B_{kj}\right)\delta_{ji} \\
&= \sum_{k=1}^{n_2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} \delta_{ji}A_{ik}B_{kj}\delta_{ji} \\
&= \sum_{k=1}^{n_2} A_{jk}B_{ki} = \sum_{k=1}^{n_2} B^{\mathrm{T}}_{ik}A^{\mathrm{T}}_{kj}.
\end{aligned}$$

The fourth property we verify from $A^{\mathrm{T}}(A^{\mathrm{T}})^{-1} = I$ and transposing both sides of the equality $A^{-1}A = I$, when we equal the left sides of these two expressions, arriving at the equality $A^{\mathrm{T}}(A^{\mathrm{T}})^{-1} = A^{\mathrm{T}}(A^{-1})^{\mathrm{T}}$. □

A more generic case of matrix transposition involves complex scalars and their conjugates. Thus, let's consider H and F matrices with dimensions $n_1 \times n_2$ and $n_2 \times n_1$ respectively, whose elements are complex numbers. We say that these matrices are **conjugate transpose**, or that one is the conjugate transpose of the other if $H_{ij} = \overline{F_{ji}}$. In other words, F is the conjugate transpose $H^\dagger$ of matrix H if it is the transpose of the complex conjugates of the elements of H. For example, if matrix

$$H = \begin{bmatrix} 1+3i & -1+i & 2i \\ 2+i & 4 & -i \end{bmatrix} \qquad \text{then} \qquad H^\dagger = \begin{bmatrix} 1-3i & 2-i \\ -1-i & 4 \\ -2i & i \end{bmatrix}.$$

When the elements involved are all real scalars, we have the equality $H^\dagger = H^{\mathrm{T}}$. Moreover, it is convenient to say that the four properties presented above for transpose matrices are also valid for conjugate transposes, namely, $(H_1^\dagger)^\dagger = H_1$, $(H_1+H_2)^\dagger = H_1^\dagger + H_2^\dagger$,

$(H_1H_2)^\dagger = H_2^\dagger H_1^\dagger$ and $(H_1^\dagger)^{-1} = (H_1^{-1})^\dagger$, if H_1 is invertible, where matrices $H_1, H_2 \in \mathrm{M}$.

In the context of invertible matrices, a matrix whose inverse equals its conjugate transpose, that is, when $U^{-1} = U^\dagger$ where $U \in \bar{\mathrm{M}}$, is called **unitary**. Considering the property $\det(H)^\dagger = \overline{\det(H)}$, we have

$$1 = \det(UU^{-1}) = \det(UU^\dagger) = \det(U)\overline{\det(U)} = |\det(U)|^2, \tag{1.38}$$

from where we can conclude that $\det(U) = \pm 1$. When this determinant is positive, U is called a **proper unitary** matrix. Given a matrix $A \in \bar{\mathrm{M}}$, we have $\det(UA) = \pm\det(A)$; which reveals, except for an eventual sign, the *neutrality* of the unitary matrix in the determinant of matrix products. If the unitary matrix elements are reals, it results that $U^\dagger = U^\mathrm{T} = U^{-1}$, or that it is orthogonal. Moreover, for an arbitrary invertible matrix A, the determinant of the inverse is the inverse of the determinant, according to the following expression:

$$1 = \det(AA^{-1}) = \det(A)\cdot\det(A^{-1}) \implies \det(A^{-1}) = [\det(A)]^{-1}. \tag{1.39}$$

A given square matrix A that equals its conjugate transpose $A^\dagger$ is called **Hermitian**. Real symmetric matrices are examples of Hermitian matrices: if the elements of the square matrix S are real, then $S^\dagger = S^\mathrm{T}$, and since S is symmetric, $S^\dagger = S$. Similarly to antisymmetric matrices, an **anti-Hermitian** matrix A equals the negative of its conjugate transpose, that is, $A = -A^\dagger$. Thus, let $B \in \bar{\mathrm{M}}$ be a matrix from which we write the following development:

$$B = \frac{1}{2}(B+B) = \frac{1}{2}\left(B+B+B^\dagger - B^\dagger\right) = \underbrace{\frac{1}{2}\left(B+B^\dagger\right)}_{B_1} + \underbrace{\frac{1}{2}\left(B-B^\dagger\right)}_{B_2}. \tag{1.40}$$

Considering the properties of conjugate transposes, we can obtain that the matrix B_1 equals its conjugate transpose and matrix B_2 equals the negative of its conjugate transpose; thereby, we say that they are respectively the **Hermitian** and **anti-Hermitian** parts of B. This result is generalized by saying that every square matrix can be decomposed additively in a Hermitian and an anti-Hermitian parts.

Any square matrix N is called **normal** when it commutes with its conjugate transpose, that is, when $N^\dagger N = NN^\dagger$. These matrices, which make the product $N^\dagger N$ an Hermitian matrix, *are always susceptible of diagonalization by a unitary matrix*. In order to understand what this means, let's consider firstly the matrices $A, B \in \bar{\mathrm{M}}$ and say that they are called **similar** when an invertible matrix Q exists such that

$$A = Q^{-1}BQ. \tag{1.41}$$

It is important to note that A and B positions in the equality do not affect the definition, since by adopting $P := Q^{-1}$, we arrive at $B = P^{-1}AP$, where the concept of similarity is maintained. Moreover, similar matrices have the same determinant value as can be verified in the following equalities:

$$\det(A) = \det(Q^{-1})\det(B)\det(Q) = (\det(Q))^{-1}\det(Q)\det(B) = \det(B). \tag{1.42}$$

A mapping $\mathrm{q} : \bar{\mathrm{M}} \mapsto \bar{\mathrm{N}}$, where $\bar{\mathrm{N}} \subset \bar{\mathrm{M}}$, is a **similarity transformation** if

$$\mathrm{q}(X) = Q^{-1}XQ. \tag{1.43}$$

An important example of this transformation is called **diagonalization**, defined from the concept of **diagonal matrix**, which is a square matrix whose elements in positions $\mathrm{i} \neq \mathrm{j}$ are null. Thereby, if there exists a similarity transformation in a square matrix A which $\mathrm{q}(A)$ results diagonal, this transformation is called a diagonalization of A or matrix A is said to be **diagonalizable**. Finally, it is now possible to understand the diagonalization of normal matrices by unitary.

Theorem 1 – Spectral Diagonalization

For any normal matrix N, there is always an unitary matrix U such that

$$\widetilde{N} = U^\dagger N U, \tag{1.44}$$

where $\widetilde{\mathrm{N}}$ *is a diagonal matrix whose elements constitute the spectrum of N.*

Proof. Firstly we must say that, considering any square matrix A and a unitary matrix U, it is always possible to obtain an **upper triangular matrix** $T = U^\dagger A U$, whose elements in positions $\mathrm{i} > \mathrm{j}$ are null. This statement is known as **Schur's Lemma**, whose tedious proof can be seen in Strang[48]. Now we need to show that if A is normal, T is diagonal: if A is normal,

$$\begin{aligned}
AA^\dagger &= A^\dagger A \\
UTU^\dagger \left(UTU^\dagger\right)^\dagger &= \left(UTU^\dagger\right)^\dagger UTU^\dagger \\
UTT^\dagger U^\dagger &= UT^\dagger TU^\dagger \\
TT^\dagger &= T^\dagger T,
\end{aligned}$$

from where we conclude that T is also normal. From the last equality, we have to each position ij

$$\begin{aligned}
\sum_{\mathrm{k}=1}^{n} T_{\mathrm{ik}} T^\dagger_{\mathrm{kj}} &= \sum_{\mathrm{k}=1}^{n} T^\dagger_{\mathrm{ik}} T_{\mathrm{kj}} \\
\sum_{\mathrm{k}=1}^{n} T_{\mathrm{ik}} \overline{T_{\mathrm{jk}}} &= \sum_{\mathrm{k}=1}^{n} \overline{T_{\mathrm{ki}}} T_{\mathrm{kj}}\,.
\end{aligned}$$

For the element in position $\mathrm{i} = \mathrm{j} = 1$, this last equality results $\sum_{\mathrm{k}=1}^{n} |T_{1\mathrm{k}}|^2 = |T_{11}|^2$, from where we can say that $\sum_{\mathrm{k}=2}^{n} |T_{1\mathrm{k}}|^2 = 0$. Since the terms in this sum are nonnegative, they can only be null, that is, $|T_{1\mathrm{k}}|^2 = 0$ when $\mathrm{k} > 1$. By induction, when we run through all the positions $\mathrm{i} = \mathrm{j}$, we verify that not only the elements of T in $\mathrm{i} > \mathrm{j}$ are null but also in $\mathrm{i} < \mathrm{j}$; which proves T diagonal. □

Now let's clarify the term "spectrum" cited by the theorem above; but before that, we need some important definitions. The function in $\mathrm{tr} : \bar{\mathrm{M}} \mapsto \mathbb{F}$ is called **trace** when its rule is defined by

$$\mathrm{tr}(X) = \sum_{i=1}^{n} X_{ii}, \tag{1.45}$$

that is, the trace of a matrix is the sum of its diagonal elements. Considering the matrices $A, B \in \bar{M}$, it is clear that $\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$, where sum is preserved; thereby, the trace function results a group homomorphism on $\bar{M} \subset Y$ because Y is an additive group. Another evident property of traces is that $\mathrm{tr}(A^{\mathrm{T}}) = \mathrm{tr}(A)$. Moreover, since the respective diagonal elements of AB and BA are identical, it is true that $\mathrm{tr}(AB) = \mathrm{tr}(BA)$. Thence, if A and B are similar,

$$\mathrm{tr}(A) = \mathrm{tr}(Q^{-1}BQ) = \mathrm{tr}(QQ^{-1}B) = \mathrm{tr}(B), \tag{1.46}$$

and then we can state generically that all similar matrices have the same trace.

Among other benefits, the trace function is a convenient tool to develop what is known by **characteristic polynomial** of a matrix: it is the function in operation $g : \mathbb{F} \mapsto \mathbb{F}$, whose rule is

$$g(x) = \det(H - xI), \tag{1.47}$$

where H is a square matrix of size n. By developing the right term, we arrive at

$$g(x) = (-1)^n x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_{n-1} x + a_n, \tag{1.48}$$

which is a polynomial of order n, called characteristic polynomial of H, whose coefficients are

$$a_1 = (-1)^{n+1}\,\mathrm{tr}(H); \tag{1.49}$$

$$a_2 = -\frac{1}{2}\left[a_1 \mathrm{tr}(H) + (-1)^n\,\mathrm{tr}(H^2)\right]; \tag{1.50}$$

$$\cdots$$

$$a_n = -\frac{1}{n}\left[a_{n-1}\mathrm{tr}(H) + a_{n-2}\mathrm{tr}(H^2) + \cdots + a_1 \mathrm{tr}(H^{n-1}) + (-1)^n\,\mathrm{tr}(H^n)\right]. \tag{1.51}$$

Any scalar $\lambda \in \mathbb{F}$ that is root of this polynomial is called a **characteristic root** of H. The collection $\lambda_1, \cdots, \lambda_n$ of characteristic roots of H is called the **spectrum**[10] of H, represented by $\lambda(H)$. In the case of a normal matrix N submitted to spectral diagonalization, its spectrum is the set of scalars that constitute the diagonal matrix $\widetilde{N}$, which results from the diagonalization of N by unitary matrix. At this point, it is convenient to present a very important theorem involving similar matrices.

Theorem 2 – Isospectral Matrices

Similar matrices are isospectral, that is, they have the same spectrum.

[10] The spectrum is not a set because there can be repeated elements.

Proof. We already know that the determinant of the inverse is the inverse of the determinant. Let A an B be similar matrices through invertible matrix Q. Then, it can be said that the characteristic polynomial $\det(A - xI) = \det(Q^{-1}BQ - xI)$. From the following equalities

$$Q^{-1}xIQ = xQ^{-1}IQ = xQ^{-1}Q = xI,$$

it is possible to say that

$$\begin{aligned}\det(A - xI) &= \det\left(Q^{-1}BQ - Q^{-1}xIQ\right)\\ &= \det\left(Q^{-1}(B - xI)Q\right)\\ &= (\det(Q))^{-1}\det(B - xI)\det(Q)\\ &= \det(B - xI),\end{aligned}$$

from where we conclude A and B isospectral. □

The spectrum of a matrix has a close relation with determinant and trace functions, which *respectively preserve the operations of multiplication and addition.* If $\lambda_1, \cdots, \lambda_n$ is the spectrum of A, this close relation is expressed through the following equalities:

$$\det(A) = \prod_{i=1}^{n}\lambda_i \qquad \text{and} \qquad \operatorname{tr}(A) = \sum_{i=1}^{n}\lambda_i. \tag{1.52}$$

When two matrices have the same pair of determinant and trace, we can say that these matrices have a kind of equivalence if we interpret the two functions as scalar measures whose absolute values express a quantitative aspect and whose signs express a qualitative aspect of the matrix in question. Therefore, isospectral matrices are said to be equivalent in this sense; *something that, in a certain way, confers a metric feature for the spectrum of matrices.*

Proof. Let's verify the two equalities above. For the first one, let the rule of the characteristic polynomial of A be described in its factorized form by

$$g(x) = (-1)^n(x - \lambda_n)(x - \lambda_{n-1})\cdots(x - \lambda_2)(x - \lambda_1),$$

from where the following equality results:

$$\det((A - xI)) = (\lambda_n - x)(\lambda_{n-1} - x)\cdots(\lambda_2 - x)(\lambda_1 - x),$$

valid for all $x \in \mathbb{F}$. Thus, if $x = 0$, we have

$$\det(A) = \lambda_n\lambda_{n-1}\cdots\lambda_2\lambda_1.$$

In order to prove the second equality, one of the so called Viète Formulae, according to Vinberg[54], states that

$$-\frac{a_1}{(-1)^n} = \lambda_n + \lambda_{n-1} + \cdots + \lambda_2 + \lambda_1,$$

valid for polynomials having the same format of (1.48). Thence we can substitute the coefficient a_1 by the right term of (1.49), resulting

$$\operatorname{tr}(A) = \lambda_n + \lambda_{n-1} + \cdots + \lambda_2 + \lambda_1.$$

□

A matrix $A \in \bar{\mathrm{M}}$ is said to be **nonnegative** or **positive-semidefinite** if

$$\Re\left(X^{\dagger}AX\right)_{11} \geqslant 0, \tag{1.53}$$

where X is a non zero $n \times 1$ matrix. When the inequality imposes the left side to be always positive, A is called **positive-definite**. A necessary and sufficient condition that ensures positivity of a matrix is that the spectrum of its hermitian part be constituted of nonnegative elements[11], when this matrix results nonnegative; similarly, if these same elements are all positive, the matrix is positive-definite. Thereby, from the two equalities (1.52) of the previous paragraph, it can be concluded that the determinant and trace of a nonnegative Hermitian matrix are always nonnegative, while for a positive-definite Hermitian matrix both are positive. Therefore, in the context of Hermitian matrices, the determinant value indicates whether a matrix is nonnegative or not. From this conclusion and the definition of invertibility, we can state that *every positive-definite hermitian matrix is invertible*. Moreover, when a positive-definite hermitian matrix H pre or post multiplies an arbitrary matrix $B \in \bar{\mathrm{M}}$, we have

$$\operatorname{sgn}\left(\det(BH)\right) = \operatorname{sgn}\left(\det(HB)\right) = \operatorname{sgn}\left(\det(B)\right)\operatorname{sgn}\left(\det(H)\right) = \operatorname{sgn}\left(\det(B)\right), \tag{1.54}$$

where the sign "sgn" of the determinant of B defines the sign of the product of the determinants, making the concept of matrix positivity similar to scalar positivity, in which the sign of a product is not defined by an eventual positive number.

We shall finish this chapter by presenting the following fundamental equality, a property of any square matrix that will make further developments feasible.

Theorem 3 – Cayley-Hamilton

Let $\bar{\mathrm{M}}$ be the set of all square matrices and $\hat{g} : \bar{\mathrm{M}} \mapsto \bar{\mathrm{M}}$ an operation whose function rule is described by

$$\hat{g}(X) = (-1)^n X^n + a_1 X^{n-1} + a_2 X^{n-2} + \cdots + a_{n-1} X + a_n I. \tag{1.55}$$

If the coefficients on this rule equal the coefficients of the characteristic polynomial of a square matrix H, then the matrix $\hat{g}(H)$ is null.

Proof. In order to verify this theorem, we need to present some preliminary definitions. There is an algorithm [a], called **Laplace Expansion** or **Cofactorial Expansion**, that is used in certain cases to find determinants. Here it is: given a square matrix A, we can obtain for an arbitrary line i that

$$\det(A) = \sum_{j=1}^{n} A_{ij} \underbrace{(-1)^{i+j}\det\left(M_{(ij)}\right)}_{C_{ij}},$$

where C is called **cofactor matrix** of A and $M_{(ij)}$ is a square matrix of dimension $n-1$ which results

[11] This condition will be verified on section 2.5.

from removing the line i and the column j from A. The matrix C^{T} is called the **adjugate matrix** of A, represented by $\operatorname{adj} A$. Now, considering $A = H - xI$ and the property $\operatorname{adj}(X)\,X = (\det(X))I$, we have that

$$\operatorname{adj}(A)\,(H - xI) = g(x)I = I(-1)^n x^n + Ia_1 x^{n-1} + \cdots + Ia_{n-1}x + Ia_n\,.$$

Through a tedious development, it can be obtained that the adjugate of A results a polynomial of order q, with matrix coefficients H_i, described by

$$\operatorname{adj}(A) = H_1 x^q + H_2 x^{q-1} + \cdots + H_{n-1}x + H_n\,.$$

The product

$$\operatorname{adj}(A)\,(H - xI) = -H_1 x^{q+1} + (H_1 H - H_2)x^q + \cdots + (H_{n-2}H - H_{n-1})x + H_n H\,.$$

Comparing the two expressions on the right that equal $\operatorname{adj}(A)\,(H - xI)$, we can conclude that integer $n = q + 1$ and the following equalities:

$$\begin{aligned} I(-1)^n &= -H_1 \\ Ia_1 &= H_1 H - H_2 \\ &\vdots \\ Ia_{n-1} &= H_{n-2}H - H_{n-1} \\ Ia_n &= H_n H\,. \end{aligned}$$

If we post-multiply the sequence of equalities successively by $H^n, H^{n-1}, \cdots, H, H^0$ and adding all of them, we arrive at

$$(-1)^n H^n + a_1 H^{n-1} + a_2 H^{n-2} + \cdots + a_{n-1}H + a_n I = 0\,,$$

that is, $\hat{g}(H) = 0$. □

[a] According to Knuth[31], the word “algorithm” has the same etymological source of “algebra”.

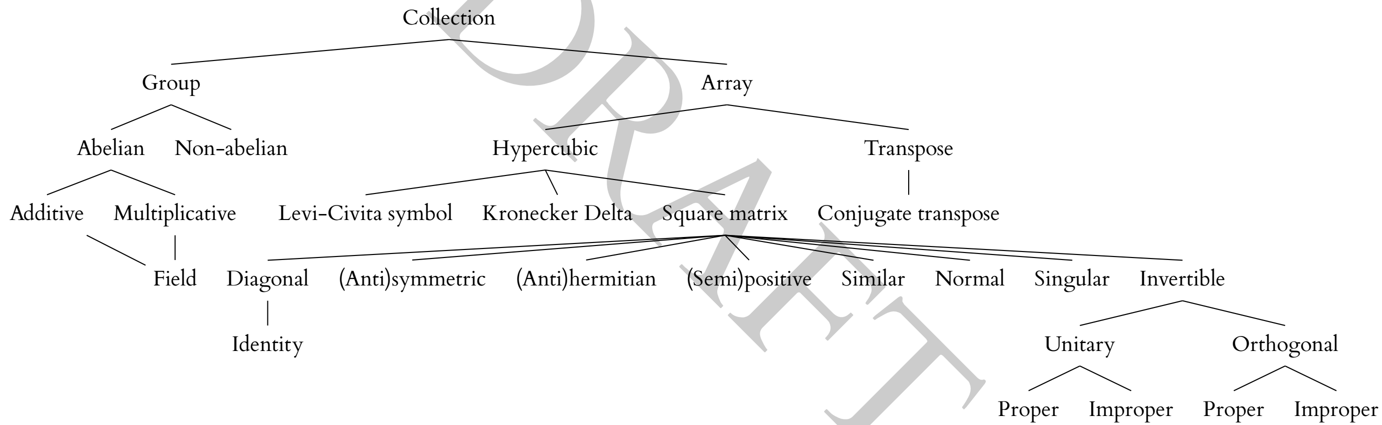

Figure 1.5 – Conceptual hierarchy of groups and arrays.

CHAPTER

2 Basic Linear Algebra

A set is called **space** when it is structured by another set, by an operation or by some relevant property to which all of its elements are subjected. In the previous chapter, we created a space with an additive structure, called an additive group, and cumulatively assigned to this space a multiplicative structure, when it became a field. In this chapter, the cumulative structuring of these specific spaces is developed, now using fields, norms, metrics and inner products as structural entities. Firstly, we shall gather the concepts of additive group and field in such a way that, from this interaction, scalars end up assigning certain multiplicative properties to group elements, namely, abbreviation of repetitive additions, positivity and negativity. Regarding the relationships between these spaces, Linear Algebra deals mainly with specific homomorphic functions in which scalars take part and structures are preserved.

2.1 Structuring by Field

The group-field space is the fundamental object of Linear Algebra and the interaction between these two sets is subjected to restrictions. In order to present them, we shall mathematically describe and complement what has been said so far. Let V be an additive group structured by a field $\mathbb{F}$ through the function p in mapping $p : \mathbb{F} \times V \mapsto V$. This function, whose values $p(\alpha, x)$ are represented by αx or $x\alpha$, must obey the following axioms:

i. $\alpha(x + y) = \alpha x + \alpha y$;

ii. $(\alpha + \beta)x = \alpha x + \beta x$;

iii. $(\alpha\beta)x = \alpha(\beta x)$;

iv. $1\mathbf{x} = \mathbf{x}$, where 1 is the multiplicative identity of $\mathbb{F}$;

v. $0\mathbf{x} = \mathbf{0}$, where $\mathbf{0}$ is the zero element of V;

for all $\alpha, \beta \in \mathbb{F}$ and $\mathbf{x}, \mathbf{y} \in V$. Under these conditions, an element of V is called **vector** and the triple $(V, \mathbb{F}, p)$ is a **vector space** of V in $\mathbb{F}$, whose representation is abbreviated by the symbol $V_{\mathbb{F}}$, which will be treated from now on as a set, in order to simplify notation. Moreover, if the field $\mathbb{F}$ is complex, vector space $V_{\mathbb{C}}$ is said to be **complex**, for an arbitrary group V; similarly, $V_{\mathbb{R}}$ is called a **real vector space**. From the previous definitions, we can conclude that if a field is a group, then $\mathbb{F}_{\mathbb{F}}$ or $\mathbb{F}$ is also a vector space. Now, when we consider a mapping $\overline{p} : \mathbb{F} \times V \mapsto V$ whose function rule is $\overline{p}(\alpha, \mathbf{x}) = p(\overline{\alpha}, \mathbf{x}) = \overline{\alpha}\mathbf{x}$, vector space $(V, \mathbb{F}, \overline{p})$, represented by $\overline{V_{\mathbb{F}}}$, is called the **conjugate vector space** of $V_{\mathbb{F}}$. It is interesting to note that only the rule of multiplication by scalar distinguishes a vector space from its conjugate; both being defined by the same additive group and field, that is, they have the same elements which can be multiplied by the same scalars. Since the definer group and field are the same, in order to avoid any dubieties on which vector space is being considered, an arbitrary vector $\boldsymbol{v} \in V_{\mathbb{F}}$ is represented by $\boldsymbol{v}^{c}$ when V defines $\overline{V_{\mathbb{F}}}$, that is, $\boldsymbol{v} \in V_{\mathbb{F}}$ and $\boldsymbol{v}^{c} \in \overline{V_{\mathbb{F}}}$ are representations of the same vector, which are related through

$$\alpha \boldsymbol{v}^{c} = \overline{\alpha}\boldsymbol{v}, \forall \alpha \in \mathbb{F}. \tag{2.1}$$

Note that when $\Im(\alpha) = 0$, the arbitrary vector $\boldsymbol{v}$ is equal to $\boldsymbol{v}^{c}$, which is valid since they refer to the same element of V. From previous equality, we conclude the following property:

$$(\beta_1 \mathbf{x} + \beta_2 \mathbf{y})^{c} = \overline{\beta_1}\mathbf{x}^{c} + \overline{\beta_2}\mathbf{y}^{c} \tag{2.2}$$

for all $\beta_1, \beta_2 \in \mathbb{F}$. Moreover, we can state that functions which define $V_{\mathbb{F}} \times \overline{V_{\mathbb{F}}} \mapsto V$ or $\overline{V_{\mathbb{F}}} \times V_{\mathbb{F}} \mapsto V$ are binary operators, in the terms of section 1.5, because the domain is ultimately V^2.

Proof. Let's verify the last equality. If we assume that $\boldsymbol{v} = \beta_1 \mathbf{x} + \beta_2 \mathbf{y}$ in $\alpha \boldsymbol{v}^{c} = \overline{\alpha}\boldsymbol{v}$, then

$$\alpha(\beta_1 \mathbf{x} + \beta_2 \mathbf{y})^{c} = \overline{\alpha}(\beta_1 \mathbf{x} + \beta_2 \mathbf{y}) = \overline{\alpha}\beta_1 \mathbf{x} + \overline{\alpha}\beta_2 \mathbf{y} = \alpha\overline{\beta_1}\mathbf{x}^{c} + \alpha\overline{\beta_2}\mathbf{y}^{c} = \alpha(\overline{\beta_1}\mathbf{x}^{c} + \overline{\beta_2}\mathbf{y}^{c}),$$

which proves the property. □

Since a set admits a subset under the conditions already presented, spaces admit subspaces. A **vector subspace**, structured by the field $\mathbb{F}$, is actually a vector space in $\mathbb{F}$ whose elements also belong to a set that defines a vector space in $\mathbb{F}$. In more precise terms, we say that the vector space $(S, \mathbb{F}, \tilde{p})$, where $\tilde{p} : \mathbb{F} \times S \mapsto S$, is a vector subspace of $V_{\mathbb{F}}$ if the set $S \subseteq V$; or, in a detailed notation, if the space $S_{\mathbb{F}} \subseteq V_{\mathbb{F}}$. It is important to say that since all vector spaces are defined to have a zero element $\mathbf{0}$, then $\mathbf{0}$ must belong to any vector subspace.

The possibility of multiplication by scalars, according to the mapping defined by p, enables us to combine the vectors of $\tilde{U} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n\} \subset V_{\mathbb{F}}$ as in

$$\alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2 + \cdots + \alpha_n \boldsymbol{v}_n \,, \tag{2.3}$$

where α_i are arbitrary scalars of $\mathbb{F}$. Thereby, this expression is called *the* **linear combination** of $\tilde{U}$ in $\mathbb{F}$ and, when the scalars are given, the vector $\sum_{i=1}^{n} \alpha_i \boldsymbol{v}_i$ is said to be *a* linear combination of $\tilde{U}$ in $\mathbb{F}$. Considering $n > 1$, if the zero vector is a linear combination of $\tilde{U}$ when at least one of the scalars $\alpha_1, \cdots, \alpha_n$ is not zero, then we classify $\tilde{U}$ as **linearly dependent**. In this case, admitting that $\alpha_1 \neq 0$, from the equality $\sum_{i=1}^{n} \alpha_i \boldsymbol{v}_i = \mathbf{0}$, we can write that $\boldsymbol{v}_1 = \sum_{i=2}^{n} (\alpha_i / \alpha_1) \boldsymbol{v}_i$, where $\boldsymbol{v}_1$ is said to be a linear combination of the other vectors. However, this linear combination of vectors can not be written when a sequence of zero scalars is the only possible sequence to make an arbitrary linear combination of $\tilde{U}$ equals the zero vector. In this context, if the vectors of $\tilde{U}$ are not zero, this set is called **linearly independent**.

Recalling our definition of vector space, it is important to observe that the multiplication by scalar defined in mapping $p : \mathbb{F} \times V \mapsto V$ together with the operation $+ : V^2 \mapsto V$, typical of additive groups, assure that every linear combination of arbitrary vectors of $V_{\mathbb{F}}$ is also a vector of $V_{\mathbb{F}}$; that is, if n vectors $\boldsymbol{v}_i \in V_{\mathbb{F}}$, then the vector $\sum_{i=1}^{n} \alpha_i \boldsymbol{v}_i \in V_{\mathbb{F}}$. In this context, let U be a non empty subset of $V_{\mathbb{F}}$, described the following way:

$$U = \bigcup_{i=1}^{\infty} \tilde{U}_i \,, \tag{2.4}$$

where each set $\tilde{U}_i \subset U$ is finite. Thereby, the subset of $V_{\mathbb{F}}$ constituted by all linear combinations of the subsets $\tilde{U}_i$ is called a **span** of U, whose representation is $\text{span}(U)$. In other words,

$$\text{span}(U) := \left\{ \sum_{i=1}^{n} \alpha_i \boldsymbol{v}_i : \forall n \in \mathbb{N},\, \forall \alpha_i \in \mathbb{F},\, \forall \boldsymbol{v}_i \in U \right\} . \tag{2.5}$$

If $\text{span}(U)$ is spanned or generated by U, then we can also say that U spans or generates $\text{span}(U)$. Now, let's take two arbitrary elements of the subset spanned by U, namely the vectors $\boldsymbol{x} = \sum_{i=1}^{n} \varphi_i \boldsymbol{v}_i$ and $\boldsymbol{y} = \sum_{i=1}^{n} \beta_i \boldsymbol{v}_i$, where $\varphi_i, \beta_i \in \mathbb{F}$. Adding these two vectors results the vector $\boldsymbol{x} + \boldsymbol{y} = \sum_{i=1}^{n} (\varphi_i + \beta_i) \boldsymbol{v}_i$, which is also an element of $\text{span}(U)$, since $\varphi_i + \beta_i \in \mathbb{F}$ and $\boldsymbol{v}_i \in U$; that is, the operation of addition can be defined by the mapping $+ : \text{span}(U)^2 \mapsto \text{span}(U)$. Moreover, the product of an arbitrary scalar $\alpha \in \mathbb{F}$ and $\boldsymbol{x}$ results $\alpha \boldsymbol{x} = \sum_{i=1}^{n} \alpha \varphi_i \boldsymbol{v}_i \in \text{span}(U)$, since $\alpha \varphi_i \in \mathbb{F}$; which proves the multiplication $p : \mathbb{F} \times \text{span}(U) \mapsto \text{span}(U)$. From these facts, it is easily verified that $\text{span}(U)$ observes the five axioms of vectors spaces presented above; which permits us to conclude that the subset spanned by U defines a vector space $\text{span}(U)_{\mathbb{F}} \subseteq V_{\mathbb{F}}$. Therefore, we can state generically that every spanned subset defines a **spanned subspace**.

Considering the previous conditions in the case where U spans the space $V_{\mathbb{F}}$ as a whole, we define the following: a) if U is finite, $V_{\mathbb{F}}$ is said to be a **finite-dimensional** vector space; b) if U is linearly independent, it is called a **basis** of $V_{\mathbb{F}}$. Gathering these two definitions, when U is a basis with n elements that spans a vector space $V_{\mathbb{F}}$, an arbitrary vector $\boldsymbol{w} \in V_{\mathbb{F}}$ is generated by one and only one linear combination $\sum_{i=1}^{n} \alpha_i \boldsymbol{v}_i$. Therefore, in the context of the basis U, there is a biunivocal relationship between the vector $\boldsymbol{w}$ and the n-tuple $(\alpha_i, \cdots, \alpha_n)$, whose ordering follows the sequence of the basis vectors. The n-tuple of scalars that defines vector $\boldsymbol{w}$ on the basis U is called the **coordinates** of $\boldsymbol{w}$ on U. Since the idea of ordering is required for defining coordinates, it is important to point out that when working with vector coordinates, a basis must be implicitly understood as an "ordered set" whose elements are sequenced; and thereby, in this case, two bases are equal not only if they have the same elements but also if these elements are ordered the same way.

Proof. Let's verify if it is true that $\sum_{i=1}^{n} \alpha_i \boldsymbol{v}_i$ is the only linear combination that defines $\boldsymbol{w}$ on U. If there were another linear combination $\sum_{i=1}^{n} \beta_i \boldsymbol{v}_i$ defining $\boldsymbol{w}$, then the difference between them would be $\sum_{i=1}^{n} (\alpha_i - \beta_i) \boldsymbol{v}_i = \mathbf{0}$. As U does not have a zero element, from the previous equality we have $\alpha_i - \beta_i = 0$, or $\alpha_i = \beta_i$. □

Now, let's consider $U_1 = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ a basis of $V_{\mathbb{F}}$ and $U_2 = \{\boldsymbol{w}_1, \cdots, \boldsymbol{w}_m\}$ a linearly independent set such that $m \geqslant n$. If U_1 spans $V_{\mathbb{F}}$, then the linearly dependent set $\{\boldsymbol{w}_1\} \cup U_1 = \{\boldsymbol{w}_1, \boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ also spans $V_{\mathbb{F}}$. When an element $\boldsymbol{v}_k$ is removed from U_1, the resulting set $(\{\boldsymbol{w}_1\} \cup U_1) \setminus \{\boldsymbol{v}_k\}$ also spans $V_{\mathbb{F}}$ because $\boldsymbol{w}_1$ is a linear combination of U_1. If we proceed including elements of U_2 and removing elements of U_1, we shall obtain the set $\{\boldsymbol{w}_1, \cdots, \boldsymbol{w}_n\}$, which is a basis of $V_{\mathbb{F}}$. Thereby, linearly independent sets which span the same finite-dimensional vector space have the same number of vectors. From this general statement, we can say that every basis of $V_{\mathbb{F}}$ has n elements, or that the **dimension** of $V_{\mathbb{F}}$ is n, written $\dim(V_{\mathbb{F}}) = n$. Therefore, we can also state that *every subset of* $V_{\mathbb{F}}$ *having* n *linearly independent vectors is a basis of* $V_{\mathbb{F}}$, from which results the following: if $W_{\mathbb{F}} \subset V_{\mathbb{F}}$ then $\dim(W_{\mathbb{F}}) < \dim(V_{\mathbb{F}})$.

There is an important type of vector space whose group is additionally structured by what is called a **norm**, which assigns to each one of the group elements a non-negative real number that enables the concept of vector size or vector intensity. Like the case of structuring by field, structuring by norm also occurs according to some restrictions. Thereby, we say that a **normed space** is defined by the double $(V_{\mathbb{F}}, \eta)$, where $V_{\mathbb{F}}$ is a vector space and the function in $\eta : V_{\mathbb{F}} \mapsto \mathbb{R}^+$, called norm, observes the axioms

i. Of definition: $\eta(\boldsymbol{v}) = 0 \Leftrightarrow \boldsymbol{v} = \mathbf{0}$;

ii. Of homogeneity: $\eta(\alpha \boldsymbol{x}) = |\alpha| \eta(\boldsymbol{x})$ and

iii. Of triangular inequality: $\eta(\boldsymbol{x} + \boldsymbol{y}) \leqslant \eta(\boldsymbol{x}) + \eta(\boldsymbol{y})$;

where $\alpha \in \mathbb{F}$ and $\mathbf{x}, \mathbf{y} \in V_{\mathbb{F}}$ are arbitrary elements of their respective sets. On the last item, the triangular inequality axiom imposes that the size of a vector sum is never greater than the sum of vector sizes. In notational terms, as the use of η is not very common, $\|\mathbf{x}\|$ is also written to represent the value $\eta(\mathbf{x})$. If a function like norm η fails to obey the axiom of definition, it is called a **seminorm**.

We define that two vectors have an **incidence interrelationship** when there is a vector, multiple of one of them, that is function of both. In more precise terms, given two non zero vectors $\mathbf{u}, \mathbf{v} \in \mathsf{U}_{\mathbb{F}}$, it is said that $\mathbf{u}$ has an incidence on $\mathbf{v}$ if there is a vector multiple of $\mathbf{v}$ which is a function of $(\mathbf{u}, \mathbf{v})$, that is, if there is a mapping $f : \mathsf{U}_{\mathbb{F}}^2 \mapsto \mathsf{U}_{\mathbb{F}}$ where the vector $\alpha\mathbf{v} = f(\mathbf{u}, \mathbf{v})$, $\alpha \in \mathbb{F}$. Moreover, incidence is defined to be commutative: if $\mathbf{u}$ has an incidence on $\mathbf{v}$, the converse is also true. Therefore, if the previous assumptions assure that $\mathbf{v}$ has an incidence on $\mathbf{u}$ also, then there is a vector $\beta\mathbf{u} = f(\mathbf{v}, \mathbf{u})$, $\beta \in \mathbb{F}$. The incidence interrelationship of two vectors is usually measured by real or complex values, where zero value means that there is no incidence between these vectors or that they are **orthogonal**. Let the function in $\xi : \mathsf{U}_{\mathbb{F}}^2 \mapsto \mathbb{F}$ be a measure of incidence between any pair of vectors of $\mathsf{U}_{\mathbb{F}}$ by observing the axioms

i. Of nonnegativity: $\xi(\mathbf{u}, \mathbf{u}) \in \mathbb{R}^+$;

ii. Of definition: $\xi(\mathbf{u}, \mathbf{u}) = 0 \Leftrightarrow \mathbf{u} = \mathbf{0}$;

iii. Of conjugate symmetry: $\xi(\mathbf{u}_1, \mathbf{u}_2) = \overline{\xi(\mathbf{u}_2, \mathbf{u}_1)}$;

iv. Of linearity[1] in the first argument:

$$\xi(\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2, \mathbf{u}_3) = \alpha_1\xi(\mathbf{u}_1, \mathbf{u}_3) + \alpha_2\xi(\mathbf{u}_2, \mathbf{u}_3) \text{ and}$$

v. Of antilinearity or conjugate linearity in the second argument:

$$\xi(\mathbf{u}_1, \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3) = \overline{\alpha_2}\xi(\mathbf{u}_1, \mathbf{u}_2) + \overline{\alpha_3}\xi(\mathbf{u}_1, \mathbf{u}_3);$$

where $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3 \in \mathsf{U}_{\mathbb{F}}$ and $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{F}$ are arbitrary elements of their respective sets. In this context, function ξ is called a **positive-definite inner product** because the incidence of a vector on itself is a nonnegative real number, as described by the first axiom. In our study, ξ is simply called **inner product**, and the double $(\mathsf{U}_{\mathbb{F}}, \xi)$ an **inner product space**. From this double, we conclude that ξ structures the group U in such a way that an incidence interrelationship of any pair of its elements can be obtained. Henceforth, in order to shorten notation, $\mathbf{x} \cdot \mathbf{y}$ will also be used to represent the inner product $\xi(\mathbf{x}, \mathbf{y})$. From equality (2.1), if $\mathbf{x} = \alpha\mathbf{u}_1$ and $\mathbf{y} = \beta\mathbf{u}_2$, we can write that

$$\overline{\mathbf{x} \cdot \mathbf{y}} = \overline{\alpha}\beta\overline{\mathbf{u}_1 \cdot \mathbf{u}_2} = \overline{\alpha}\mathbf{u}_2 \cdot \overline{\beta}\mathbf{u}_1 = \alpha\mathbf{u}_2^c \cdot \beta\mathbf{u}_1^c = \overline{\beta}\mathbf{u}_2^c \cdot \overline{\alpha}\mathbf{u}_1^c = \mathbf{y}^c \cdot \mathbf{x}^c . \tag{2.6}$$

[1]See definition at p. 43.

Moreover, in an inner product space, it is valid the so called **Cauchy-Schwartz Inequality**, where

$$|\mathbf{u}\cdot\mathbf{v}| \leqslant \sqrt{(\mathbf{u}\cdot\mathbf{u})(\mathbf{v}\cdot\mathbf{v})},\ \forall \mathbf{u},\mathbf{v}\in U_{\mathbb{F}}. \tag{2.7}$$

Proof. Let's prove the above inequality. If vectors $\mathbf{u}$ or $\mathbf{v}$ are zero, proof is trivial. Now, considering $\mathbf{v}\neq \mathbf{0}$ and a scalar $\lambda = (\mathbf{u}\cdot\mathbf{v})/(\mathbf{v}\cdot\mathbf{v})$, from the first axiom above,

$$\begin{aligned}
0 &\leqslant (\mathbf{u}-\lambda\mathbf{v})\cdot(\mathbf{u}-\lambda\mathbf{v})\\
0 &\leqslant \mathbf{u}\cdot\mathbf{u}-\bar{\lambda}\mathbf{u}\cdot\mathbf{v}-\lambda\mathbf{v}\cdot\mathbf{u}+\lambda\bar{\lambda}\mathbf{v}\cdot\mathbf{v}\\
0 &\leqslant \mathbf{u}\cdot\mathbf{u}-\bar{\lambda}\lambda\mathbf{v}\cdot\mathbf{v}-\lambda\bar{\lambda}\mathbf{v}\cdot\mathbf{v}+\lambda\bar{\lambda}\mathbf{v}\cdot\mathbf{v}\\
|\lambda|^2\mathbf{v}\cdot\mathbf{v} &\leqslant \mathbf{u}\cdot\mathbf{u}\\
|\mathbf{u}\cdot\mathbf{v}|^2 &\leqslant (\mathbf{u}\cdot\mathbf{u})(\mathbf{v}\cdot\mathbf{v}),
\end{aligned}$$

since $\mathbf{v}\cdot\mathbf{v}$ is a positive real number. □

Considering the inner product space $(U_{\mathbb{F}},\xi)$, it is now possible to present in more mathematical terms the definition of orthogonality: arbitrary vectors $\mathbf{u}_1,\mathbf{u}_2\in U_{\mathbb{F}}$ are said to be orthogonal, or $\mathbf{u}_1\perp\mathbf{u}_2$, when $\mathbf{u}_1\cdot\mathbf{u}_2=0$. Given the subspaces $Z_{\mathbb{F}}$ and $W_{\mathbb{F}}$ of $U_{\mathbb{F}}$, if $\mathbf{u}\in U_{\mathbb{F}}$ is orthogonal to every vector of $Z_{\mathbb{F}}$, we write $\mathbf{u}\perp Z_{\mathbb{F}}$, and if every vector of $Z_{\mathbb{F}}$ is orthogonal to every vector of $W_{\mathbb{F}}$, we write $Z_{\mathbb{F}}\perp W_{\mathbb{F}}$. A set $\{\mathbf{u}_1,\cdots,\mathbf{u}_n\}\subset U_{\mathbb{F}}$ is called orthogonal if $\mathbf{u}_i\perp\mathbf{u}_j$, $i\neq j$. Thereby, when the vectors of this set are non zero, the inner product of each side of $\alpha\mathbf{u}_j=\mathbf{u}_i$ and $\mathbf{u}_j$, where $\alpha\in\mathbb{F}$ and $i\neq j$, results $\alpha(\mathbf{u}_j\cdot\mathbf{u}_j)=0$, from where we conclude that $\alpha=0$ or that every orthogonal set is linearly independent. This conclusion permits us to state that in a n-*dimensional inner product space, every orthogonal subset of* n *elements is a basis*. Moreover, any two vectors $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$ of a normed inner product space $(U_{\mathbb{F}},\eta,\xi)$ are said to be orthonormal if they are orthogonal and each one is **unitary**, where $\|\hat{\mathbf{u}}_i\|=1$. If the vector space $U_{\mathbb{F}}$ is n-dimensional, an arbitrary subset $\{\hat{\mathbf{u}}_1,\hat{\mathbf{u}}_2,\cdots,\hat{\mathbf{u}}_n\}$ of orthonormal vectors is called an **orthonormal basis** of $U_{\mathbb{F}}$. Thereby, to every orthogonal basis $\{\mathbf{u}_1,\mathbf{u}_2,\cdots,\mathbf{u}_n\}$, where $\mathbf{u}_i\cdot\mathbf{u}_j=0$ when $i\neq j$, there is always an orthonormal basis $\{\hat{\mathbf{u}}_1,\hat{\mathbf{u}}_2,\cdots,\hat{\mathbf{u}}_n\}$ where each $\hat{\mathbf{u}}_i:=\mathbf{u}_i/\|\mathbf{u}_i\|$ and then $\|\hat{\mathbf{u}}_i\|=\|\mathbf{u}_i/\|\mathbf{u}_i\|\|=\|\mathbf{u}_i\|/\|\mathbf{u}_i\|=1$, when we say that the orthonormal basis results from **normalizing** the orthogonal basis.

2.2 Structuring by Metrics

If the group-field interaction assigns to the group certain multiplicative features, a set that is structured by metrics carries with it the concept of distance. In other words, in a set-metrics space or a **metric space**, there is always a distance between two elements, measured in scalar values. This idea of distance is fundamental in Mathematics, making, for example, the usual notion of derivative viable and consequently of elementary Differential Calculus as a whole.

Like structuring by field, the structure of metrics in a set is also subjected to restrictions, described as follows. Let A be a set and $\rho: A\times A\mapsto\mathbb{R}$ a mapping. Given arbitrary elements $a_1, a_2, a_3 \in A$, the double (A, ρ) is said to be a metric space and the function ρ a **metric** or a **distance function** if it observes the axioms

i. Of nonnegativity: $\rho(a_1, a_2) \geqslant 0$;
ii. Of definition: $\rho(a_1, a_2) = 0 \Leftrightarrow a_1 = a_2$;
iii. Of commutativity: $\rho(a_1, a_2) = \rho(a_2, a_1)$ and
iv. Of triangular inequality: $\rho(a_1, a_2) \leqslant \rho(a_1, a_3) + \rho(a_3, a_2)$.

If distances are intuitively seen as paths, from the last axiom we can state that the distance from a_1 to a_2 always establishes the shortest path between these two elements. Moreover, given metric spaces (A, ρ_A) and (B, ρ_B), we call the function in a bijective mapping $f : A \mapsto B$ an **isometry** when $\rho_A(a_1, a_2) = \rho_B(f(a_1), f(a_2))$. In other words, an isometry preserves distances between the elements of its domain.

From the above definitions, many new concepts arise concerning the study of spaces structured by metrics. Among these concepts, we shall present hereafter those involved in the definition of "continuum", a space of fundamental relevance in our study. Let's start by considering a metric space (A, ρ), an element $a \in A$ and a scalar $r \in \mathbb{R}$, from which the set

$$\overline{B}_{a,r} := \{x \in A : \rho(a, x) \leqslant r\} \tag{2.8}$$

is said to be a **closed ball** with center a and radius r. It is then a subset of A delimited by a "spheric" set whose elements belong to this subset. When such sphere is not included in the subset, as is the case with

$$B_{a,r} := \{x \in A : \rho(a, x) < r\}, \tag{2.9}$$

we call $B_{a,r}$ an **open ball** with center a and radius r. Thereby, the sphere itself, also with center a and radius r, can be defined as follows:

$$\partial B_{a,r} := \{x \in A : \rho(a, x) = r\}. \tag{2.10}$$

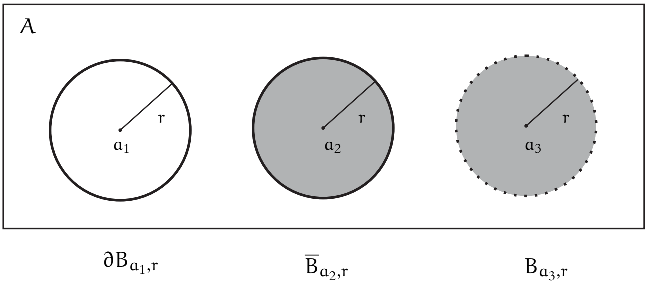

Figure 2.1 – Sphere, closed and open balls.

A subset A_1 of A is said to be **open** in A if any of its elements is the center of an open ball subset of A_1, that is, for every $a \in A_1$, there is always a scalar $r \in \mathbb{R}$ such that $B_{a,r} \subset A_1$. A set $A_2 \subset A$ is **closed** if its complement is open in A. Thereby, we can say that the complement of the open set A_1 is closed in A. In general terms, open sets, being a generalization of open intervals, are devoid of elements in borders, which refers to the idea of boundaries and interiors. A subset that results from the union of a boundary and an interior is closed because its complement is open. In mathematical terms, considering a set $A_3 \subset A$, there is an **interior** $\widehat{A}_3$ of A_3 defined by

$$\widehat{A}_3 = \{x \in A_3 : \exists r \in \mathbb{R} \text{ where } B_{x,r} \subset A_3\} \tag{2.11}$$

and a **closure** $\overline{A}_3$ of A_3 defined by

$$\overline{A}_3 = \{x \in A : A_3 \cap B_{x,r} \neq \emptyset, \forall r \in \mathbb{R}\}, \tag{2.12}$$

such that $\partial A_3 := \overline{A}_3 \setminus \widehat{A}_3$ is the **boundary** of A_3. From these definitions, we can conclude that A_3 is open when $A_3 = \widehat{A}_3$ and closed when $A_3 = \overline{A}_3$. An open subset is called closed-open or **clopen** when its complement is also open. As an example, the sets $W_1 = (1,2)$ and $W_2 = (3,4)$, defined by open intervals of real values, are clopen subsets of set $W_1 \cup W_2$. It is important to say that the empty set and the set A, to which the

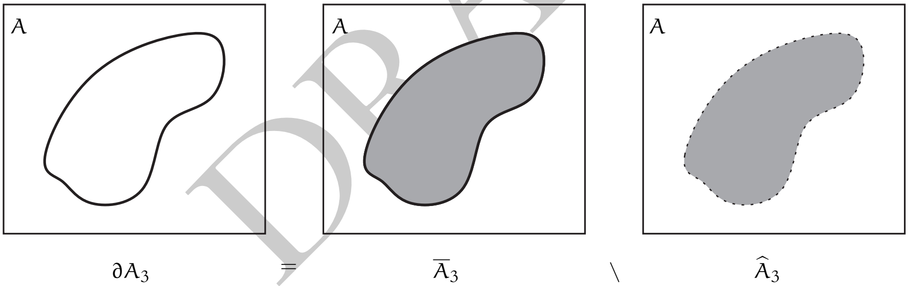

Figure 2.2 – Boundary, closure and interior of A_3.

elements of the subsets A_i belong, are defined to be clopen.

> *Proof.* Let's prove that W_1 and W_2 are clopen in $W := W_1 \cup W_2$. Let $(w - 1/2, w + 1/2)$ be an open interval in $\mathbb{R}$ where $w \in W$. This interval centered in $w = 2$ results $(3/2, 2)$, which is also open since there are no elements greater than 2 and less than $5/2$. Through this same process, it is always possible to find an open interval centered in an arbitrary $w \in W_1$; when we conclude that W_1 is open in W. By this same reasoning, W_2 is also open. But W_1 and W_2 are also closed because they are each other's open complement in W. □

Still considering the conditions above, the space (A, ρ) is called **connected** when there is no proper non empty subset that is clopen in A; otherwise, the space is said to be **disconnected**, as is the case of a metric space defined by $W_1 \cup W_2$. In other words, a disconnected space results from the union of disjoint open non empty subsets.

Intuitively, we can say that this space is fragmented, constituted by scattered collections of elements. Therefore, it is possible to conclude that vector spaces, as defined in this chapter, can not be disconnected because their defining fields $\mathbb{R}$ or $\mathbb{C}$ are connected.

A metric space (U,ρ), subspace of (Z,ρ), is called **bounded** if there are an element $u \in U$ and a scalar $r \in \mathbb{R}$ such that $U \subset B_{u,r}$. Now we shall restrict this condition a little more, but firstly let $C = \{U_1, U_2, \cdots\}$ be an infinite set constituted by subsets of Z. We say that C covers U or that C is a **cover** of U when $U \subseteq \bigcup_{i=1}^{\infty} U_i$. If there is a finite set $\{B_{u_1,r}, B_{u_2,r}, \cdots, B_{u_n,r}\}$, $u_i \in U$ and $r \in \mathbb{R}$, that covers U, we say that (U,ρ) is a **totally bounded** space. Boundedness and, more strongly, total boundedness are restrictions that impose on the space in question a feature of being delimited, from which it is possible to attain the concept of size.

Considering a sequence where distances between its elements decrease as it progresses, there are metric spaces in which every such sequence is convergent. In simple terms, we may say that these spaces result devoid of "voids" or completely "filled". Mathematically, given a metric space (V,ρ), in a sequence of elements $v_1, v_2, \cdots \in V$ where

$$\lim_{\min(i,j)\to\infty} \rho(v_i, v_j) = 0, \tag{2.13}$$

called **Cauchy Sequence**, the infinite decrease of distances is assured. If any Cauchy Sequence in V is convergent, that is, in addition to the limit above, if there is a $v \in V$ where

$$\lim_{i\to\infty} \rho(v_i, v) = 0, \tag{2.14}$$

the metric space in question is said to be **complete**. When a complete metric space is also connected and totally bounded, it is called a **continuum**. Thereby, for the purposes of our study, *every continuum is a metric space defined by a delimited set that is devoid of "voids" and not "fragmented".*

Theorem 4 – Isometry Preserves Completeness

If an isometry has a complete domain then its image is also complete.

Proof. Considering the isometric mapping $f : V \mapsto W$, where the domain V is complete, and $v_1, v_2, \cdots, v_n$ an arbitrary Cauchy Sequence in V, from the definition of isometry, the equalities

$$\lim_{\min(i,j)\to\infty} \rho_V(v_i, v_j) = \lim_{\min(i,j)\to\infty} \rho_W(f(v_i), f(v_j)) = 0$$

show that $f(v_1), f(v_2), \cdots, f(v_n) \in R_f$ is also a Cauchy Sequence. Moreover, if $v_1, v_2, \cdots, v_n$ converges to v, the equalities

$$\lim_{i\to\infty} \rho_V(v_i, v) = \lim_{i\to\infty} \rho_W(f(v_i), f(v)) = 0$$

show that every Cauchy Sequence in R_f is convergent. □

Now, let's bring the concept of distance to the context of vector spaces, since they are the main object of this chapter. A vector space $V_{\mathbb{F}}$ that is structured by a metric

ρ is defined to be a **metric vetor space** $(V_{\mathbb{F}}, \rho)$. From this definition, specific types of metric and vector spaces already presented can be combined, and then three important spaces arise: a normed complete space or, more briefly, a **Banach space**, represented by $(V_{\mathbb{F}}, \rho, \eta)$, where

$$\rho(\boldsymbol{v}_1, \boldsymbol{v}_2) := \eta(\boldsymbol{v}_1 - \boldsymbol{v}_2), \forall \boldsymbol{v}_1, \boldsymbol{v}_2 \in V_{\mathbb{F}}; \tag{2.15}$$

a Banach space with an inner product $(V_{\mathbb{F}}, \rho, \eta, \xi)$, called a **Hilbert space**, whose inner product induces the norm through $\eta(\boldsymbol{x}) = \sqrt{\xi(\boldsymbol{x}, \boldsymbol{x})}$; and a real finite dimensional Hilbert space $(V_{\mathbb{R}}, \rho, \eta, \xi)$, called **Euclidean space**. In order to avoid notational abuse, all metric vector spaces will henceforth be identified only by the definer vector space: for example, the quadruple $(V_{\mathbb{F}}, \rho, \eta, \xi)$ will be described by "the Hilbert space $V_{\mathbb{F}}$", where the functions are implied. From the concept of continuum already presented, note that *totally bounded Banach spaces are continua*.

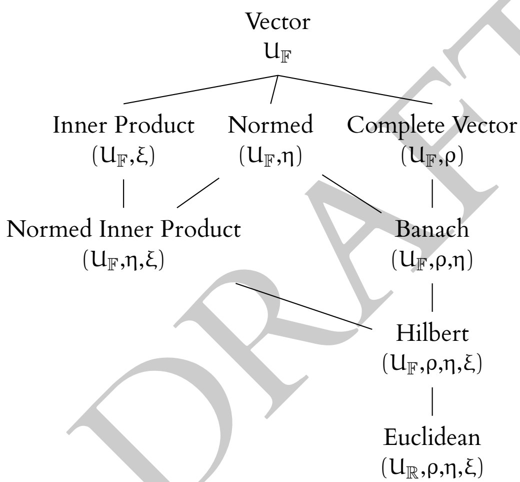

Figure 2.3 – Notable combinations of vector spaces.

Theorem 5 – Orthogonal Basis in Hilbert Spaces

Every Hilbert space has orthogonal basis.

Proof. Through a long and tedious proof, starting from the so called **Zorn's Lemma**, it is possible to obtain that every Hilbert space has a basis (see Kreyszig[32]). Once the existence of a basis is assured, the **Gram-Schmidt Algorithm** is able to find an orthogonal set from any other set as follows. Let $U = \{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ be a basis and $X = \{\mathbf{x}_1, \cdots, \mathbf{x}_n\}$ a set where $\mathbf{x}_1 = \mathbf{u}_1$. If $n = 2$, the goal is to find a $\mathbf{x}_2 \perp \mathbf{x}_1$ that makes X orthogonal. The algorithm proposes that $\mathbf{x}_2 = p_{21}\mathbf{x}_1 + \mathbf{u}_2$, where $p_{21} := -(\mathbf{u}_2 \cdot \mathbf{x}_1)/\|\mathbf{x}_1\|^2$. It is evident that an arbitrary vector $\mathbf{u} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2$, and then $\mathbf{u} = (\alpha_1 - p_{21})\mathbf{x}_1 + \alpha_2\mathbf{x}_2$; therefore, $\operatorname{span}(U) = \operatorname{span}(X)$. When $n = 3$, vector $\mathbf{x}_3 \perp \{\mathbf{x}_1, \mathbf{x}_2\}$ is found from $\mathbf{x}_3 = p_{31}\mathbf{x}_1 + p_{32}\mathbf{x}_2 + \mathbf{u}_3$, where scalar $p_{31} := -(\mathbf{u}_3 \cdot \mathbf{x}_1)/\|\mathbf{x}_1\|^2$ and scalar $p_{32} := -(\mathbf{u}_3 \cdot \mathbf{x}_2)/\|\mathbf{x}_2\|^2$. A vector $\mathbf{u} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3$ can be rewritten as the vector

> $\mathbf{u} = (\alpha_1 - \alpha_2 p_{21} - \alpha_3 p_{31})\mathbf{x}_1 + (\alpha_2 - \alpha_3 p_{32})\mathbf{x}_2 + \alpha_3 \mathbf{x}_3$, from which results $\operatorname{span}(U) = \operatorname{span}(X)$. This same process can be done for any $n > 3$. $\square$

The theorem above also assures the existence of orthonormal bases because they can be obtained from normalization of orthogonal bases. Thereby, for an arbitrary orthogonal basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ of a Hilbert space, it is clear that inner products

$$\mathbf{u}_i \cdot \mathbf{u}_j / \|\mathbf{u}_i\| \overline{\|\mathbf{u}_j\|} = \hat{\mathbf{u}}_i \cdot \hat{\mathbf{u}}_j = \delta_{ij}\,. \tag{2.16}$$

Now, let $\boldsymbol{x}$ and $\boldsymbol{y}$ be arbitrary vectors of Euclidean space $E_{\mathbb{R}}$, of which $\hat{B} = \{\hat{\boldsymbol{v}}_1, \cdots, \hat{\boldsymbol{v}}_n\}$ is an orthonormal basis. Then, we can say that inner product

$$\boldsymbol{x} \cdot \boldsymbol{y} = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \beta_j \hat{\boldsymbol{v}}_i \cdot \hat{\boldsymbol{v}}_j = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \beta_j \delta_{ij} = \sum_{i=1}^{n} \alpha_i \beta_i\,, \tag{2.17}$$

where $(\alpha_1, \cdots, \alpha_n)$ and $(\beta_1, \cdots, \beta_n)$ are the coordinates of $\boldsymbol{x}$ and $\boldsymbol{y}$ on $\hat{B}$ respectively. As a consequence of this equality, where inner products of basis vectors do not contribute numerically, a standard orthonormal basis $O = \{\hat{\mathbf{e}}_1, \cdots, \hat{\mathbf{e}}_n\}$, called **natural basis**, is defined in Euclidean spaces. From this basis, it is possible to say that the scalars $x_i := \boldsymbol{x} \cdot \hat{\mathbf{e}}_i$ constitute the **natural coordinates** $(x_1, \cdots, x_n)$ of $\boldsymbol{x}$.

The presence of **reciprocal sets** is another consequence of the existence of orthogonal sets in Hilbert spaces. We say that $U = \{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ and $W = \{\boldsymbol{w}_1, \cdots, \boldsymbol{w}_n\}$, subsets of the Hilbert space $V_{\mathbb{F}}$, are reciprocal or **biorthogonal** if their vectors are non zero and $\mathbf{u}_i \cdot \boldsymbol{w}_j = \delta_{ij}$. As the pair of reciprocal sets is unique, notations relative to one of these sets are usually defined: for the present case, set $U^{\perp} := W$ and vectors $\mathbf{u}^i := \boldsymbol{w}_i$. It is interesting to note that if the subset U is orthonormal, its reciprocal set $U^{\perp} = U$. Now, considering B a basis of $V_{\mathbb{F}}$ and $B^{\perp}$ its reciprocal set, let $\mathbf{u} = \sum_{i=1}^{n} \gamma_i \mathbf{u}^i$. If this vector $\mathbf{u}$ is zero, then

$$\Big(\sum_{j=1}^{n} \gamma_j \mathbf{u}^j\Big) \cdot \mathbf{u}_i = \sum_{j=1}^{n} \gamma_j \delta_{ij} = \gamma_i = 0\,.$$

This result shows that $B^{\perp}$ is linearly independent since the scalars γ_i are zero when $\mathbf{u} = 0$. Moreover, as both reciprocal sets have the same number of elements, we can conclude that if one of them is a basis of $V_{\mathbb{F}}$, so is the other. Thereby, if $(\alpha_1, \cdots, \alpha_n)$ are the coordinates of a vector on basis B, we usually use $(\alpha^1, \cdots, \alpha^n)$ to represent the coordinates of this same vector on basis $B^{\perp}$.

> *Proof.* Let's verify the uniqueness and existence of reciprocal sets on the context above. From theorem 5, we can admit an orthogonal subset $Z = \{\boldsymbol{z}_1, \cdots, \boldsymbol{z}_n\}$. Thereby, let $\{\tilde{\boldsymbol{z}}_1, \cdots, \tilde{\boldsymbol{z}}_n\}$ be a subset where $\tilde{\boldsymbol{z}}_i := \boldsymbol{z}_i / \|\boldsymbol{z}_i\| \overline{\|\boldsymbol{z}_i\|}$. Therefore, $\boldsymbol{z}_i \cdot \tilde{\boldsymbol{z}}_j = (\boldsymbol{z}_i \cdot \boldsymbol{z}_j) / \|\boldsymbol{z}_j\| \overline{\|\boldsymbol{z}_j\|} = \delta_{ij}$, which proves the existence. Now, supposing that there exists another subset $\{\mathbf{x}_1, \cdots, \mathbf{x}_n\}$ reciprocal to Z, we can say that $\boldsymbol{z}_i \cdot (\tilde{\boldsymbol{z}}_j - \mathbf{x}_j) = 0$. As vectors $\boldsymbol{z}_i$, $\tilde{\boldsymbol{z}}_j$ and $\mathbf{x}_j$ can not be zero, then $\tilde{\boldsymbol{z}}_j = \mathbf{x}_j$, which proves the uniqueness. $\square$

2.3 Linear Functions

The most fundamental relationships studied in Linear Algebra have the feature of preserving the group structures involved, including those defined by fields. Such relationships are expressed by homomorphisms whose main property is to keep structures of vector spaces unaltered. Moreover, if these vector spaces are metric, it is required this additional structure to remain unchanged as well. In practical terms, this means that if the homomorphism domain is a metric vector space, so must be its image. Selected through a criteria of defining the same mapping, we study these functions by gathering them in a vector space, where there are additional restrictions concerning the relations to their arguments. Moreover, from then on, we call a function whose domain is a vector space a **vector function**. In particular, if this domain is a field, its function is said to be scalar. Concerning its codomain, the vector function is usually called scalar or vector valued: for example, a function that maps a vector to a scalar is a scalar valued vector function.

Let's start the study of linear functions considering first an additive group $V_{\mathbb{F}}^{U}$ constituted of generic functions that define mappings of the type $U_{\mathbb{F}} \mapsto V_{\mathbb{F}}$, where $U_{\mathbb{F}}$ and $V_{\mathbb{F}}$ are complete vector spaces. The elements of this additive group This group is said to define a vector space $V_{\mathbb{F}}^{U}$, usually called a **function space**, if for arbitrary vector valued vector functions $\mathbf{f},\mathbf{g} \in V_{\mathbb{F}}^{U}$ and scalar $\alpha \in \mathbb{F}$ the following restrictions are observed:

i. $\mathbf{0}(\mathbf{x}) = 0$;

ii. $[\alpha\mathbf{f}](\mathbf{x}) = \alpha\mathbf{f}(\mathbf{x})$;

iii. $[\mathbf{f}+\mathbf{g}](\mathbf{x}) = \mathbf{f}(x)+\mathbf{g}(\mathbf{x})$.

The domain U may eventually be a cartesian product $W^{\times q}$, where an arbitrary function $\mathbf{f}$ of the function space has a q-tuple of vectors or q vectors as arguments, and its value is represented by $\mathbf{f}(\boldsymbol{w}_1,\cdots,\boldsymbol{w}_q)$, where the tuple $(\boldsymbol{w}_1,\cdots,\boldsymbol{w}_q) \in W^{\times q}$ or the vectors $\boldsymbol{w}_i \in W_i$. Since they are vector spaces, we can adopt for function spaces the same nomenclature described in figure 2.3 if inner product, norm and metric are defined accordingly.

An important example of function space is the space constituted by continuous functions. In order to define these type of functions, we need to say firstly that a set $S \subset U_{\mathbb{F}}$ is called a **neighborhood** of an element $\mathbf{u} \in S$, represented by $\{u\}_v$, when there is a real number $r > 0$ that defines an open ball $B_{\mathbf{u},r} \subset S$. In this context, the function in $g : U_{\mathbb{F}} \mapsto V_{\mathbb{F}}$ is said to be **continuous** on an element $\mathbf{u} \in S$ if for any neighborhood $\{\mathbf{g}(\mathbf{u})\}_v$ in the codomain there is a neighborhood $\{\mathbf{u}\}_v$ in the domain where every element $x \in \{\mathbf{u}\}_v$ is related to a value $\mathbf{g}(x) \in \{\mathbf{g}(\mathbf{u})\}_v$. In more direct terms, $\mathbf{g}$ is continuous on u when

$$\lim_{\mathbf{x}\to\mathbf{u}} \mathbf{g}(\mathbf{x}) = \mathbf{g}(\mathbf{u}),\ \forall \mathbf{x} \in U_{\mathbb{F}}, \tag{2.18}$$

that is, $x \to u$ implies $\mathbf{g}(x) \to \mathbf{g}(u)$. In the case of a function that is continuous on every element of the domain, it is called continuous on the domain or simply continuous and the space constituted by all continuous functions that map $U_{\mathbb{F}}$ to $V_{\mathbb{F}}$ we represent $\mathcal{C}_{\mathbb{F}}(U,V)$. Moreover, if a bijection and its inverse function are continuous on their respective domains, each one is called a **homeomorphism**[2]. There is also a particular type of function continuity that has a stronger restriction than that presented above: a function $\mathbf{g}$ is said to be **Lipschitz continuous** on $\mathbf{u}$ if there exists a scalar $\vartheta \in \mathbb{R}_*^+$, called **Lipschitz constant**, where

$$\vartheta \geqslant \frac{\rho(\mathbf{g}(x),\mathbf{g}(\mathbf{u}))}{\rho(x,\mathbf{u})}, \forall x \in \{U_{\mathbb{F}} \setminus \{\mathbf{u}\}\}. \tag{2.19}$$

From this definition we can conclude that every Lipschitz continuous function is also continuous, with the additional property of presenting upper limited distance ratios relative to every element u of its domain.

Now, let a vector function $\mathbf{h} \in V_{\mathbb{F}}^{U}$ be a homomorphism that keeps the additive structure of $U_{\mathbb{F}}$ unaltered. To our purposes, this function must also preserve the structure created by the field $\mathbb{F}$ in such a way that

$$\mathbf{h}(\alpha x + \beta y) = \alpha \mathbf{h}(x) + \beta \mathbf{h}(y), \tag{2.20}$$

for all $\alpha, \beta \in \mathbb{F}$, $x, y \in U_{\mathbb{F}}$. Thereby, we call $\mathbf{h}$ a **linear function** and the corresponding mapping a **linear transformation**. If the function space $V_{\mathbb{F}}^{U}$ has only linear functions, its usual representation is $\mathcal{L}_{\mathbb{F}}(U,V)$.

Now, given $\{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ a basis of $U_{\mathbb{F}}$, a bijective linear operator $\mathbf{g} \in \mathcal{L}_{\mathbb{F}}(U,U)$ and an arbitrary vector $\mathbf{u} \in U_{\mathbb{F}}$, we can say that since $\mathbf{g}(\mathbf{u})$ is an arbitrary vector of $U_{\mathbb{F}}$, then set $\{\mathbf{g}(\mathbf{u}_1), \cdots, \mathbf{g}(\mathbf{u}_n)\}$ is also a basis of $U_{\mathbb{F}}$ because linearity leads to $\mathbf{g}(\mathbf{u}) = \sum_{i=1}^{n} \alpha_i \mathbf{g}(\mathbf{u}_i)$.

Still considering the previous conditions, in certain cases where domain $U_{\mathbb{F}} = W_{\mathbb{F}}^{\times q}$, a function $\mathbf{k}$ is said to be **multilinear**, or **bilinear** if $q = 2$, when

$$\begin{aligned} &\mathbf{k}(w_1, \cdots, \alpha w_i + \beta w, \cdots, w_q) = \\ &\qquad \alpha \mathbf{k}(w_1, \cdots, w_i, \cdots, w_q) + \beta \mathbf{k}(w_1, \cdots, w, \cdots, w_q), \end{aligned} \tag{2.21}$$

for all $\alpha, \beta \in \mathbb{F}$ and $w, w_i \in W_i$. In this context, when domain $U_{\mathbb{F}} = V_{\mathbb{F}}^{q}$, the vectors of $V_{\mathbb{F}}^{V^q}$ are called **multilinear operators** and the mappings they define are **multilinear operations**. Moreover, for the purposes of our study, it is useful to define a "kind of" multilinear function $\mathbf{c} \in V_{\mathbb{F}}^{W^{\times q}}$, called **conjugate multilinear** or **multiantilinear** function, where

$$\begin{aligned} &\mathbf{c}(w_1, \cdots, \alpha w_i + \beta w, \cdots, w_q) = \\ &\qquad \overline{\alpha} \mathbf{c}(w_1, \cdots, w_i, \cdots, w_q) + \overline{\beta} \mathbf{c}(w_1, \cdots, w, \cdots, w_q), \end{aligned} \tag{2.22}$$

[2]Not to be confused with homomorphism, without "e".

for all $\alpha, \beta \in \mathbb{F}$ and $\boldsymbol{w}, \boldsymbol{w}_i \in W_i$. In the context of vector spaces defined by real fields, a multiantilinear function results multilinear. If $q = 1$ or $q = 2$, function $\mathbf{c}$ is called antilinear or biantilinear respectively.

Linear functions, as we presented, can also be continuous and constitute a normed function space if a norm is defined. In our study, given $Z_{\mathbb{F}}$ and $Y_{\mathbb{F}}$ Banach spaces, we define that *a vector space $\mathcal{L}_{\mathbb{F}}(Z,Y)$ of continuous linear functions, represented hereafter by $\mathcal{CL}_{\mathbb{F}}(Z,Y)$, is normed and metric inner product where the norm and the inner product are related through $\|\mathbf{g}\| = \sqrt{\mathbf{g}\cdot\mathbf{g}}$ and the distance $\rho(\mathbf{g},\mathbf{f}) := \|\mathbf{g}-\mathbf{f}\|$ for all $\mathbf{g},\mathbf{f} \in \mathcal{CL}_{\mathbb{F}}(Z,Y)$.* Thereby, for any linear function $\mathbf{h}$ in $Y_{\mathbb{F}}^{Z}$ to be continuous, a necessary and sufficient condition requires that it is **bounded**[3], namely, that there exists a number $\nu \in \mathbb{R}^+$ where

$$\nu \geqslant \|\mathbf{h}(\boldsymbol{z})\|/\|\boldsymbol{z}\|, \forall \boldsymbol{z} \in \{Z_{\mathbb{F}} \setminus \{\mathbf{0}\}\}. \tag{2.23}$$

Here, we consider the minimum of all these values ν to be the norm of $\mathbf{h}$. In more precise terms, the following rule is defined:

$$\eta(\mathbf{x}) = \sup\{\|\mathbf{x}(\boldsymbol{z})\|/\|\boldsymbol{z}\|, \forall \boldsymbol{z} \in \{Z_{\mathbb{F}} \setminus \{\mathbf{0}\}\}\}. \tag{2.24}$$

Since any field $\mathbb{F}$ is also a vector space, we call **functional** an element of the function space $\mathbb{F}_{\mathbb{F}}^{U_{\mathbb{F}}}$ or simply $\mathbb{F}^{U_{\mathbb{F}}}$, whose domain and codomain are complete vector spaces. Therefore, it can be stated that *a mapping defined by a functional maps a complete vector space to its structuring field*. In other words, if the domain of a functional has a complex field, for example, then its values are complex. By gathering the concepts of functional and linear function, the coordinates of a vector on a certain basis can be obtained from a sequence of values of linear functionals whose argument is the vector in question. In this sense, given a basis $B = \{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ of the complete vector space $U_{\mathbb{F}}$, the elements of the n-tuple $(\mathbf{f}_1^B, \cdots, \mathbf{f}_n^B)$ are called **coordinate functionals** of basis B if each $\mathbf{f}_i^B$ belongs to $\mathcal{L}_{\mathbb{F}}(U,\mathbb{F})$ and

$$\mathbf{u} = \sum_{i=1}^{n} \mathbf{f}_i^B(\mathbf{u})\mathbf{u}_i, \forall\, \mathbf{u} \in U_{\mathbb{F}}, \tag{2.25}$$

where $(\mathbf{f}_1^B(\mathbf{u}), \cdots, \mathbf{f}_n^B(\mathbf{u}))$ are the coordinates of $\mathbf{u}$ on basis B. In the particular case of an arbitrary vector $\mathbf{u}^c \in \overline{U_{\mathbb{F}}}$, from equality (2.2), we have

$$\mathbf{u}^c = (\sum_{i=1}^{n} \mathbf{f}_i^B(\mathbf{u})\mathbf{u}_i)^c = \sum_{i=1}^{n} \overline{\mathbf{f}_i^B}(\mathbf{u})\mathbf{u}_i^c \implies \overline{\mathbf{f}_i^B}(\mathbf{u}) = \mathbf{f}_i^{B^c}(\mathbf{u}^c). \tag{2.26}$$

where $B^c := \{\mathbf{u}_1^c, \cdots, \mathbf{u}_n^c\}$ is a basis of $\overline{U_{\mathbb{F}}}$. Now, since any basis vector $\mathbf{u}_i = \sum_{j=1}^{n} \mathbf{f}_j^B(\mathbf{u}_i)\mathbf{u}_j$, we conclude that

$$\mathbf{f}_j^B(\mathbf{u}_i) = \delta_{ij}. \tag{2.27}$$

[3]See theorem on Kreyszig[32], p.97.

Moreover, considering the set $\hat{B} = \{\hat{\mathbf{u}}_1, \cdots, \hat{\mathbf{u}}_n\}$ an orthonormal basis of the Hilbert space $U_{\mathbb{F}}$, we can say that

$$\mathbf{x} \cdot \hat{\mathbf{u}}_i = \sum_{j=1}^{n} \mathbf{f}_j^{\hat{B}}(\mathbf{x}) \hat{\mathbf{u}}_j \cdot \hat{\mathbf{u}}_i = \sum_{j=1}^{n} \mathbf{f}_j^{\hat{B}}(\mathbf{x}) \delta_{ji} = \mathbf{f}_i^{\hat{B}}(\mathbf{x}), \forall \mathbf{x} \in U_{\mathbb{F}}, \tag{2.28}$$

from which we obtain the following rule that assures the existence of coordinate functionals:

$$\mathbf{f}_i^{\hat{B}}(\mathbf{x}) = \mathbf{x} \cdot \hat{\mathbf{u}}_i. \tag{2.29}$$

Still in this context, for orthonormal basis vectors, it is important to point out that

$$\hat{\mathbf{u}}_j \cdot \hat{\mathbf{u}}_i^c = \hat{\mathbf{u}}_j^c \cdot \hat{\mathbf{u}}_i = \delta_{ji}. \tag{2.30}$$

Proof. Multiplying $\hat{\mathbf{u}}_j \cdot \hat{\mathbf{u}}_i = \delta_{ji}$ by a non zero scalar α we have $\hat{\mathbf{u}}_j \cdot \overline{\alpha}\hat{\mathbf{u}}_i = \hat{\mathbf{u}}_j \cdot \alpha\hat{\mathbf{u}}_i^c = \overline{\alpha}\hat{\mathbf{u}}_j \cdot \hat{\mathbf{u}}_i^c = \alpha\delta_{ji}$ and by $\overline{\alpha}$ we arrive at $\alpha\hat{\mathbf{u}}_j \cdot \hat{\mathbf{u}}_i^c = \overline{\alpha}\delta_{ji}$. When $j \neq i$, it is obvious that $\hat{\mathbf{u}}_j \cdot \hat{\mathbf{u}}_i^c$ is zero. If $j = i$, we obtain that $\hat{\mathbf{u}}_j \cdot \hat{\mathbf{u}}_i^c = \alpha/\overline{\alpha} = \overline{\alpha}/\alpha$ resulting $\alpha^2 = \overline{\alpha}^2$. From this equality, if $\alpha = a + bi$, we arrive at $b = -b$, which means that α is a real value and thus $\hat{\mathbf{u}}_j \cdot \hat{\mathbf{u}}_i^c = 1$. Verification of $\hat{\mathbf{u}}_j^c \cdot \hat{\mathbf{u}}_i = \delta_{ji}$ follows the same procedure. □

Considering a complete vector space $V_{\mathbb{F}}$, the linear function space $\mathcal{L}_{\mathbb{F}}(V, \mathbb{F})$ is said to be the **dual space** of $V_{\mathbb{F}}$, represented by $V_{\mathbb{F}}^*$, whose elements are called the **dual vectors** of $V_{\mathbb{F}}$. Intuitively, a dual vector is a scalar measure of vectors that keeps the structure of its domain unaltered; a feature that the norm, being a scalar measure, does not assure, since it is a non linear functional. The coordinate functional $\mathbf{f}_i^B$, in turn, is indeed a dual vector and, in a certain way, "measures" its argument in relation to the i-th vector of basis B, when we call the value of this measure a coordinate. As already presented, in the context of orthonormal bases, the measurement of a coordinate functional is expressed through the incidence interrelationship between its argument and a basis vector, that is, the inner product of both. A subset $\{\mathbf{g}_1, \cdots, \mathbf{g}_m\}$ of $V_{\mathbb{F}}^*$ is a **dual set** of $\{\mathbf{w}_1, \cdots, \mathbf{w}_m\} \subset V_{\mathbb{F}}$ if $\mathbf{g}_i(\mathbf{w}_j) = \delta_{ij}$. When such a subset is the dual set of a basis $B = \{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ of $V_{\mathbb{F}}$, its elements will be the coordinate functionals of B, as can be verified by the following development:

$$\mathbf{g}_i(\mathbf{x}) = \mathbf{g}_i(\sum_{j=1}^{n} \mathbf{f}_j^B(\mathbf{x}) \mathbf{u}_j) = \sum_{j=1}^{n} \mathbf{f}_j^B(\mathbf{x}) \mathbf{g}_i(\mathbf{u}_j) = \sum_{j=1}^{n} \mathbf{f}_j^B(\mathbf{x}) \delta_{ij} = \mathbf{f}_i^B(\mathbf{x}).$$

But a set constituted of coordinate functionals, in this case $B^* := \{\mathbf{f}_1^B, \cdots, \mathbf{f}_n^B\}$, is itself a basis of its dual space, that is, dual set B^* of a basis B is a basis of dual space $V_{\mathbb{F}}^*$, when we call it a **dual basis** of $V_{\mathbb{F}}$. Let's verify if this is true: given an arbitrary dual vector $\mathbf{h} \in V_{\mathbb{F}}^*$, equalities

$$\mathbf{h}(\mathbf{x}) = \mathbf{h}[\sum_{i=1}^{n} \mathbf{f}_i^B(\mathbf{x}) \mathbf{u}_i] = \sum_{i=1}^{n} \mathbf{h}(\mathbf{u}_i) \mathbf{f}_i^B(\mathbf{x}) = [\sum_{i=1}^{n} \mathbf{h}(\mathbf{u}_i) \mathbf{f}_i^B](\mathbf{x}) \tag{2.31}$$

enable us to conclude that $\mathbf{f}_i^B$ spans $V_{\mathbb{F}}^*$ and if $\mathbf{h} = \mathbf{0}$, then functions $\mathbf{f}_i^B$ result zero, which proves that B^* is linearly independent. Therefore, we can state that scalars $\mathbf{h}(\mathbf{u}_i)$ result the coordinates of $\mathbf{h}$ in B^* and a complete vector space have the same dimension of its dual or that $\dim(V_{\mathbb{F}}) = \dim(V_{\mathbb{F}}^*)$ in the present context. It is convenient that this strong correspondence between a complete vector space and its dual becomes even stronger, in such a way that vectors and dual vectors are biunivocally related, when dual vectors are renamed to **covectors**. By defining a rule similar to (2.29), the following theorem, the most important of our study, establishes this one-to-one relationship.

Theorem 6 – Riesz-Fréchet Representation

Let $\Phi : U_{\mathbb{F}} \mapsto U_{\mathbb{F}}^*$ be a mapping where $U_{\mathbb{F}}$ is a Hilbert space and $U_{\mathbb{F}}^*$ its dual space. If for every $\mathbf{u} \in U_{\mathbb{F}}$, a covector $\mathbf{u}^* := \Phi(\mathbf{u})$ is described by the rule $\mathbf{u}^*(\mathbf{x}) = \mathbf{x} \cdot \mathbf{u}$, then Φ results an antilinear bijection. When $\mathbf{u}^*$ is continuous, $\|\mathbf{u}^*\|_{U_{\mathbb{F}}^*} = \|\mathbf{u}\|_{U_{\mathbb{F}}}$.

Proof. Considering $\mathbf{u}$ and $\mathbf{v}$ vectors of $U_{\mathbb{F}}$, function Φ is proved antilinear from the following equalities:

$$[\Phi(\alpha\mathbf{u} + \beta\mathbf{v})](\mathbf{x}) = \mathbf{x} \cdot (\alpha\mathbf{u} + \beta\mathbf{v}) = \overline{\alpha}\mathbf{u}^*(\mathbf{x}) + \overline{\beta}\mathbf{v}^*(\mathbf{x}) = \left[\overline{\alpha}\Phi(\mathbf{u}) + \overline{\beta}\Phi(\mathbf{v})\right](\mathbf{x}).$$

If Φ were not an injection, there would be different covectors $\mathbf{u}^*$ and $\mathbf{v}^*$ where $\mathbf{u}^*(\mathbf{x}) = \mathbf{v}^*(\mathbf{x})$ or $\mathbf{x} \cdot \mathbf{u} = \mathbf{x} \cdot \mathbf{v}$. From this supposition, the following equalities $\mathbf{x} \cdot (\mathbf{u} - \mathbf{v}) = (\mathbf{u} - \mathbf{v})^*(\mathbf{x}) = 0$ do not confirm $\mathbf{u}^* \neq \mathbf{v}^*$ since $(\mathbf{u} - \mathbf{v})^* = \mathbf{u}^* - \mathbf{v}^*$. In order to prove that Φ is a surjection, we need to obtain for an arbitrary functional $\mathbf{g} \in U_{\mathbb{F}}^*$ a vector $\mathbf{u} \in U_{\mathbb{F}}$ such that $\Phi(\mathbf{u}) = \mathbf{g}$. Considering the rule (2.29) and an orthonormal basis $\hat{B} = \{\hat{\mathbf{u}}_1, \cdots, \hat{\mathbf{u}}_n\}$ of $U_{\mathbb{F}}$ whose coordinate functionals span $U_{\mathbb{F}}^*$, we can say that

$$\mathbf{g}(\mathbf{x}) = [\sum_{i=1}^{n} \alpha_i \mathbf{f}_i^{\hat{B}}](\mathbf{x}) = \sum_{i=1}^{n} \alpha_i \mathbf{f}_i^{\hat{B}}(\mathbf{x}) = \sum_{i=1}^{n} \alpha_i (\mathbf{x} \cdot \hat{\mathbf{u}}_i) = \mathbf{x} \cdot (\sum_{i=1}^{n} \overline{\alpha_i} \hat{\mathbf{u}}_i).$$

Since $\mathbf{g}(\mathbf{x}) = [\Phi(\mathbf{u})](\mathbf{x}) = \mathbf{x} \cdot \mathbf{u}$, the existence of $\mathbf{u} = \sum_{i=1}^{n} \overline{\alpha_i} \hat{\mathbf{u}}_i$ is verified. Finally, using definition (2.24) on the present conditions, leaving space representation on norms implicit, we obtain equality $\|\mathbf{u}^*\| = \sup\{|\mathbf{x} \cdot \mathbf{u}|/\|\mathbf{x}\|\}$ for all non zero $\mathbf{x}$. If $\mathbf{u}$ is zero, it is evident that $\|\mathbf{u}^*\| = \|\mathbf{u}\|$; otherwise, $\mathbf{u}^*$ is non zero and we conclude that $\|\mathbf{u}^*\| \geqslant |\mathbf{x} \cdot \mathbf{u}|/\|\mathbf{x}\|$. Cauchy-Schwarz Inequality states that $|\mathbf{x} \cdot \mathbf{u}| \leqslant \|\mathbf{x}\| \, \|\mathbf{u}\|$. Subtracting these two previous inequalities, we arrive at inequality $(\|\mathbf{u}^*\| - \|\mathbf{u}\|)\|\mathbf{x}\| \geqslant 0$, whose left side we assume to be zero. In this context, we conclude $\|\mathbf{u}^*\| = \|\mathbf{u}\|$ because $\mathbf{u}^*$, $\mathbf{u}$ and $\mathbf{x}$ are not zero. □

Considering the conditions of the theorem, if $\mathbf{f}$ is a vector function whose domain is a Hilbert space, it is obvious that $\mathbf{f}(\mathbf{x}) = \mathbf{f} \circ \Phi^{-1}(\mathbf{x}^*)$. From this equality and by defining a **covector function** as a function whose domain is constituted by covectors, we state that every linear vector function $\mathbf{g}$ on the context of Hilbert spaces has a unique associated linear covector function $\mathbf{g}_\Phi := \mathbf{g} \circ \Phi^{-1}$ and vice-versa. The following corollary explains this important statement more formally.

Corollary 6.1 – Linear Vector and Covector Functions

Given a Hilbert space $V_{\mathbb{F}}$, a mapping $\varphi : \mathcal{L}_{\mathbb{F}}(U, V) \mapsto \mathcal{L}_{\mathbb{F}}(U^*, V)$ is a linear bijective transformation if its function rule is $\varphi(\mathbf{x}) = \mathbf{x}_\Phi = \mathbf{x} \circ \Phi^{-1}$.

Proof. If $\mathbf{g}$ and $\mathbf{h}$ are elements of $\mathcal{L}_{\mathbb{F}}(\mathsf{U},\mathsf{V})$, then φ is linear because

$$(\alpha\mathbf{g}+\beta\mathbf{h})_{\Phi} = (\alpha\mathbf{g}+\beta\mathbf{h})\circ\Phi^{-1} = [\alpha\mathbf{g}]\circ\Phi^{-1}+[\beta\mathbf{h}]\circ\Phi^{-1} = \alpha\mathbf{g}_{\Phi}+\beta\mathbf{h}_{\Phi}\,,$$

according to restrictions on page 42. If φ weren't an injection there'd be two distinct functions $\mathbf{g}$ and $\mathbf{h}$ where $\mathbf{g}\circ\Phi^{-1}(\mathbf{x}^*) = \mathbf{h}\circ\Phi^{-1}(\mathbf{x}^*)$; but this equality leads to $\mathbf{g}(\mathbf{x}) = \mathbf{h}(\mathbf{x})$, which shows that functions $\mathbf{g}$ and $\mathbf{h}$ are not distinct. Moreover, function φ is also a surjection because for every covector function $\mathbf{z}\in\mathcal{L}_{\mathbb{F}}(\mathsf{U}^*,\mathsf{V})$ we can find a vector function $\mathbf{z}\circ\Phi\in\mathcal{L}_{\mathbb{F}}(\mathsf{U},\mathsf{V})$. □

In the context of a field $\mathbb{F}$, which is also a Hilbert space, the Riesz-Fréchet Representation theorem leads to another interesting property that biunivocally relates $\mathbb{F}$ and the dual space of a one dimensional Hilbert space $\mathsf{U}_{\mathbb{F}}$. Since this property will be relevant for future purposes, we shall present it as follows.

Corollary 6.2 – Scalars and Covectors

In the case of a one dimensional $\mathsf{U}_{\mathbb{F}}$, considering set $\{\hat{\mathbf{u}}\}$ its orthonormal basis, the mapping $\omega:\mathbb{F}\mapsto\mathsf{U}_{\mathbb{F}}^*$ is a linear bijective transformation if $\omega(x) = x\hat{\mathbf{u}}^*$.

Proof. The proof of ω linearity is trivial. Now let's verify if it is an injection: considering arbitrary nonzero and distinct $\alpha,\beta\in\mathbb{F}$, since $x\hat{\mathbf{u}}^*(\mathbf{x}) = x(\mathbf{x}\cdot\hat{\mathbf{u}}) = \mathbf{x}\cdot\overline{x}\hat{\mathbf{u}} = (\overline{x}\hat{\mathbf{u}})^*(\mathbf{x})$, the inequality $\omega(\alpha) = (\overline{\alpha}\hat{\mathbf{u}})^* \neq (\overline{\beta}\hat{\mathbf{u}})^* = \omega(\beta)$ is always valid because $\overline{\alpha}\hat{\mathbf{u}}\neq\overline{\beta}\hat{\mathbf{u}}$. Injection ω is also a surjection because, for an arbitrary $\mathbf{v}^*\in\mathsf{U}_{\mathbb{F}}^*$, we have $\mathbf{v}^* = (\alpha\hat{\mathbf{u}})^* = \overline{\alpha}\hat{\mathbf{u}}^* = \omega(\overline{\alpha})$. □

The Riesz-Fréchet Representation also enables us to define a rule for coordinate functionals on bases not necessarily orthogonal. Let's see how this happens. Considering the conditions of the theorem, let $\mathsf{B} = \{\mathbf{u}_1,\cdots,\mathbf{u}_n\}$ be a basis of $\mathsf{U}_{\mathbb{F}}$ and basis $\mathsf{B}^* = \{\mathbf{f}_1^{\mathsf{B}},\cdots,\mathbf{f}_n^{\mathsf{B}}\}$ its dual correspondent. In this context, functionals $\mathbf{f}_i^{\mathsf{B}}\in\mathsf{U}_{\mathbb{F}}^*$ are biunivocally related to vectors $\mathbf{v}_i\in\mathsf{U}_{\mathbb{F}}$ in such a way that, given a vector $\mathbf{u}\in\mathsf{U}_{\mathbb{F}}$, the following equalities are valid:

$$\mathbf{f}_i^{\mathsf{B}}(\mathbf{u}) = \mathbf{u}\cdot\mathbf{v}_i = \sum_{j=1}^{n}\mathbf{f}_j^{\mathsf{B}}(\mathbf{u})\mathbf{u}_j\cdot\mathbf{v}_i\,.$$

From these equalities, in order to obtain identity $\mathbf{f}_i^{\mathsf{B}}(\mathbf{u}) = \mathbf{f}_i^{\mathsf{B}}(\mathbf{u})$, inner product $\mathbf{u}_j\cdot\mathbf{v}_i$ must be δ_{ji}. Thereby, we can say that subset $\{\mathbf{v}_1,\cdots,\mathbf{v}_n\}$, whose elements are biunivocally related to the vectors of B^*, is the reciprocal basis $\mathsf{B}^{\perp} = \{\mathbf{u}^1,\cdots,\mathbf{u}^n\}$, that is, covectors $\mathbf{v}_i^* = (\mathbf{u}^i)^* = \mathbf{f}_i^{\mathsf{B}}$. Then, from the previous theorem, the rule for coordinate functionals can be described by

$$\mathbf{f}_i^{\mathsf{B}}(\mathbf{x}) = (\mathbf{u}^i)^*(\mathbf{x}) = \mathbf{x}\cdot\mathbf{u}^i\,. \tag{2.32}$$

If $\mathsf{B}^{\perp}$ is the reciprocal basis of B, the inverse is also true; thus, we can affirm that

$$\mathbf{f}_i^{\mathsf{B}^{\perp}}(\mathbf{x}) = (\mathbf{u}_i)^*(\mathbf{x}) = \mathbf{x}\cdot\mathbf{u}_i\,. \tag{2.33}$$

In this context, given an arbitrary vector $\mathbf{u} \in \mathrm{U}_{\mathbb{F}}$, from development

$$
\begin{aligned}
\mathbf{u}^*(x) &= x \cdot \sum_{i=1}^{n} \mathbf{f}_i^{B}(\mathbf{u})\mathbf{u}_i \\
&= \sum_{i=1}^{n} \overline{\mathbf{f}_i^{B}}(\mathbf{u}) x \cdot \mathbf{u}_i \\
&= \sum_{i=1}^{n} \overline{\mathbf{f}_i^{B}}(\mathbf{u}) \mathbf{f}_i^{B^{\perp}}(x) \\
&= [\sum_{i=1}^{n} \overline{\mathbf{f}_i^{B}}(\mathbf{u}) \mathbf{f}_i^{B^{\perp}}](x),
\end{aligned}
$$

we conclude that each coodinate of covector $\mathbf{u}^*$ on basis $(B^*)^{\perp}$ is the complex conjugate of its corresponding vector coordinate basis B, that is,

$$
\mathbf{u}^* = \sum_{i=1}^{n} \overline{\mathbf{f}_i^{B}}(\mathbf{u}) \mathbf{f}_i^{B^{\perp}}. \tag{2.34}
$$

If basis B is orthornomal, that is, $B^{\perp} = B$, and field $\mathbb{F}$ real then the coordinates of $\mathbf{u}$ on B and of $\mathbf{u}^*$ on B^* are the same. Now, the inner product between two arbitrary vectors $\mathbf{u}$ and $\mathbf{v}$ leads to the following equalities:

$$
\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{n} \sum_{j=1}^{n} \mathbf{f}_i^{B}(\mathbf{u}) \overline{\mathbf{f}_j^{B^{\perp}}(\mathbf{v}) \mathbf{u}_i \cdot \mathbf{u}^j} = \sum_{i=1}^{n} \mathbf{f}_i^{B}(\mathbf{u}) \overline{\mathbf{f}_i^{B^{\perp}}}(\mathbf{v}). \tag{2.35}
$$

Again, if B is orthonormal and $\mathbb{F}$ real, we have $\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{n} \mathbf{f}_i^{B}(\mathbf{u}) \mathbf{f}_i^{B}(\mathbf{v})$. Figure 2.4 summarizes the relationships between basis B and the reciprocal bases it induces. In the particular case of an orthonormal basis $\hat{B}$, vectors $\hat{\mathbf{u}}_i = \hat{\mathbf{u}}^i$, from which we conclude that the elements of $\hat{B}$ and of $\hat{B}^*$ have a biunivocal relationship in terms of the previous theorem. Thereby, equalities $\mathbf{f}_i^{\hat{B}}(x) = (\hat{\mathbf{u}}_i)^*(x) = x \cdot \hat{\mathbf{u}}_i$ enable us to say that (2.32) is a generalization of (2.29) in the context of Hilbert spaces.

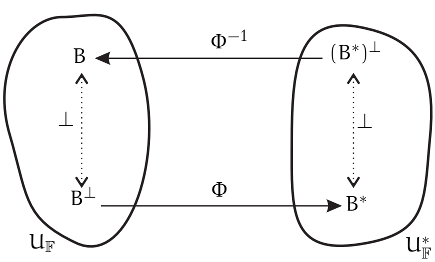

Figure 2.4 – Relationships between bases induced by B in terms of theorem 6.

Still considering the terms of the previous theorem, from equality between norms of vectors and norms of continuous covectors it is possible to obtain that the dual space

$\mathsf{U}^*_{\mathbb{F}}$ is also a Hilbert space. Since this feature is very important, we shall present it in a more formal way, through the following corollary.

Corollary 6.3 – Hilbert Dual Space

If the elements of $\mathsf{U}^*_{\mathbb{F}}$ are continuous, then $\mathsf{U}^*_{\mathbb{F}}$ is a Hilbert space.

Proof. By the rule (2.24), $\mathsf{U}^*_{\mathbb{F}}$ is a normed space that is also defined to be metric inner product (See p. 44). Thus, it remains here to verify that $\mathsf{U}^*_{\mathbb{F}}$ is complete. This condition is assured if the bijection Φ is isometric, according to theorem 4. From the rule for covectors established by Riesz-Fréchet Representation, it is trivial to obtain that $(\boldsymbol{v}-\boldsymbol{w})^* = \boldsymbol{v}^* - \boldsymbol{w}^*$, or that $\Phi(\boldsymbol{v}-\boldsymbol{w}) = \Phi(\boldsymbol{v}) - \Phi(\boldsymbol{w})$. Thereby, knowing that $\|\Phi(\boldsymbol{x})\|_{\mathsf{U}^*_{\mathbb{F}}} = \|\boldsymbol{x}\|_{\mathsf{U}_{\mathbb{F}}}$,

$$\rho_{\mathsf{U}^*_{\mathbb{F}}}(\Phi(\boldsymbol{v}),\Phi(\boldsymbol{w})) = \|\Phi(\boldsymbol{v})-\Phi(\boldsymbol{w})\|_{\mathsf{U}^*_{\mathbb{F}}} = \|\Phi(\boldsymbol{v}-\boldsymbol{w})\|_{\mathsf{U}^*_{\mathbb{F}}} = \|\boldsymbol{v}-\boldsymbol{w}\|_{\mathsf{U}_{\mathbb{F}}} = \rho_{\mathsf{U}_{\mathbb{F}}}(\boldsymbol{v},\boldsymbol{w}).$$

□

Now, if $\mathsf{U}_{\mathbb{F}}$ and $\mathsf{V}_{\mathbb{F}}$ are Hilbert spaces that define $\mathcal{CL}_{\mathbb{F}}(\mathsf{U},\mathsf{V})$, we call $\mathbf{g}^\dagger \in \mathcal{CL}_{\mathbb{F}}(\mathsf{V},\mathsf{U})$ the **Hilbert-adjoint** function of $\mathbf{g} \in \mathcal{CL}_{\mathbb{F}}(\mathsf{U},\mathsf{V})$, when, given arbitrary vectors $\boldsymbol{u} \in \mathsf{U}_{\mathbb{F}}$ and $\boldsymbol{v} \in \mathsf{V}_{\mathbb{F}}$ we have

$$\mathbf{g}^\dagger(\boldsymbol{u})\cdot\boldsymbol{v} = \boldsymbol{u}\cdot\mathbf{g}(\boldsymbol{v}). \tag{2.36}$$

In the particular case of real fields $\mathbf{g}^\dagger$ is also called the **transpose**of $\mathbf{g}$, represented by $\mathbf{g}^{\mathrm{T}}$. In the case of Hilbert-adjoint functions, the following properties are valid for all $\alpha \in \mathbb{F}$ and $\mathbf{k} \in \mathcal{CL}_{\mathbb{F}}(\mathsf{V},\mathsf{U})$:

i. $(\alpha\mathbf{g})^\dagger = \overline{\alpha}\mathbf{g}^\dagger$;

ii. $(\mathbf{g}\circ\mathbf{k})^\dagger = \mathbf{k}^\dagger \circ \mathbf{g}^\dagger$;

iii. If $\mathbf{g}$ is a bijection, there is a function $\mathbf{g}^{-\dagger} := (\mathbf{g}^{-1})^\dagger = (\mathbf{g}^\dagger)^{-1}$.

It is important to point out that these three properties are also valid for the case of transpose functions, the first property being slightly different: $(\alpha\mathbf{g})^{\mathrm{T}} = \alpha\mathbf{g}^{\mathrm{T}}$.

Proof. Firstly, we need to prove the existence and uniqueness of Hilbert-adjoint functions. On equality (2.36), let's consider $\boldsymbol{u} = \hat{\boldsymbol{u}}_k$ and $\boldsymbol{v} = \hat{\boldsymbol{v}}_k$, where vectors on the right sides belong to orthonormal bases $\hat{B}_1$ and $\hat{B}_2$ of n-dimensional Hilbert spaces $\mathsf{U}_{\mathbb{F}}$ and $\mathsf{V}_{\mathbb{F}}$ respectively. Thereby, we can develop the following:

$$\hat{\boldsymbol{u}}_k\cdot\mathbf{g}^\dagger(\hat{\boldsymbol{v}}_k) = \mathbf{g}(\hat{\boldsymbol{u}}_k)\cdot\hat{\boldsymbol{v}}_k$$

$$\hat{\boldsymbol{u}}_k\cdot\sum_{i=1}^{n}\mathbf{f}_i^{\hat{B}_1}[\mathbf{g}^\dagger(\hat{\boldsymbol{v}}_k)]\hat{\boldsymbol{u}}_i = \sum_{i=1}^{n}\mathbf{f}_i^{\hat{B}_2}[\mathbf{g}(\hat{\boldsymbol{u}}_k)]\hat{\boldsymbol{v}}_i\cdot\hat{\boldsymbol{v}}_k$$

$$\overline{\mathbf{f}_k^{\hat{B}_1}[\mathbf{g}^\dagger(\hat{\boldsymbol{v}}_k)]} = \mathbf{f}_k^{\hat{B}_2}[\mathbf{g}(\hat{\boldsymbol{u}}_k)].$$

From this result, it can be said that if $\mathbf{g}$ exists, so does $\mathbf{g}^\dagger$. Now, supposing the existence of two Hilbert-adjoint functions $\mathbf{g}_1^\dagger$ and $\mathbf{g}_2^\dagger$ of $\mathbf{g}$, there are two equalities similar to (2.36). Subtracting one from the other, we obtain $\boldsymbol{u}\cdot(\mathbf{g}_1^\dagger(\boldsymbol{v}) - \mathbf{g}_2^\dagger(\boldsymbol{v})) = 0$, which is valid for all $\boldsymbol{u}$ and $\boldsymbol{v}$; thereby, $\mathbf{g}_1^\dagger = \mathbf{g}_2^\dagger$. Considering the properties now, the first can be verified from $(\alpha\mathbf{g})^\dagger(\boldsymbol{u})\cdot\boldsymbol{v} = \boldsymbol{u}\cdot\alpha\mathbf{g}(\boldsymbol{v})$ and from $\overline{\alpha}\mathbf{g}^\dagger(\boldsymbol{u})\cdot\boldsymbol{v} = \boldsymbol{u}\cdot\alpha\mathbf{g}(\boldsymbol{v})$, which results from the antilinearity of the inner product. The second can be proved from $\mathbf{g}\circ\mathbf{k}(\boldsymbol{u})\cdot\boldsymbol{v} = \mathbf{k}(\boldsymbol{u})\cdot\mathbf{g}^\dagger(\boldsymbol{v}) = \boldsymbol{u}\cdot\mathbf{k}^\dagger\circ\mathbf{g}^\dagger(\boldsymbol{u})$. In order to prove the equality on the third property, we need to know that an identity operator always equals its transpose; which is not difficult to verify. Thereby, "adjoiting" both sides of equality $\mathbf{i}_V = \mathbf{g} \circ \mathbf{g}^{-1}$, the result is that identity function $\mathbf{i}_V = (\mathbf{g}^{-1})^\dagger \circ \mathbf{g}^\dagger$, from the second property. We also know that $(\mathbf{g}^\dagger)^{-1} \circ \mathbf{g}^\dagger = \mathbf{i}_V$, which proves $(\mathbf{g}^{-1})^\dagger = (\mathbf{g}^\dagger)^{-1}$. □

A vector $\mathbf{h}$ of function space $\mathcal{CL}_{\mathbb{F}}(V,V)$ is said to be a **Hermitian** or **self-adjoint** operator when $\mathbf{h} = \mathbf{h}^\dagger$; but, if $\mathbf{h} = -\mathbf{h}^\dagger$, it is called **anti-Hermitian**. Moreover, in the context of real fields, (anti-)Hermitian operators are called **(anti)symmetric**, that is, $\mathbf{h} = \mathbf{h}^T$ or $\mathbf{h} = -\mathbf{h}^T$. An element of the function space $\mathcal{CL}_{\mathbb{F}}(V,U)$ constituted by invertible functions is said to be **unitary** when it is a bijection whose inverse equals the Hilbert-adjoint. Thereby, we can say that an unitary operator $\mathbf{q} \in \mathcal{CL}_{\mathbb{F}}(V,V)$ preserves inner products because, for all $\mathbf{u},\mathbf{v} \in V_{\mathbb{F}}$, equalities

$$\mathbf{q}(\mathbf{u}) \cdot \mathbf{q}(\mathbf{v}) = \mathbf{u} \cdot \mathbf{q}^\dagger \circ \mathbf{q}(\mathbf{v}) = \mathbf{u} \cdot \mathbf{q}^{-1} \circ \mathbf{q}(\mathbf{v}) = \mathbf{u} \cdot \mathbf{v} \tag{2.37}$$

are valid. Conversely, if an operator $\mathbf{g} \in \mathcal{CL}_{\mathbb{F}}(V,V)$ has its Hilbert-adjoint correspondent and shows the property $\mathbf{g}(\mathbf{u}) \cdot \mathbf{g}(\mathbf{v}) = \mathbf{u} \cdot \mathbf{v}$, then from definition (2.36), it is possible to affirm that every operator which preserves inner product is unitary. In this context, we can state that unitary operators also preserve norms, since

$$\|\mathbf{q}(\mathbf{u})\|^2 = \mathbf{q}(\mathbf{u}) \cdot \mathbf{q}(\mathbf{u}) = \mathbf{u} \cdot \mathbf{u} = \|\mathbf{u}\|^2.$$

Another consequence of $\mathbf{q}$ preserving inner products is the following: if $\{\hat{\mathbf{u}}_1, \cdots, \hat{\mathbf{u}}_n\}$ is an orthonormal basis of $V_{\mathbb{F}}$, then $\{\mathbf{q}(\hat{\mathbf{u}}_1), \cdots, \mathbf{q}(\hat{\mathbf{u}}_n)\}$ is also an orthonormal basis of $V_{\mathbb{F}}$. Still about this relation between bases and unitary operators, given $B = \{\hat{\mathbf{u}}_1, \cdots, \hat{\mathbf{u}}_n\}$ and $C = \{\hat{\mathbf{v}}_1, \cdots, \hat{\mathbf{v}}_n\}$ arbitrary orthonormal bases of $V_{\mathbb{F}}$, an operator $\mathbf{g} \in \mathcal{CL}_{\mathbb{F}}(V,V)$ defined by

$$\mathbf{g}(\mathbf{x}) = \sum_{i=1}^{m} \mathbf{f}_i^B(\mathbf{x}) \hat{\mathbf{v}}_i \tag{2.38}$$

results unitary and clearly specifies $\mathbf{g}(\hat{\mathbf{u}}_i) = \hat{\mathbf{v}}_i$. Thereby, *every pair of orthonormal bases can be related through a unitary operator*. In the context of real fields, unitary functions are usually called **orthogonal**, that is, $\mathbf{h}^T = \mathbf{h}^{-1}$.

Proof. Let's prove that $C = \{\mathbf{q}(\hat{\mathbf{u}}_1), \cdots, \mathbf{q}(\hat{\mathbf{u}}_n)\}$ is an orthonormal basis. Since operator $\mathbf{q}$ preserves inner product, C is clearly orthonormal because $\mathbf{q}(\hat{\mathbf{u}}_i) \cdot \mathbf{q}(\hat{\mathbf{u}}_j) = \hat{\mathbf{u}}_i \cdot \hat{\mathbf{u}}_j = \delta_{ij}$. For an arbitrary vector $\mathbf{u} \in V_{\mathbb{F}}$ there is always a $\mathbf{v} \in V_{\mathbb{F}}$ such that $\mathbf{u} = \mathbf{q}(\mathbf{v})$ because $\mathbf{q}$ is a bijection. Thereby, equality $\mathbf{u} = \sum_{i=1}^n \mathbf{f}_i^C(\mathbf{v}) \mathbf{q}(\hat{\mathbf{u}}_i)$ prove that $V_{\mathbb{F}} = \operatorname{span} C$. Now, we'll verify if $\mathbf{g}$ is indeed unitary. Expression $\mathbf{g} \circ \mathbf{g}^{-1} = \mathbf{i}$ is verified for rule $\mathbf{g}^{-1}(\mathbf{x}) = \sum_{i=1}^m \mathbf{f}_i^C(\mathbf{x}) \hat{\mathbf{u}}_i$ through the following equalities:

$$\mathbf{g} \circ \mathbf{g}^{-1}(\mathbf{x}) = \sum_{j=1}^{n} \sum_{i=1}^{n} \mathbf{f}_i^C(\mathbf{x}) \underbrace{\mathbf{f}_j^B(\hat{\mathbf{u}}_i)}_{\delta_{ij}} \hat{\mathbf{v}}_j = \sum_{i=1}^{n} \mathbf{f}_i^C(\mathbf{x}) \hat{\mathbf{v}}_i = \mathbf{x}.$$

Now, we prove that $\mathbf{g}^\dagger = \mathbf{g}^{-1}$ because the following development results an identity.

$$\mathbf{g}^{-1}(\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{g}(\mathbf{v})$$

$$\sum_{i=1}^{m} \mathbf{f}_i^C(\mathbf{u}) \hat{\mathbf{u}}_i \cdot \mathbf{v} = \sum_{i=1}^{m} \overline{\mathbf{f}_i^B(\mathbf{v})} \mathbf{u} \cdot \hat{\mathbf{v}}_i$$

$$\sum_{i=1}^{m} \mathbf{f}_i^C(\mathbf{u})\overline{\mathbf{f}_i^B(\mathbf{v})} = \sum_{i=1}^{m} \overline{\mathbf{f}_i^B(\mathbf{v})}\mathbf{f}_i^C(\mathbf{u}).$$

□

In the previous chapter we said that the set of all unary invertible operators defines a group on the operation of composition. Thereby, let $\mathcal{G}_{\mathbb{F}}(V) := (\mathcal{CL}_{\mathbb{F}}(V,V), \circ)$ be a group constituted by all invertible continuous linear operators on $V_{\mathbb{F}}$. We shall verify now if the set $X \subset \mathcal{G}_{\mathbb{F}}(V)$ of all unitary operators on $V_{\mathbb{F}}$, since they are unary and invertible, constitutes the group $\mathcal{N}_{\mathbb{F}}(V) := (X, \circ)$ called **unitary group** on $V_{\mathbb{F}}$. Equalities

$$(\mathbf{q}_1 \circ \mathbf{q}_2)^{-1} = \mathbf{q}_2^{-1} \circ \mathbf{q}_1^{-1} = \mathbf{q}_2^{\dagger} \circ \mathbf{q}_1^{\dagger} = (\mathbf{q}_1 \circ \mathbf{q}_2)^{\dagger} \tag{2.39}$$

show, for all $\mathbf{q}_1, \mathbf{q}_2 \in X$, that bijection[4] $\mathbf{q}_1 \circ \mathbf{q}_2$ indeed belongs to the set X. In the context of real fields, we have the **orthogonal group** $\mathcal{O}(V) := \mathcal{N}_{\mathbb{R}}(V)$. Moreover, there are linear operators which preserve metrics or distance, being called **isometric operators**. In more rigorous mathematical terms, $\mathbf{k} \in \mathcal{L}_{\mathbb{F}}(V,V)$ is an isometric operator if it is an injection where $\rho\,[\mathbf{k}(\mathbf{u}), \mathbf{k}(\mathbf{v})] = \rho(\mathbf{u}, \mathbf{v})$, for all $\mathbf{u}, \mathbf{v} \in V_{\mathbb{F}}$. Similarly to unitary operators, the set of all isometric operators on $V_{\mathbb{F}}$ defines a group $\mathcal{I}_{\mathbb{F}}(V) \subset \mathcal{L}_{\mathbb{F}}(V,V)$ called **isometry group**, since the composition of isometric operators is also an isometric operator, that is, considering any $\mathbf{k}, \mathbf{g} \in \mathcal{I}_{\mathbb{F}}(V)$,

$$\rho\,[\mathbf{g} \circ \mathbf{k}(\mathbf{u}), \mathbf{g} \circ \mathbf{k}(\mathbf{v})] = \rho\,[\mathbf{k}(\mathbf{u}), \mathbf{k}(\mathbf{v})] = \rho(\mathbf{u}, \mathbf{v}) \tag{2.40}$$

for all $\mathbf{u}, \mathbf{v} \in V_{\mathbb{F}}$. *In the context of Euclidean spaces, an orthogonal group is also an isometry group because any operator that preserves inner product is isometric. Moreover, the operators of an isometry group always preserve inner product.*

Proof. Let's verify these so categorical last statements. If $V_{\mathbb{R}}$ is an Euclidean space and if an operator $\mathbf{k} \in \mathcal{L}_{\mathbb{R}}(V,V)$ preserves inner product, then

$$\begin{aligned}
\rho\,[\mathbf{k}(\mathbf{u}), \mathbf{k}(\mathbf{v})]^2 &= \|\mathbf{k}(\mathbf{u}) - \mathbf{k}(\mathbf{v})\|^2 \\
&= \mathbf{k}(\mathbf{u}) \cdot \mathbf{k}(\mathbf{u}) - 2\mathbf{k}(\mathbf{u}) \cdot \mathbf{k}(\mathbf{v}) + \mathbf{k}(\mathbf{v}) \cdot \mathbf{k}(\mathbf{v}) \\
&= \mathbf{u} \cdot \mathbf{u} - 2\mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{v} \\
&= \|\mathbf{u} - \mathbf{v}\|^2 \\
&= \rho(\mathbf{u}, \mathbf{v})^2,
\end{aligned}$$

from which we prove that $\mathbf{k}$ is an isometry. Now, let $\mathbf{g} \in \mathcal{L}_{\mathbb{R}}(V,V)$ be an isometry. Raising both sides of equality $\rho[\mathbf{g}(\mathbf{u}), \mathbf{g}(\mathbf{v})] = \rho(\mathbf{u}, \mathbf{v})$ to the square, we have

$$\begin{aligned}
\mathbf{g}(\mathbf{u}) \cdot \mathbf{g}(\mathbf{u}) - 2\mathbf{g}(\mathbf{u}) \cdot \mathbf{g}(\mathbf{v}) + \mathbf{g}(\mathbf{v}) \cdot \mathbf{g}(\mathbf{v}) &= \mathbf{u} \cdot \mathbf{u} - 2\mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{v} \\
\rho[\mathbf{g}(\mathbf{u}), 0]^2 - 2\mathbf{g}(\mathbf{u}) \cdot \mathbf{g}(\mathbf{v}) + \rho[\mathbf{g}(\mathbf{v}), 0]^2 &= \rho(\mathbf{u}, 0)^2 - 2\mathbf{u} \cdot \mathbf{v} + \rho(\mathbf{v}, 0)^2 \\
\rho[\mathbf{g}(\mathbf{u}), \mathbf{g}(0)]^2 - 2\mathbf{g}(\mathbf{u}) \cdot \mathbf{g}(\mathbf{v}) + \rho[\mathbf{g}(\mathbf{v}), \mathbf{g}(0)]^2 &= \rho(\mathbf{u}, 0)^2 - 2\mathbf{u} \cdot \mathbf{v} + \rho(\mathbf{v}, 0)^2 \\
\rho(\mathbf{u}, 0)^2 - 2\mathbf{g}(\mathbf{u}) \cdot \mathbf{g}(\mathbf{v}) + \rho(\mathbf{v}, 0)^2 &= \rho(\mathbf{u}, 0)^2 - 2\mathbf{u} \cdot \mathbf{v} + \rho(\mathbf{v}, 0)^2 \\
\mathbf{g}(\mathbf{u}) \cdot \mathbf{g}(\mathbf{v}) &= \mathbf{u} \cdot \mathbf{v}\,.
\end{aligned}$$

□

[4] See properties on page 14.

2.4 Matrix Representations

We already know that a vector can be identified through its coordinates, represented by a tuple, on a specific basis, in such a way that distinct tuples never imply equal vectors; this vector-coordinates relationship is thus biunivocal. Thereby, mathematical expressions with vectors may include their coordinate scalars, gathered conveniently in matrices, when all the functional-arithmetic apparatus presented in the previous chapter, applicable to this type of collection, becomes available. Given a certain basis, the resulting relationship between vectors and matrices is also biunivocal, as in the case of tuples representing coordinates. In practical terms, the matrix representation of a vector occurs the following way: let $\mathbf{u}$ be an arbitrary element of vector space $U_{\mathbb{F}}$, of which subset $B=\{\mathbf{u}_1,\cdots,\mathbf{u}_n\}$ is a basis. Then, we define a $n\times 1$ matrix $[\mathbf{u}]^B$ as being the **representative matrix** of $\mathbf{u}$ on B, whose elements $[\mathbf{u}]^B_{i1} := \mathbf{f}^B_i(\mathbf{u})$. This definition makes it obvious that the representative matrix of the zero vector is always the zero matrix, on any basis. Moreover, the linearity of coordinate functionals enables the following development:

$$\alpha\sum_{i=1}^{n}\mathbf{f}^B_i(\mathbf{x})\mathbf{u}_i+\beta\sum_{i=1}^{n}\mathbf{f}^B_i(\mathbf{y})\mathbf{u}_i=\sum_{i=1}^{n}\left[\alpha\mathbf{f}^B_i(\mathbf{x})+\beta\mathbf{f}^B_i(\mathbf{y})\right]\mathbf{u}_i=\sum_{i=1}^{n}\mathbf{f}^B_i(\alpha\mathbf{x}+\beta\mathbf{y})\mathbf{u}_i,$$

from where it is possible to conclude that

$$\alpha[\mathbf{x}]^B+\beta[\mathbf{y}]^B=[\alpha\mathbf{x}+\beta\mathbf{y}]^B\,, \tag{2.41}$$

for all vectors $\mathbf{x},\mathbf{y}\in U_{\mathbb{F}}$ and all scalars $\alpha,\beta\in\mathbb{F}$. Now, concerning the vectors of basis B, if we represent them relative to B itself, we have $[\mathbf{u}_j]^B_{i1}=\mathbf{f}^B_i(\mathbf{u}_j)=\delta_{ij}$. This representation strategy is frequent for the natural basis $O=\{\hat{\mathbf{e}}_1,\cdots,\hat{\mathbf{e}}_n\}$ of Euclidean spaces, where $[\hat{\mathbf{e}}_j]^O_{i1}=\delta_{ij}$.

Considering $V_{\mathbb{F}}$ a m-dimensional complete vector space, let $\mathbf{g}$ be an element of the function space $\mathcal{L}_{\mathbb{F}}(U,V)$. If coordinates are considered for vector identification, by mapping elements of U to elements of V the function $\mathbf{g}$ ends up defining indirectly a relationship between two distinct bases, in such a way that the matrix representation of this relationship needs to evidence them both. Thereby, if $C=\{\mathbf{v}_1,\cdots,\mathbf{v}_m\}$ is a basis of $V_{\mathbb{F}}$ and $\mathbf{u}\in U_{\mathbb{F}}$, we have

$$\begin{aligned}
\mathbf{g}(\mathbf{u}) &= \sum_{i=1}^{m}\mathbf{f}^C_i(\mathbf{g}(\mathbf{u}))\mathbf{v}_i\\
\mathbf{g}(\mathbf{u}) &= \sum_{i=1}^{m}\mathbf{f}^C_i[\sum_{j=1}^{n}\mathbf{f}^B_j(\mathbf{u})\mathbf{g}(\mathbf{u}_j)]\mathbf{v}_i\\
\mathbf{g}(\mathbf{u}) &= \sum_{i=1}^{m}\sum_{j=1}^{n}\mathbf{f}^C_i[\mathbf{g}(\mathbf{u}_j)]\mathbf{f}^B_j(\mathbf{u})\mathbf{v}_i\,,
\end{aligned}$$

when we can state that

$$[\mathbf{g}(\mathbf{u})]^{C} = [\mathbf{g}_{B}]^{C}[\mathbf{u}]^{B}, \tag{2.42}$$

where $[\mathbf{g}_B]^C$ is a $m \times n$ matrix, with elements $[\mathbf{g}_B]^C_{ij} := \mathbf{f}^C_i\,[\mathbf{g}(\mathbf{u}_j)]$, that represents the linear function $\mathbf{g} \in \mathcal{L}_{\mathbb{F}}(U,V)$ on bases B and C. In other words, for bases B and C with dimensions n and m respectively, the $m \times 1$ representative matrix of vector $\mathbf{g}(\mathbf{u})$ on C results from the $m \times n$ representative matrix of function $\mathbf{g}$ on B and C multiplied by the $n \times 1$ representative matrix of vector $\mathbf{u}$ on B. Using a similar development that led to this previous result, given $\mathbf{h} \in \mathcal{L}_{\mathbb{F}}(U,V)$ and $\alpha, \beta \in \mathbb{F}$, we obtain the following equality:

$$[[\alpha\mathbf{g} + \beta\mathbf{h}](\mathbf{u})]^{C} = (\alpha[\mathbf{g}_{B}]^{C} + \beta[\mathbf{h}_{B}]^{C})[\mathbf{u}]^{B}. \tag{2.43}$$

It is possible that U equals V, when the functions involved result in linear operators. Moreover, bases B and C may also be equal, which enables us to write, for example, $[\mathbf{g}(\mathbf{u})]^B = [\mathbf{g}_B]^B[\mathbf{u}]^B$, from which $[\mathbf{g}_B]^B$ is said to be the representative matrix of linear operator $\mathbf{g}$ on B. Composite linear functions can also be represented in matrix form. Let's see how. Considering $\mathbf{l}$ a function in space $\mathcal{L}_{\mathbb{F}}(V,W)$, set Z a basis of q-dimensional complete vector space $W_{\mathbb{F}}$ and $\mathbf{v} = \mathbf{g}(\mathbf{u})$ a vector of $V_{\mathbb{F}}$, we can write that matrix $[\mathbf{l}(\mathbf{v})]^Z = [\mathbf{l}_C]^Z[\mathbf{v}]^C$. From this equality, we obtain

$$[\mathbf{l} \circ \mathbf{g}(\mathbf{u})]^{Z} = [\mathbf{l}_{C}]^{Z}[\mathbf{g}(\mathbf{u})]^{C} = [\mathbf{l}_{C}]^{Z}[\mathbf{g}_{B}]^{C}[\mathbf{u}]^{B}, \tag{2.44}$$

which can be used to write

$$[\mathbf{u}]^{B} = [\mathbf{g}^{-1} \circ \mathbf{g}(\mathbf{u})]^{B} = [\mathbf{g}^{-1}_{C}]^{B}[\mathbf{g}_{B}]^{C}[\mathbf{u}]^{B} \implies [\mathbf{g}^{-1}_{C}]^{B} = [\mathbf{g}_{B}]^{C^{-1}} \tag{2.45}$$

if linear function $\mathbf{g}$ is invertible and dimension $m = n$. Moreover, as we know that matrix $[\mathbf{l} \circ \mathbf{g}(\mathbf{u})]^Z = [(\mathbf{l} \circ \mathbf{g})_B]^Z[\mathbf{u}]^B$, from equalities (2.44), we can conclude that

$$[(\mathbf{l} \circ \mathbf{g})_{B}]^{Z} = [\mathbf{l}_{C}]^{Z}[\mathbf{g}_{B}]^{C}. \tag{2.46}$$

The inner product of vectors $\mathbf{x}, \mathbf{y} \in X_{\mathbb{F}}$, where $X_{\mathbb{F}}$ is a Hilbert space, can be represented through the following equality:

$$\mathbf{x} \cdot \mathbf{y} = [\mathbf{x}]^{B} : [\mathbf{y}]^{B^{\perp}}, \tag{2.47}$$

according to (2.35). The rule (2.32) enables us to obtain an important property involving Hilbert-adjoint functions and conjugate transpose matrices: considering that a function $\mathbf{g} \in \mathcal{CL}_{\mathbb{F}}(X,Y)$, where $Y_{\mathbb{F}}$ is also a Hilbert space, has an Hilbert-adjoint counterpart, from orthonormal bases $B = \{\hat{\mathbf{x}}_i, \cdots, \hat{\mathbf{x}}_n\}$ of $X_{\mathbb{F}}$ and $D = \{\hat{\mathbf{y}}_j, \cdots, \hat{\mathbf{y}}_m\}$ of $Y_{\mathbb{F}}$, we have

$$\mathbf{g}(\hat{\mathbf{x}}_i) \cdot \hat{\mathbf{y}}_j = \overline{\mathbf{g}^{\dagger}(\hat{\mathbf{y}}_j) \cdot \hat{\mathbf{x}}_i}$$
$$\mathbf{f}^{D}_{j}\,[\mathbf{g}(\hat{\mathbf{x}}_i)] = \overline{\mathbf{f}^{B}_{i}[\mathbf{g}^{\dagger}(\hat{\mathbf{y}}_j)]}$$

$$\overline{\mathbf{f}_i^{D}[\mathbf{g}(\hat{x}_j)]^{T}} = \mathbf{f}_i^{B}[\mathbf{g}^{\dagger}(\hat{y}_j)]$$
$$[\mathbf{g}_B]^{D\,\dagger} = [\mathbf{g}_D^{\dagger}]^{B}, \tag{2.48}$$

where left and right matrices have dimension $n \times m$. When we consider an operator $\mathbf{h} \in \mathcal{CL}_{\mathbb{F}}(X, X)$ and only one orthogonal basis B, from previous equality we can state that the matrix representation of operator $\mathbf{h}^{\dagger}$ is the conjugate transpose of the representation matrix of $\mathbf{h}$.

There are concepts that arise from all these matrix representations of vectors and linear functions. One of them has a fundamental importance for us and we call it **change of coordinates**. In this study, *to change coordinates of a vector or linear operator from a basis* B *to a basis* C *means to relate biunivocally the representative matrix of this vector or linear operator on* B *with its representative matrix on* C. In order to mathematically substantiate this idea, here is the following theorem.

Theorem 7 – Change of Vector Coordinates

If $U(B)_{\mathbb{F}}$ *and* $U(C)_{\mathbb{F}}$ *are vector spaces constituted by representative matrices of the elements of vector space* $U_{\mathbb{F}}$ *on its bases* B *and* C *respectively, there is one and only one linear bijective transformation* $\Gamma : U(B)_{\mathbb{F}} \mapsto U(C)_{\mathbb{F}}$*, called change of coordinates from* B *to* C*, where* $\Gamma([\mathbf{x}]^{B}) = [\mathbf{x}]^{C}$ *for all* $\mathbf{x} \in U_{\mathbb{F}}$.

Proof. If $B = \{\mathbf{x}_1, \mathbf{x}_2\}$ and $C = \{\alpha_1\mathbf{x}_1, \alpha_2\mathbf{x}_2\}$, the existence of Γ is assured by the rule

$$\Gamma(X) = \begin{bmatrix} 1/\alpha_1 & 0 \\ 0 & 1/\alpha_2 \end{bmatrix} X.$$

The uniqueness of Γ is the trivial result from supposing $\Gamma_1([\mathbf{x}]^{B}) = [\mathbf{x}]^{C}$ and $\Gamma_2([\mathbf{x}]^{B}) = [\mathbf{x}]^{C}$. Now, equality (2.42) enables us to write that $[\alpha\mathbf{u} + \beta\mathbf{v}]^{B} = \alpha[\mathbf{u}]^{B} + \beta[\mathbf{v}]^{B}$ for all $\mathbf{u}, \mathbf{v} \in U_{\mathbb{F}}$ and $\alpha, \beta \in \mathbb{F}$, from which we verify the linearity of Γ, that is,

$$\begin{aligned} \Gamma(\alpha[\mathbf{u}]^{B} + \beta[\mathbf{v}]^{B}) &= \Gamma([\alpha\mathbf{u} + \beta\mathbf{v}]^{B}) \\ &= [\alpha\mathbf{u} + \beta\mathbf{v}]^{C} \\ &= \alpha[\mathbf{u}]^{C} + \beta[\mathbf{v}]^{C} \\ &= \alpha\Gamma([\mathbf{u}]^{B}) + \beta\Gamma([\mathbf{v}]^{B}). \end{aligned}$$

As the representative matrices of $\mathbf{x} \in U_{\mathbb{F}}$ on B and C are unique, it is evident that Γ is an injection. Moreover, Γ results a bijection because there is no matrix in $U(C)_{\mathbb{F}}$ that does not have a correspondent in $U(B)_{\mathbb{F}}$, since every vector of $U_{\mathbb{F}}$ can be described in terms of B and C. □

Considering now $B = \{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ and $C = \{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ distinct bases of vector space $U_{\mathbb{F}}$, they enable distinct matrix representations for an arbitrary vector $\mathbf{u} \in U_{\mathbb{F}}$. Thereby, from equalities

$$\mathbf{f}_i^{C}(\mathbf{u}) = \mathbf{f}_i^{C}[\sum_{j=1}^{n} \mathbf{f}_j^{B}(\mathbf{u})\mathbf{u}_j] = \sum_{j=1}^{n} \mathbf{f}_i^{C}(\mathbf{u}_j)\mathbf{f}_j^{B}(\mathbf{u})$$

together with expression (2.42), we can write that

$$[\mathbf{u}]^{C} = [\mathbf{i}_{B}]^{C}[\mathbf{u}]^{B}, \tag{2.49}$$

where $[\mathbf{i}_B]^C_{ij} = \mathbf{f}^C_i(\mathbf{i}(\mathbf{u}_j))$ and $\mathbf{i}$ is the identity function of $U_{\mathbb{F}}$. Therefore, according to the previous theorem, we can write the rule

$$\Gamma(X) = [\mathbf{i}_B]^C X \tag{2.50}$$

for the change of coordinates $\Gamma : U(B)_{\mathbb{F}} \mapsto U(C)_{\mathbb{F}}$ and the rule $\Gamma^{-1}(X) = [\mathbf{i}_C]^B X$ for the inverse mapping. In the particular case of Hilbert spaces, there is a rule for coordinate functionals, described by equality (2.32), which enables us to specify matrix elements $[\mathbf{i}_B]^C_{ij} = \mathbf{u}_j \cdot \boldsymbol{v}^i$.

According to the conditions of equality (2.49), as $[\mathbf{i}_B]^C$ is a square matrix of size n and the identity function equals its inverse and its conjugate transpose, we can affirm from (2.45) that

$$[\mathbf{i}_B]^C = [\mathbf{i}_C]^{B^{-1}} \tag{2.51}$$

and from (2.48), whose context is Hilbert spaces,

$$[\mathbf{i}_B]^C = [\mathbf{i}_C]^{B^{-1}} = [\mathbf{i}_{C^\perp}]^{B^{\perp\dagger}}. \tag{2.52}$$

In other words, for this last equality, valid for the specific case of Hilbert spaces, the matrix that enables a change of coordinates from C to B has as its inverse the matrix that enables a change of coordinates from B to C, whose conjugate transpose enables the change of coordinates from $C^\perp$ to $B^\perp$.

Still considering the vector space $U_{\mathbb{F}}$, the elements of base C can be written on basis B according to the following expressions:

$$\boldsymbol{v}_j = \sum_{i=1}^{n} \mathbf{u}_i \mathbf{f}^B_i(\boldsymbol{v}_j) = \sum_{i=1}^{n} \mathbf{u}_i [\mathbf{i}_C]^B_{ij}, \tag{2.53}$$

from which, given an arbitrary basis Z of $U_{\mathbb{F}}$ used to describe numerically the elements of bases B and C, we can say that

$$\mathbf{f}^Z_k(\boldsymbol{v}_j) = \sum_{i=1}^{n} \mathbf{f}^Z_k(\mathbf{u}_i)[\mathbf{i}_C]^B_{ij}.$$

Arranging the components on Z all the elements of bases B and C according to this equality, the following matrix expression results:

$$\begin{bmatrix} [\boldsymbol{v}_1]^Z_{11} & [\boldsymbol{v}_2]^Z_{11} & \cdots & [\boldsymbol{v}_n]^Z_{11} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ [\boldsymbol{v}_1]^Z_{n1} & [\boldsymbol{v}_2]^Z_{n1} & \cdots & [\boldsymbol{v}_n]^Z_{n1} \end{bmatrix} = \begin{bmatrix} [\mathbf{u}_1]^Z_{11} & [\mathbf{u}_2]^Z_{11} & \cdots & [\mathbf{u}_n]^Z_{11} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ [\mathbf{u}_1]^Z_{n1} & [\mathbf{u}_2]^Z_{n1} & \cdots & [\mathbf{u}_n]^Z_{n1} \end{bmatrix} [\mathbf{i}_C]^B, \tag{2.54}$$

from which we say that $[\mathbf{i}_C]^B$ changes basis B to basis C, that is, this matrix performs a **change of basis.** Concerning the changes involving generic bases B and C, when a matrix that changes coordinates from B to C is the inverse of the matrix that changes basis B to C, the coordinates of elements of $U_{\mathbb{F}}$ are called **contravariant**, since these coordinates are submitted to a "contrary" transformation relative to the change of basis B to C, as is the case of the values of Γ in (2.50). Now, let's see what happens when we change coordinates of dual vectors. If $\mathbf{h} \in U_{\mathbb{F}}^*$ is a dual vector, we already know that $[\mathbf{h}]^{B^*}$ is its representative matrix on dual basis B^*. Thereby, equalities (2.31) and (2.53) enable the following development:

$$[\mathbf{h}]_j^{C^*} = \mathbf{h}(\boldsymbol{v}_j) = \sum_{i=1}^{n} \mathbf{h}(\mathbf{u}_i)[\mathbf{i}_C]_{ij}^B = \sum_{i=1}^{n} [\mathbf{h}]_i^{B^*} [\mathbf{i}_C]_{ij}^B , \tag{2.55}$$

from which we conclude that the change of coordinates of $\mathbf{h}$ from B^* to C^* is performed by the same matrix $[\mathbf{i}_C]^B$ responsible for the change of basis, when the coordinates of dual vectors are then called **covariant**. Thereby, we can define a change of coordinates $\Gamma^* : U^*(B^*)_{\mathbb{F}} \mapsto U^*(C^*)_{\mathbb{F}}$ with a rule

$$\Gamma^*(X) = X [\mathbf{i}_C]^B . \tag{2.56}$$

From these two classifications of coordinates, considering $U_{\mathbb{F}}$ a Hilbert space, the scalars $\mathbf{f}_i^B(\mathbf{u}) = \mathbf{u} \cdot \mathbf{u}^i$ constitute the contravariant coordinates of vector $\mathbf{u} \in U_{\mathbb{F}}$, whose covector

$$\mathbf{u}^* = \sum_{i=1}^{n} \mathbf{u}^*(\mathbf{u}_i)\mathbf{f}_i^B = \sum_{i=1}^{n} (\mathbf{u}_i \cdot \mathbf{u})(\mathbf{u}^i)^* .$$

Since $\mathbf{u}$ and $\mathbf{u}^*$ are biunivocally related through Riesz-Fréchet Representation, the coordinates $(\mathbf{u}_1 \cdot \mathbf{u}, \cdots, \mathbf{u}_n \cdot \mathbf{u})$ of $\mathbf{u}^*$ on B^* are usually called the covariant coordinates of vector $\mathbf{u}$ on $B^\perp$. If the Hilbert space in question is real, considering $\hat{B} = \{\hat{\mathbf{u}}_1, \cdots, \hat{\mathbf{u}}_n\}$ an orthonormal basis of $U_{\mathbb{R}}$, we can say that covariant and contravariant coordinates of an arbitrary vector $\mathbf{u}$ are always equal because expressions $\hat{\mathbf{u}}_i = \hat{\mathbf{u}}^i$ and the commutativity of inner product imply $\mathbf{u} \cdot \mathbf{u}^i = \mathbf{u}_i \cdot \mathbf{u}$, that is, $\mathbf{u}^*(\mathbf{u}_i) = \mathbf{f}_i^{\hat{B}}(\mathbf{u})$. This also occurs in a tridimensional Euclidean space, when natural basis $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$ is called **cartesian basis** and all the linear combinations of its elements are **cartesian vectors.**

Now, an interesting consequence of equalities (2.52) is that matrix $[\mathbf{i}_C]^B$ results unitary if the bases involved are orthonormal, since they are reciprocal to themselves. Thereby, we present the following corollary, that describes an important relationship between representative matrices of linear operators.

Corollary 7.1 – Change of Linear Operator Coordinates

If $Y(B)_{\mathbb{F}}$ *and* $Y(C)_{\mathbb{F}}$ *are vector spaces constituted by the representative matrices of linear operators that belong to* $\mathcal{L}_{\mathbb{F}}(Y,Y)$*, described respectively on bases* B *and* C *of vector space* $Y_{\mathbb{F}}$*,* *the change of coordinates* $\Theta : \mathsf{Y}(\mathsf{B})_{\mathbb{F}} \mapsto \mathsf{Y}(\mathsf{C})_{\mathbb{F}}$ *is always a similarity transformation where* $\Theta(X) = [\mathbf{i}_{\mathrm{B}}]^{\mathrm{C}} X [\mathbf{i}_{\mathrm{C}}]^{\mathrm{B}}$.

Proof. Considering a function $\mathbf{g} \in \mathcal{L}_{\mathbb{F}}(\mathsf{Y},\mathsf{Y})$ and a vector $\mathbf{u} \in \mathsf{Y}_{\mathbb{F}}$, the last equality of the following development

$$\begin{aligned}
[\mathbf{g}(\mathbf{u})]^{\mathrm{B}} &= [\mathbf{g}_{\mathrm{B}}]^{\mathrm{B}} [\mathbf{u}]^{\mathrm{B}} \\
[\mathbf{i}_{\mathrm{C}}]^{\mathrm{B}} [\mathbf{g}(\mathbf{u})]^{\mathrm{C}} &= [\mathbf{g}_{\mathrm{B}}]^{\mathrm{B}} [\mathbf{i}_{\mathrm{C}}]^{\mathrm{B}} [\mathbf{u}]^{\mathrm{C}} \\
[\mathbf{g}(\mathbf{u})]^{\mathrm{C}} &= [\mathbf{i}_{\mathrm{B}}]^{\mathrm{C}} [\mathbf{g}_{\mathrm{B}}]^{\mathrm{B}} [\mathbf{i}_{\mathrm{C}}]^{\mathrm{B}} [\mathbf{u}]^{\mathrm{C}} \\
[\mathbf{g}_{\mathrm{C}}]^{\mathrm{C}} [\mathbf{u}]^{\mathrm{C}} &= [\mathbf{i}_{\mathrm{B}}]^{\mathrm{C}} [\mathbf{g}_{\mathrm{B}}]^{\mathrm{B}} [\mathbf{i}_{\mathrm{C}}]^{\mathrm{B}} [\mathbf{u}]^{\mathrm{C}}
\end{aligned}$$

enables us to affirm that

$$[\mathbf{g}_{\mathrm{C}}]^{\mathrm{C}} = [\mathbf{i}_{\mathrm{B}}]^{\mathrm{C}} [\mathbf{g}_{\mathrm{B}}]^{\mathrm{B}} [\mathbf{i}_{\mathrm{C}}]^{\mathrm{B}}.$$

As $[\mathbf{i}_{\mathrm{B}}]^{\mathrm{C}} = [\mathbf{i}_{\mathrm{C}}]^{\mathrm{B}\,-1}$, we conclude that matrices $[\mathbf{g}_{\mathrm{C}}]^{\mathrm{C}}$ and $[\mathbf{g}_{\mathrm{B}}]^{\mathrm{B}}$ are similar. □

Until the end of this chapter, we shall be working with linear functions that admit change of coordinates, that is, linear operators. In this context, some typical concepts of matrices can be applied to such operators. Let's see which and why. Given the conditions of the previous corollary, matrices $\Theta(X)$ and X are similar, from which it can be concluded that *the trace and determinant of any representative matrix of any linear operator are immune or indifferent to any change of basis.* Because of this indifference, we consider scalars

$$\det(\mathbf{g}) := \det([\mathbf{g}]) \tag{2.57}$$

and

$$\operatorname{tr}(\mathbf{g}) := \operatorname{tr}([\mathbf{g}]) \tag{2.58}$$

as being respectively the **determinant** and the **trace** of operator $\mathbf{g} \in \mathcal{L}_{\mathbb{F}}(\mathsf{Y},\mathsf{Y})$, where $[\mathbf{g}]$ is the representative matrix of this operator on any basis of $\mathsf{Y}_{\mathbb{F}}$. Considering this definition and the linear groups presented in the previous section, it is now possible to define a **proper** $\mathcal{U}^{+}{}_{\mathbb{F}}(\mathsf{Y}) \subset \mathcal{N}_{\mathbb{F}}(\mathsf{Y})$ and **improper** $\mathcal{U}^{-}{}_{\mathbb{F}}(\mathsf{Y}) \subset \mathcal{N}_{\mathbb{F}}(\mathsf{Y})$ **unitary groups**[5] constituted by invertible continuous linear operators $\mathbf{h}$ whose determinants are respectively positive, $\det\mathbf{h} = 1$, and negative negative, $\det\mathbf{h} = -1$, according to (1.38). Another concept that linear operators inherit from matrices is positivity. Considering B a basis of $\mathsf{Y}_{\mathbb{F}}$, if the following condition is observed

$$\Re(\,[\mathbf{g}_{\mathrm{B}}]^{\mathrm{B}} [\mathbf{y}]^{\mathrm{B}} : [\mathbf{y}]^{\mathrm{B}}\,) \geqslant 0, \tag{2.59}$$

for all non zero $\mathbf{y} \in \mathsf{Y}_{\mathbb{F}}$, matrix $[\mathbf{g}_{\mathrm{B}}]^{\mathrm{B}}$ is said to be nonnegative, according to the definition presented in the previous chapter. This inequality is still valid if $\mathsf{Y}_{\mathbb{F}}$ is a Hilbert space and B is orthonormal, when it is possible to develop, from (2.35),

$$\Re(\,[\mathbf{g}_{\mathrm{B}}]^{\mathrm{B}} [\mathbf{y}]^{\mathrm{B}} : [\mathbf{y}]^{\mathrm{B}}\,) = \Re(\,[\mathbf{g}(\mathbf{y})]^{\mathrm{B}} : [\mathbf{y}]^{\mathrm{B}}\,) = \Re(\mathbf{g}(\mathbf{y}) \cdot \mathbf{y}) \geqslant 0, \tag{2.60}$$

[5] These groups are called **proper and improper orthogonal groups** in the context of real fields.

where $\mathbf{g}$ is called a **nonnegative linear operator**[6] or a **positive-definite linear operator** if the real value $\Re(\mathbf{g}(\mathbf{y})\cdot\mathbf{y})$ is always positive.

Example 2.1. The subject "change of coordinates" is very important for our study and it is convenient to finish this section dealing with something less abstract. What follows is a naive little story. Once upon a time, Brenda, a teacher, living on a bank of a large width river, hands over to a boatman a package containing a gift to be delivered at a place on the other bank, where the mechanical engineer Calvin, beloved consignee of her order, lives. At a certain point while crossing the river, the boatman sees himself forced to modify his velocity, which will change the exact time and place where Calvin anxiously waits for his gift. Fully aware of his task's relevance, the boatman, after reading his instruments, sends a text message to Calvin's phone with the following information: *It is now 2:00 p.m. and after crossing, relative to Brenda, 30km inside the river, that is, perpendicular to her bank, and 5km upstream, against the river flow, I was at 49km/h inside the river and at 13.1km/h upstream when I saw a dangerously shallow part of the riverbed, an then was forced to reduce 20% my north velocity and 40% my east velocity. As I won't be able to rectify the route, I ask you to meet me at the new time and on the new point of the border where I'll deliver your package. And please, don't forget that Brenda's place, in relation to yours, is located 30km downstream and 55km inside the river.* Although extremely frustrated after reading this complicated message, Calvin took a deep breadth, kept himself calm and remembered his enjoyable and outstanding classes of Linear Algebra. He came back home, took pencil, paper and a calculator, sat at his desk and started reasoning to discover the new time and place of arrival: "First of all, I'll use a bidimensional Euclidean space with a natural basis $O=\{\hat{\mathbf{e}}_1,\hat{\mathbf{e}}_2\}$, where $\hat{\mathbf{e}}_1$ represents the east and $\hat{\mathbf{e}}_2$ the north. So, in relation to this basis, I already know that matrices $[\mathbf{c}_1]^O=[-1\ \ 0]^T$ and $[\mathbf{c}_2]^O=[-0,42\ \ 0,91]^T$ are representative bases of the elements of my point of view $C=\{\mathbf{c}_1,\mathbf{c}_2\}$, where $\mathbf{c}_1$ represents the 'inside the river' direction while $\mathbf{c}_2$ is related to the upstream direction. Similarly, in the case of Brenda, matrices $[\mathbf{b}_1]^O=[0,87\ \ 0,5]^T$ and $[\mathbf{b}_2]^O=[0\ \ 1]^T$ represent the elements of her point of view $B=\{\mathbf{b}_1,\mathbf{b}_2\}$. From what the boatman said, $[\mathbf{v}]^B=[49\ \ 13.1]^T$ represented his velocity $\mathbf{v}$ when he needed to make a change $\mathbf{f}$, described by

$$[\mathbf{f}_O]^O=\begin{bmatrix}0.6 & 0\\ 0 & 0.8\end{bmatrix}.$$

In order to obtain what I need, I'll rewrite $\mathbf{v}$ and $\mathbf{f}$ representations on my point of view. I still remember equalities $[\mathbf{v}]^C=[\mathbf{i}_B]^C[\mathbf{v}]^B$ and $[\mathbf{f}_C]^C=[\mathbf{i}_O]^C[\mathbf{f}_O]^O[\mathbf{i}_C]^O$, where

$$[\mathbf{i}_B]^C_{ij}=\mathbf{b}_j\cdot\mathbf{c}^i \qquad \text{and} \qquad [\mathbf{i}_O]^C_{ij}=\hat{\mathbf{e}}_j\cdot\mathbf{c}^i\,.$$

Now I need to discover my reciprocal basis $C^\perp=\{\mathbf{c}^1,\mathbf{c}^2\}$, where $\mathbf{c}_i\cdot\mathbf{c}^j=\delta_{ij}$. Using this last equality to solve for $[\mathbf{c}^1]^O=[x_1\ \ y_1]^T$ and $[\mathbf{c}^2]^O=[x_2\ \ y_2]^T$, I can write the systems

$$\begin{cases}-x_1=1\\ -0.42x_1+0.91y_1=0\end{cases} \qquad \text{and} \qquad \begin{cases}-x_2=0\\ -0.42x_2+0.91y_2=1\end{cases},$$

[6] Or positive-semidefinite.

whose solutions are $[\mathbf{c}^1]^O = [-1 \;\; -0.46]^T$ and $[\mathbf{c}^2]^O = [0 \;\; 1.1]^T$, when I arrive at matrices

$$[\mathbf{i}_B]^C = \begin{bmatrix} -1.1 & -0.46 \\ 0.55 & 1.1 \end{bmatrix} \qquad \text{and} \qquad [\mathbf{i}_O]^C = \begin{bmatrix} -1 & 0 \\ -0.42 & 0.91 \end{bmatrix}.$$

Oh, and finally I can have the representations on my perspective:

$$[\boldsymbol{v}]^C = \begin{bmatrix} -59.93 \\ 41.36 \end{bmatrix} \qquad \text{and} \qquad [\mathbf{f}_C]^C = \begin{bmatrix} 0.6 & 0.25 \\ 0.25 & 0.77 \end{bmatrix}.$$

After deviating from original route, boat velocity representative matrix $[\boldsymbol{v}]^C$ becomes $[\mathbf{f}(\boldsymbol{v})]^C = [-25.62 \;\; 16.87]^T$. At the moment of this deviation, dislocation $\mathbf{u}$ of the boat was $[\mathbf{u}]^B = [30 \;\; 5]^T$, or on my point of view, $[\mathbf{u}]^C = [\mathbf{i}_B]^C[\mathbf{u}]^B = [-35.3 \;\; 22]^T$. Since he informed Brenda's position relative to mine, the boat is now performing the remaining displacement $[\boldsymbol{z}]^C = [-24.7 \;\; 8]^T$ at a 'inside the river' speed of -25.62km/h, which will require 0.96h. During this period, at a speed of 16.87km/h upstream, he will travel 16.2km upstream, which means 8.2km upstream from where I am. He will arrive at this new point around 3:00 p.m. Since it is now 2:30 p.m., I'll arrive on time by car."

2.5 Eigenvalues and Eigenvectors

Considering a n-dimensional Hilbert space $\mathcal{U}_{\mathbb{F}}$, we call $\alpha\mathbf{u} \in \mathcal{U}_{\mathbb{F}}$, where $\alpha \in \mathbb{F}$, a **multiple scalar** of vector $\mathbf{u}$. The scalar α, which specifies this multiplicity, on the context of norm $\|\alpha\mathbf{u}\| = |\alpha|\|\mathbf{u}\|$, performs a "resizing" of $\mathbf{u}$, that is, if $|\alpha| < 1$, there is a decrease in its size or intensity; conversely, if $|\alpha| > 1$, there is an increase. Given a linear operator $\mathbf{l} \in \mathcal{L}_{\mathbb{F}}(\mathcal{U}, \mathcal{U})$, a non zero vector $\mathbf{u}$ is called an **eigenvector** of $\mathbf{l}$ if value $\mathbf{l}(\mathbf{u})$ is its multiple scalar, that is, if $\mathbf{l}(\mathbf{u}) = \alpha\mathbf{u}$, where multiplicity α is said to be an **eigenvalue** of $\mathbf{l}$. In other words, *a vector resized by a linear operator is its eigenvector and the sign-magnitude of this resizing is its eigenvalue.*

Now let's suppose that, for a known operator $\mathbf{l}$, we want to discover all of its eigenvalues and eigenvectors from the unknowns of the equation $\mathbf{l}(\mathbf{x}) = \lambda\mathbf{x}$ or rather, from λ and $\mathbf{x}$ in

$$(\mathbf{l} - \lambda\mathbf{i})(\mathbf{x}) = \mathbf{0}, \tag{2.61}$$

where $\mathbf{i}$ is the identity function in $\mathcal{U}_{\mathbb{F}}$. On an arbitrary basis of this space, matrix $[\mathbf{i}] = I$ and the matrix representation of previous equation results $([\mathbf{l}] - \lambda I)[\mathbf{x}] = 0$. As we know, matrix $[\mathbf{x}]$ cannot be zero because there are no zero eigenvectors; a restriction that prevents matrix $([\mathbf{l}] - \lambda I)$ from being invertible, that is, it has to be singular:

$$\det([\mathbf{l}] - \lambda I) = 0. \tag{2.62}$$

Recalling our matrix theory in the previous chapter, the left side of this equation is the characteristic polynomial[7] of $[\mathbf{l}]$ on variable λ, whose n characteristic roots solve (2.62), that is, there are n eigenvalues of $\mathbf{l}$, which are not necessarily distinct and whose collection is represented by $\lambda(\mathbf{l})$. Once we have these eigenvalues, it it possible to determine each correspondent eigenvector from equation $([\mathbf{l}]-\lambda I)[\mathbf{x}]=0$. As this processes is valid and leads to identical results for any chosen basis, expression $\det(\mathbf{l}-\lambda\mathbf{i})$ is sometimes called the characteristic polynomial of linear operator $\mathbf{l}$ also.

Eigenvalues and eigenvectors subsidize the forthcoming Polar Decomposition Theorem in such a way that they reveal three relevant properties of every Hermitian operator:

i. the composition of an operator with its Hilbert-adjoint is always a nonnegative Hermitian operator;
ii. the eigenvalues of a Hermitian operator are always real;
iii. two distinct eigenvectors of a Hermitian operator are always orthogonal.

This last property implies that the n eigenvectors of a Hermitian operator are distinct from each other and the resulting orthogonal set constituted by them is obviously a basis of $U_{\mathbb{F}}$, usually called the **eigenbasis** of this Hermitian operator.

Proof. For the first property, let $\mathbf{g}\circ\mathbf{g}^\dagger$ be an Hermitian operator where $\mathbf{g}\in\mathcal{L}_{\mathbb{F}}(U,U)$. We need to show that it is nonnegative the real number $\Re([\mathbf{y}]^\dagger[\mathbf{g}][\mathbf{g}]^\dagger[\mathbf{y}])_{11}$, whose matrices are described on any orthonormal basis. If a matrix $A:=[\mathbf{g}]^\dagger[\mathbf{y}]$, the previous real number becomes $\Re(A^\dagger A)_{11}$, which is always nonnegative because $(A^\dagger A)_{11}=\sum_{i=1}^n\overline{A_{i1}}A_{i1}=\sum_{i=1}^n|A_{i1}|^2$. Now, let $\mathbf{h}\in\mathcal{L}_{\mathbb{F}}(U,U)$ be a Hermitian operator with eigenvalues λ_i and eigenvectors $\mathbf{x}_i$, from which we can write $\mathbf{h}(\mathbf{x}_i)=\lambda_i\mathbf{x}_i$. Performing the inner product of an eigenvector $\mathbf{x}_j$ and each side of the previous equality, we obtain $\mathbf{x}_j\cdot\mathbf{h}(\mathbf{x}_i)=\mathbf{x}_j\cdot\lambda_i\mathbf{x}_i$ (*). Similarly, the inner product of $\mathbf{h}(\mathbf{x}_j)=\lambda_i\mathbf{x}_j$ and $\mathbf{x}_i$ results $\mathbf{h}(\mathbf{x}_j)\cdot\mathbf{x}_i=\lambda_j\mathbf{x}_j\cdot\mathbf{x}_i$ (**). Since $\mathbf{h}$ is Hermitian, if we subtract (**) from (*),

$$\begin{aligned}\mathbf{x}_j\cdot\mathbf{h}(\mathbf{x}_i)-\mathbf{h}(\mathbf{x}_j)\cdot\mathbf{x}_i &= \mathbf{x}_j\cdot\lambda_i\mathbf{x}_i-\lambda_j\mathbf{x}_j\cdot\mathbf{x}_i\\ 0 &= (\overline{\lambda_i}-\lambda_j)\mathbf{x}_j\cdot\mathbf{x}_i\,.\end{aligned}$$

Since the last equality is valid for an arbitrary pair (i,j), when $i=j$ the scalar $\overline{\lambda_i}=\lambda_i$, which proves property b). Moreover, as a pair of distinct eigenvalues imply a pair of distinct eigenvectors, we can say that $\lambda_i\neq\lambda_j$ implies $\mathbf{x}_i\neq\mathbf{x}_j$. Thus, from the last equality of the development, we conclude $\mathbf{x}_i\perp\mathbf{x}_j$, which proves c). □

In addition to properties above, an arbitrary Hermitian operator $\mathbf{h}\in\mathcal{L}_{\mathbb{F}}(U,U)$, where $U_{\mathbb{F}}$ is a Hilbert space, has a Hermitian representative matrix because, for any orthonormal basis $\hat{B}$ of $U_{\mathbb{F}}$, equality (2.48) enables us to write

$$[\mathbf{h}_{\hat{B}}]^{\hat{B}\,\dagger}=[\mathbf{h}^\dagger_{\hat{B}}]^{\hat{B}}=[\mathbf{h}_{\hat{B}}]^{\hat{B}}\,. \tag{2.63}$$

Since they are Hermitian, such representative matrices are normal, that is, they are susceptible of spectral diagonalization[8]. Thereby, considering $\mathbf{x}_i$ the mutually orthogonal

[7] See definition on p. 25.
[8] See definition of normal matrix at p. 23.

n eigenvectors of $\mathbf{h}$, if $\widetilde{H}$ is the resultant diagonal matrix of the spectral diagonalization of $[\mathbf{h}_{\hat{B}}]^{\hat{B}}$ and set $\hat{X} = \{\hat{x}_1, \cdots, \hat{x}_n\}$ is the normalized orthogonal basis of its eigenvectors, each element

$$\widetilde{H}_{ij} = \lambda_j \delta_{ij} = \lambda_j \hat{\mathbf{x}}_j \cdot \hat{\mathbf{x}}^i = \lambda_j \mathbf{f}_i^{\hat{X}}(\hat{\mathbf{x}}_j) = \mathbf{f}_i^{\hat{X}}(\lambda_j \hat{\mathbf{x}}_j) = \mathbf{f}_i^{\hat{X}}(\mathbf{h}(\hat{\mathbf{x}}_j)) = [\mathbf{h}_{\hat{X}}]_{ij}^{\hat{X}}. \tag{2.64}$$

From these equalities, we can state that, in the context of Hermitian operators, the result of the spectral diagonalization of a representative matrix $[\mathbf{h}_{\hat{B}}]^{\hat{B}}$ on any orthonormal basis is the representative matrix $[\mathbf{h}_{\hat{X}}]^{\hat{X}}$ on the normalized basis of eigenvectors. This resultant matrix is usually called the **spectral representation** of operator $\mathbf{h}$. In this context, we now want to change the coordinates of $\mathbf{h}$ from basis $\hat{X}$ of its eigenvectors to any orthonormal basis $\hat{C}$ of $U_{\mathbb{F}}$. From corollary 7.1, we can write the following:

$$[\mathbf{h}_{\hat{C}}]^{\hat{C}} = [\mathbf{i}_{\hat{X}}]^{\hat{C}} [\mathbf{h}_{\hat{X}}]^{\hat{X}} [\mathbf{i}_{\hat{C}}]^{\hat{X}} = [\mathbf{i}_{\hat{X}}]^{\hat{C}} [\mathbf{h}_{\hat{X}}]^{\hat{X}} [\mathbf{i}_{\hat{X}}]^{\hat{C}\dagger}, \tag{2.65}$$

since $[\mathbf{i}_{\hat{X}}]^{\hat{C}}$ results a unitary matrix from equalities (2.52). Moreover, since the representative matrix of $\mathbf{h}$ on any orthonormal basis is normal, this change of coordinates from $\hat{X}$ to $\hat{C}$ results a **spectral decomposition**[9] of $[\mathbf{h}_{\hat{C}}]^{\hat{C}}$. Let's consider now that $\mathbf{h}$ is a nonnegative operator. From the definition presented at the end of the previous section, we can write that $\Re(\mathbf{x}_i \cdot \mathbf{h}(\mathbf{x}_i)) \geqslant 0$. Since the eigenvalues of $\mathbf{h}$ are real, we have $\lambda_i(\mathbf{x}_i \cdot \mathbf{x}_i) \geqslant 0$, where scalars λ_i result nonnegative, since the inner product $\mathbf{x}_i \cdot \mathbf{x}_i$ is always positive. Therefore, *we can state that if operator* $\mathbf{h}$ *is nonnegative, so are its eigenvalues.* In this context, if $\mathbf{h}$ is a positive-definite operator, its eigenvalues are all positive and then, considering $\alpha_1, \cdots, \alpha_n$ the eigenvalues of $\mathbf{h}^{-1}$, equalities

$$0 = ([\mathbf{h}^{-1}] - \alpha_i I)[\mathbf{x}] = -\alpha_i [\mathbf{h}^{-1}]([\mathbf{h}^{-1}]^{-1} - \alpha_i^{-1} I)[\mathbf{x}] = ([\mathbf{h}] - \alpha_i^{-1} I)[\mathbf{x}]$$

are valid, from which we conclude that $\alpha_i = \lambda_i^{-1}$.

Before we deal with the theorem that will end this chapter, an additional definition for Hermitian operators needs to be done. In our study, a nonnegative hermitian operator $\mathbf{h}$ of $\mathcal{L}_{\mathbb{F}}(U, U)$ can be decomposed according to equality $\mathbf{h} = \mathbf{h}^{1/2} \circ \mathbf{h}^{1/2}$, where operator $\mathbf{h}^{1/2} \in \mathcal{L}_{\mathbb{F}}(U, U)$, unique and also nonnegative hermitian, is called the **square root operator** of $\mathbf{h}$. We already know that function composition is expressed, in matrix terms, as the product of the representative matrices of these functions. Thereby, the name "square root" is due to equality $[\mathbf{h}] = [\mathbf{h}^{1/2}][\mathbf{h}^{1/2}]$, which refers to the same concept applied to scalars.

> *Proof.* We need to show that a square root operator exists and is unique. The following proof is adapted from Gurtin[23], pp. 13-14. For equality $[\mathbf{h}] = [\mathbf{h}^{1/2}][\mathbf{h}^{1/2}]$, which is valid for an arbitrary basis, let's

[9] In the terms of theorem 1, equalities $\widetilde{N} = U^\dagger N U$ and $N = U \widetilde{N} U^\dagger$ are said to be the spectral diagonalization and the spectral decomposition of N respectively.

choose an orthonormal basis $\hat{X} = \{\hat{\mathbf{x}}_1, \cdots, \hat{\mathbf{x}}_n\}$, constituted by the eigenvectors of $\mathbf{h}$. Thereby,

$$[\mathbf{h}_{\hat{X}}]^{\hat{X}} = [\mathbf{h}^{1/2}{}_{\hat{X}}]^{\hat{X}} [\mathbf{h}^{1/2}{}_{\hat{X}}]^{\hat{X}},$$

and we already know that $[\mathbf{h}_{\hat{X}}]^{\hat{X}}_{ij} = \lambda_i \delta_{ij}$, where $\lambda_i \geqslant 0$. If we admit $[\mathbf{h}^{1/2}{}_{\hat{X}}]^{\hat{X}}_{ij} = \delta_{ij}\sqrt{\lambda_i}$, matrix $[\mathbf{h}^{1/2}{}_{\hat{X}}]^{\hat{X}}$ results nonnegative Hermitian, from which it can be concluded that $\mathbf{h}^{1/2}$ is also nonnegative Hermitian according to equalities (2.63). Thereby, the existence of a square root for $\mathbf{h}$ is verified. Now, in order to prove the uniqueness of $\mathbf{h}^{1/2}$, let's suppose that $\mathbf{c}^{1/2} \circ \mathbf{c}^{1/2} = \mathbf{h}$. Adopting a basis B, a vector $\mathbf{u} \in V$ and the equality (2.61), we can do the following development:

$$\begin{aligned} 0 &= \left(\left[\mathbf{h}_B^{1/2}\right]^B \left[\mathbf{h}_B^{1/2}\right]^B - \lambda_i I \right) [\mathbf{x}_i]^B \\ &= \left(\left[\mathbf{h}_B^{1/2}\right]^B + \sqrt{\lambda_i}\, I \right) \underbrace{\left(\left[\mathbf{h}_B^{1/2}\right]^B - \sqrt{\lambda_i}\, I \right) [\mathbf{x}_i]^B}_{[\mathbf{u}]^B}, \end{aligned}$$

from which we conclude that

$$-\sqrt{\lambda_i}[\mathbf{u}]^B = \left[\mathbf{h}_B^{1/2}\right]^B [\mathbf{u}]^B.$$

Matrix $[\mathbf{u}]^B$, that shortens the highlighted term, has to be zero, otherwise there would be an impossible situation of a negative eigenvalue related to a nonnegative Hermitian operator $\mathbf{h}^{1/2}$. In the case of $\lambda_i = 0$, there is no restriction for matrix $[\mathbf{u}]^B$, that can be zero, for example. So, this highlighted term becomes

$$\sqrt{\lambda_i}\, [\mathbf{x}_i]^B = \left[\mathbf{h}_B^{1/2}\right]^B [\mathbf{x}_i]^B.$$

This same process can be used in the case of operator $\mathbf{c}^{1/2}$, from which we conclude that

$$\left[\mathbf{h}_B^{1/2}\right]^B [\mathbf{x}_i]^B = \left[\mathbf{c}_B^{1/2}\right]^B [\mathbf{x}_i]^B$$

for all eigenvectors of $\mathbf{h}$. Since they are not zero, $\mathbf{h}^{1/2}$ is unique. □

An operator of unitary group $(O, \circ)$, whose elements have the Hilbert space $U_{\mathbb{F}}$ as a domain, can be represented by a unitary matrix if the basis in question is orthonormal. In other words, if $\mathbf{q} \in O$ and $\hat{B}$ is an orthonormal basis of $U_{\mathbb{F}}$, through equalities (2.45) and (2.48), matrix

$$[\mathbf{q}_{\hat{B}}]^{\hat{B}^{-1}} = [\mathbf{q}_{\hat{B}}]^{\hat{B}^{\dagger}}. \tag{2.66}$$

It is convenient to recall that a unitary operator preserves norm, that is, it does not alter "size" or "intensity" of the vectors in its domain. For our study, it is important to discriminate this characteristic in any linear operator through a decomposition, in such a way that there results a unitary part and a non unitary part, which is exclusively responsible for altering norm. The following theorem meets this demand. The term "polar" on its name refers to a feature that is similar to the polar form of a complex number, where there is a nonnegative real part that describes the magnitude and another part whose magnitude is unitary.

Theorem 8 – Polar Decomposition

Considering $\mathsf{U}_{\mathbb{F}}$ a Hilbert space, a bijection $\mathbf{g} \in \mathcal{L}_{\mathbb{F}}(\mathsf{U},\mathsf{U})$ has one and only one decomposition of the form $\mathbf{g} = \mathbf{q} \circ \mathbf{h}^{1/2}$, where $\mathbf{q} \in \mathcal{L}_{\mathbb{F}}(\mathsf{U},\mathsf{U})$ is unitary and $\mathbf{h} = \mathbf{g} \circ \mathbf{g}^{\dagger}$ is nonnegative Hermitian.

Proof. Let $\mathbf{g} \in \mathcal{L}_{\mathbb{F}}(\mathsf{U},\mathsf{U})$ be a bijection and function $\mathbf{h} = \mathbf{g} \circ \mathbf{g}^{\dagger}$ a Hermitian nonnegative operator. We know that decomposition $\mathbf{h} = \mathbf{h}^{1/2} \circ \mathbf{h}^{1/2}$ exists, and therefore

$$\begin{aligned} \mathbf{h}^{1/2} \circ \mathbf{h}^{1/2} &= \mathbf{g} \circ \mathbf{g}^{\dagger} \\ \mathbf{i} &= \underbrace{\mathbf{h}^{-1/2} \circ \mathbf{g}}_{\mathbf{q}} \circ \underbrace{\mathbf{g}^{\dagger} \circ \mathbf{h}^{-1/2}}_{\mathbf{q}^{\dagger}}, \end{aligned}$$

from which we can conclude that $\mathbf{q}$ is indeed unitary and thus a polar decomposition exists. As the square root $\mathbf{h}^{1/2}$ is unique for $\mathbf{h} = \mathbf{g} \circ \mathbf{g}^{\dagger}$, the unitary operator $\mathbf{q} = \mathbf{g} \circ \mathbf{h}^{-1/2}$ is also unique; which results a unique polar decomposition of $\mathbf{g}$. $\square$

Corollary 8.1 – Left and Right Polar Decompositions

Considering the polar decomposition $\mathbf{g} = \mathbf{q} \circ \mathbf{h}_1^{1/2}$, called ***right polar decomposition****, equality $\mathbf{g} = \mathbf{h}_2^{1/2} \circ \mathbf{q}$, where $\mathbf{h}_2 = \mathbf{g}^{\dagger} \circ \mathbf{g}$, is unique and hence called* ***left polar decomposition****.*

Proof. Existence and uniqueness of right and left polar decomposition follows are verified similarly. Let's prove now that the unitary operators in both decompositions are really the same. From the left polar decomposition, whose unitary operator is $\mathbf{q}_1$, we need to verify that the part $\mathbf{c} = \mathbf{q}_1^{-1} \circ \mathbf{h}_2^{1/2} \circ \mathbf{q}_1$ of equality $\mathbf{g} = \mathbf{q}_1 \circ \mathbf{q}_1^{-1} \circ \mathbf{h}_2^{1/2} \circ \mathbf{q}_1$ is nonnegative Hermitian because, if this is true, we'll have a right polar decomposition and thus $\mathbf{q}_1 = \mathbf{q}$. Proving that $\mathbf{c}$ is Hermitian, or that $\mathbf{c} = \mathbf{c}^{\dagger}$, is trivial. Now, let's verify if, given an arbitrary orthonormal basis and a vector $\mathbf{x} \in \mathsf{U}_{\mathbb{F}}$, the scalar $\Re([\mathbf{x}]^{\dagger}[\mathbf{c}][\mathbf{x}])_{11}$ is nonnegative. From equality $[\mathbf{x}]^{\dagger}[\mathbf{c}][\mathbf{x}] = [\mathbf{x}]^{\dagger}[\mathbf{q}_1]^{-1}[\mathbf{h}_2^{1/2}][\mathbf{q}_1][\mathbf{x}]$, we can conclude that $\Re(A^{\dagger}[\mathbf{h}_2^{1/2}]A)_{11} \geqslant 0$, where $A = [\mathbf{q}_1][\mathbf{x}]$, because $\mathbf{h}_2^{1/2}$ is nonnegative. $\square$

Chapter

3 Basic Tensor Algebra

Continuing our study of Linear Algebra, whose foundations we presented in the previous chapter, let's deal now with something a bit more sophisticated, although extremely useful for describing physical phenomena concerning deformable bodies. The form and the content which we shall present this topic differs from the usual approach available on the literature of Continuum Mechanics, where the concept of tensor arises suddenly, out of nowhere, and through simplistic definitions. Formally, we shall proceed the same way as we have been doing so far, in a continuous flow of successively dependent concepts, and present tensors as ordinary elements of a vector space; on the content side, we shall not restrict ourselves to the strictly necessary, typical of pragmatic approaches, since our objective, in this and future chapters, is to provide the reader with a generous framework of concepts related to tensors, firstly because of their mathematical beauty and secondly because of their wide application.

3.1 Historic Preamble

The word "tensor" was used for the first time in Mathematics in 1846 by Hamilton[27] in order to designate the modulus of numbers that he called quaternions: a generalization of complex numbers. In 1881, J. Willard Gibbs(1839–1903) introduced the notion of dyad as being an "undefined product" of two vectors, starting, in a rather obscure way, the concept that nowadays we understand as tensor product. He used this concept to express cause and effect relations between three dimensional real vectors such that each coordinate of the effect vector was described by linear contributions from all the coordinates of the cause vector. In order to illustrate this idea, considering a three dimensional Euclidean space $\mathsf{U}_{\mathbb{R}}$, from an orthonormal basis $\mathsf{B} = \{\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2, \hat{\mathbf{u}}_3\} \subset \mathsf{U}_{\mathbb{R}}$, GIBBS[22] described the function in a linear mapping $\mathbf{\Phi} : \mathbb{U}_{\mathbb{R}} \mapsto \mathbb{U}_{\mathbb{R}}$ as the following:

$$\mathbf{\Phi} = \sum_{i=1}^{3}\sum_{j=1}^{3} \Phi_{ij}\hat{\mathbf{u}}_i\hat{\mathbf{u}}_j\,,$$

where $\Phi_{ij} \in \mathbb{R}$ and $\hat{\mathbf{u}}_i\hat{\mathbf{u}}_j$ is his dyad such that $\hat{\mathbf{u}}_i\hat{\mathbf{u}}_j(\mathbf{x}) = (\hat{\mathbf{u}}_j \cdot \mathbf{x})\hat{\mathbf{u}}_i$.Therefore, vector

$$\mathbf{y} = \mathbf{\Phi}(\mathbf{x}) = [\sum_{i=1}^{3}\sum_{j=1}^{3} \Phi_{ij}\hat{\mathbf{u}}_i\hat{\mathbf{u}}_j](\sum_{k=1}^{3} \mathbf{f}_k^B(\mathbf{x})\hat{\mathbf{u}}_k) = \sum_{i=1}^{3}\sum_{j=1}^{3} \Phi_{ij}\mathbf{f}_j^B(\mathbf{x})\hat{\mathbf{u}}_i\,,$$

where each coordinate $\mathbf{f}_i^B(\mathbf{y})$ is indeed defined by linear contributions from all the coordinates of $\mathbf{x}$. Since the term "tensor" has the latin word *tendere* – which means "to stretch", as its etymological origin – VOIGT[55] adopted a meaning closer to the current mechanical approach, when he used tensor to designate the combination of a vector measure, representing the forces and stretches involved, with another vector measure, normal to the surface in question.

At the end of XIX century, after ten years conceiving his "Absolute Differential Calculus", the italian mathematician Gregorio Ricci-Curbastro(1853-1925) published a seminal paper entitled *Méthodes de Calcul Différentiel Absolu et Leurs Applications*[1], together with his pupil Tullio Levi-Civita(1873-1941), in which they established the foundations of what is known today as Tensor Calculus. This work is an "evolution" of the works of Carl F. Gauss(1777-1855) and his pupil Bernhard Riemann(1826-1866) on positive-definite quadratic polynomial metrics defined on spaces tangent to curved hypersurfaces. Heavily inspired by the methodology created by CHRISTOFFEL[11] for differentating functions along hypersurfaces in the context of variable bases, Ricci and Levi-Civita developed what they called covariant and contravariant derivatives in such a formulation that is immune to choices or changes of bases. In this sense, they concluded that *"after having surmounted the initial difficulties, one will readily find that the generality and independence of choice of coordinates leads not only to elegance, but also to agility and insight into proofs and conclusions"*[2]. In simple terms, a tensor or "system", as they called, is a measure that arises from the coefficients of polynomials like

$$\sum_{i_1=1}^{n}\cdots\sum_{i_m=1}^{n} \underbrace{(\alpha_1)_{i_1}\cdots(\alpha_m)_{i_m}}_{A_{i_1\cdots i_m}}(x_1)_{i_1}\cdots(x_m)_{i_m}\,,$$

where $n \times \cdots \times n$ (m times) array A represents this measure numerically. Moreover, the values of such polynomials must be immune to changes of variables, that is,

$$\sum_{i_1=1}^{n}\cdots\sum_{i_m=1}^{n} A_{i_1\cdots i_m}(x_1)_{i_1}\cdots(x_m)_{i_m} = \sum_{i_1=1}^{n}\cdots\sum_{i_m=1}^{n} B_{i_1\cdots i_m}(y_1)_{i_1}\cdots(y_m)_{i_m}\,,$$

[1] See RICCI & LEVI-CIVITA[44] and HERMANN[28].

[2] HERMMAN[28], p. vi.

where the following is valid:

$$(y_k)_{i_k} = \sum_{j_k=1}^{n} C_{i_k j_k} (x_k)_{j_k},$$

from which the transformation law

$$A_{i_1 \cdots i_m} = \sum_{j_1=1}^{n} \cdots \sum_{j_m=1}^{n} C_{j_1 i_1} \cdots C_{j_m i_m} B_{j_1 \cdots j_m}$$

can be obtained. Similarly to vector measures, a tensor must have such an abstraction level that its definition results independent of the choice of variables, that is, in expressions above, the tensor represented by A is the same tensor represented by B. Thereby, in rather generic terms, a tensor was initially seen as an *absolute measure* that arises from the interaction of different sets of variables, arranged in an scalar-valued expression of linear character.

Ricci's Absolute Differential Calculus was presented by the hungarian mathematician Marcel Grossman (1878-1936) to his old school fellow Albert Einstein (1879-1955). When Einstein was developing his General Theory of Relativity, Grossman used to update him with the most recent mathematical tools available; in this case, non Euclidean geometry. During his work, Einstein noticed that there is not a preferred reference frame, inertial or not, that describes a certain physical phenomenon more adequately, that is, changes of frame do not promote qualitative modification on the description of phenomena. Since a tensor is immune to changes of variables, its use became a rather convenient mathematical tool for enabling the theory, in a concise and elegant development. About this tool and its mathematical foundations, Einstein[19] wrote the following compliment: *"The charm of this theory will hardly scape from anyone who really understands it; it means the real triumph of the methodology of Absolute Differential Calculus, established by GAUSS, RIEMANN, CHRISTOFFEL, RICCI and LEVI-CIVITA"*[3]. From then on, Tensor Calculus raised to prominence: its development intensified and its application spread out to other branches of Mathematical Physics.

3.2 Tensor Spaces

Let's recall vector spaces of functions that have a tuple of vectors as argument, that is, functions whose domain is a cartesian product of many sets, each one defining a vector space. Among these function spaces, we have a particular interest in those

[3] Translated by me from the original *"Dem Zauber dieser Theorie wird sich kaum jemand entziehen können, der sie wirklich erfaßt hat; sie bedeutet einen wahren Triumph der durch GAUSS, RIEMANN, CHRISTOFFEL, RICCI und LEVI-CIVITER begründeten Methode des allgemeinen Differentialkalkuls".*

constituted by multilinear functions, on the conditions presented in section 2.3. In order to proceed with the theory, based on the framework already studied, we chose to adopt an approach that takes tensor spaces as vector spaces of multilinear functionals, called tensors. In more mathematical terms, given a cartesian product $U^{\times m}$ where each set U_i defines a finite dimensional vector space $(U_i)_{\mathbb{F}}$, a function space $\mathcal{L}_{\mathbb{F}}(U^{\times m}, \mathbb{F})$, constituted by multilinear functionals, is said to be a **tensor space** of **order** m, whose representation is here shortened by $\mathcal{L}_{\mathbb{F}}(U^{\times m})$, and each functional $\mathbf{t} \in \mathcal{L}_{\mathbb{F}}(U^{\times m})$ is said to be a **tensor** of **order** m, represented by $\mathbf{T}$. In the particular case of $m = 1$, a tensor space is considered to be a dual space, that is, $\mathcal{L}_{\mathbb{F}}(U) = U^*_{\mathbb{F}}$, where a tensor is a linear functional. When all dimensions $\dim(U_i)_{\mathbb{F}} = n$, it is convenient to make this fact explicit by using representation $\mathcal{L}^{(n)}_{\mathbb{F}}(U^{\times m})$. Moreover, like any other vector space, a tensor space $\mathcal{L}_{\mathbb{F}}(U^{\times m})$ has its **conjugate tensor space** $\overline{\mathcal{L}_{\mathbb{F}}(U^{\times m})}$, whose arbitrary element is represented by $\mathbf{T}^c$ where $\alpha\mathbf{T}^c = \overline{\alpha}\mathbf{T}$ for all $\alpha \in \mathbb{F}$ according to what was presented in section 2.1 for vector spaces.

It is possible to build particular tensor spaces from catersian powers together with dual spaces: a tensor space $\mathcal{L}_{\mathbb{F}}(V^p \times V^{*q})$, represented by $\mathcal{L}_{\mathbb{F}}(V^{(p,q)})$, where $V^*_{\mathbb{F}}$ is the dual space of $V_{\mathbb{F}}$, is said to have a **contravariant order** p and a **covariant order** q or we can call it a type (p,q) tensor space. Moreover, a tensor can be defined from the dual spaces of its constituent vector spaces, that is, a linear functional in dual vector space $(U^*_i)_{\mathbb{F}}$ defines a multilinear functional (tensor) in $\mathcal{L}_{\mathbb{F}}(U^{\times m})$. Let's see how. The tuple $(\mathbf{g}_1, \cdots, \mathbf{g}_m)$, which is an element of cartesian product $(U^*)^{\times m} = U^*_1 \times \cdots \times U^*_m$, specifies a tensor in $\mathcal{L}_{\mathbb{F}}(U^{\times m})$ when the following rule is defined:

$$\mathbf{G}(x_1, \cdots, x_m) = \prod_{i=1}^{m} \mathbf{g}_i(x_i), \tag{3.1}$$

where tensor $\mathbf{G} \in \mathcal{L}_{\mathbb{F}}(U^{\times m})$ is classified dyadic if $m = 2$, **triadic** if $m = 3$ and **polyadic** if $m > 3$. Under these conditions, each linear functional $\mathbf{g}_i$ is called **dyad**, **triad** and **polyad** respectively. In order to handle and make its polyads explict, the polyadic tensor that $(\mathbf{g}_1, \cdots, \mathbf{g}_m)$ defines we represent it by $\mathbf{g}_1 \otimes \cdots \otimes \mathbf{g}_m$, where symbol $\otimes$ is a function of two arguments, definer of a mapping called tensor product. In more formal and generic terms, considering tensor spaces $\mathcal{L}_{\mathbb{F}}(V^{\times p})$ and $\mathcal{L}_{\mathbb{F}}(W^{\times q})$, a mapping $\otimes : \mathcal{L}_{\mathbb{F}}(V^{\times p}) \times \mathcal{L}_{\mathbb{F}}(W^{\times q}) \mapsto \mathcal{L}_{\mathbb{F}}(V^{\times p} \times W^{\times q})$ is called a **tensor product** if, for all tensors $\mathbf{V} \in \mathcal{L}_{\mathbb{F}}(V^{\times p})$ and $\mathbf{W} \in \mathcal{L}_{\mathbb{F}}(W^{\times q})$,

$$\mathbf{V} \otimes \mathbf{W}(x_1, \cdots, x_p, y_1, \cdots, y_q) = \mathbf{V}(x_1, \cdots, x_p)\mathbf{W}(y_1, \cdots, y_q), \tag{3.2}$$

where tensor $\mathbf{V} \otimes \mathbf{W} := \otimes(\mathbf{V}, \mathbf{W})$ belongs to $\mathcal{L}_{\mathbb{F}}(V^{\times p} \times W^{\times q})$. Items

i. Zero tensor: $\mathbf{U} \otimes \mathbf{V} = \mathbf{0} \implies \mathbf{U} = \mathbf{0}$ or $\mathbf{V} = \mathbf{0}$,

ii. Multiplication by scalar: $\alpha\mathbf{U} \otimes \beta\mathbf{V} = \alpha\beta(\mathbf{U} \otimes \mathbf{V}) = \beta\mathbf{U} \otimes \alpha\mathbf{V}$,

iii. Associativity: $(\mathbf{U} \otimes \mathbf{V}) \otimes \mathbf{W} = \mathbf{U} \otimes (\mathbf{V} \otimes \mathbf{W})$,

iv. Transposition: if $\mathbf{U}\otimes\mathbf{V}=\mathbf{U}_1\otimes\mathbf{U}_2+\mathbf{V}_1\otimes\mathbf{V}_2$ then $\mathbf{V}\otimes\mathbf{U}=\mathbf{U}_2\otimes\mathbf{U}_1+\mathbf{V}_2\otimes\mathbf{V}_1$,

v. Left distributivity: $(\mathbf{U}+\mathbf{V})\otimes\mathbf{W}=\mathbf{U}\otimes\mathbf{W}+\mathbf{V}\otimes\mathbf{W}$,

vi. Right distributivity: $\mathbf{U}\otimes(\mathbf{V}+\mathbf{W})=\mathbf{U}\otimes\mathbf{V}+\mathbf{U}\otimes\mathbf{W}$ and

vii. Conjugacy: $(\mathbf{U}\otimes\mathbf{V})^c=\mathbf{U}^c\otimes\mathbf{V}=\mathbf{U}\otimes\mathbf{V}^c$

are valid properties for all $\alpha,\beta\in\mathbb{F}$ and $\mathbf{U},\mathbf{U}_i,\mathbf{V},\mathbf{V}_i,\mathbf{W}\in\mathcal{L}_{\mathbb{F}}(U^{\times m})$. From the previous associativity property, it is valid to represent a polyadic tensor by a succession of tensor products, as we did for $\mathbf{g}_1\otimes\cdots\otimes\mathbf{g}_m$.

Now, the following theorem states that a set of specific polyadic tensors, defined by tuples of coordinate functionals, is able to span a tensor space.

Theorem 9 – Basis of Polyadic Tensors

Given a tensor space $\mathcal{L}_{\mathbb{F}}(U^{\times m})$ *and a basis* $B_i=\{\mathbf{u}_1^{(i)},\cdots,\mathbf{u}_{n_i}^{(i)}\}$ *of each constituent vector space* $(U_i)_{\mathbb{F}}$, $i=1,\cdots,m$, *the set* B *constituted by all polyadic tensors of the form* $\mathbf{f}_{j_1}^{B_1}\otimes\cdots\otimes\mathbf{f}_{j_m}^{B_m}$ *is a basis of* $\mathcal{L}_{\mathbb{F}}(U^{\times m})$, *where coordinate functionals* $\mathbf{f}_{j_i}^{B_i}\in(U_i)^*_{\mathbb{F}}$, $j_i=1,\cdots,n_i$.

Proof. Firstly, concerning the above properties of tensor products, the first six can be verified from the definition (3.2) and from trivial addition and scalar multiplication properties. The last property is actually a corollary of the second: if we consider that $\mathbf{T}=\mathbf{U}\otimes\mathbf{V}$ in equality $\alpha\mathbf{T}^c=\overline{\alpha}\mathbf{T}$, we can write that $\alpha(\mathbf{U}\otimes\mathbf{V})^c=\overline{\alpha}\mathbf{U}\otimes\mathbf{V}=\alpha\mathbf{U}^c\otimes\mathbf{V}$ or that $\alpha(\mathbf{U}\otimes\mathbf{V})^c=\mathbf{U}\otimes(\overline{\alpha}\mathbf{V})=\alpha\mathbf{U}\otimes\mathbf{V}^c$. Concerning the theorem, we need to prove firstly that if equalities $\mathbf{T}_1(\mathbf{u}_{j_1}^{(1)},\cdots,\mathbf{u}_{j_m}^{(m)})=\mathbf{T}_2(\mathbf{u}_{j_1}^{(1)},\cdots,\mathbf{u}_{j_m}^{(m)})$, where $\mathbf{T}_1,\mathbf{T}_2\in\mathcal{L}_{\mathbb{F}}(U^{\times m})$, are valid, it results that $\mathbf{T}_1=\mathbf{T}_2$. In order to do that, we write

$$\begin{aligned}\mathbf{T}_1(\mathbf{x}_1,\cdots,\mathbf{x}_m)&=\sum_{j_1=1}^{n_1}\cdots\sum_{j_m=1}^{n_m}\mathbf{f}_{j_1}^{B_1}(\mathbf{x}_1)\cdots\mathbf{f}_{j_m}^{B_m}(\mathbf{x}_m)\mathbf{T}_1(\mathbf{u}_{j_1}^{(1)},\cdots,\mathbf{u}_{j_m}^{(m)})\\&=\sum_{j_1=1}^{n_1}\cdots\sum_{j_m=1}^{n_m}\mathbf{f}_{j_1}^{B_1}(\mathbf{x}_1)\cdots\mathbf{f}_{j_m}^{B_m}(\mathbf{x}_m)\mathbf{T}_2(\mathbf{u}_{j_1}^{(1)},\cdots,\mathbf{u}_{j_m}^{(m)})\\&=\mathbf{T}_2(\mathbf{x}_1,\cdots,\mathbf{x}_m),\end{aligned}$$

from which we conclude $\mathbf{T}_1=\mathbf{T}_2$. Now, we consider $\mathbf{T}\in\mathcal{L}_{\mathbb{F}}(U^{\times m})$ to be a tensor whose scalars $\beta_{i_1\cdots i_m}=\mathbf{T}(\mathbf{u}_{i_1}^{(1)},\cdots,\mathbf{u}_{i_m}^{(m)})$ define a tensor

$$\mathbf{X}=\sum_{i_1=1}^{n_1}\cdots\sum_{i_m=1}^{n_m}\beta_{i_1\cdots i_m}\mathbf{f}_{i_1}^{B_1}\otimes\cdots\otimes\mathbf{f}_{i_m}^{B_m}.$$

Thus we can write the following:

$$\mathbf{X}(\mathbf{u}_{j_1}^{(1)},\cdots,\mathbf{u}_{j_m}^{(m)})=\sum_{i_1=1}^{n_1}\cdots\sum_{i_m=1}^{n_m}\beta_{i_1\cdots i_m}\prod_{k=1}^{m}\delta_{i_k j_k}=\beta_{j_1\cdots j_m}=\mathbf{T}(\mathbf{u}_{j_1}^{(1)},\cdots,\mathbf{u}_{j_m}^{(m)}).$$

Such equalities enable us to conclude that $\mathbf{X}=\mathbf{T}$ and then $\operatorname{span}(B)=\mathcal{L}_{\mathbb{F}}(U^{\times m})$, since $\mathbf{T}$ is an arbitrary tensor of $\mathcal{L}_{\mathbb{F}}(U^{\times m})$. Now we need to verify if B is linearly independent. This is true when coefficients $\alpha_{i_1\cdots i_m}$ on equality

$$\sum_{i_1=1}^{n_1}\cdots\sum_{i_m=1}^{n_m}\alpha_{i_1\cdots i_m}\mathbf{f}_{i_1}^{B_1}\otimes\cdots\otimes\mathbf{f}_{i_m}^{B_m}=\mathbf{0}$$

are zero, which can be verified through the following development:

$$\sum_{i_1=1}^{n_1}\cdots\sum_{i_m=1}^{n_m}\alpha_{i_1\cdots i_m}\mathbf{f}_{i_1}^{B_1}\otimes\cdots\otimes\mathbf{f}_{i_m}^{B_m}(\mathbf{u}_{j_1}^{(1)},\cdots,\mathbf{u}_{j_m}^{(m)})=\mathbf{0}(\mathbf{u}_{j_1}^{(1)},\cdots,\mathbf{u}_{j_m}^{(m)})$$

$$\sum_{i_1=1}^{n_1}\cdots\sum_{i_m=1}^{n_m}\alpha_{i_1\cdots i_m}\prod_{k=1}^{m}\delta_{i_k j_k}=0$$

$$\alpha_{j_1\cdots j_m}=0$$

□

In the context of the above theorem, it can be noted that the number of polyadic tensors $\mathbf{f}_{j_1}^{B_1}\otimes\cdots\otimes\mathbf{f}_{j_m}^{B_m}$ of basis B is given by $n_1\cdots n_m$. Moreover, from the previous chapter, we already know that a vector space and its dual have the same dimension, that is, $\dim((U_i)_{\mathbb{F}})=\dim((U_i^*)_{\mathbb{F}})$. Thereby, we can state that the dimension of tensor space $\mathcal{L}_{\mathbb{F}}(U^{\times m})$ is the product of dimensions of each of vector spaces $(U_i)_{\mathbb{F}}$. In other words,

$$\dim(\mathcal{L}_{\mathbb{F}}(U^{\times m}))=\prod_{i=1}^{m}\dim((U_i)_{\mathbb{F}}). \tag{3.3}$$

Like any other basis of a vector space, each element of basis B of tensor space $\mathcal{L}_{\mathbb{F}}(U^{\times m})$ also have a related coordinate functional, here represented by $\mathbf{f}_{j_1\cdots j_m}^{B}$. Thereby, given an arbitrary tensor $\mathbf{T}\in\mathcal{L}_{\mathbb{F}}(U^{\times m})$, the tuple constituted by scalars $\mathbf{f}_{j_1\cdots j_m}^{B}(\mathbf{T})$, $j_i=1,\cdots,n_i$ is said to be the coordinates of $\mathbf{T}$ on B. Using this definition, tensor $\mathbf{T}$ can be described by expression

$$\mathbf{T}=\sum_{k_1=1}^{n_1}\cdots\sum_{k_m=1}^{n_m}\mathbf{f}_{k_1\cdots k_m}^{B}(\mathbf{T})\mathbf{f}_{k_1}^{B_1}\otimes\cdots\otimes\mathbf{f}_{k_m}^{B_m}, \tag{3.4}$$

from which the equalities

$$\mathbf{T}(\mathbf{u}_{j_1}^{(1)},\cdots,\mathbf{u}_{j_m}^{(m)})=\sum_{k_1=1}^{n_1}\cdots\sum_{k_m=1}^{n_m}\mathbf{f}_{k_1\cdots k_m}^{B}(\mathbf{T})\prod_{i=1}^{m}\mathbf{f}_{k_i}^{B_i}(\mathbf{u}_{j_i}^{(i)})=\mathbf{f}_{j_1\cdots j_m}^{B}(\mathbf{T})$$

enable to define the following rule for coordinate functionals on basis B:

$$\mathbf{f}_{j_1\cdots j_m}^{B}(\mathbf{X})=\mathbf{X}(\mathbf{u}_{j_1}^{(1)},\cdots,\mathbf{u}_{j_m}^{(m)}). \tag{3.5}$$

In type (p,q) tensor spaces, the coordinates of a tensor $\mathbf{V}\in\mathcal{L}_{\mathbb{F}}(V^{(p,q)})$ are defined by scalars usually represented by

$$(\mathbf{f}^{Z})_{k_1\cdots k_q}^{j_1\cdots j_p}(\mathbf{V}),\ j_i,k_l=1,\cdots,n, \tag{3.6}$$

where Z is a basis of $\mathcal{L}_{\mathbb{F}}(V^{(p,q)})$, $i=1,\cdots,p$ and $l=1,\cdots,q$. The superscript indexes j_i are related to space $V_{\mathbb{F}}$ of vectors whose coordinates are contravariant and the subscript indexes k_l related to $V_{\mathbb{F}}^*$ of vectors whose coordinates are covariant. Tensors are called

covariant or **contravariant** when they exclusively have covariant or contravariant indexes respectively and called **mixed** when they have both. For (p,q) tensor spaces, given a basis $W=\{\boldsymbol{w}_1,\cdots,\boldsymbol{w}_n\}$ of $V_{\mathbb{F}}$, rule (3.5) for coordinate functionals results

$$(\mathbf{f}^Z)^{j_1\cdots j_p}_{k_1\cdots k_q}(\mathbf{X})=\mathbf{X}(\boldsymbol{w}_{j_1},\cdots,\boldsymbol{w}_{j_p},\mathbf{f}^W_{k_1},\cdots,\mathbf{f}^W_{k_q}),\tag{3.7}$$

where bases $W^*=\{\mathbf{f}^W_1,\cdots,\mathbf{f}^W_n\}$ of $V^*_{\mathbb{F}}$ and $W^{**}=\{\mathbf{f}^{W^*}_1,\cdots,\mathbf{f}^{W^*}_n\}$ of $V^{**}_{\mathbb{F}}$ define basis Z, in such a way that its elements have the format

$$\mathbf{f}^W_{j_1}\otimes\cdots\otimes\mathbf{f}^W_{j_p}\otimes\mathbf{f}^{W^*}_{k_1}\otimes\cdots\otimes\mathbf{f}^{W^*}_{k_q},\tag{3.8}$$

according to theorem 9. In the context of Hilbert spaces, from the schematic figure 2.4, we can affirm that $\mathbf{f}^W_{j_i}=(\boldsymbol{w}^{j_i})^*$ and $\mathbf{f}^{W^*}_{k_i}=(\mathbf{f}^W)^{k_i}=\mathbf{f}^{W^\perp}_{k_i}=\boldsymbol{w}^*_{k_i}$. Thereby, the elements of Z in this case have the classic format

$$(\boldsymbol{w}^{j_1})^*\otimes\cdots\otimes(\boldsymbol{w}^{j_p})^*\otimes\boldsymbol{w}^*_{k_1}\otimes\cdots\otimes\boldsymbol{w}^*_{k_q}.\tag{3.9}$$

Since we just learned the rule to obtain the coordinates of a tensor on a certain basis, let's see how these coordinates behave when there is a change of basis. In the context of theorem 9, let $C_i=\{\boldsymbol{v}^{(i)}_1,\cdots,\boldsymbol{v}^{(i)}_{n_i}\}$ be a basis of each vector space $(U_i)_{\mathbb{F}}$ and $[\boldsymbol{i}_{C_i}]^{B_i}$ a matrix that changes basis B_i to basis C_i. From the rule (3.5), given a tensor $\mathbf{T}\in\boldsymbol{\mathcal{L}}_{\mathbb{F}}(U^{\times m})$, we can develop the following:

$$\begin{aligned}\mathbf{f}^C_{j_1\cdots j_m}(\mathbf{T})&=\mathbf{T}(\boldsymbol{v}^{(1)}_{j_1},\cdots,\boldsymbol{v}^{(m)}_{j_m})\\&=\sum_{k_1=1}^{n_1}\cdots\sum_{k_m=1}^{n_m}\mathbf{f}^B_{k_1\cdots k_m}(\mathbf{T})\,\mathbf{f}^{B_1}_{k_1}\otimes\cdots\otimes\mathbf{f}^{B_m}_{k_m}(\boldsymbol{v}^{(1)}_{j_1},\cdots,\boldsymbol{v}^{(m)}_{j_m})\\&=\sum_{k_1=1}^{n_1}\cdots\sum_{k_m=1}^{n_m}\mathbf{f}^B_{k_1\cdots k_m}(\mathbf{T})\prod_{i=1}^{m}\mathbf{f}^{B_i}_{k_i}(\boldsymbol{v}^{(i)}_{j_i})\\&=\sum_{k_1=1}^{n_1}\cdots\sum_{k_m=1}^{n_m}\mathbf{f}^B_{k_1\cdots k_m}(\mathbf{T})\prod_{i=1}^{m}[\boldsymbol{i}_{C_i}]^{B_i}_{k_ij_i},\end{aligned}\tag{3.10}$$

from which the last equality describes a change of coordinates of tensor $\mathbf{T}$ from basis B to basis C. Through this same process, it is possible to change the coordinates of a type (p,q) tensor.

A consequence of describing an arbitrary tensor $\mathbf{T}$ as a linear combination of polyadic tensor elements of a certain basis is the possibility to arrange its coordinates in such a way that $\mathbf{T}$ results an additive decomposition of polyadic tensors. In order to obtain this, let's take the property of multiplication by scalars of tensor products[4] and

[4]See p. 68.

"move", during the k_i-th sum in (3.4), the coordinates $\mathbf{f}^{B}_{k_1\cdots k_i\cdots k_m}(\mathbf{T})$ to multiply the coordinate functional $\mathbf{f}^{B_i}_{k_i}$ of the i-th polyadic basis tensor, resulting in

$$\mathbf{T} = \sum_{i=1}^{m}\sum_{k_i=1}^{n_i} \underbrace{\mathbf{f}^{B_1}_{k_1} \otimes \cdots \otimes (\mathbf{f}^{B}_{k_1\cdots k_i\cdots k_m}(\mathbf{T})\mathbf{f}^{B_i}_{k_i}) \otimes \cdots \otimes \mathbf{f}^{B_m}_{k_m}}_{(\mathbf{t}_1)_i \otimes \cdots \otimes (\mathbf{t}_m)_i}, \tag{3.11}$$

where linear functional $(\mathbf{t}_j)_i \in (U_j)^*_{\mathbb{F}}$ is the j-th polyad of the i-th polyadic tensor member of the additive decomposition of $\mathbf{T}$. Moreover, if an arbitrary tensor can be decomposed as a sum of polyadic tensors and a sum of polyadic tensors results in a polyadic tensor because every tensor is a vector, we can state that *every tensor can be described as a polyadic tensor*. For the previous case,

$$\mathbf{T} = \sum_{i=1}^{m} (\mathbf{t}_1)_i \otimes \cdots \otimes (\mathbf{t}_m)_i = \mathbf{t}_1 \otimes \cdots \otimes \mathbf{t}_m . \tag{3.12}$$

It is important to note that, in the context of Hilbert spaces, where Riesz-Fréchet Representation Theorem is valid, *every polyad of a polyadic tensor is a covector*. Moreover, also from this theorem, if a tensor space has a unitary order, its tensors are covectors because every linear functional whose domain is a Hilbert space is biunivocally related to a vector. Therefore, given an arbitrary vector $\mathbf{u} \in U_{\mathbb{F}}$, its covector $\mathbf{u}^*$ is a first order tensor, or $\mathbf{u}^* \in \mathcal{L}_{\mathbb{F}}(U)$. Moreover, if a scalar α belongs to a Hilbert space $\mathbb{F}_{\mathbb{F}}$, then covector or **coscalar** α^*, where $\alpha^*(x) = \alpha x$, is defined to be a tensor of order zero. In other words, *concerning tensor spaces defined by Hilbert spaces, covectors are first order tensors and coscalars are defined to be zeroth order tensors*. Now, given a tensor space $\mathcal{L}_{\mathbb{R}}(E^m)$, where $E_{\mathbb{R}}$ is a three-dimensional Euclidean space, let $\hat{B}$ be a basis whose 3^m elements have the format $\hat{\mathbf{e}}^*_{j_1} \otimes \cdots \otimes \hat{\mathbf{e}}^*_{j_m}$, $j_i = 1,2,3$. Thereby, we call **cartesian tensor** any linear combination of the elements of $\hat{B}$.

The tensors of space $\mathcal{L}_{\mathbb{F}}(V^m)$, $m > 1$, can have the positions of their arguments interchanged, that is, given a tensor of $\mathcal{L}_{\mathbb{F}}(V^2)$, for example, arguments in $(\mathbf{x},\mathbf{y})$ can be changed to $(\mathbf{y},\mathbf{x})$, where $\mathbf{x},\mathbf{y} \in V_{\mathbb{F}}$. In this context, we say that a tensor $\mathbf{T}_1$ is a permutation of tensor $\mathbf{T}_2$ if $\mathbf{T}_1(\mathbf{x},\mathbf{y}) = \mathbf{T}_2(\mathbf{y},\mathbf{x})$. In general terms, given a bijective mapping $\pi : \{1,\cdots,n\} \mapsto \{1,\cdots,n\}$, where the sets involved are constituted by ordinals, and a function π called permutation of order m, in definition

$$\mathbf{T}_\pi(\mathbf{v}_1,\cdots,\mathbf{v}_m) := \mathbf{T}(\mathbf{v}_{\pi(1)},\cdots,\mathbf{v}_{\pi(m)}), \forall \mathbf{v}_i \in V_{\mathbb{F}}, \tag{3.13}$$

the tensor $\mathbf{T}_\pi \in \mathcal{L}_{\mathbb{F}}(V^m)$ is called the π **permutation** of $\mathbf{T} \in \mathcal{L}_{\mathbb{F}}(V^m)$. There are two important properties: a) the permutation of sum is the sum of permutations, that is, if $\mathbf{T} = \mathbf{A} + \mathbf{B}$ then $(\mathbf{A}+\mathbf{B})_\pi = \mathbf{A}_\pi + \mathbf{B}_\pi$; b) tensor $(\alpha\mathbf{T})_\pi = \alpha\mathbf{T}_\pi$, $\forall \alpha \in \mathbb{F}$. Now, considering the specific rule $\pi(x) = a\delta_{xb} + b\delta_{xa} + x(1-\delta_{xa}-\delta_{xb})$ that interchanges indexes a and b, we call $\mathbf{T}_\pi$ the (a,b) **transposition** of $\mathbf{T}$, represented by $\mathbf{T}_{(a,b)}$, where $a \neq b$. In order to represent p successive transpositions of $\mathbf{T}$, we have

$$\mathbf{T}_{(a_1,b_1)(a_p,b_p)} := [\cdots[\mathbf{T}_{(a_1,b_1)}]_{(a_2,b_2)}\cdots]_{(a_p,b_p)}. \tag{3.14}$$

Proof. Let's verify the two properties above. As any other vector, a tensor can always be the result of a sum of tensors; thereby, from definition (3.13) and considering $\mathbf{T} = \mathbf{A} + \mathbf{B}$, we have the following development:

$$\begin{aligned}(\mathbf{A}+\mathbf{B})_\pi(\mathbf{u}_1,\cdots,\mathbf{u}_m) &= (\mathbf{A}+\mathbf{B})\,(\mathbf{u}_{\pi(1)},\cdots,\mathbf{u}_{\pi(m)})\\ &= \mathbf{A}\,(\mathbf{u}_{\pi(1)},\cdots,\mathbf{u}_{\pi(m)}) + \mathbf{B}\,(\mathbf{u}_{\pi(1)},\cdots,\mathbf{u}_{\pi(m)})\\ &= \mathbf{A}_\pi(\mathbf{u}_1,\cdots,\mathbf{u}_m) + \mathbf{B}_\pi(\mathbf{u}_1,\cdots,\mathbf{u}_m)\\ &= (\mathbf{A}_\pi + \mathbf{B}_\pi)(\mathbf{u}_1,\cdots,\mathbf{u}_m),\end{aligned}$$

from which we conclude that the permutation of sum is the sum of permutations. The proof of the second property follows this same process. □

The concept of permutation enables the concept of symmetry as follows: a tensor $\mathbf{S} \in \mathcal{L}_{\mathbb{F}}(V^m)$ is said to be **symmetric** if equality $\mathbf{S}_\pi = \mathbf{S}$ is valid for any permutation π. In contrast, an element $\mathbf{P} \in \mathcal{L}_{\mathbb{F}}(V^m)$ is said to be **antisymmetric** if a tensor $\mathbf{P}_\pi = \epsilon_{\pi(1)\cdots\pi(m)}\mathbf{P}$ for any π. An antisymmetric tensor is usually called **alternating** when $m = \dim(V_{\mathbb{F}})$, that is, the order of tensor space equals the dimension of its definer vector space. If tensor $\mathbf{P}$ is antisymmetric, we can conclude that scalar $\mathbf{P}(x_1,\cdots,x_a,\cdots,x_b,\cdots,x_m) = 0$ if $x_a = x_b$, where $x_i \in V_{\mathbb{F}}$ and $a,b \in \{1,\cdots,m\}$. Thereby,

$$\mathbf{P}(x_{i_1},\cdots,x_{i_m}) = \epsilon_{i_1\cdots i_m}\mathbf{P}(x_1,\cdots,x_m). \tag{3.15}$$

Moreover, since the sum of two (anti)symmetric tensors is an (anti)symmetric tensor and multiplication by scalar also preserves (anti)symmetry, the set of all (anti)symmetric tensors defines a tensor subspace of $\mathcal{L}_{\mathbb{F}}(V^m)$: if this subspace is constituted by all symmetric tensors, it is represented by $\mathcal{LS}_{\mathbb{F}}(V^m)$; if it is constituted by antisymmetric tensors, we have $\mathcal{LA}_{\mathbb{F}}(V^m)$. Then, it is possible, and intriguing, to obtain that *every tensor subspace $\mathcal{LA}_{\mathbb{F}}^{(m)}(V^m)$ defined by alternating tensors is one-dimensional*. Less intriguing and quite convenient for future uses are the following properties.

i. For an arbitrary basis $B \subset V_{\mathbb{F}}$, a set $\{\mathbf{A}_B\}$ is a basis of alternating tensor space $\mathcal{LA}_{\mathbb{F}}^{(m)}(V^m)$ when

$$\mathbf{A}_B(x_1,\cdots,x_m) := \sum_{i_1=1}^{m}\cdots\sum_{i_m=1}^{m}\mathbf{f}_{i_1}^B(x_1)\cdots\mathbf{f}_{i_m}^B(x_m)\epsilon_{i_1\cdots i_m}; \tag{3.16}$$

ii. If basis $B = \{\mathbf{u}_1,\cdots,\mathbf{u}_m\}$, then $\mathbf{A}_{B\,(a,b)} = -\mathbf{A}_B$ and $\mathbf{A}_B(\mathbf{u}_1,\cdots,\mathbf{u}_m) = 1$;

iii. If tensor $\mathbf{A}$ is an arbitrary non-zero element of tensor space $\mathcal{LA}_{\mathbb{F}}^{(m)}(V^m)$ and subset $B = \{\mathbf{u}_1,\cdots,\mathbf{u}_m\}$ is a basis of $V_{\mathbb{F}}$, then $\mathbf{A} = \mathbf{A}(\mathbf{u}_1,\cdots,\mathbf{u}_m)\mathbf{A}_B$;

iv. A subset $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_m\}\subset V_{\mathbb{F}}$ is linearly dependent if and only if $\mathbf{A}(\boldsymbol{v}_1,\cdots,\boldsymbol{v}_m)=0$, where $\mathbf{A}\in\{\boldsymbol{\mathcal{LA}}_{\mathbb{F}}^{(m)}(V^m)\setminus\mathbf{0}\}$ is arbitrary;

v. Given an arbitrary operator $\mathbf{g}\in\mathcal{L}_{\mathbb{F}}(V,V)$, an alternating tensor $\mathbf{A}\in\boldsymbol{\mathcal{LA}}_{\mathbb{F}}^{(m)}(V^m)$ and a basis $B=\{\mathbf{u}_1,\cdots,\mathbf{u}_m\}$ of $V_{\mathbb{F}}$, then

$$\mathbf{A}(\mathbf{g}(\mathbf{u}_1),\cdots,\mathbf{g}(\mathbf{u}_m))=\det(\mathbf{g})\mathbf{A}(\mathbf{u}_1,\cdots,\mathbf{u}_m)\,. \tag{3.17}$$

vi. Given an arbitrary operator $\mathbf{g}\in\mathcal{N}_{\mathbb{F}}(V)$, then

$$\mathbf{A}(\mathbf{g}(\mathbf{u}_1),\cdots,\mathbf{g}(\mathbf{u}_m))=\pm\mathbf{A}(\mathbf{u}_1,\cdots,\mathbf{u}_m)\,. \tag{3.18}$$

Proof. First, let's prove (3.15). If $\boldsymbol{x}_a=\boldsymbol{x}_b$, then

$$\begin{aligned}\mathbf{P}_\pi(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_a,\cdots,\boldsymbol{x}_b,\cdots,\boldsymbol{x}_m)&=\mathbf{P}_\pi(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_b,\cdots,\boldsymbol{x}_a,\cdots,\boldsymbol{x}_m)\\ \epsilon_{i_1\cdots i_a\cdots i_b\cdots i_m}\mathbf{P}(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_a,\cdots,\boldsymbol{x}_b,\cdots,\boldsymbol{x}_m)&=\epsilon_{i_1\cdots i_b\cdots i_a\cdots i_m}\mathbf{P}(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_b,\cdots,\boldsymbol{x}_a,\cdots,\boldsymbol{x}_m)\\ \mathbf{P}(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_a,\cdots,\boldsymbol{x}_b,\cdots,\boldsymbol{x}_m)&=-\mathbf{P}(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_b,\cdots,\boldsymbol{x}_a,\cdots,\boldsymbol{x}_m)\,,\end{aligned}$$

from which we conclude that $\mathbf{P}(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_a,\cdots,\boldsymbol{x}_b,\cdots,\boldsymbol{x}_m)=0$. From this result, if there is some index $i_a=i_b$ in (3.15), we obtain identity $0=0$. If all indexes are different, we can define the rule $\pi(x)=i_x$, when equality results in the definition of antisymmetric tensor, that is, scalar

$$\begin{aligned}\mathbf{P}(\boldsymbol{x}_{\pi(1)},\cdots,\boldsymbol{x}_{\pi(m)})&=\epsilon_{\pi(1)\cdots\pi(m)}\mathbf{P}(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_m)\\ \mathbf{P}_\pi(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_m)&=\epsilon_{\pi(1)\cdots\pi(m)}\mathbf{P}(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_m)\\ \mathbf{P}_\pi&=\epsilon_{\pi(1)\cdots\pi(m)}\mathbf{P}\,.\end{aligned}$$

Let's verify if (anti)symmetric tensors really define tensor subspaces. Let $\mathbf{X}$ and $\mathbf{Y}$ be symmetric tensors of $\boldsymbol{\mathcal{L}}_{\mathbb{F}}(V^m)$, where $\mathbf{X}=\mathbf{X}_\pi$ and $\mathbf{Y}=\mathbf{Y}_\pi$. Adding these equalities, we obtain $\mathbf{X}+\mathbf{Y}=\mathbf{X}_\pi+\mathbf{Y}_\pi=(\mathbf{X}+\mathbf{Y})_\pi$, which shows that $\mathbf{X}+\mathbf{Y}$ is symmetric. This same process can be used for multiplication by scalar. Thereby, it is thus trivial to verify that a set of symmetric tensors observes the five definer axioms of vector spaces, presented in section 2.1. In the case of antisymmetric tensors $\mathbf{Z}$ and $\boldsymbol{W}$, we have

$$\epsilon_{\pi(1)\cdots\pi(m)}(\mathbf{Z}+\boldsymbol{W})=\mathbf{Z}_\pi+\boldsymbol{W}_\pi=(\mathbf{Z}+\boldsymbol{W})_\pi\,,$$

which shows $\mathbf{Z}+\boldsymbol{W}$ antisymmetric. Thereby, we can state that a set of antisymmetric tensors also defines a tensor space. Now, we'll verify the curious one-dimensionality of alternating tensors by adapting to our case the idea of BACKUS[5], pp. 15-17. Let $\boldsymbol{\mathcal{LA}}_{\mathbb{F}}^{(m)}(V^m)\subset\boldsymbol{\mathcal{L}}_{\mathbb{F}}^{(m)}(V^m)$ be a tensor subspace of alternating tensors, $B=\{\mathbf{u}_1,\cdots,\mathbf{u}_m\}$ a basis of vector space $V_{\mathbb{F}}$ and an alternating tensor $\mathbf{A}_B\in\boldsymbol{\mathcal{LA}}_{\mathbb{F}}^{(m)}(V^m)$ defined by

$$\mathbf{A}_B(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_m)=\sum_{i_1=1}^{m}\cdots\sum_{i_m=1}^{m}\mathbf{f}_{i_1}^B(\boldsymbol{x}_1)\cdots\mathbf{f}_{i_m}^B(\boldsymbol{x}_m)\epsilon_{i_1\cdots i_m}\,,$$

where $\boldsymbol{x}_i\in V_{\mathbb{F}}$. This tensor $\mathbf{A}_B$ is verified to be antisymmetric through the following development:

$$\begin{aligned}\mathbf{A}_{B(a,b)}(\cdots,\boldsymbol{x}_a,\cdots,\boldsymbol{x}_b,\cdots)&=\cdots\sum_{i_a=1}^{m}\cdots\sum_{i_b=1}^{m}\cdots\mathbf{f}_{i_a}^B(\boldsymbol{x}_b)\cdots\mathbf{f}_{i_b}^B(\boldsymbol{x}_a)\cdots\epsilon_{\dots i_a\cdots i_b\cdots}\\ &=\cdots\sum_{i_b=1}^{m}\cdots\sum_{i_a=1}^{m}\cdots\mathbf{f}_{i_b}^B(\boldsymbol{x}_b)\cdots\mathbf{f}_{i_a}^B(\boldsymbol{x}_a)\cdots\epsilon_{\dots i_b\cdots i_a\cdots}\\ &=-\cdots\sum_{i_b=1}^{m}\cdots\sum_{i_a=1}^{m}\cdots\mathbf{f}_{i_b}^B(\boldsymbol{x}_b)\cdots\mathbf{f}_{i_a}^B(\boldsymbol{x}_a)\cdots\epsilon_{\dots i_a\cdots i_b\cdots}\end{aligned}$$

$$= \epsilon_{1\cdots b\cdots a\cdots m} \mathbf{A}_B(\cdots, \boldsymbol{x}_a, \cdots, \boldsymbol{x}_b, \cdots)$$

where terms with indexes i_1 and i_m are omitted and a, b are arbitrary ordinals between 1 and m. Considering an arbitrary tensor $\mathbf{P} \in \mathcal{LA}^{(m)}_{\mathbb{F}}(V^m)$, we know that

$$\begin{aligned}
\mathbf{P}(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_m) &= \sum_{i_1=1}^{m} \cdots \sum_{i_m=1}^{m} \mathbf{f}^B_{i_1}(\boldsymbol{x}_1) \cdots \mathbf{f}^B_{i_m}(\boldsymbol{x}_m) \mathbf{P}(\boldsymbol{u}_{i_1}, \cdots, \boldsymbol{u}_{i_m}) \\
&= \underbrace{\sum_{i_1=1}^{m} \cdots \sum_{i_m=1}^{m} \mathbf{f}^B_{i_1}(\boldsymbol{x}_1) \cdots \mathbf{f}^B_{i_m}(\boldsymbol{x}_m) \epsilon_{i_1 \cdots i_m}}_{\mathbf{A}_B(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_m)} \underbrace{\mathbf{P}(\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m)}_{\alpha}
\end{aligned}$$

from which we conclude $\mathbf{P} = \alpha \mathbf{A}_B$, that is, $\{\mathbf{A}_B\}$ is a basis of one-dimensional $\mathcal{LA}^{(m)}_{\mathbb{F}}(V^m)$ and simultaneously property i is proved. Let's verify now item ii. From the definition of $\mathbf{A}_B$ it is clear that

$$\mathbf{A}_{B(a,b)}(\cdots, \boldsymbol{x}_a, \cdots, \boldsymbol{x}_b, \cdots) = \mathbf{A}_B(\cdots, \boldsymbol{x}_b, \cdots, \boldsymbol{x}_a, \cdots) = -\mathbf{A}_B(\cdots, \boldsymbol{x}_a, \cdots, \boldsymbol{x}_b, \cdots)$$

and from (2.32), we have

$$\mathbf{A}_B(\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m) = \sum_{i_1=1}^{m} \cdots \sum_{i_m=1}^{m} \delta_{1 i_1} \cdots \delta_{m i_m} \epsilon_{i_1 \cdots i_m} = \epsilon_{1 \cdots m} = 1.$$

Now, considering arbitrary vectors $\boldsymbol{w}_i \in V_{\mathbb{F}}$, from (3.15) and (3.16), the following development proves property iii.

$$\begin{aligned}
\mathbf{A}(\boldsymbol{w}_1, \cdots, \boldsymbol{w}_m) &= \sum_{i_1=1}^{m} \cdots \sum_{i_m=1}^{m} \mathbf{f}^B_{i_1}(\boldsymbol{w}_1) \cdots \mathbf{f}^B_{i_m}(\boldsymbol{w}_m) \mathbf{A}(\boldsymbol{u}_{i_1}, \cdots, \boldsymbol{u}_{i_m}) \\
&= \sum_{i_1=1}^{m} \cdots \sum_{i_m=1}^{m} \mathbf{f}^B_{i_1}(\boldsymbol{w}_1) \cdots \mathbf{f}^B_{i_m}(\boldsymbol{w}_m) \epsilon_{i_1 \cdots i_m} \mathbf{A}(\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m) \\
&= \mathbf{A}_B(\boldsymbol{w}_1, \cdots, \boldsymbol{w}_m) \mathbf{A}(\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m) \\
&= [\mathbf{A}(\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m) \mathbf{A}_B](\boldsymbol{w}_1, \cdots, \boldsymbol{w}_m).
\end{aligned}$$

Let's prove property iv. First, if $\mathbf{A}(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m) = 0$ were true for $C = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m\}$ linearly independent (basis), then from property iii, $\mathbf{A} = \mathbf{A}(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m) \mathbf{A}_C = \mathbf{0}$, which is in contradiction to condition $\mathbf{A} \neq \mathbf{0}$. Therefore, $\mathbf{A}(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m) = 0$ implies C linear dependent. We now prove the inverse implication. Let's say that C is linearly dependent because vector $\boldsymbol{v}_1 = \sum_{j=2}^{m} \beta_j \boldsymbol{v}_j$. Then, in scalar $\sum_{j=2}^{m} \beta_j \mathbf{A}(\boldsymbol{v}_j, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m)$ there is always a pair of equal arguments of the antisymmetric tensor $\mathbf{A}$, which makes $\mathbf{A}(\boldsymbol{v}_j, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m) = 0$ as we already know. Now, in order to verify item v, from (1.35) and (3.15), we develop as follows:

$$\begin{aligned}
\mathbf{A}(\mathbf{g}(\boldsymbol{u}_1), \cdots, \mathbf{g}(\boldsymbol{u}_m)) &= \sum_{i_1=1}^{m} \cdots \sum_{i_m=1}^{m} \mathbf{f}^B_{i_1}(\mathbf{g}(\boldsymbol{u}_1)) \cdots \mathbf{f}^B_{i_m}(\mathbf{g}(\boldsymbol{u}_m)) \mathbf{A}(\boldsymbol{u}_{i_1}, \cdots, \boldsymbol{u}_{i_m}) \\
&= \sum_{i_1=1}^{m} \cdots \sum_{i_m=1}^{m} [\mathbf{g}_B]^B_{i_1 1} \cdots [\mathbf{g}_B]^B_{i_m m} \mathbf{A}(\boldsymbol{u}_{i_1}, \cdots, \boldsymbol{u}_{i_m}) \\
&= \sum_{i_1=1}^{m} \cdots \sum_{i_m=1}^{m} \epsilon_{i_1 \cdots i_n} \prod_{k=1}^{m} [\mathbf{g}_B]^B_{i_k k} \mathbf{A}(\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m) \\
&= \det(\mathbf{g}) \mathbf{A}(\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m).
\end{aligned}$$

Item vi is an obvious consequence of item v. □

A tensor $\mathbf{A} \in \mathcal{LA}_{\mathbb{F}}^{(m)}(\mathcal{U}^m)$, where $\mathcal{U}_{\mathbb{F}}$ is a Hilbert space, is called **unimodular** if scalar $|\mathbf{A}(\hat{\mathbf{u}}_1, \cdots, \hat{\mathbf{u}}_m)| = 1$ for every orthonormal basis $\mathrm{B} = \{\hat{\mathbf{u}}_1, \cdots, \hat{\mathbf{u}}_m\}$ of $\mathcal{U}_{\mathbb{F}}$. If $\mathbf{X}$ and $\mathbf{Y}$ are arbitrary unimodular tensors, we have $|\mathbf{X}(\hat{\mathbf{u}}_1, \cdots, \hat{\mathbf{u}}_m)| = |\mathbf{Y}(\hat{\mathbf{u}}_1, \cdots, \hat{\mathbf{u}}_m)|$ and then $\mathbf{X} = \pm\mathbf{Y}$. Thereby, we can affirm that if tensor $\mathbf{A}$ is unimodular, so is $-\mathbf{A}$ and then conclude that a unique pair of unimodultar tensors $(\mathbf{A}, -\mathbf{A})$ in $\mathcal{LA}_{\mathbb{F}}^{(m)}(\mathcal{U}^m)$ is possible. Since this pair exists, if a tensor $\mathbf{A} \in \mathcal{LA}_{\mathbb{F}}^{(m)}(\mathcal{U}^m)$ is unimodular, the double $(\mathcal{U}_{\mathbb{F}}, \mathbf{A})$ is said to be an **oriented vector space**, where arbitrary basis $W = \{\boldsymbol{w}_1, \cdots, \boldsymbol{w}_m\}$ is called **positively** or **negatively oriented** if scalar $\mathbf{A}(\boldsymbol{w}_1, \cdots, \boldsymbol{w}_m)$ is greater than zero or less than zero respectively. Particularly, if W is orthonormal, it is positively or negatively oriented if $\mathbf{A}(\boldsymbol{w}_1, \cdots, \boldsymbol{w}_m)$ is 1 or -1. In this context, for an arbitrary unitary operator $\mathbf{h} \in \mathcal{U}^{+}{}_{\mathbb{F}}(V)$, a basis $\{\mathbf{h}(\hat{\mathbf{u}}_1), \cdots, \mathbf{h}(\hat{\mathbf{u}}_m)\}$ is positively (negatively) oriented if $\{\hat{\mathbf{u}}_1, \cdots, \hat{\mathbf{u}}_m\}$ positively (negatively) oriented; and then we can say that *basis orientation is immune to proper unitary operators*. Now, since basis B is here orthonormal, considering previous property ii, *the alternating tensor* $\mathbf{A}_{\mathrm{B}}$ *defined in* (3.16) *results unimodular and then the oriented space* $(\mathcal{U}_{\mathbb{F}}, \mathbf{A}_{\mathrm{B}})$ *specifies that* B *is a positively oriented orthonormal basis*. Therefore, from (3.15), given an arbitrary alternating tensor $\mathbf{P} \in \mathcal{LA}_{\mathbb{F}}^{(m)}(\mathcal{U}^m)$,

$$\mathbf{f}^{X}_{i_1 \cdots i_m}(\mathbf{P}) = \underbrace{\alpha \mathbf{A}_{\mathrm{B}}}_{\mathbf{P}}(\hat{x}_{i_1}, \cdots, \hat{x}_{i_m}) = \alpha \epsilon_{i_1 \cdots i_m} \mathbf{A}_{\mathrm{B}}(\hat{x}_1, \cdots, \hat{x}_m) = \pm \alpha \epsilon_{i_1 \cdots i_m} \tag{3.19}$$

where basis $\{\hat{x}_1, \cdots, \hat{x}_m\}$ of $(\mathcal{U}_{\mathbb{F}}, \mathbf{A}_{\mathrm{B}})$ is positively or negatively oriented. In other words, the coordinates of an alternating tensor are identical for any positively (negatively) oriented basis. Therefore, we can conclude that $\mathbf{f}^{X}_{i_1 \cdots i_m}(\mathbf{A}_{\mathrm{B}}) = \pm \epsilon_{i_1 \cdots i_m}$.

Proof. First, let's verify if unimodular tensors really exist. Given $\mathrm{B} = \{\hat{\mathbf{u}}_1, \cdots, \hat{\mathbf{u}}_m\}$ and $\mathrm{C} = \{\hat{\boldsymbol{v}}_1, \cdots, \hat{\boldsymbol{v}}_m\}$ arbitrary orthonormal bases of $\mathcal{U}_{\mathbb{F}}$, we know that there is always a unitary operator $\mathbf{g} \in \mathcal{N}_{\mathbb{F}}(\mathcal{U})$ defined by (2.38) where $\mathbf{g}(\hat{\mathbf{u}}_i) = \hat{\boldsymbol{v}}_i$. Considering the basis tensor $\mathbf{A}_{\mathrm{B}}$ defined by (3.16) and expression (1.35),

$$\begin{aligned}
\mathbf{A}_{\mathrm{B}}(\hat{\boldsymbol{v}}_1, \cdots, \hat{\boldsymbol{v}}_m) &= \sum_{i_1=1}^{m} \cdots \sum_{i_m=1}^{m} \mathbf{f}^{\mathrm{B}}_{i_1}(\mathbf{g}(\hat{\mathbf{u}}_1)) \cdots \mathbf{f}^{\mathrm{B}}_{i_m}(\mathbf{g}(\hat{\mathbf{u}}_m)) \epsilon_{i_1 \cdots i_m} \\
&= \sum_{i_1=1}^{m} \cdots \sum_{i_m=1}^{m} \epsilon_{i_1 \cdots i_m} \prod_{k=1}^{m} [\mathbf{g}_{\mathrm{B}}]^{\mathrm{B}}_{i_k k} \\
&= \det(\mathbf{g}) \\
&= \pm 1.
\end{aligned}$$

Suposing an arbitrary unimodular tensor $\mathbf{H} = \alpha \mathbf{A}_{\mathrm{B}}$, we prove that pair $(\mathbf{A}_{\mathrm{B}}, -\mathbf{A}_{\mathrm{B}})$ is unique by

$$\pm 1 = \mathbf{H}(\hat{\mathbf{u}}_1, \cdots, \hat{\mathbf{u}}_m) = \alpha \mathbf{A}_{\mathrm{B}}(\hat{\mathbf{u}}_1, \cdots, \hat{\mathbf{u}}_m) = \pm \alpha.$$

Now we prove that proper unitary operators do not change basis orientation. From equalities (3.18), we obtain for the present case that $\mathbf{A}_{\mathrm{B}}(\mathbf{h}(\hat{\mathbf{u}}_1), \cdots, \mathbf{h}(\hat{\mathbf{u}}_m)) = \mathbf{A}_{\mathrm{B}}(\hat{\mathbf{u}}_1, \cdots, \hat{\mathbf{u}}_m) = 1$. Similarly, $\mathbf{A}_{\mathrm{B}}(\mathbf{h}(\hat{\boldsymbol{v}}_1), \cdots, \mathbf{h}(\hat{\boldsymbol{v}}_m)) = \mathbf{A}_{\mathrm{B}}(\hat{\boldsymbol{v}}_1, \cdots, \hat{\boldsymbol{v}}_m) = -1$. □

Additionally to symmetric and antisymmetric, there is another notable tensor that will enable very important concepts for our study, like isotropy and anti-isotropy for example: I'm referring to the identity tensor, which is defined in tensor spaces of even order. We call $\mathbf{I} \in \mathcal{L}_{\mathbb{F}}(V^{\times p} \times V^{\times p})$ an **identity tensor** of order $2p$ if, considering $\mathbf{x}_i$ and $\mathbf{y}_i$ arbitrary vectors of $(V_{\mathbb{F}})_i$, scalar

$$\mathbf{I}(\mathbf{x}_1, \cdots, \mathbf{x}_p, \mathbf{y}_1, \cdots, \mathbf{y}_p) = \prod_{k=1}^{p} \mathbf{x}_k \cdot \mathbf{y}_k^c. \tag{3.20}$$

From this definition, if $(V_{\mathbb{F}})_i$ is a Hilbert space and $B_i = \{\hat{\mathbf{u}}_1^{(i)}, \cdots, \hat{\mathbf{u}}_{n_i}^{(i)}\}$ any of its orthonormal bases, the following is valid for $\alpha = \mathbf{I}(\mathbf{x}_1, \cdots, \mathbf{x}_p, \mathbf{y}_1, \cdots, \mathbf{y}_p)$:

$$\begin{aligned}
\alpha &= \sum_{i_1=1}^{n_1} \cdots \sum_{i_p=1}^{n_p} \sum_{j_1=1}^{n_1} \cdots \sum_{j_p=1}^{n_p} \prod_{k=1}^{p} \mathbf{f}_{i_k}^{B_k}(\mathbf{x}_k) \mathbf{f}_{j_k}^{B_k}(\mathbf{y}_k) \mathbf{I}(\hat{\mathbf{u}}_{i_1}^{(1)}, \cdots, \hat{\mathbf{u}}_{i_p}^{(p)}, \hat{\mathbf{u}}_{j_1}^{(1)}, \cdots, \hat{\mathbf{u}}_{j_p}^{(p)}) \\
&= \sum_{i_1=1}^{n_1} \cdots \sum_{i_p=1}^{n_p} \sum_{j_1=1}^{n_1} \cdots \sum_{j_p=1}^{n_p} \prod_{k=1}^{p} \mathbf{f}_{i_k}^{B_k}(\mathbf{x}_k) \mathbf{f}_{j_k}^{B_k}(\mathbf{y}_k) \delta_{i_1 j_1} \cdots \delta_{i_p j_p} \\
&= \sum_{i_1=1}^{n_1} \cdots \sum_{i_p=1}^{n_p} \prod_{k=1}^{p} \mathbf{f}_{i_k}^{B_k}(\mathbf{x}_k) \mathbf{f}_{i_k}^{B_k}(\mathbf{y}_k),
\end{aligned}$$

where equalities (2.26), (2.27), (2.30) were considered. Thereby, from the definition of tensor product, we can write that tensor

$$\mathbf{I} = \sum_{i_1=1}^{n_1} \cdots \sum_{i_p=1}^{n_p} \mathbf{f}_{i_1}^{B_1} \otimes \cdots \otimes \mathbf{f}_{i_p}^{B_p} \otimes \mathbf{f}_{i_1}^{B_1} \otimes \cdots \otimes \mathbf{f}_{i_p}^{B_p} \tag{3.21}$$

or, more completely,

$$\mathbf{I} = \sum_{i_1=1}^{n_1} \cdots \sum_{i_p=1}^{n_p} \sum_{j_1=1}^{n_1} \cdots \sum_{j_p=1}^{n_p} \delta_{i_1 j_1} \cdots \delta_{i_p j_p} \mathbf{f}_{i_1}^{B_1} \otimes \cdots \otimes \mathbf{f}_{i_p}^{B_p} \otimes \mathbf{f}_{j_1}^{B_1} \otimes \cdots \otimes \mathbf{f}_{j_p}^{B_p}. \tag{3.22}$$

From this result, we conclude that

$$\mathbf{f}_{i_1 \cdots i_p j_1 \cdots j_p}^{B}(\mathbf{I}) = \prod_{k=1}^{p} \delta_{i_k j_k}, \tag{3.23}$$

where B is a basis of $\mathcal{L}_{\mathbb{F}}(V^{\times p} \times V^{\times p})$ constructed by B_i.

3.3 Tensor Functions

Since it is not our purpose to work with first order tensors that are not covectors, from now on we shall be dealing only with tensors defined by Hilbert spaces, or **Hilbert tensors**. In this scope of study, a function whose domain is constitued by tensors we call it a tensor function: $\boldsymbol{\psi}$ in mapping $\boldsymbol{\psi} : \mathcal{L}_{\mathbb{F}}(X^{\times m}) \mapsto A$ is then a tensor function, regardless of codomain A, which also contributes to classify $\boldsymbol{\psi}$ as a scalar, (co)vector or tensor valued function: a functional whose domain is a tensor space is a scalar valued tensor function. A simple example of a tensor function is the associated linear covector function as described on corollary 6.1 because its domain is constituted by covectors, which are first order tensors. It is important to say that, when conditions described at the beginning of section 2.3 are considered, tensor functions can constitute function spaces, which may be structured to follow the same nomenclature presented in figure 2.3. In this sense, it is possible to consider a function space $V_{\mathbb{F}}^{U}$, where domain $U=\mathcal{L}_{\mathbb{F}}(X^{\times m})$ and codomain $V_{\mathbb{F}}$ are complete spaces, with a metric properly defined. Thereby, similarly to vector function spaces, as presented in section 2.3, tensor function space $V_{\mathbb{F}}^{U}$ can be composed by (anti)linear tensor functions or by multi(anti)linear tensor functions if U is the cartesian product of tensor spaces. A linear tensor function space $\mathcal{CL}_{\mathbb{F}}(Z,Y)$ constituted by continuous tensor functions is considered here also normed and metric, where the domain and codomain involved are Banach tensor spaces, equality $\|\boldsymbol{\psi}\|=\sqrt{\boldsymbol{\psi}\cdot\boldsymbol{\psi}}$ is valid, distance $\rho(\boldsymbol{\psi},\boldsymbol{\phi}):=\|\boldsymbol{\psi}-\boldsymbol{\phi}\|$ and norm

$$\eta(\boldsymbol{\psi}):=\sup\{\|\boldsymbol{\psi}(\mathbf{X})\|/\|\mathbf{X}\|,\forall\mathbf{X}\in\{\mathcal{L}_{\mathbb{F}}(Z)\setminus\{\mathbf{0}\}\}, \tag{3.24}$$

for arbitrary continuous linear tensor functions $\boldsymbol{\psi},\boldsymbol{\phi}\in\mathcal{CL}_{\mathbb{F}}(Z,Y)$. It is important to say that the concepts of continuity and Lipschitz continuity as described by (2.18) and (2.19) respectively are also valid for the present case of tensor functions.

It is still possible to consider $V_{\mathbb{F}}=\mathbb{F}_{\mathbb{F}}$, when the elements of tensor function space $V_{\mathbb{F}}^{U}$ are called tensor functionals. In the terms of Theorem 9, if B is a basis of $\mathcal{L}_{\mathbb{F}}(X^{\times m})$, coordinate functionals $\mathbf{f}^{B}_{j_1\cdots j_m}$, since they are linear according to definition (3.5), are elements of the function space $\mathcal{L}_{\mathbb{F}}(\mathcal{L}_{\mathbb{F}}(X^{\times m}),\mathbb{F})$, represented by $\mathcal{L}^{*}_{\mathbb{F}}(X^{\times m})$ because it is the dual space of $\mathcal{L}_{\mathbb{F}}(X^{\times m})$. Again, similarly to vector function spaces, every linear tensor functional that belongs to this dual space is then called a **dual tensor** of $\mathcal{L}_{\mathbb{F}}(X^{\times m})$. Moreover, the finite set $B^{*}\subset\mathcal{L}^{*}_{\mathbb{F}}(X^{\times m})$ constituted by the coordinate functionals on B is the dual set of this basis because

$$\mathbf{f}^{B}_{j_1\cdots j_m}(\mathbf{f}^{B_1}_{k_1}\otimes\cdots\otimes\mathbf{f}^{B_m}_{k_m})=\mathbf{f}^{B_1}_{k_1}\otimes\cdots\otimes\mathbf{f}^{B_m}_{k_m}(\mathbf{u}^{(1)}_{j_1},\cdots,\mathbf{u}^{(m)}_{j_m})=\prod_{i=1}^{m}\delta_{k_i j_i}\,,$$

and it is also a basis of $\mathcal{L}^{*}_{\mathbb{F}}(X^{\times m})$. Let's see if this last statement is true: given an arbitrary dual tensor $\boldsymbol{\psi}\in\mathcal{L}^{*}_{\mathbb{F}}(X^{\times m})$, development

$$\begin{aligned}\boldsymbol{\psi}(\mathbf{X})&=\sum_{i_1=1}^{n_1}\cdots\sum_{i_m=1}^{n_m}\boldsymbol{\psi}(\mathbf{f}^{B_1}_{i_1}\otimes\cdots\otimes\mathbf{f}^{B_m}_{i_m})\,\mathbf{f}^{B}_{i_1\cdots i_m}(\mathbf{X})\\&=[\sum_{i_1=1}^{n_1}\cdots\sum_{i_m=1}^{n_m}\underbrace{\boldsymbol{\psi}(\mathbf{f}^{B_1}_{i_1}\otimes\cdots\otimes\mathbf{f}^{B_m}_{i_m})}_{\alpha_{i_1\cdots i_m}}\,\mathbf{f}^{B}_{i_1\cdots i_m}](\mathbf{X})\end{aligned} \tag{3.25}$$

shows that the elements of B^{*} span $\mathcal{L}^{*}_{\mathbb{F}}(X^{\times m})$. Moreover, if $\boldsymbol{\psi}$ is a zero dual tensor, then functions $\mathbf{f}^{B}_{i_1\cdots i_m}$ result zero, which proves that B^{*} is a linearly independent set. From this verification, we can affirm that scalars $\alpha_{i_1\cdots i_m}$ are the coordinates of $\boldsymbol{\psi}$ on B^* and that $\dim(\boldsymbol{\mathcal{L}}^*_{\mathbb{F}}(X^{\times m})) = \dim(\boldsymbol{\mathcal{L}}_{\mathbb{F}}(X^{\times m}))$.

This context of tensors defined by Hilbert spaces enables us to always relate a multiantilinear function and a linear tensor function, which is a fundamental feature to ensure uniqueness and existence of an important linear tensor function called contraction. The following theorem[5], borrowed from the realm of Multilinear Algebra and generalized in order to attain our purposes, describes such feature by establishing a one to one correspondence between a multiantilinear function and a linear tensor function. From this biunivocal relationship, the multiantilinear function is said to be "lifted" to a linear tensor function.

Theorem 10 – Lifting of Multiantilinear Functions

Considering function spaces $V_{\mathbb{F}}^{U}$ *and* $\mathcal{L}_{\mathbb{F}}(\boldsymbol{\mathcal{L}}_{\mathbb{F}}(U), V_{\mathbb{F}})$*, where* $V_{\mathbb{F}}$ *and* $U_{\mathbb{F}} = X_{\mathbb{F}}^{\times m}$ *are Hilbert spaces, for every multiantilinear function* $\mathbf{c} \in V_{\mathbb{F}}^{U}$ *there is a unique linear tensor function* $\mathbf{c}^{\otimes} \in \mathcal{L}_{\mathbb{F}}(\boldsymbol{\mathcal{L}}_{\mathbb{F}}(U), V)$ *such that*

$$\mathbf{c}^{\otimes}(\mathbf{x}_1^* \otimes \cdots \otimes \mathbf{x}_m^*) = \mathbf{c}(\mathbf{x}_1, \cdots, \mathbf{x}_m),\ \forall \mathbf{x}_i \in X_i.$$

Proof. The following proof is adapted from Backus[5], pp. 43–45. According to theorem 9, let element $\mathbf{X} \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(U)$ be an arbitrary tensor described by

$$\mathbf{X} = \sum_{i_1=1}^{n_1} \cdots \sum_{i_m=1}^{n_m} \mathbf{f}^{B}_{i_1\cdots i_m}(\mathbf{X}) \mathbf{f}^{B_1}_{i_1} \otimes \cdots \otimes \mathbf{f}^{B_m}_{i_m},$$

where B is a basis of $\boldsymbol{\mathcal{L}}_{\mathbb{F}}(U)$ defined by orthonormal bases $B_i = \{\hat{\mathbf{u}}_1^{(i)}, \cdots, \hat{\mathbf{u}}_{n_i}^{(i)}\}$ of X_i. According to rule (2.29), a function $\mathbf{c}^{\otimes}$ that observes $\mathbf{c}^{\otimes}(\mathbf{x}_1^* \otimes \cdots \otimes \mathbf{x}_m^*) = \mathbf{c}(\mathbf{x}_1, \cdots, \mathbf{x}_m)$ really exists and is uniquely determined by $\mathbf{c}$, as can be verified in

$$\begin{aligned}
\mathbf{c}^{\otimes}(\mathbf{X}) &= \sum_{i_1=1}^{n_1} \cdots \sum_{i_m=1}^{n_m} \mathbf{f}^{B}_{i_1\cdots i_m}(\mathbf{X}) \mathbf{c}^{\otimes}(\mathbf{f}^{B_1}_{i_1} \otimes \cdots \otimes \mathbf{f}^{B_m}_{i_m}) \\
&= \sum_{i_1=1}^{n_1} \cdots \sum_{i_m=1}^{n_m} \mathbf{f}^{B}_{i_1\cdots i_m}(\mathbf{X}) \mathbf{c}(\hat{\mathbf{u}}_{i_1}^{(1)}, \cdots, \hat{\mathbf{u}}_{i_m}^{(m)}),
\end{aligned}$$

because the last equality results a rule for $\mathbf{c}^{\otimes}$ since $\mathbf{X}$ is arbitrary. Now, let's verify if equality $\mathbf{c}^{\otimes}(\mathbf{x}_1^* \otimes \cdots \otimes \mathbf{x}_m^*) = \mathbf{c}(\mathbf{x}_1, \cdots, \mathbf{x}_m)$ is valid. Considering $\mathbf{X} = \mathbf{x}_1^* \otimes \cdots \otimes \mathbf{x}_m^*$, equality (2.22), the properties of complex conjugates, the conjugate property of the inner product, rule (2.29) and rule (3.5), the previous development can proceed as follows:

$$\begin{aligned}
\mathbf{c}^{\otimes}(\mathbf{x}_1^* \otimes \cdots \otimes \mathbf{x}_m^*) &= \sum_{i_1=1}^{n_1} \cdots \sum_{i_m=1}^{n_m} \mathbf{f}^{B}_{i_1\cdots i_m}(\mathbf{X}) \mathbf{c}(\hat{\mathbf{u}}_{i_1}^{(1)}, \cdots, \hat{\mathbf{u}}_{i_m}^{(m)}) \\
&= \sum_{i_1=1}^{n_1} \cdots \sum_{i_m=1}^{n_m} \mathbf{X}(\hat{\mathbf{u}}_{i_1}^{(1)}, \cdots, \hat{\mathbf{u}}_{i_m}^{(m)}) \mathbf{c}(\hat{\mathbf{u}}_{i_1}^{(1)}, \cdots, \hat{\mathbf{u}}_{i_m}^{(m)}) \\
&= \sum_{i_1=1}^{n_1} \cdots \sum_{i_m=1}^{n_m} \prod_{k=1}^{m} \mathbf{x}_k^*(\hat{\mathbf{u}}_{i_k}^{(k)})\, \mathbf{c}(\hat{\mathbf{u}}_{i_1}^{(1)}, \cdots, \hat{\mathbf{u}}_{i_m}^{(m)})
\end{aligned}$$

[5] Adapted from Backus[5] after generalizing a property from Marcus[40], p.14.

$$
\begin{aligned}
&= \sum_{i_1=1}^{n_1}\cdots\sum_{i_m=1}^{n_m}\prod_{k=1}^{m}\overline{\mathbf{f}_{i_k}^{B_k}(x_k)}\,\mathbf{c}(\hat{\mathbf{u}}_{i_1}^{(1)},\cdots,\hat{\mathbf{u}}_{i_m}^{(m)}) \\
&= \mathbf{c}(\sum_{i_1=1}^{n_1}\mathbf{f}_{i_1}^{B_1}(x_1)\hat{\mathbf{u}}_{i_1}^{(1)},\cdots,\sum_{i_m=1}^{n_m}\mathbf{f}_{i_m}^{B_m}(\hat{x}_m)\hat{\mathbf{u}}_{i_m}^{(m)}) \\
&= \mathbf{c}(x_1,\cdots,x_m).
\end{aligned}
$$

□

As we mentioned in the previous paragraph, a special linear tensor function whose existence and uniqueness is assured by lifting multiantilinear functions is called contraction. In simple words, this function relates a tensor of a higher order and a tensor of a lower order, when we say that it "contracts" the order of its argument. Mathematically, in the context of the previous theorem, the tensor function $\mathbf{c}_p^{\otimes} \in \mathcal{L}_{\mathbb{F}}(\boldsymbol{\mathcal{L}}_{\mathbb{F}}(U), V_{\mathbb{F}})$, where $V_{\mathbb{F}} = \boldsymbol{\mathcal{L}}_{\mathbb{F}}(X^{m-p})$, is said to be a **contraction** of order p if it is lifted from the antilinear function in $\mathbf{c}_p : U_{\mathbb{F}} \mapsto \boldsymbol{\mathcal{L}}_{\mathbb{F}}(V)$ in such a way that the $(m-p)$-th order tensor

$$\mathbf{c}_p^{\otimes}(\mathbf{X}) = \mathbf{c}_p(\mathbf{x}_1,\cdots,\mathbf{x}_m), \tag{3.26}$$

where $\mathbf{X}$ is an arbitrary element of $\boldsymbol{\mathcal{L}}_{\mathbb{F}}(U)$ that can be described as a polyadic tensor $\mathbf{x}_1^* \otimes\cdots\otimes \mathbf{x}_m^*$, according to equalities (3.12).

In our study, we deal with two different types of contraction, expressed by two different rules for $\mathbf{c}_p$. The first type is called the trace of tensor, valid for tensor spaces of order greater than one and defined as follows. Still considering the above conditions, if $U = X_1 \times\cdots\times X_r \times\cdots\times X_s \times\cdots\times X_m$ where $X_r = X_s$ and $m > 1$, the tensor $\mathbf{c}_2^{\otimes}(\mathbf{X})$ of order $m-2$ is a contraction called the r,s-**trace** of $\mathbf{X}$, represented by $\mathrm{tr}_{r,s}(\mathbf{X})$, when

$$\mathbf{c}_2(\mathbf{x}_1,\cdots,\mathbf{x}_m) := (\mathbf{x}_s^c\cdot\mathbf{x}_r)\,\mathbf{x}_1^*\otimes\cdots\otimes\mathbf{x}_{r-1}^*\otimes\mathbf{x}_{r+1}^*\otimes\cdots\otimes\mathbf{x}_{s-1}^*\otimes\mathbf{x}_{s+1}^*\otimes\cdots\otimes\mathbf{x}_m^*,$$

where x_r^c belongs to $\overline{X_{r\mathbb{F}}}$. In other words, the lifting of this function $\mathbf{c}_2$ results the r,s-trace function and thus

$$\mathrm{tr}_{r,s}(\mathbf{X}) = (\mathbf{x}_s^c\cdot\mathbf{x}_r)\,\mathbf{x}_1^*\otimes\cdots\otimes\mathbf{x}_{r-1}^*\otimes\mathbf{x}_{r+1}^*\otimes\cdots\otimes\mathbf{x}_{s-1}^*\otimes\mathbf{x}_{s+1}^*\otimes\cdots\otimes\mathbf{x}_m^*. \tag{3.27}$$

It is important to observe that the product $x_r^c\cdot x_s$ is antilinear also on the left term since $\mathbf{c}_2$ has x_r as one of its arguments[6]. Moreover, in the context of Euclidean spaces, from equality (2.35), we can write that $x_s^c\cdot x_r = x_s\cdot x_r$. Now, considering the set $U = X^{\times q}\times X^{\times q}$, if we happen to properly compose q successive r,s-trace functions where $s = r+q$, the resulting scalar

$$\mathrm{tr}(\mathbf{X}) := \underbrace{\mathrm{tr}_{r,s}\circ\cdots\circ\mathrm{tr}_{r,s}}_{q\text{ times}}(\mathbf{X}) = \prod_{k=1}^{q} x_{k+q}^c\cdot x_k \tag{3.28}$$

[6] See property (2.2).

is said to be the **trace** of tensor $\mathbf{X}$. Still in this context, following theorem 9, if a basis B of tensor space $\mathcal{L}_{\mathbb{F}}(X^{\times q}\times X^{\times q})$ of order $m=2q$ is constructed from bases $B_i=\{\hat{\mathbf{u}}_1,\cdots,\hat{\mathbf{u}}_{n_i}\}$ of $(X_i)_{\mathbb{F}}$, $i=1,\cdots,q$, we can calculate the value of the trace of $\mathbf{X}$ from the last equality of the following development, where (2.30) and (3.4) are considered:

$$\begin{aligned}
\operatorname{tr}(\mathbf{X}) &= \sum_{j_1=1}^{n_1}\cdots\sum_{j_m=1}^{n_m}\mathbf{f}^{B}_{j_1\cdots j_m}(\mathbf{X})\operatorname{tr}(\mathbf{f}^{B_1}_{j_1}\otimes\cdots\otimes\mathbf{f}^{B_q}_{j_q}\otimes\mathbf{f}^{B_1}_{j_1}\otimes\cdots\otimes\mathbf{f}^{B_q}_{j_q})\\
&= \sum_{j_1=1}^{n_1}\cdots\sum_{j_m=1}^{n_m}\mathbf{f}^{B}_{j_1\cdots k_m}(\mathbf{X})\prod_{k=1}^{q}(\hat{\mathbf{u}}^{(k+q)}_{j_{k+q}})^{c}\cdot(\hat{\mathbf{u}}^{(k)}_{j_k})_k\\
&= \sum_{j_1=1}^{n_1}\cdots\sum_{j_m=1}^{n_m}\mathbf{f}^{B}_{j_1\cdots k_m}(\mathbf{X})\prod_{k=1}^{q}\delta_{j_{k+q}j_k}\\
&= \sum_{j_1=1}^{n_1}\cdots\sum_{j_q=1}^{n_q}\mathbf{f}^{B}_{j_1\cdots j_q j_1\cdots j_q}(\mathbf{X}).
\end{aligned}\tag{3.29}$$

For the purposes of our study, the contraction that follows will take part in the majority of forthcoming definitions in this chapter. In order to define this important contraction using conditions above, let's consider that $X^{\times m}=A^{\times q}\times W^{\times p}\times W^{\times p}\times B^{\times n}$, where $m=q+2p+n$, that tensor

$$\mathbf{A}:=\mathbf{a}_1^*\otimes\cdots\otimes\mathbf{a}_q^*\otimes\mathbf{x}_1^*\otimes\cdots\otimes\mathbf{x}_p^*$$

belongs to tensor space $\mathcal{L}_{\mathbb{F}}(A^{\times q}\times W^{\times p})$ and tensor

$$\mathbf{B}:=\mathbf{y}_1^*\otimes\cdots\otimes\mathbf{y}_p^*\otimes\mathbf{b}_1^*\otimes\cdots\otimes\mathbf{b}_n^*$$

belongs to $\mathcal{L}_{\mathbb{F}}(W^{\times p}\times B^{\times n})$. When

$$\begin{aligned}
\mathbf{c}_{2p}(\mathbf{a}_1,\cdots,\mathbf{a}_q,\mathbf{x}_1,\cdots,\mathbf{x}_p,\mathbf{y}_1,\cdots,\mathbf{y}_p,\mathbf{b}_1,\cdots,\mathbf{b}_n) &:=\\
&(\prod_{i=1}^{p}\mathbf{y}_i^c\cdot\mathbf{x}_i)\,\mathbf{a}_1^*\otimes\cdots\otimes\mathbf{a}_q^*\otimes\mathbf{b}_1^*\otimes\cdots\otimes\mathbf{b}_n^*,
\end{aligned}$$

we call tensor $\mathbf{c}^{\otimes}_{2p}(\mathbf{A}\otimes\mathbf{B})$ the p-th order **contractive product** of $\mathbf{A}$ and $\mathbf{B}$, represented by $\mathbf{A}\diamond_p\mathbf{B}$. Thereby, we can write in simpler terms that

$$\mathbf{A}\diamond_p\mathbf{B}=(\prod_{i=1}^{p}\mathbf{y}_i^c\cdot\mathbf{x}_i)\mathbf{a}_1^*\otimes\cdots\otimes\mathbf{a}_q^*\otimes\mathbf{b}_1^*\otimes\cdots\otimes\mathbf{b}_n^*.\tag{3.30}$$

For the sake of generality, when manipulating contractive products we consider that $\mathbf{A}\diamond_0\mathbf{B}:=\mathbf{A}\otimes\mathbf{B}$. Moreover, it is trivial to obtain that these type of products are also associative:

$$\mathbf{A}\diamond_p(\mathbf{C}\diamond_r\mathbf{D})=(\mathbf{A}\diamond_p\mathbf{C})\diamond_r\mathbf{D},\tag{3.31}$$

where $\mathbf{C} \in \mathcal{L}_{\mathbb{F}}(W^{\times p} \times C^{\times n} \times Z^{\times r})$ and $\mathbf{D} \in \mathcal{L}_{\mathbb{F}}(Z^{\times r} \times D^{\times s})$. It is also important to say that if both tensors $\mathbf{A}$ and $\mathbf{B}$ belong to $\mathcal{L}_{\mathbb{F}}(W^{\times p})$, scalar $\mathbf{A} \diamond \mathbf{B} := \mathbf{A} \diamond_p \mathbf{B}$ is called simply the contractive product of the tensors in question. Thereby, if $\alpha_{i_1 \cdots i_p}$ and $\beta_{j_1 \cdots j_p}$ are the coordinates of $\mathbf{A}$ and $\mathbf{B}$ on a basis B suitably build from bases $B_i = \{\hat{\mathbf{u}}_1^{(i)}, \cdots, \hat{\mathbf{u}}_{n_i}^{(i)}\}$, we can develop the following:

$$
\begin{aligned}
\mathbf{A} \diamond \mathbf{B} &= \sum_{i_1=1}^{n_1} \cdots \sum_{i_p=1}^{n_p} \sum_{j_1=1}^{n_1} \cdots \sum_{j_p=1}^{n_p} \alpha_{i_1 \cdots i_p} \beta_{j_1 \cdots j_p} (\mathbf{f}_{i_1}^{B_1} \otimes \cdots \otimes \mathbf{f}_{i_p}^{B_p}) \diamond (\mathbf{f}_{j_1}^{B_1} \otimes \cdots \otimes \mathbf{f}_{j_p}^{B_p}) \\
&= \sum_{i_1=1}^{n_1} \cdots \sum_{i_p=1}^{n_p} \sum_{j_1=1}^{n_1} \cdots \sum_{j_p=1}^{n_p} \alpha_{i_1 \cdots i_p} \beta_{j_1 \cdots j_p} \prod_{k=1}^{p} \hat{\mathbf{u}}_{j_k}^{c(k)} \cdot \hat{\mathbf{u}}_{i_k}^{(k)} \\
&= \sum_{i_1=1}^{n_1} \cdots \sum_{i_p=1}^{n_p} \alpha_{i_1 \cdots i_p} \beta_{i_1 \cdots i_p} ,
\end{aligned} \tag{3.32}
$$

where equality (2.30) was considered.

Proof. Let's prove that the r, s trace and the contractive product are indeed linear. From property ii on page 68, verification of equalities $\mathrm{tr}_{r,s}\, \alpha \mathbf{X} = \alpha \mathrm{tr}_{r,s} \mathbf{X}$ and $(\alpha \mathbf{A}) \diamond_p (\beta \mathbf{B}) = \alpha\beta \mathbf{A} \diamond_p \mathbf{B}$ is straightforward (ESTA FRASE ESTÁ ERRADA!!!). Henceforth, in order to shorten expression sizes, without losing generality, we shall consider that a contractive product always results a scalar, that is, the order of the contractive product equals the order of its terms. Considering the above conditions, let's first verify the additive property of linear functions for a first order contractive product, that is, set $X^{\times 2} = W \times W$ and tensors $\mathbf{A}, \mathbf{B}, \boldsymbol{W} \in \mathcal{L}_{\mathbb{F}}(W)$. Since $\mathbf{x}_1^* + \mathbf{y}_1^* = (\mathbf{x}_1 + \mathbf{y}_1)^*$, equalities

$$
(\mathbf{A} + \mathbf{B}) \diamond_1 \boldsymbol{W} = (\mathbf{x}_1 + \mathbf{y}_1)^* \diamond_1 \boldsymbol{w}_1^* = \boldsymbol{w}_1^c \cdot (\mathbf{x}_1 + \mathbf{y}_1) = \boldsymbol{w}_1^c \cdot \mathbf{x}_1 + \boldsymbol{w}_1^c \cdot \mathbf{y}_1 = \mathbf{A} \diamond_1 \boldsymbol{W} + \mathbf{B} \diamond_1 \boldsymbol{W}
$$

assure linearity of first order contractive products. Now, we shall consider a second order contractive product where $\mathbf{A}, \mathbf{B}, \boldsymbol{W} \in \mathcal{L}_{\mathbb{F}}(W^{\times 2})$. From the Riesz-Fréchet Representation, the linearity of first order contractive products and property iv on page 68, we have the following development:

$$
\begin{aligned}
\mathbf{A} \diamond_2 \boldsymbol{W} + \mathbf{B} \diamond_2 \boldsymbol{W} &= (\mathbf{x}_1^* \otimes \mathbf{x}_2^*) \diamond_2 (\boldsymbol{w}_1^* \otimes \boldsymbol{w}_2^*) + (\mathbf{y}_1^* \otimes \mathbf{y}_2^*) \diamond_2 (\boldsymbol{w}_1^* \otimes \boldsymbol{w}_2^*) \\
&= (\boldsymbol{w}_1^c \cdot \mathbf{x}_1)(\boldsymbol{w}_2^c \cdot \mathbf{x}_2) + (\boldsymbol{w}_1^c \cdot \mathbf{y}_1)(\boldsymbol{w}_2^c \cdot \mathbf{y}_2) \\
&= \boldsymbol{w}_2^c \cdot (\mathbf{x}_1 \cdot \boldsymbol{w}_1^c)\mathbf{x}_2 + \boldsymbol{w}_2^c \cdot (\mathbf{y}_1 \cdot \boldsymbol{w}_1^c)\mathbf{y}_2 \\
&= \boldsymbol{w}_2^c \cdot [(\mathbf{x}_1 \cdot \boldsymbol{w}_1^c)\mathbf{x}_2 + (\mathbf{y}_1 \cdot \boldsymbol{w}_1^c)\mathbf{y}_2] \\
&= [(\mathbf{x}_1 \cdot \boldsymbol{w}_1^c)\mathbf{x}_2 + (\mathbf{y}_1 \cdot \boldsymbol{w}_1^c)\mathbf{y}_2]^* \diamond_1 \boldsymbol{w}_2^* \\
&= [(\boldsymbol{w}_1^c \cdot \mathbf{x}_1)\mathbf{x}_2^* + (\boldsymbol{w}_1^c \cdot \mathbf{y}_1)\mathbf{y}_2^*] \diamond_1 \boldsymbol{w}_2^* \\
&= [(\mathbf{x}_2^* \otimes \mathbf{x}_1^*) \diamond_1 \boldsymbol{w}_1^* + (\mathbf{y}_2^* \otimes \mathbf{y}_1^*) \diamond_1 \boldsymbol{w}_1^*] \diamond_1 \boldsymbol{w}_2^* \\
&= \underbrace{[(\mathbf{x}_2^* \otimes \mathbf{x}_1^*) + (\mathbf{y}_2^* \otimes \mathbf{y}_1^*)]}_{\mathbf{u}^* \otimes \boldsymbol{v}^*} \diamond_1 \boldsymbol{w}_1^* \diamond_1 \boldsymbol{w}_2^* \\
&= \boldsymbol{v}^* \otimes \mathbf{u}^* \diamond_2 \boldsymbol{w}_1^* \otimes \boldsymbol{w}_2^* \\
&= [(\mathbf{x}_1^* \otimes \mathbf{x}_2^*) + (\mathbf{y}_1^* \otimes \mathbf{y}_2^*)] \diamond_2 \boldsymbol{w}_1^* \otimes \boldsymbol{w}_2^* \\
&= (\mathbf{A} + \mathbf{B}) \diamond_2 \boldsymbol{W}.
\end{aligned}
$$

By the same strategy of this development, we can find $\boldsymbol{W} \diamond_2 \mathbf{A} + \boldsymbol{W} \diamond_2 \mathbf{B} = \boldsymbol{W} \diamond_2 (\mathbf{A} + \mathbf{B})$. Therefore, it results that second order contractive products are also linear. In the case of arbitrary tensors $\mathbf{A}, \mathbf{B}, \boldsymbol{W} \in \mathcal{L}_{\mathbb{F}}(W^{\times p})$, for any $p > 2$, this procedure can also be applied, leading to linearity for higher order contractive products. Linearity for r, s trace can be verified by this procedure as well. $\square$

From this concept of contractive product, it is now convenient to define a new product that is linear on the first argument and has some form of antilinearity on the second with an obvious purpose of generalizing the inner product. In order to do this, let's consider the same conditions that led to definition (3.30) and state that the tensor

$$\mathbf{A} \odot_{p} \mathbf{B} := \mathbf{A} \diamond_{p} ((\mathbf{y}_1^c)^* \otimes \cdots \otimes (\mathbf{y}_p^c)^* \otimes \mathbf{b}_1^* \otimes \cdots \otimes \mathbf{b}_n^*) \tag{3.33}$$

is called the p-th order **partial inner product** of $\mathbf{A}$ and $\mathbf{B}$, which is antilinear on the first p constituent covectors of the second argument through $\mathbf{y}_i^*$. Thereby, we have

$$\mathbf{A} \odot_{p} \mathbf{B} = (\prod_{i=1}^{p} \mathbf{y}_i \cdot \mathbf{x}_i)\, \mathbf{a}_1^* \otimes \cdots \otimes \mathbf{a}_q^* \otimes \mathbf{b}_1^* \otimes \cdots \otimes \mathbf{b}_n^* . \tag{3.34}$$

For practical reasons, whenever the partial inner product results a covector, that is $q+n=1$, it is convenient to have a vector

$$\mathbf{A} \hat{\odot}_{p} \mathbf{B} := [\Phi^{-1}(\mathbf{A} \odot_{p} \mathbf{B})]^c , \tag{3.35}$$

where Φ is defined in theorem 6 . In the context of Euclidean spaces, it is not unusual to see a first order partial inner product of a dyadic tensor and a covector being used to define the so called "tensor product of vectors"[7]. Although we shall not adopt this rather obscure concept, its frequent use by renowned authors forces us to clarify it in terms of our particular approach as follows. Firstly, since vectors and covectors have a one to one correspondence, these authors use the term "vector" to indistinctly refer to an element of a vector space or to a covector, depending on the context: in $\mathbf{u}(\mathbf{v}) = \mathbf{v} \cdot \mathbf{u}$, for example, $\mathbf{u}$ is a covector on the left-hand side and a vector on the right. Thereby, they get rid of the burden of defining a covector and all of its preliminary concepts. In this sense, our representation of a dyadic tensor $\mathbf{u}^* \otimes \mathbf{v}^*$ is subverted to $\mathbf{u} \otimes \mathbf{v}$ and the partial inner product operator is suppressed from our precise representation $(\mathbf{u}^* \otimes \mathbf{v}^*) \odot_1 \mathbf{x}^* = (\mathbf{x} \cdot \mathbf{v})\mathbf{u}^*$ in order to avoid defining a contraction and its supporting concepts. Finally, they define the "tensor product" $\mathbf{u} \otimes \mathbf{v}$ as a "tensor" which causes $(\mathbf{u} \otimes \mathbf{v})\mathbf{x} = (\mathbf{v} \cdot \mathbf{x})\mathbf{u}$, for an arbitrary $\mathbf{x}$. Again, *this pragmatic approach of considering vectors as first order tensors and all of its consequences will not be adopted nor implicit in this text.*

When the order of a partial inner product equals the order of its terms, the result is a scalar. In this context, considering arbitrary tensors $\mathbf{A}, \mathbf{B} \in \mathcal{L}_{\mathbb{F}}(U^{\times m})$, where tensor $\mathbf{A} = \mathbf{a}_1^* \otimes \cdots \otimes \mathbf{a}_m^*$ and $\mathbf{B} = \mathbf{b}_1^* \otimes \cdots \otimes \mathbf{b}_m^*$, the scalar

$$\mathbf{A} \cdot \mathbf{B} := \mathbf{A} \odot_{m} \mathbf{B} = \prod_{i=1}^{m} \mathbf{b}_i \cdot \mathbf{a}_i \tag{3.36}$$

is called the **inner product of tensors A** and **B** because it observes the axioms of ordinary inner products. If an inner product can be defined on the elements of $\mathcal{L}_{\mathbb{F}}(U^{\times m})$,

[7]See Gurtin[23], p.4; Truesdell & Noll[52], p.15; Ogden[43], p.16.

then this space results an inner product tensor space. From the above definition, the following interesting equalities are valid:

$$\mathbf{A}\cdot\mathbf{B}=\prod_{i=1}^{m}\mathbf{a}_i^*(\mathbf{b}_i)=\mathbf{A}(\mathbf{b}_1,\cdots,\mathbf{b}_m). \tag{3.37}$$

Thereby, if $\mathcal{B}$ is a basis constructed from the bases $\mathcal{B}_i=\{\hat{\mathbf{u}}_1^{(i)},\cdots,\hat{\mathbf{u}}_{n_i}^{(i)}\}$, considering scalars $\alpha_{i_1\cdots i_m}$ and $\beta_{j_1\cdots j_m}$ the coordinates of $\mathbf{A}$ and $\mathbf{B}$ on $\mathcal{B}$ respectively, we can develop the following:

$$\begin{aligned}
\mathbf{A}\cdot\mathbf{B}&=\sum_{i_1=1}^{n_1}\cdots\sum_{i_m=1}^{n_m}\sum_{j_1=1}^{n_1}\cdots\sum_{j_m=1}^{n_m}\alpha_{i_1\cdots i_m}\overline{\beta_{j_1\cdots j_m}}(\mathbf{f}_{i_1}^{\mathcal{B}_1}\otimes\cdots\otimes\mathbf{f}_{i_m}^{\mathcal{B}_m})\cdot(\mathbf{f}_{j_1}^{\mathcal{B}_1}\otimes\cdots\otimes\mathbf{f}_{j_m}^{\mathcal{B}_m})\\
&=\sum_{i_1=1}^{n_1}\cdots\sum_{i_m=1}^{n_m}\sum_{j_1=1}^{n_1}\cdots\sum_{j_m=1}^{n_m}\alpha_{i_1\cdots i_m}\overline{\beta_{j_1\cdots j_m}}\prod_{k=1}^{m}\mathbf{f}_{i_k}^{\mathcal{B}_k}(\hat{\mathbf{u}}_{j_k}^{(k)})\\
&=\sum_{i_1=1}^{n_1}\cdots\sum_{i_m=1}^{n_m}\alpha_{i_1\cdots i_m}\overline{\beta_{i_1\cdots i_m}},
\end{aligned} \tag{3.38}$$

which is a similar result to (2.35). Given arbitrary tensors $\mathbf{C},\mathbf{D}\in\mathcal{L}_{\mathbb{F}}(V^{\times p})$, it is straightforward to see that

$$(\mathbf{A}\otimes\mathbf{C})\cdot(\mathbf{B}\otimes\mathbf{D})=(\mathbf{A}\cdot\mathbf{B})(\mathbf{C}\cdot\mathbf{D}). \tag{3.39}$$

Moreover, if tensor $\mathbf{X}=\mathbf{x}_1^*\otimes\cdots\otimes\mathbf{x}_m^*\otimes\mathbf{y}_1^*\otimes\cdots\otimes\mathbf{y}_m^*$ belongs to $\mathcal{L}_{\mathbb{F}}(U^{\times m}\otimes U^{\times m})$,

$$\mathbf{I}\cdot\mathbf{X}=\mathbf{I}(\mathbf{x}_1,\cdots,\mathbf{x}_m,\mathbf{y}_1,\cdots,\mathbf{y}_m)=\prod_{i=1}^{m}\mathbf{x}_i\cdot\mathbf{y}_i^c=\overline{\mathrm{tr}}\mathbf{X}, \tag{3.40}$$

according to definitions (3.20) and (3.28).

Proof. First we prove that the partial inner product is indeed antilinear on the first p constituent covectors of second argument through the following equalities:

$$\begin{aligned}
\mathbf{A}\odot_p\alpha\mathbf{B}&=\mathbf{A}\odot_p\mathbf{y}_1^*\otimes\cdots\otimes(\overline{\alpha}\mathbf{y}_i)^*\otimes\cdots\otimes\mathbf{y}_p^*\otimes\mathbf{b}_1^*\otimes\cdots\otimes\mathbf{b}_n^*\\
&=\mathbf{A}\diamond_p(\mathbf{y}_1^c)^*\otimes\cdots\otimes((\overline{\alpha}\mathbf{y}_i)^c)^*\otimes\cdots\otimes(\mathbf{y}_p^c)^*\otimes\mathbf{b}_1^*\otimes\cdots\otimes\mathbf{b}_n^*\\
&=\mathbf{A}\diamond_p\overline{\alpha}(\mathbf{y}_1^c)^*\otimes\cdots\otimes(\mathbf{y}_p^c)^*\otimes\mathbf{b}_1^*\otimes\cdots\otimes\mathbf{b}_n^*\\
&=\overline{\alpha}(\mathbf{A}\odot_p\mathbf{B}).
\end{aligned}$$

Now, let's verify if definition (3.36) really observes the axioms of inner products (see p. 35). We do not compromise generality if we assume $\mathbf{A},\mathbf{B}\in\mathcal{L}_{\mathbb{F}}(U^{\times 2})$ in order to shorten expression sizes. Equality $\mathbf{A}\cdot\mathbf{A}=(\mathbf{a}_1\cdot\mathbf{a}_1)(\mathbf{a}_2\cdot\mathbf{a}_2)$ assures nonnegativity for the left-hand side since each inner product on the right-hand side is nonnegative. Now, let's say that $(\mathbf{a}_1\cdot\mathbf{a}_1)(\mathbf{a}_2\cdot\mathbf{a}_2)=0$. If this is true, one of the the inner products on the left-hand side has to be zero. Therefore, according to the axiom of definition of ordinary inner products, vector $\mathbf{a}_1=\mathbf{0}$ or vector $\mathbf{a}_2=\mathbf{0}$ and as a consequence, $\mathbf{A}=\mathbf{0}$, according to property i on page 68. Equalities $\mathbf{A}\cdot\mathbf{B}=(\mathbf{b}_1\cdot\mathbf{a}_1)(\mathbf{b}_2\cdot\mathbf{a}_2)=\overline{(\mathbf{a}_1\cdot\mathbf{b}_1)(\mathbf{a}_2\cdot\mathbf{b}_2)}=\overline{\mathbf{B}\cdot\mathbf{A}}$ show that the inner product of tensors is conjugate symmetric. Finally, according to property ii on page 68, we have equalities

$$(\alpha\mathbf{A})\cdot(\beta\mathbf{B}) = (\overline{\alpha}\mathbf{a}_1)^*\otimes\mathbf{a}_2^*\cdot(\overline{\beta}\mathbf{b}_1)^*\otimes\mathbf{b}_2^* = \\ ((\overline{\beta}\mathbf{b}_1)\cdot(\overline{\alpha}\mathbf{a}_1))(\mathbf{b}_2\cdot\mathbf{a}_2) = \alpha\overline{\beta}(\mathbf{b}_1\cdot\mathbf{a}_1)(\mathbf{b}_2\cdot\mathbf{a}_2) = \alpha\overline{\beta}\mathbf{A}\cdot\mathbf{B}$$

and development

$$\begin{aligned}
\mathbf{A}\cdot\boldsymbol{W}+\mathbf{B}\cdot\boldsymbol{W} &= (\mathbf{a}_1^*\otimes\mathbf{a}_2^*)\cdot(\boldsymbol{w}_1^*\otimes\boldsymbol{w}_2^*)+(\mathbf{b}_1^*\otimes\mathbf{b}_2^*)\cdot(\boldsymbol{w}_1^*\otimes\boldsymbol{w}_2^*)\\
&= (\boldsymbol{w}_1\cdot\mathbf{a}_1)(\boldsymbol{w}_2\cdot\mathbf{a}_2)+(\boldsymbol{w}_1\cdot\mathbf{b}_1)(\boldsymbol{w}_2\cdot\mathbf{b}_2)\\
&= \boldsymbol{w}_1\cdot(\mathbf{a}_2\cdot\boldsymbol{w}_2)\mathbf{a}_1+\boldsymbol{w}_1\cdot(\mathbf{b}_2\cdot\boldsymbol{w}_2)\mathbf{b}_1\\
&= \boldsymbol{w}_1\cdot[(\mathbf{a}_2\cdot\boldsymbol{w}_2)\mathbf{a}_1+(\mathbf{b}_2\cdot\boldsymbol{w}_2)\mathbf{b}_1]\\
&= \boldsymbol{w}_1\cdot[(\mathbf{a}_1^*\otimes\mathbf{a}_2^*)\diamond_1(\boldsymbol{w}_2^c)^*+(\mathbf{b}_1^*\otimes\mathbf{b}_2^*)\diamond_1(\boldsymbol{w}_2^c)^*]\\
&= \boldsymbol{w}_1\cdot\{[(\mathbf{a}_1^*\otimes\mathbf{a}_2^*)+(\mathbf{b}_1^*\otimes\mathbf{b}_2^*)]\diamond_1(\boldsymbol{w}_2^c)^*\}\\
&= (\mathbf{A}+\mathbf{B})\diamond_1(\boldsymbol{w}_2^c)^*\diamond_1(\boldsymbol{w}_1^c)^*\\
&= (\mathbf{A}+\mathbf{B})\diamond_2(\boldsymbol{w}_1^c)^*\otimes(\boldsymbol{w}_2^c)^*\\
&= (\mathbf{A}+\mathbf{B})\cdot\boldsymbol{W},
\end{aligned}$$

whose strategy can also prove that $\boldsymbol{W}\cdot\mathbf{A}+\boldsymbol{W}\cdot\mathbf{B}=\boldsymbol{W}\cdot(\mathbf{A}+\mathbf{B})$. Therefore, the inner product of tensors also observes linearity-antilinearity axioms. □

In this study, a tensor function is treated like an ordinary function, on which the same functional definitions and nomenclature already presented can be applied properly. For instance, it is possible to consider a tensor function that is the composition of other tensor functions and consequently to define an inverse tensor function. Considering a tensor function $\boldsymbol{\psi}\in\mathcal{CL}_{\mathbb{F}}(\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{U}^{\times p}),\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{V}^{\times q}))$, where domain $\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{U}^{\times p})$ and codomain $\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{V}^{\times p})$ are Hilbert tensor spaces, the existence of an inner product for tensors enables us to write, for arbitrary tensors $\mathbf{U}\in\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{U}^{\times p})$ and $\mathbf{V}\in\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{V}^{\times p})$,

$$\boldsymbol{\psi}^\dagger(\mathbf{V})\cdot\mathbf{U}=\mathbf{V}\cdot\boldsymbol{\psi}(\mathbf{U}), \tag{3.41}$$

where function $\boldsymbol{\psi}^\dagger\in\mathcal{CL}_{\mathbb{F}}(\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{V}^{\times q}),\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{U}^{\times p}))$ is called the Hilbert-adjoint of $\boldsymbol{\psi}$, as defined on section 2.3. Similarly to this section, if the fields involved are real, $\boldsymbol{\psi}^\dagger$ is usually called the transpose of $\boldsymbol{\psi}$, represented by $\boldsymbol{\psi}^{\mathrm{T}}$. All the properties of ordinary adjoint functions presented on page 49 are also valid and operator $\boldsymbol{\varphi}\in\mathcal{CL}_{\mathbb{F}}(\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{U}^{\times p}),\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{U}^{\times p}))$ is said to be Hermitian when $\boldsymbol{\varphi}=\boldsymbol{\varphi}^\dagger$ and anti-Hermitian when $\boldsymbol{\varphi}=-\boldsymbol{\varphi}^\dagger$. Moreover, considering real fields, operator $\boldsymbol{\varphi}$ is called symmetric if $\boldsymbol{\varphi}=\boldsymbol{\varphi}^{\mathrm{T}}$ and antisymmetric when $\boldsymbol{\varphi}=-\boldsymbol{\varphi}^{\mathrm{T}}$. As defined previously, if tensor function $\boldsymbol{\psi}$ is an injection whose inverse equals the adjoint, it is called unitary; or orthogonal, in real fields. Similarly to unitary and orthogonal vector operators, the set of all unitary tensor operators on domain $\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{U}^{\times p})$ defines a unitary group of tensors $\mathcal{N}_{\mathbb{F}}(\mathsf{U}^{\times p})\subset\mathcal{G}_{\mathbb{F}}(\mathsf{U}^{\times p})$ on this same domain and if real fields are considered the set of all orthogonal tensor operators on $\boldsymbol{\mathcal{L}}_{\mathbb{R}}(\mathsf{U}^{\times p})$ defines an orthogonal group $\mathcal{O}(\mathsf{U}^{\times p}):=\mathcal{N}_{\mathbb{R}}(\mathsf{U}^{\times p})$. Again, like ordinary unitary operators, unitary tensor operators clearly preserve inner product. An operator $\boldsymbol{\kappa}\in\mathcal{L}_{\mathbb{F}}(\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{U}^{\times p}),\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{U}^{\times p}))$ is called isometric if $\rho[\boldsymbol{\kappa}(\mathbf{U}_1),\boldsymbol{\kappa}(\mathbf{U}_2)]=\rho(\mathbf{U}_1,\mathbf{U}_2)$, where tensors $\mathbf{U}_1$ and $\mathbf{U}_2$ are arbitrary. Thereby, the set of all isometric tensor operators defines an isometry group $\mathfrak{I}_{\mathbb{F}}(\mathsf{U}^{\times p})$. For Euclidean tensor spaces, an isometry group is also an unitary group whose elements preserve inner product; it is also true that adjoint, (anti-)Hermitian and unitary functions equal respectively to transpose, (anti)symmetric and orthogonal functions.

Theorem 11 – Bases and Unitary Tensor Operators

Let $\boldsymbol{\psi} \in \mathcal{N}_{\mathbb{F}}(\mathsf{U}^{\times p})$ be an arbitrary operator and B_i an orthonormal basis of $(\mathsf{U}_i)_{\mathbb{F}}$. If tensor $\mathbf{f}_{i_1}^{\mathsf{B}_1} \otimes \cdots \otimes \mathbf{f}_{i_p}^{\mathsf{B}_p}$ is an arbitrary element of a basis of $\mathcal{L}_{\mathbb{F}}(\mathsf{U}^{\times p})$, then $\boldsymbol{\psi}(\mathbf{f}_{i_1}^{\mathsf{B}_1} \otimes \cdots \otimes \mathbf{f}_{i_p}^{\mathsf{B}_p})$ is also an element of a basis of $\mathcal{L}_{\mathbb{F}}(\mathsf{U}^{\times p})$.

Proof. Let's verify if $\boldsymbol{\psi}(\mathbf{f}_{i_1}^{\mathsf{B}_1} \otimes \cdots \otimes \mathbf{f}_{i_p}^{\mathsf{B}_p})$ really builds a basis of $\mathcal{L}_{\mathbb{F}}(\mathsf{U}^{\times p})$. As $\boldsymbol{\psi}$ preserves inner products,

$$\boldsymbol{\psi}(\mathbf{f}_{i_1}^{\mathsf{B}_1} \otimes \cdots \otimes \mathbf{f}_{i_p}^{\mathsf{B}_p}) \cdot \boldsymbol{\psi}(\mathbf{f}_{j_1}^{\mathsf{B}_1} \otimes \cdots \otimes \mathbf{f}_{j_p}^{\mathsf{B}_p}) = \mathbf{f}_{i_1}^{\mathsf{B}_1} \otimes \cdots \otimes \mathbf{f}_{i_p}^{\mathsf{B}_p} \cdot \mathbf{f}_{j_1}^{\mathsf{B}_1} \otimes \cdots \otimes \mathbf{f}_{j_p}^{\mathsf{B}_p} = \prod_{k=1}^{p} \delta_{i_k j_k}$$

prove that a set C built from all $\boldsymbol{\psi}(\mathbf{f}_{i_1}^{\mathsf{B}_1} \otimes \cdots \otimes \mathbf{f}_{i_p}^{\mathsf{B}_p})$ is linearly independent. We know that operator $\boldsymbol{\psi}$ is a bijection and then an arbitrary tensor $\mathbf{X} \in \mathcal{L}_{\mathbb{F}}(\mathsf{U}^{\times p})$ equals to some $\boldsymbol{\psi}(\mathbf{T})$. Thereby,

$$\mathbf{X} = \boldsymbol{\psi}(\mathbf{T}) = \sum_{i_1=1}^{n_1} \cdots \sum_{i_p=1}^{n_p} \underbrace{\mathbf{f}_{i_1 \cdots i_p}^{\mathsf{B}}(\mathbf{T})}_{\mathbf{f}_{i_1 \cdots i_p}^{\mathsf{C}}(\mathbf{X})} \boldsymbol{\psi}(\mathbf{f}_{i_1}^{\mathsf{B}_1} \otimes \cdots \otimes \mathbf{f}_{i_p}^{\mathsf{B}_p})$$

from which we conclude that C spans $\mathcal{L}_{\mathbb{F}}(\mathsf{U}^{\times p})$. □

Considering the context of previous theorem, for a tensor space $\mathcal{L}_{\mathbb{F}}^{(n)}(\mathsf{U}^p)$, we have $\mathbf{f}_{i_1}^{\mathsf{C}} \otimes \cdots \otimes \mathbf{f}_{i_p}^{\mathsf{C}} := \boldsymbol{\psi}(\mathbf{f}_{i_1}^{\mathsf{B}} \otimes \cdots \otimes \mathbf{f}_{i_p}^{\mathsf{B}})$. As we already know, the pair of bases B and C can be related through a unitary operator $\mathbf{g} \in \mathcal{N}_{\mathbb{F}}(\mathsf{U})$, that is, if $\mathsf{B} = \{\hat{\mathbf{u}}_1, \cdots, \hat{\mathbf{u}}_n\}$ then basis $\mathsf{C} = \{\mathbf{g}(\hat{\mathbf{u}}_1), \cdots, \mathbf{g}(\hat{\mathbf{u}}_n)\}$. Thereby, it is possible to write that

$$\mathbf{g}(\hat{\mathbf{u}}_{i_1})^* \otimes \cdots \otimes \mathbf{g}(\hat{\mathbf{u}}_{i_p})^* := \boldsymbol{\psi}(\hat{\mathbf{u}}_{i_1}^* \otimes \cdots \otimes \hat{\mathbf{u}}_{i_p}^*),$$

and then, for all $\mathbf{x}_i \in \mathsf{U}_{\mathbb{F}}$, we can develop the following:

$$\begin{aligned}
\mathbf{g}(\mathbf{x}_1)^* \otimes \cdots \otimes \mathbf{g}(\mathbf{x}_p)^* &= \sum_{i_1=1}^{n} \cdots \sum_{i_p=1}^{n} \overline{\mathbf{f}_{i_1}^{\mathsf{B}}}(\mathbf{x}_1) \cdots \overline{\mathbf{f}_{i_n}^{\mathsf{B}}}(\mathbf{x}_n)\, \mathbf{g}(\hat{\mathbf{u}}_{i_1})^* \otimes \cdots \otimes \mathbf{g}(\hat{\mathbf{u}}_{i_p})^* \\
&= \sum_{i_1=1}^{n} \cdots \sum_{i_p=1}^{n} \overline{\mathbf{f}_{i_1}^{\mathsf{B}}}(\mathbf{x}_1) \cdots \overline{\mathbf{f}_{i_n}^{\mathsf{B}}}(\mathbf{x}_n)\, \boldsymbol{\psi}(\hat{\mathbf{u}}_{i_1}^* \otimes \cdots \otimes \hat{\mathbf{u}}_{i_p}^*) \\
&= \boldsymbol{\psi}(\mathbf{x}_1^* \otimes \cdots \otimes \mathbf{x}_p^*),
\end{aligned} \tag{3.42}$$

where an arbitrary unitary tensor operator is related to a unitary vector operator.

The concept of tensor function, up to now, has enabled us to arrive at dual and inner product tensor spaces from the definitions of linear tensor functional and contractive product respectively. In comparative terms, we are now in a similar situation to that which preceded the essential Riesz-Fréchet Theorem in section 2.3. Then, an interesting question naturally arises: "Considering a Hilbert tensor space, is it possible for each of its tensors to have a corresponding 'cotensor' just like a vector in Hilbert spaces has a covector?" An affirmative answer to this question would be convenient for our purposes because we could properly replace in suitable cases the multilinearity of tensors by the linearity of the so called cotensors. The remaining text of this section shows that the concept of **cotensor** is indeed feasible.

Theorem 12 – Riesz-Fréchet Representation of Tensors

Let

$$\boldsymbol{\Phi} : \boldsymbol{\mathcal{L}}_{\mathbb{F}}(W^{\times p} \times A^{\times q}) \mapsto \mathcal{L}_{\mathbb{F}}(\boldsymbol{\mathcal{L}}_{\mathbb{F}}(W^{\times p}), \boldsymbol{\mathcal{L}}_{\mathbb{F}}(A^{\times q}))$$

be a mapping where the tensor spaces involved are Hilbert spaces. If for every tensor $\mathbf{T}$ *of* $\boldsymbol{\mathcal{L}}_{\mathbb{F}}(W^{\times p} \times A^{\times q})$*, a* p*-cotensor* $\mathbf{T}^{\circledast}_p := \boldsymbol{\Phi}(\mathbf{T})$ *is described by the rule* $\mathbf{T}^{\circledast}_p(\mathbf{X}) = \mathbf{X} \odot_p \mathbf{T}$ *then* $\boldsymbol{\Phi}$ *is an antilinear bijection.*

Proof. Considering $\mathbf{A}, \mathbf{B} \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(W^{\times p} \times A^{\times q})$, function $\boldsymbol{\Phi}$ is proved antilinear from the following:

$$[\boldsymbol{\Phi}(\alpha\mathbf{A} + \beta\mathbf{B})](\mathbf{X}) = \mathbf{X} \odot_p (\alpha\mathbf{A} + \beta\mathbf{B}) = \overline{\alpha}\mathbf{A}^{\circledast}_p(\mathbf{X}) + \overline{\beta}\mathbf{B}^{*}_p(\mathbf{X}) = \left[\overline{\alpha}\boldsymbol{\Phi}(\mathbf{A}) + \overline{\beta}\boldsymbol{\Phi}(\mathbf{B})\right](\mathbf{X}).$$

If $\boldsymbol{\Phi}$ were not an injection, there would be different p-cotensors $\mathbf{A}^{\circledast}_p$ and $\mathbf{B}^{*}_p$ where $\mathbf{X} \odot_p \mathbf{A} = \mathbf{X} \odot_p \mathbf{B}$. From this supposition, equality $\mathbf{X} \odot_p (\mathbf{A} - \mathbf{B}) = \mathbf{0}$ do not confirm $\mathbf{A}^{\circledast}_p \neq \mathbf{B}^{*}_p$ because it is valid for arbitrary $\mathbf{X}$. In order to prove that $\boldsymbol{\Phi}$ is a surjection, we need to obtain for an arbitrary tensor function $\boldsymbol{\psi} \in \mathcal{L}_{\mathbb{F}}(\boldsymbol{\mathcal{L}}_{\mathbb{F}}(W^{\times p}), \boldsymbol{\mathcal{L}}_{\mathbb{F}}(A^{\times q}))$ a tensor $\mathbf{A} \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(W^{\times p} \times A^{\times q})$ such that $\boldsymbol{\Phi}(\mathbf{A}) = \boldsymbol{\psi}$. Using a strategy that we could call mathematical sagacity, let's consider a tensor

$$\mathbf{A} = \sum_{i_1=1}^{n_1} \cdots \sum_{i_p=1}^{n_p} \mathbf{f}^{B_1}_{i_1} \otimes \cdots \otimes \mathbf{f}^{B_p}_{i_p} \otimes \boldsymbol{\psi}(\mathbf{f}^{B_1}_{i_1} \otimes \cdots \otimes \mathbf{f}^{B_p}_{i_p})$$

where polyadic tensors are defined by orthonormal bases according to theorem 9. Thereby, we can write the following development:

$$\begin{aligned}
[\boldsymbol{\Phi}(\mathbf{A})](\mathbf{X}) &= \mathbf{X} \odot_p \mathbf{A} \\
&= \Big(\sum_{j_1=1}^{n_1} \cdots \sum_{j_p=1}^{n_p} \mathbf{f}^{B}_{j_1\cdots j_p}(\mathbf{X})\mathbf{f}^{B_1}_{j_1} \otimes \cdots \otimes \mathbf{f}^{B_p}_{j_p}\Big) \odot_p \mathbf{A} \\
&= \sum_{j_1=1}^{n_1} \cdots \sum_{j_p=1}^{n_p} \sum_{i_1=1}^{n_1} \cdots \sum_{i_p=1}^{n_p} \mathbf{f}^{B}_{j_1\cdots j_p}(\mathbf{X}) \prod_{k=1}^{p} \underbrace{\hat{\mathbf{u}}^{(k)}_{i_k} \cdot \hat{\mathbf{u}}^{(k)}_{j_k}}_{\delta_{i_k j_k}} \boldsymbol{\psi}(\mathbf{f}^{B_1}_{i_1} \otimes \cdots \otimes \mathbf{f}^{B_p}_{i_p}) \\
&= \sum_{j_1=1}^{n_1} \cdots \sum_{j_p=1}^{n_p} \mathbf{f}^{B}_{j_1\cdots j_p}(\mathbf{X})\boldsymbol{\psi}(\mathbf{f}^{B_1}_{j_1} \otimes \cdots \otimes \mathbf{f}^{B_p}_{j_p}) \\
&= \boldsymbol{\psi}\Big(\sum_{j_1=1}^{n_1} \cdots \sum_{j_p=1}^{n_p} \mathbf{f}^{B}_{j_1\cdots j_p}(\mathbf{X})\mathbf{f}^{B_1}_{j_1} \otimes \cdots \otimes \mathbf{f}^{B_p}_{j_p}\Big) = \boldsymbol{\psi}(\mathbf{X}).
\end{aligned}$$

□

Considering the conditions of the theorem, let's see what is the p-cotensor of identity tensor $\mathbf{I} \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(W^{\times p} \times W^{\times p})$. If equality (3.21) and its definer conditions are considered, development

$$
\begin{aligned}
\mathbf{I}_p^{\circledast}(\mathbf{X}) &= \mathbf{X} \odot_p \mathbf{I} \\
&= \sum_{i_1=1}^{n_1} \cdots \sum_{i_p=1}^{n_p} \prod_{k=1}^{p} \hat{\mathbf{u}}_{i_k} \cdot \mathbf{x}_k \hat{\mathbf{u}}_{i_1}^* \otimes \cdots \otimes \hat{\mathbf{u}}_{i_p}^* \\
&= (\sum_{i_1=1}^{n_1} \mathbf{f}_{i_1}^{B_1}(\mathbf{x}_1)\hat{\mathbf{u}}_{i_1})^* \otimes \cdots \otimes (\sum_{i_p=1}^{n_1} \mathbf{f}_{i_p}^{B_p}(\mathbf{x}_p)\hat{\mathbf{u}}_{i_p})^* \\
&= \mathbf{x}_1^* \otimes \cdots \otimes \mathbf{x}_p^* \\
&= \mathbf{i}(\mathbf{X})
\end{aligned}
\tag{3.43}
$$

shows that the identity operator $\mathbf{i} \in \mathcal{L}_{\mathbb{F}}(\boldsymbol{\mathcal{L}}_{\mathbb{F}}(W^{\times p}), \boldsymbol{\mathcal{L}}_{\mathbb{F}}(W^{\times p}))$ is the p-cotensor of tensor $\mathbf{I}$. Moreover, if scalar field $\mathbb{F} = \mathbb{R}$, then $\mathbf{I}_p^{\circledast}(\mathbf{X}) = \mathbf{X} \odot_p \mathbf{I} = \mathbf{I} \odot_p \mathbf{X} = \mathbf{i}(\mathbf{X})$. Now, considering an arbitrary tensor $\mathbf{Z} \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(A^{\times q} \times Z^{\times r})$, the associative property (3.31) and previous conditions, we can write that

$$
\mathbf{Z}_q^{\circledast} \circ \mathbf{T}_p^{\circledast}(\mathbf{X}) = (\mathbf{X} \odot_p \mathbf{T}) \odot_q \mathbf{Z} = \mathbf{X} \odot_p (\mathbf{T} \odot_q \mathbf{Z}) = (\mathbf{T} \odot_q \mathbf{Z})_p^{\circledast}(\mathbf{X}), \tag{3.44}
$$

from which we conclude that the p-cotensor of $(\mathbf{T} \odot_q \mathbf{Z}) \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(W^{\times p} \times Z^{\times r})$ is the composite function $\mathbf{Z}_q^{\circledast} \circ \mathbf{T}_p^{\circledast}$. In the particular case of $\mathbf{Z} \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(A^{\times q} \times W^{\times p})$ and $\mathbf{T} \odot_q \mathbf{Z} = \mathbf{I}$, we obtain from previous equalities that $\mathbf{Z}_q^{\circledast} \circ \mathbf{T}_p^{\circledast} = \mathbf{I}_p^{\circledast} = \mathbf{i}$, that is, the q-cotensor of $\mathbf{Z}$ is the inverse of the p-cotensor of $\mathbf{T}$, or vice-versa.

Corollary 12.1 – Cotensors

Given $\boldsymbol{\Phi} : \boldsymbol{\mathcal{L}}_{\mathbb{F}}(W^{\times p}) \mapsto \boldsymbol{\mathcal{L}}_{\mathbb{F}}^*(W^{\times p})$*, the rule of tensor functional* $\mathbf{T}^{\circledast} := \mathbf{T}_p^{\circledast}$*, called the cotensor of* $\mathbf{T}$*, results* $\mathbf{T}^{\circledast}(\mathbf{X}) = \mathbf{X} \cdot \mathbf{T}$ *and when* $\mathbf{T}^{\circledast}$ *is continuous, norm* $\|\mathbf{T}^{\circledast}\|_{\boldsymbol{\mathcal{L}}_{\mathbb{F}}^*(W^{\times p})} = \|\mathbf{T}\|_{\boldsymbol{\mathcal{L}}_{\mathbb{F}}(W^{\times p})}$.

Proof. The verification of previous theorem is not compromised by the conditions of this corollary. Now, let's prove the equality of norms. Using definition (3.24) on the present conditions, leaving space representation on norms implicit, we obtain equality $\|\mathbf{A}^{\circledast}\| = \sup\{|\mathbf{X} \cdot \mathbf{A}|/\|\mathbf{X}\|\}$ for all non zero $\mathbf{X}$. If $\mathbf{A}$ is zero, it is evident that $\|\mathbf{A}^{\circledast}\| = \|\mathbf{A}\|$; otherwise, $\mathbf{A}^{\circledast}$ is non zero and we conclude that $\|\mathbf{A}^{\circledast}\| \geqslant |\mathbf{X} \cdot \mathbf{A}|/\|\mathbf{X}\|$. Cauchy-Schwarz Inequality states that $|\mathbf{X} \cdot \mathbf{A}| \leqslant \|\mathbf{X}\| \, \|\mathbf{A}\|$. Subtracting these two previous inequalities, we arrive at $(\|\mathbf{A}^{\circledast}\| - \|\mathbf{A}\|)\|\mathbf{X}\| \geqslant 0$, whose left side can be zero for arbitrary non zero $\mathbf{A}^{\circledast}$, $\mathbf{A}$ and $\mathbf{X}$; therefore, $\|\mathbf{A}^{\circledast}\| = \|\mathbf{A}\|$. □

In the case of $p = 1$ on the previous theorem, the tensors involved are covectors and then, for an arbitary $\boldsymbol{w}^* \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(W)$ and according to equality (3.36), the tensor functional $\boldsymbol{w}^{\circledast} := (\boldsymbol{w}^*)^{\circledast}$ is described by

$$
\boldsymbol{w}^{\circledast}(\boldsymbol{x}^*) = \boldsymbol{x}^* \cdot \boldsymbol{w}^* = \boldsymbol{w} \cdot \boldsymbol{x} = \boldsymbol{x}^*(\boldsymbol{w}). \tag{3.45}
$$

Given the conditions of this last corollary and a function space of tensor functionals $\mathbb{F}_{\mathbb{F}}^{V}$, where the set $V = \boldsymbol{\mathcal{L}}_{\mathbb{F}}(U^{\times m})$, it is convenient to define an antilinear tensor functional $\overline{\mathbf{T}^{\circledast}} \in \mathbb{F}_{\mathbb{F}}^{V}$ described by the rule

$$\overline{\mathbf{T}^{\circledast}}(\mathbf{X}) = \mathbf{T} \cdot \mathbf{X}, \tag{3.46}$$

whose existence and uniqueness are straightforward consequences of the previous theorem. Consdering the conditions of (3.45), where the tensors involved are covectors, scalar

$$\overline{\boldsymbol{w}^{\circledast}}(\boldsymbol{x}^*) = \boldsymbol{w}^*(\boldsymbol{x}). \tag{3.47}$$

From rule (3.46), it is obvious that every tensor $\mathbf{T}$ has not only a correspondent cotensor $\mathbf{T}^{\circledast}$ but also a correspondent **conjugate cotensor** $\overline{\mathbf{T}^{\circledast}}$. If the variable $\mathbf{X}$ is expressed in its polyadic form $\boldsymbol{x}_1^* \otimes \cdots \otimes \boldsymbol{x}_m^*$, the conjugate cotensor $\overline{\mathbf{T}^{\circledast}}$ is convenient because it is linear on each variable $\boldsymbol{x}_i$, just like its correspondent tensor $\mathbf{T}$. In this context, from the above definition and equalities (3.37), we can write that scalar

$$\overline{\mathbf{T}^{\circledast}}(\boldsymbol{x}_1^* \otimes \cdots \otimes \boldsymbol{x}_m^*) = \mathbf{T} \cdot (\boldsymbol{x}_1^* \otimes \cdots \otimes \boldsymbol{x}_m^*) = \mathbf{T}(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_m), \tag{3.48}$$

where $\overline{\mathbf{T}^{\circledast}}$ results a **representative tensor functional** of tensor $\mathbf{T}$ since their values are the same when variables $\boldsymbol{x}_i \in (U_i)_{\mathbb{F}}$ are arranged as above. Note that if the scalar fields involved are real, cotensor $\mathbf{T}^{\circledast}$ is the representative tensor functional of $\mathbf{T}$. Now, it is possible to study tensors through their conjugate cotensors, which substitute them numerically and have only one argument. If we consider the context of expression (3.40), *equalities* $\overline{\mathbf{I}^{\circledast}}(\mathbf{X}) = \mathbf{I} \cdot \mathbf{X} = \overline{\mathrm{tr}}\,\mathbf{X}$ *show that the conjugate trace represents the identity tensor.* Moreover, if the identity tensor described by (3.21) is the argument of a representative function $\overline{(\mathbf{A} \otimes \mathbf{B})^{\circledast}}$ where $\mathbf{A}, \mathbf{B} \in \mathcal{L}_{\mathbb{F}}(V^{\times p})$, we have the following:

$$\begin{aligned}
\overline{(\mathbf{A} \otimes \mathbf{B})^{\circledast}}(\mathbf{I}) &= \sum_{i_1=1}^{n_1} \cdots \sum_{i_p=1}^{n_p} \overline{(\mathbf{A} \otimes \mathbf{B})^{\circledast}}(\mathbf{f}_{i_1}^{B_1} \otimes \cdots \otimes \mathbf{f}_{i_p}^{B_p} \otimes \mathbf{f}_{i_1}^{B_1} \otimes \cdots \otimes \mathbf{f}_{i_p}^{B_p}) \\
&= \sum_{i_1=1}^{n_1} \cdots \sum_{i_p=1}^{n_p} \mathbf{A} \otimes \mathbf{B}(\hat{\mathbf{u}}_{i_1}^{(1)}, \cdots, \hat{\mathbf{u}}_{i_p}^{(p)}, \hat{\mathbf{u}}_{i_1}^{(1)}, \cdots, \hat{\mathbf{u}}_{i_p}^{(p)}) \\
&= \sum_{i_1=1}^{n_1} \cdots \sum_{i_p=1}^{n_p} \mathbf{A}(\hat{\mathbf{u}}_{i_1}^{(1)}, \cdots, \hat{\mathbf{u}}_{i_p}^{(p)}) \mathbf{B}(\hat{\mathbf{u}}_{i_1}^{(1)}, \cdots, \hat{\mathbf{u}}_{i_p}^{(p)}) \\
&= \sum_{i_1=1}^{n_1} \cdots \sum_{i_p=1}^{n_p} \mathbf{f}_{i_1 \cdots i_p}(\mathbf{A}) \mathbf{f}_{i_1 \cdots i_p}(\mathbf{B}) = \mathbf{A} \diamond \mathbf{B},
\end{aligned} \tag{3.49}$$

where equality (3.32) was considered.

In the context of Hilbert spaces, we already know that corollary 6.1 establishes a biunivocal relationship between a linear vector function space and a linear covector function space, that is, if $\mathbf{g}$ is a linear vector function, there is always a linear covector function $\mathbf{g}_{\Phi}$, which is a linear tensor function whose domain is constituted by first order tensors (covectors). But the theorem above establishes a biunivocal relationship between a linear tensor function space and a tensor space, when we conclude that covector function $\mathbf{g}_\Phi$ has its correspondent second order tensor $\mathbf{G}$. Therefore, if tensor function $\mathbf{g}_\Phi = \mathbf{G}_1^{\circledast}$ has a correspondent linear function $\mathbf{g}$ through corollary 6.1 and a correspondent second order tensor $\mathbf{G}$ through theorem above, we can conclude that $\mathbf{g}$ and $\mathbf{G}$ are correspondents. Let's now rewrite this rationale in a more concise form.

Corollary 12.2 – Second Order Tensors and Linear Functions

Considering Hilbert spaces $\mathsf{U}_\mathbb{F}$ and $\mathsf{V}_\mathbb{F}$, a mapping $\boldsymbol{\Upsilon} : \boldsymbol{\mathcal{L}}_\mathbb{F}(\mathsf{U}\times\mathsf{V}) \mapsto \mathcal{L}_\mathbb{F}(\mathsf{U},\mathsf{V})$ is a bijective transformation if its function rule is $\boldsymbol{\Upsilon}(\mathbf{X}) = \Phi^{-1}\circ\mathbf{X}_1^{\circledast}\circ\Phi$, where Φ is defined on theorem 6.

Proof. If function $\boldsymbol{\Upsilon}$ were not an injection, there would be two distinct tensors $\mathbf{T}$ and $\boldsymbol{W}$ where $\Phi^{-1}\circ\mathbf{T}_1^{\circledast}\circ\Phi(\boldsymbol{x}) = \Phi^{-1}\circ\boldsymbol{W}_1^{\circledast}\circ\Phi(\boldsymbol{x})$; but this equality leads to $\mathbf{T}_1^{\circledast}(\boldsymbol{x}^*) = \boldsymbol{W}_1^{\circledast}(\boldsymbol{x}^*)$, which shows that $\mathbf{T}_1^{\circledast} = \boldsymbol{W}_1^{\circledast}$. Therefore, since p-cotensors are biunivocally related to tensors, $\mathbf{T}$ and $\boldsymbol{W}$ cannot be distinct. Function $\boldsymbol{\Upsilon}$ is also a surjection because, through corollary 6.1, for every linear function $\mathbf{g}\in\mathcal{L}_\mathbb{F}(\mathsf{U},\mathsf{V})$ a covector function $\mathbf{g}_\Phi$ can be defined and then, for every tensor function $\mathbf{g}_\Phi$ there exists a unique second order tensor $\mathbf{G}\in\boldsymbol{\mathcal{L}}_\mathbb{F}(\mathsf{U}\times\mathsf{V})$ where 1-cotensor $\mathbf{G}_1^{\circledast} = \Phi\circ\mathbf{g}_\Phi$. □

From this corollary, the linear bijection

$$g := \boldsymbol{\Upsilon}(\mathbf{G}) \tag{3.50}$$

is called the **representative function** of second order tensor $\mathbf{G}\in\boldsymbol{\mathcal{L}}_\mathbb{F}(\mathsf{U}\times\mathsf{V})$. Considering $\mathbf{g}_1^*\otimes\mathbf{g}_2^*$ the dyadic description of $\mathbf{G}$, arbitrary vectors $\mathbf{u}\in\mathsf{U}_\mathbb{F}$ and $\boldsymbol{v}\in\mathsf{V}_\mathbb{F}$, we can write that

$$g(\mathbf{u}) = \Phi^{-1}\circ\mathbf{G}_1^*\circ\Phi(\mathbf{u}) = \Phi^{-1}(\mathbf{u}^*\odot_1\mathbf{G}) = [\mathbf{u}^*\hat{\odot}_1\mathbf{G}]^c = (\mathbf{g}_2^*\otimes\mathbf{g}_1^*)\hat{\odot}_1\mathbf{u}^* = (\mathbf{u}\cdot\mathbf{g}_1)\mathbf{g}_2 \tag{3.51}$$

because Φ^{-1} is antilinear, and from this result scalar

$$\mathbf{G}(\mathbf{u},\boldsymbol{v}) = \mathbf{g}_1^*(\mathbf{u})\mathbf{g}_2^*(\boldsymbol{v}) = \boldsymbol{v}\cdot\overline{\mathbf{g}_1^*(\mathbf{u})}\mathbf{g}_2 = \boldsymbol{v}\cdot\mathbf{g}_1^*(\mathbf{u}^c)\mathbf{g}_2 = \boldsymbol{v}\cdot[(\mathbf{u}^c\cdot\mathbf{g}_1)\mathbf{g}_2] = \boldsymbol{v}\cdot g(\mathbf{u}^c). \tag{3.52}$$

Corollary 12.3 – Symmetric Second Order Tensors and Operators

The set of symmetric operators, subset of $\mathcal{CL}_\mathbb{R}(\mathsf{U},\mathsf{U})$, is biunivocally related to the space of symmetric second order tensors $\boldsymbol{\mathcal{LS}}_\mathbb{R}(\mathsf{U}^2)$.

Proof. The corollary is straightforwardly obtained from (3.52) for real fields and symmetric operators. □

Still considering an arbitrary second order tensor $\mathbf{G}\in\boldsymbol{\mathcal{L}}_\mathbb{F}(\mathsf{U}\times\mathsf{V})$ and its representative function g, we now sacrifice mathematical rigor in favor of mathematical practicality by specifying[8] three tensors in $\boldsymbol{\mathcal{L}}_\mathbb{F}(\mathsf{V}\times\mathsf{U})$:

[8]This is a little notational permissiveness since there are no such things as the Hermitian-adjoint, the inverse or the power of a tensor.

i. $\mathsf{G}^{\dagger} := \boldsymbol{\Upsilon}^{-1}(\boldsymbol{g}^{\dagger})$;

ii. $\mathsf{G}^{-1} := \boldsymbol{\Upsilon}^{-1}(\boldsymbol{g}^{-1})$, if $\boldsymbol{g}$ is invertible;

iii. $\mathsf{G}^{n} := \boldsymbol{\Upsilon}^{-1}(\underbrace{\boldsymbol{g} \circ \cdots \circ \boldsymbol{g}}_{n \text{ times}})$, where $\mathsf{V} = \mathsf{U}$ and $n > 0$.

They belong to $\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{V} \times \mathsf{U})$ as a straightforward consequence of the definitions of their representative functions together with equalities (3.51). From these definitions, we have the following properties:

i. $(\mathbf{a}^* \otimes \mathbf{b}^*)^{\dagger} = \mathbf{b}^* \otimes \mathbf{a}^*$ for arbitrary $\mathbf{a} \in \mathsf{U}_{\mathbb{F}}$ and $\mathbf{b} \in \mathsf{V}_{\mathbb{F}}$;

ii. For arbitrary $\mathbf{u} \in \mathsf{U}_{\mathbb{F}}$, vector

$$\boldsymbol{g}(\mathbf{u}) = [\mathbf{u}^* \widehat{\odot}_1 \mathsf{G}]^c = \mathsf{G}^{\dagger} \widehat{\odot}_1 \mathbf{u}^* ; \tag{3.53}$$

iii. If functions $\boldsymbol{g}_1$ and $\boldsymbol{g}_2$ represent $\mathsf{G}_1 \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{U} \times W)$ and $\mathsf{G}_2 \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(W \times \mathsf{V})$, then $\boldsymbol{g}_2 \circ \boldsymbol{g}_1$ represents $\mathsf{G}_1 \odot_1 \mathsf{G}_2$ or

$$\boldsymbol{g}_2 \circ \boldsymbol{g}_1(\mathbf{u}) = (\mathsf{G}_2^{\dagger} \odot_1 \mathsf{G}_1^{\dagger}) \widehat{\odot}_1 \mathbf{u}^* = (\mathsf{G}_1 \odot_1 \mathsf{G}_2)^{\dagger} \widehat{\odot}_1 \mathbf{u}^* . \tag{3.54}$$

Thereby, for $\mathsf{G} \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{U}^2)$,

$$\mathsf{G}^{n} = \underbrace{\mathsf{G} \odot_1 \cdots \odot_1 \mathsf{G}}_{n \text{ times}} ; \tag{3.55}$$

iv. If a representative function $\boldsymbol{g}$ of tensor G is hermitian then

$$\mathsf{G}^{\dagger} = \mathsf{G} ; \tag{3.56}$$

v. Considering tensors $\mathsf{G}^{-1} \odot_1 \mathsf{G} \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{V}^2)$ and $\mathsf{G} \odot_1 \mathsf{G}^{-1} \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{U}^2)$,

$$\mathsf{G}^{-1} \odot_1 \mathsf{G} = \mathbf{I} \quad \text{and} \quad \mathsf{G} \odot_1 \mathsf{G}^{-1} = \mathbf{I} . \tag{3.57}$$

Proof. From (2.36) and (3.51), if f is the representative function of $\mathbf{a}^* \otimes \mathbf{b}^*$, $\mathbf{u} \in \mathsf{U}_{\mathbb{F}}$ and $\boldsymbol{v} \in \mathsf{V}_{\mathbb{F}}$, the following procedure proves the first item.

$$\begin{gathered} \boldsymbol{f}^{\dagger}(\mathbf{u}) \cdot \boldsymbol{v} = \mathbf{u} \cdot \boldsymbol{f}(\boldsymbol{v}) \\ [\mathbf{u}^* \widehat{\odot}_1 (\mathbf{a}^* \otimes \mathbf{b}^*)^{\dagger}]^c \cdot \boldsymbol{v} = \mathbf{u} \cdot [\boldsymbol{v}^* \widehat{\odot}_1 (\mathbf{a}^* \otimes \mathbf{b}^*)]^c \\ \mathbf{u} \cdot [(\mathbf{a}^* \otimes \mathbf{b}^*)^{\dagger} \widehat{\odot}_1 \boldsymbol{v}^*] = \mathbf{u} \cdot [(\mathbf{b}^* \otimes \mathbf{a}^*) \widehat{\odot}_1 \boldsymbol{v}^*] . \end{gathered}$$

Proof of item ii is trivial. Now, let arbitrary functions $\boldsymbol{g}_1$ and $\boldsymbol{g}_2$ represent $\mathsf{G}_1 \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{U} \times W)$ and $\mathsf{G}_2 \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(W \times \mathsf{V})$. Then, from item ii, development

$$\boldsymbol{g}_2 \circ \boldsymbol{g}_1(\mathbf{u}) = \boldsymbol{g}_2(\mathsf{G}_1^{\dagger} \widehat{\odot}_1 \mathbf{u}^*) = (\mathsf{G}_2^{\dagger} \odot_1 \mathsf{G}_1^{\dagger}) \widehat{\odot}_1 \mathbf{u}^*$$

proves the first equality of (3.54). In order to prove the second equality, we will consider functions $\boldsymbol{h}_1 = \boldsymbol{g}_1^{\dagger}$ and $\boldsymbol{h}_2 = \boldsymbol{g}_2^{\dagger}$ representing tensors $\mathsf{H}_1 = \mathsf{G}_1^{\dagger}$ and $\mathsf{H}_2 = \mathsf{G}_2^{\dagger}$. It is true to say that $\boldsymbol{h}_2 \circ \boldsymbol{h}_1 = (\boldsymbol{h}_1^{\dagger} \circ \boldsymbol{h}_2^{\dagger})^{\dagger}$ represents $(\mathsf{H}_2^{\dagger} \odot_1 \mathsf{H}_1^{\dagger})^{\dagger}$ and then to conclude that, from the definition of $\mathsf{G}^{\dagger}$, function $\boldsymbol{g} = \boldsymbol{h}_1^{\dagger} \circ \boldsymbol{h}_2^{\dagger}$ represents

$\mathbf{G} = \mathbf{H}_2^\dagger \odot_1 \mathbf{H}_1^\dagger$, which proves the second equality of (3.54). Equality (3.55) is a straightforward consequence of (3.54). Verification of item iv is trivial from (3.51). Considering (3.43), if $\boldsymbol{g}$ is the representative function of $\mathbf{G}$, development

$$\begin{aligned} \boldsymbol{\imath}(\boldsymbol{v}) &= \boldsymbol{g} \circ \boldsymbol{g}^{-1}(\boldsymbol{v}) \\ [\boldsymbol{v}^* \hat{\odot}_1 \mathbf{I}]^c &= \boldsymbol{g}([\boldsymbol{v}^* \hat{\odot}_1 \mathbf{G}^{-1}]^c) \\ &= [(\boldsymbol{v}^* \odot_1 \mathbf{G}^{-1}) \hat{\odot}_1 \mathbf{G}]^c \\ &= [\boldsymbol{v}^* \hat{\odot}_1 (\mathbf{G}^{-1} \odot_1 \mathbf{G})]^c \end{aligned}$$

proves the first identity tensor of item v. Verification for $\mathbf{G} \odot_1 \mathbf{G}^{-1}$ is similar. □

Now, from equalities (3.52), we extend the concept of positivity that can be applied to the representative operator $\boldsymbol{t} \in \mathcal{L}_{\mathbb{F}}(\mathsf{U},\mathsf{U})$, described by definition (2.60), to tensor $\mathbf{T} \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{U}^2)$. Thereby, a second order tensor $\mathbf{T}$ is said to be **nonnegative** if

$$\Re(\boldsymbol{t}(\mathbf{y}) \cdot \mathbf{y}) = \Re(\mathbf{y} \cdot \boldsymbol{t}(\mathbf{y})) = \Re(\mathbf{T}(\mathbf{y},\mathbf{y})) \geqslant 0, \forall \mathbf{y} \in \mathsf{U}_{\mathbb{F}}, \tag{3.58}$$

and **positive-definite** if the real value on the left side is never zero. Moreover, we also extend to a second order tensor $\mathbf{T} \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{U}^2)$ the concepts of eigenvalues and eigenvectors of their representative operator $\boldsymbol{t} \in \mathcal{L}_{\mathbb{F}}(\mathsf{U},\mathsf{U})$, that is, the eigenvalues and eigenvectors of $\boldsymbol{t}$ are also of $\mathbf{T}$, by definition.

3.4 Array Representations

Vectors and linear functions are adequately represented by matrices, whose elements are defined from coordinates on given bases, as we already studied in section 2.4. Tensor and linear tensor functions, on the other hand, are entities of arbitrary order and then require for their arithmetic manipulation arrays of arbitrary order, also on given bases. Let's see now how this representation occurs by following the same steps we have followed previously for matrix representations. Considering the context of Theorem 9, an element $\mathbf{T}$ of a tensor space $\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{U}^{\times m})$ and a set B one of its basis, an array $[\mathbf{T}]^{\mathsf{B}}$ of dimension $n_1 \times \cdots \times n_m \times 1$ whose elements $[\mathbf{T}]^{\mathsf{B}}_{i_1 \cdots i_m 1} := \mathbf{f}^{\mathsf{B}}_{i_1 \cdots i_m}(\mathbf{T})$ is called the representative array of $\mathbf{T}$ on B. Linearity of coordinate functionals leads straightforwardly to

$$\alpha[\mathbf{X}]^{\mathsf{B}} + \beta[\mathbf{Y}]^{\mathsf{B}} = [\alpha\mathbf{X} + \beta\mathbf{Y}]^{\mathsf{B}}, \tag{3.59}$$

where $\alpha, \beta \in \mathbb{F}$ and $\mathbf{X}, \mathbf{Y} \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{U}^{\times m})$ are arbitrary. For the case of a linear tensor function $\boldsymbol{\psi} \in \mathcal{L}_{\mathbb{F}}(\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{U}^{\times m}), \boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{V}^{\times p}))$ and a basis C of $\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{V}^{\times p})$, from development

$$\begin{aligned} \boldsymbol{\psi}(\mathbf{X}) &= \sum_{i_1=1}^{n_1} \cdots \sum_{i_p=1}^{n_p} \mathbf{f}^{\mathsf{C}}_{i_1 \cdots i_p}(\boldsymbol{\psi}(\mathbf{X})) \mathbf{f}^{\mathsf{C}_1}_{i_1} \otimes \cdots \otimes \mathbf{f}^{\mathsf{C}_p}_{i_p} \\ &= \sum_{i_1=1}^{n_1} \cdots \sum_{i_p=1}^{n_p} \sum_{j_1=1}^{l_1} \cdots \sum_{j_m=1}^{l_m} \mathbf{f}^{\mathsf{C}}_{i_1 \cdots i_p}(\boldsymbol{\psi}(\mathbf{f}^{\mathsf{B}_1}_{j_1} \otimes \cdots \otimes \mathbf{f}^{\mathsf{B}_m}_{j_m})) \mathbf{f}^{\mathsf{B}}_{j_1 \cdots j_m}(\mathbf{X}) \mathbf{f}^{\mathsf{C}_1}_{i_1} \otimes \cdots \otimes \mathbf{f}^{\mathsf{C}_p}_{i_p} \end{aligned}$$

we define

$$[\boldsymbol{\psi}_B]^C{}_{i_1\cdots i_p j_1\cdots j_m} = \mathbf{f}^C_{i_1\cdots i_p}(\boldsymbol{\psi}(\mathbf{f}^{B_1}_{j_1}\otimes\cdots\otimes\mathbf{f}^{B_m}_{j_m}))\,, \tag{3.60}$$

which leads to

$$[\boldsymbol{\psi}(\mathbf{X})]^C = [\boldsymbol{\psi}_B]^C *_m [\mathbf{X}]^B\,, \tag{3.61}$$

where $[\boldsymbol{\psi}_B]^C$ is a $(p+m)$-th order array. If arbitrary scalars $\alpha,\beta\in\mathbb{F}$ and tensor function $\boldsymbol{\varphi}\in\mathcal{L}_{\mathbb{F}}(\boldsymbol{\mathcal{L}}_{\mathbb{F}}(U^{\times m}),\boldsymbol{\mathcal{L}}_{\mathbb{F}}(V^{\times p}))$ are also considered, from previous equality, it is straightforward to obtain that

$$[[\alpha\boldsymbol{\psi}+\beta\boldsymbol{\varphi}](\mathbf{X})]^C = (\alpha[\boldsymbol{\psi}_B]^C+\beta[\boldsymbol{\varphi}_B]^C) *_m [\mathbf{X}]^B\,. \tag{3.62}$$

Given an element $\boldsymbol{\gamma}$ and a basis D of a tensor space $\mathcal{L}_{\mathbb{F}}(\boldsymbol{\mathcal{L}}_{\mathbb{F}}(V^{\times p}),\boldsymbol{\mathcal{L}}_{\mathbb{F}}(W^{\times q}))$, it is possible to define a composite function $\boldsymbol{\gamma}\circ\boldsymbol{\psi}\in\mathcal{L}_{\mathbb{F}}(\boldsymbol{\mathcal{L}}_{\mathbb{F}}(U^{\times m}),\boldsymbol{\mathcal{L}}_{\mathbb{F}}(W^{\times q}))$ where

$$[\boldsymbol{\gamma}\circ\boldsymbol{\psi}(\mathbf{X})]^D = [\boldsymbol{\gamma}_C]^D *_p [\boldsymbol{\psi}(\mathbf{X})]^C = [\boldsymbol{\gamma}_C]^D *_p [\boldsymbol{\psi}_B]^C *_m [\mathbf{X}]^B\,. \tag{3.63}$$

Since array $[\boldsymbol{\gamma}\circ\boldsymbol{\psi}(\mathbf{X})]^D = [(\boldsymbol{\gamma}\circ\boldsymbol{\psi})_B]^D[\mathbf{X}]^B$, we can state that

$$[(\boldsymbol{\gamma}\circ\boldsymbol{\psi})_B]^D = [\boldsymbol{\gamma}_C]^D *_p [\boldsymbol{\psi}_B]^C\,. \tag{3.64}$$

Under conditions suitable for defining identity tensors, we already know that equality (3.23) is valid for any basis. Therefore, it is valid that

$$[\mathbf{I}]_{i_1\cdots i_m j_1\cdots j_m 1} := \prod_{k=1}^{m}\delta_{i_k j_k}\,, \tag{3.65}$$

from which we conclude that the second order identity tensor is represented by the identity matrix.

In the previous section, equalities (3.32) and (3.38), which are respectively the contractive and inner products of tensors expressed in terms of coordinates, can be rewritten here as follows:

$$\mathbf{A}\diamond\mathbf{B} = [\mathbf{A}]^B : \overline{[\mathbf{B}]^B} \qquad \text{and} \qquad \mathbf{A}\cdot\mathbf{B} = [\mathbf{A}]^B : [\mathbf{B}]^B\,. \tag{3.66}$$

Now, we shall verify if it is possible to establish a relation between a Hilbert-adjoint tensor function and a conjugate transpose array. If $\boldsymbol{\psi}\in\mathcal{L}_{\mathbb{F}}(\boldsymbol{\mathcal{L}}_{\mathbb{F}}(U^{\times m}),\boldsymbol{\mathcal{L}}_{\mathbb{F}}(V^{\times m}))$ is Hilbert-adjoint function, with sets B and D bases of $\boldsymbol{\mathcal{L}}_{\mathbb{F}}(U^{\times m})$ and $\boldsymbol{\mathcal{L}}_{\mathbb{F}}(V^{\times m})$ respectively, then, from the definition of Hilbert-adjoint tensor function,

$$\boldsymbol{\psi}(\mathbf{f}^{B_1}_{i_1}\otimes\cdots\otimes\mathbf{f}^{B_m}_{i_p})\cdot\mathbf{f}^{D_1}_{j_1}\otimes\cdots\otimes\mathbf{f}^{D_m}_{j_m} = \overline{\boldsymbol{\psi}^\dagger(\mathbf{f}^{D_1}_{j_1}\otimes\cdots\otimes\mathbf{f}^{D_m}_{j_m})\cdot\mathbf{f}^{B_1}_{i_1}\otimes\cdots\otimes\mathbf{f}^{B_m}_{i_m}}$$

$$\mathbf{f}^D_{j_1\cdots j_m}[\boldsymbol{\psi}(\mathbf{f}^{B_1}_{i_1}\otimes\cdots\otimes\mathbf{f}^{B_m}_{i_m})] = \overline{\mathbf{f}^B_{i_1\cdots i_m}[\boldsymbol{\psi}^\dagger(\mathbf{f}^{D_1}_{j_1}\otimes\cdots\otimes\mathbf{f}^{D_m}_{j_m})]}$$

$$\overline{\mathbf{f}^D_{i_1\cdots i_m}[\boldsymbol{\psi}(\mathbf{f}^{B_1}_{j_1}\otimes\cdots\otimes\mathbf{f}^{B_m}_{j_m})]^T} = \mathbf{f}^B_{i_1\cdots i_m}[\boldsymbol{\psi}^\dagger(\mathbf{f}^{D_1}_{j_1}\otimes\cdots\otimes\mathbf{f}^{D_m}_{j_m})]$$

$$[\boldsymbol{\psi}_{\mathrm{B}}]^{\mathrm{D}\,\dagger} = [\boldsymbol{\psi}^{\dagger}_{\mathrm{D}}]^{\mathrm{B}}, \tag{3.67}$$

where equalities (3.5) and (3.37) were considered.

In the context of arbitrary second order tensor $\mathbf{G} \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{U} \times \mathsf{V})$ and vector $\mathbf{u} \in \mathsf{U}_{\mathbb{F}}$, we already know from (3.53) that a linear bijection $\boldsymbol{g}$ represents $\mathbf{G}$ in such a way that $\boldsymbol{g}(\mathbf{u}) = \mathbf{G}^{\dagger}\hat{\odot}_1\mathbf{u}^*$. This equality can be rewritten in matrix terms as following:

$$[\boldsymbol{g}_{\mathrm{B}_1}]^{\mathrm{B}_2}[\mathbf{u}]^{\mathrm{B}_1} = [\mathbf{G}]^{\mathrm{B}\,\dagger}[\mathbf{u}]^{\mathrm{B}_1},$$

where basis tensor B is built from vector bases B_1 and B_2 of $\mathsf{U}_{\mathbb{F}}$ and $\mathsf{V}_{\mathbb{F}}$, according to theorem 9. Therefore, since $\mathbf{u}$ is arbitrary, matrix

$$[\mathbf{G}]^{\mathrm{B}} = [\boldsymbol{g}_{\mathrm{B}_1}]^{\mathrm{B}_2\,\dagger} = [\boldsymbol{g}^{\dagger}_{\mathrm{B}_2}]^{\mathrm{B}_1}. \tag{3.68}$$

From these equalities, if $\mathsf{V}_{\mathbb{F}} = \mathsf{U}_{\mathbb{F}}$, that is, tensor $\mathbf{G} \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{U}^2)$, it is possible to obtain that

$$\operatorname{tr}(\mathbf{G}) = \operatorname{tr}([\mathbf{G}]^{\mathrm{B}}) = \operatorname{tr}(\boldsymbol{g}). \tag{3.69}$$

Still in this context, from equality (2.63), if $\boldsymbol{g}$ is hermitian,

$$[\mathbf{G}]^{\mathrm{B}} = [\boldsymbol{g}_{\mathrm{B}_2}]^{\mathrm{B}_1} = [\boldsymbol{g}_{\mathrm{B}_1}]^{\mathrm{B}_2} = [\mathbf{G}]^{\mathrm{B}\,\dagger}. \tag{3.70}$$

Proof. Let's prove equalities (3.69). Considering basis $\mathrm{B}_1 = \mathrm{B}_2 = \{\hat{\mathbf{u}}_1, \cdots, \hat{\mathbf{u}}_n\}$, from (3.29) we can write that equalities $\operatorname{tr}(\mathbf{G}) = \sum_{j=1}^{n} \mathbf{f}^{\mathrm{B}}_{jj}(\mathbf{G}) = \operatorname{tr}([\mathbf{G}]^{\mathrm{B}})$. But since $\boldsymbol{g}$ is a linear operator on $\mathsf{U}_{\mathbb{F}}$, from (2.58) we have $\operatorname{tr}([\mathbf{G}]^{\mathrm{B}}) = \operatorname{tr}(\boldsymbol{g})$. □

It is possible to develop change of coordinates and of basis as we did for vector and linear functions. For the sake of clarity we shall reproduce here both theorem 7 and its corollary 7.1 with minor changes in order to adapt them for the current case.

Theorem 13 – Change of Tensor Coordinates

If $\mathsf{U}^{\times m}(\mathrm{B})_{\mathbb{F}}$ and $\mathsf{U}^{\times m}(\mathrm{C})_{\mathbb{F}}$ are vector spaces constituted by representative arrays of the elements of tensor space $\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{U}^{\times m})$ on its bases B and C respectively, there is one and only one linear bijective transformation $\boldsymbol{\Gamma} : \mathsf{U}^{\times m}(\mathrm{B})_{\mathbb{F}} \mapsto \mathsf{U}^{\times m}(\mathrm{C})_{\mathbb{F}}$, called change of coordinates from B to C, where $\boldsymbol{\Gamma}([\mathbf{X}]^{\mathrm{B}}) = [\mathbf{X}]^{\mathrm{C}}$ for all $\mathbf{X} \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{U}^{\times m})$.

Proof. Adaptation of theorem 7's proof to this case is trivial. □

An arbitrary tensor $\mathbf{X} \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{U}^{\times m})$ can be distinctly represented by arrays on distinct bases B and C according to theorem 9. From these two representations, it is possible to obtain the rule for $\boldsymbol{\Gamma}$, as described in the above theorem 13, by considering $\boldsymbol{\psi} = \boldsymbol{\imath}$ on equality (3.61), which leads to $[\mathbf{X}]^{\mathrm{C}} = [\boldsymbol{\imath}_{\mathrm{B}}]^{\mathrm{C}} *_m [\mathbf{X}]^{\mathrm{B}}$. Thereby, the rule for change of coordinates from B to C is

$$\boldsymbol{\Gamma}(X) = [\boldsymbol{\imath}_{\mathrm{B}}]^{\mathrm{C}} *_m X. \tag{3.71}$$

It is clear from the definition of Hilbert-adjoint tensor functions in (3.41) and from equality (3.67) that $[\mathbf{i}_{\mathsf{B}}]^{\mathsf{C}} = [\mathbf{i}_{\mathsf{C}}]^{\mathsf{B}\dagger}$. Moreover, considering equalities (2.29), (3.5) and (3.60), we can write that the elements

$$[\mathbf{i}_{\mathsf{B}}]^{\mathsf{C}}_{i_1\cdots i_m j_1\cdots j_m} = \mathbf{f}^{\mathsf{B}_1}_{j_1} \otimes\cdots\otimes \mathbf{f}^{\mathsf{B}_m}_{j_m}(\hat{\mathbf{v}}^{(1)}_{i_1},\cdots,\hat{\mathbf{v}}^{(m)}_{i_m}) = \prod_{k=1}^{m} \hat{\mathbf{v}}^{(k)}_{i_k}\cdot\hat{\mathbf{u}}^{(k)}_{j_k}, \tag{3.72}$$

where bases $\mathsf{B}_i = \{\hat{\mathbf{u}}^{(i)}_1,\cdots,\hat{\mathbf{u}}^{(i)}_{n_i}\}$ and $\mathsf{C}_i = \{\hat{\mathbf{v}}^{(i)}_1,\cdots,\hat{\mathbf{v}}^{(i)}_{n_i}\}$. In order to present the change from basis B to basis C let's consider now arbitrary bases $\mathsf{Z}_i = \{\hat{\mathbf{z}}^{(i)}_1,\cdots,\hat{\mathbf{z}}^{(i)}_{n_i}\}$ of $\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{U}^{\times m})$ such that, from previous equalities,

$$\begin{aligned}
\mathbf{f}^{\mathsf{C}_1}_{i_1}\otimes\cdots\otimes\mathbf{f}^{\mathsf{C}_m}_{i_m} &= \sum_{j_1=1}^{n_1}\cdots\sum_{j_p=1}^{n_p}\prod_{k=1}^{m}\hat{\mathbf{u}}^{(k)}_{j_k}\cdot\hat{\mathbf{v}}^{(k)}_{i_k}\,\mathbf{f}^{\mathsf{B}_1}_{j_1}\otimes\cdots\otimes\mathbf{f}^{\mathsf{B}_m}_{j_m}\\
\mathbf{f}^{\mathsf{Z}_1}_{l_1}\otimes\cdots\otimes\mathbf{f}^{\mathsf{Z}_m}_{l_m}(\mathbf{f}^{\mathsf{C}_1}_{i_1}\otimes\cdots\otimes\mathbf{f}^{\mathsf{C}_m}_{i_m}) &= \sum_{j_1=1}^{n_1}\cdots\sum_{j_p=1}^{n_p}\prod_{k=1}^{m}\hat{\mathbf{u}}^{(k)}_{j_k}\cdot\hat{\mathbf{v}}^{(k)}_{i_k}\prod_{k=1}^{m}\hat{\mathbf{z}}^{(k)}_{l_k}\cdot\hat{\mathbf{u}}^{(k)}_{j_k}\\
\underbrace{\prod_{k=1}^{m}\hat{\mathbf{z}}^{(k)}_{l_k}\cdot\hat{\mathbf{v}}^{(k)}_{i_k}}_{C_{l_1\cdots l_m i_1\cdots i_m}} &= \sum_{j_1=1}^{n_1}\cdots\sum_{j_p=1}^{n_p}\underbrace{\prod_{k=1}^{m}\hat{\mathbf{z}}^{(k)}_{l_k}\cdot\hat{\mathbf{u}}^{(k)}_{j_k}}_{B_{l_1\cdots l_m j_1\cdots j_m}}\ \underbrace{\prod_{k=1}^{m}\hat{\mathbf{u}}^{(k)}_{j_k}\cdot\hat{\mathbf{v}}^{(k)}_{i_k}}_{[\mathbf{i}_{\mathsf{C}}]^{\mathsf{B}}_{j_1\cdots j_m i_1\cdots i_m}}\\
C &= B *_m [\mathbf{i}_{\mathsf{C}}]^{\mathsf{B}},
\end{aligned}\tag{3.73}$$

where $[\mathbf{i}_{\mathsf{C}}]^{\mathsf{B}}$ performs a change of basis from B to C when the coordinates of the elements of bases B and C on an arbitrary basis Z are arranged according to arrays B and C defined in the development. Therefore, we can say that the coordinates of tensors are contravariant because if a change of coordinates is enabled by an array $[\mathbf{i}_{\mathsf{B}}]^{\mathsf{C}}$, the correspondent change of basis is performed by $[\mathbf{i}_{\mathsf{C}}]^{\mathsf{B}}$.

Corollary 13.1 – Change of Linear Tensor Operator Coordinates

If $\mathsf{V}^{\times m}(\mathsf{B})_{\mathbb{F}}$ *and* $\mathsf{V}^{\times m}(\mathsf{C})_{\mathbb{F}}$ *are vector spaces constituted by the representative arrays of linear tensor operators that belong to* $\mathcal{L}_{\mathbb{F}}(\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{V}^{\times m}),\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{V}^{\times m}))$*, described respectively on bases* B *and* C *of tensor space* $\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{V}^{\times m})$*, the change of coordinates* $\boldsymbol{\Theta}:\mathsf{V}^{\times m}(\mathsf{B})_{\mathbb{F}}\mapsto\mathsf{V}^{\times m}(\mathsf{C})_{\mathbb{F}}$ *is described by* $\boldsymbol{\Theta}(X) = [\mathbf{i}_{\mathsf{B}}]^{\mathsf{C}} *_m X *_m [\mathbf{i}_{\mathsf{C}}]^{\mathsf{B}}$.

Proof. Considering an arbitrary function $\boldsymbol{\psi}\in\mathcal{L}_{\mathbb{F}}(\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{V}^{\times m}),\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{V}^{\times m}))$ and a tensor $\mathbf{T}\in\boldsymbol{\mathcal{L}}_{\mathbb{F}}(\mathsf{V}^{\times m})$, the last equality of the following development

$$\begin{aligned}
[\boldsymbol{\psi}(\mathbf{T})]^{\mathsf{B}} &= [\boldsymbol{\psi}_{\mathsf{B}}]^{\mathsf{B}} *_m [\mathbf{T}]^{\mathsf{B}}\\
[\mathbf{i}_{\mathsf{C}}]^{\mathsf{B}} *_m [\boldsymbol{\psi}(\mathbf{T})]^{\mathsf{C}} &= [\boldsymbol{\psi}_{\mathsf{B}}]^{\mathsf{B}} *_m [\mathbf{i}_{\mathsf{C}}]^{\mathsf{B}} *_m [\mathbf{T}]^{\mathsf{C}}\\
[\mathbf{i}_{\mathsf{B}}]^{\mathsf{C}} *_m [\mathbf{i}_{\mathsf{C}}]^{\mathsf{B}} *_m [\boldsymbol{\psi}(\mathbf{T})]^{\mathsf{C}} &= [\mathbf{i}_{\mathsf{B}}]^{\mathsf{C}} *_m [\boldsymbol{\psi}_{\mathsf{B}}]^{\mathsf{B}} *_m [\mathbf{i}_{\mathsf{C}}]^{\mathsf{B}} *_m [\mathbf{T}]^{\mathsf{C}}\\
[\boldsymbol{\psi}_{\mathsf{C}}]^{\mathsf{C}} *_m [\mathbf{T}]^{\mathsf{C}} &= [\mathbf{i}_{\mathsf{B}}]^{\mathsf{C}} *_m [\boldsymbol{\psi}_{\mathsf{B}}]^{\mathsf{B}} *_m [\mathbf{i}_{\mathsf{C}}]^{\mathsf{B}} *_m [\mathbf{T}]^{\mathsf{C}}
\end{aligned}$$

enables us to affirm that
$$[\boldsymbol{\psi}_C]^C = [\mathbf{i}_B]^C *_m [\boldsymbol{\psi}_B]^B *_m [\mathbf{i}_C]^B .$$
□

Tensors and linear tensor operators can be represented in a higher aggregate form through the use of matrices. The strategy consists in considering each polyadic tensor of a given basis as an ordinary basis vector of its tensor space, which is actually a vector space of multilinear functionals, as we already know. In general terms, the labeling scheme changes from the basis vectors of $(\mathrm{U}_{\mathbb{F}})_i$ to the basis tensors of $\mathcal{L}_{\mathbb{F}}(\mathrm{U}^{\times m})$. Thereby, it is implicit in labeling an arbitrary basis $\mathbf{B}_i := \{\mathbf{f}_{i_1}^{B_1} \otimes \cdots \otimes \mathbf{f}_{i_m}^{B_m}\}$ of $\mathcal{L}_{\mathbb{F}}(\mathrm{U}^{\times m})$ the definition of a bijective mapping $\nu : (\mathbb{N}^+)^m \mapsto \mathbb{N}^+$, where integers involved are ordinals and $\nu(i_1, \cdots, i_m) = i$. The dimension $p = n_1 n_2 \cdots n_m$ of $\mathcal{L}_{\mathbb{F}}(\mathrm{U}^{\times m})$ clearly leads to $i = 1, \cdots, p$. In this context, we can write that tensor

$$\mathbf{T} = \sum_{i=1}^{p} \mathbf{f}_i^B(\mathbf{T}) \mathbf{B}_i , \tag{3.74}$$

where $\mathbf{f}_i^B(\mathbf{T}) = \mathbf{T}^{\circledast}(\mathbf{B}_i)$, according to (3.5) and (3.48), and then define $[\mathbf{T}]_{i1}^B := \mathbf{f}_i^B(\mathbf{T})$, similarly to the case of vectors. Now, considering C a basis of $\mathcal{L}_{\mathbb{F}}((\mathrm{U}^{\times m})^*)$ whose elements are also labeled by bijection ν, a tensor operator $\boldsymbol{\psi} \in \mathcal{L}_{\mathbb{F}}(\mathcal{L}_{\mathbb{F}}(\mathrm{U}^{\times m}), \mathcal{L}_{\mathbb{F}}(\mathrm{U}^{\times m}))$ can be represented by a matrix on basis B and C through definition $[\boldsymbol{\psi}_B]_{ij}^C = \mathbf{f}_i^C[\boldsymbol{\psi}(\mathbf{B}_j)]$, which results from the same development that led to (2.42). These two fundamental definitions enable matrix representation of tensors and linear tensor operators, that is performed the same way as for the case of vectors and linear functions, already presented in section 2.4. Because of this similarity, we shall not proceed here with these matrix representations.

3.5 Invariance of Tensors

The notion of invariance is of paramount importance in Classical Physics because the fundamental description of mechanical phenomena must be not only qualitatively immune to arbitrary changes of reference frame but also quantitatively indifferent when such changes are of unitary type. From a constitutive viewpoint, materials have their invariance too, when some or all of their inherent mechanical properties present themselves indistinctly on certain directions, that is, the overall mechanical behavior of a portion of matter in study is the same when this portion is subjected to specific motions. A tensor, as an algebraic entity that suitably embodies a physical magnitude, can be used to accordingly describe a mechanical measurement or a constitutive material property, including their eventual invariance features which we referred to. Thereby, it is the purpose of this section to briefly develop in mathematical terms this intuitive concept of invariance applied to tensors.

Before we develop mathematically this notion of tensor invariance, there is a preliminary definition that needs to be addressed. The feature of proper orthogonality – to which anti-isotropic tensors, defined ahead, are invariant – is also valid for tensor operators. As presented in section 2.4, an orthogonal operator is called proper if its determinant is 1, but we still don't know what a determinant of a tensor operator is. In order to present this unusual definition as concisely as possible, we shall sacrifice self-containedness and tacitly accept the approach of Spivak[50][9], adapting to our case his theorem 4-6. Firstly, let's consider $\boldsymbol{P} \in \boldsymbol{\mathcal{L}\mathcal{A}}_{\mathbb{F}}^{(p)}(\mathsf{A}^p)$ an arbitrary non-zero alternating tensor where set $\mathsf{A} = \mathsf{U}^{\times m}$, that is, the arguments of $\boldsymbol{P}$ are p tensors of order m. Let tensor function $\boldsymbol{\psi} \in \mathcal{L}_{\mathbb{F}}(\boldsymbol{\mathcal{L}}_{\mathbb{F}}^{(n)}(\mathsf{U}^{\times m}), \boldsymbol{\mathcal{L}}_{\mathbb{F}}^{(n)}(\mathsf{U}^{\times m}))$ and alternating tensor $\boldsymbol{P}^{\boldsymbol{\psi}} \in \boldsymbol{\mathcal{L}\mathcal{A}}_{\mathbb{F}}^{(p)}(\mathsf{A}^p)$ such that scalar

$$\boldsymbol{P}^{\boldsymbol{\psi}}(\mathbf{X}_1, \cdots, \mathbf{X}_p) := \boldsymbol{P}(\boldsymbol{\psi}(\mathbf{X}_1), \cdots, \boldsymbol{\psi}(\mathbf{X}_p)), \forall \mathbf{X}_i \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}^{(n)}(\mathsf{A}). \tag{3.75}$$

We already know the interesting feature that any alternating tensor space is one-dimensional and thereby the set $\{\boldsymbol{P}\}$ can be a basis of $\boldsymbol{\mathcal{L}\mathcal{A}}_{\mathbb{F}}^{(p)}(\mathsf{A}^p)$, when it is valid to write that $\boldsymbol{P}^{\boldsymbol{\psi}} = \alpha \boldsymbol{P}$, where $\alpha \in \mathbb{F}$ is the coordinate of $\boldsymbol{P}^{\boldsymbol{\psi}}$ on $\{\boldsymbol{P}\}$. If a non-zero arbitrary tensor $\mathbf{K} \in \boldsymbol{\mathcal{L}\mathcal{A}}_{\mathbb{F}}^{(p)}(\mathsf{A}^p)$ is considered, let $\beta \in \mathbb{F}$ be its coordinate on $\{\boldsymbol{P}\}$. Since the previous definition (3.75) is valid for an arbitrary tensor of $\boldsymbol{\mathcal{L}\mathcal{A}}_{\mathbb{F}}^{(p)}(\mathsf{A}^p)$, the following development

$$\begin{aligned}
\mathbf{K}^{\boldsymbol{\psi}}(\mathbf{X}_1, \cdots, \mathbf{X}_p) &= \beta \boldsymbol{P}(\boldsymbol{\psi}(\mathbf{X}_1), \cdots, \boldsymbol{\psi}(\mathbf{X}_p)) \\
&= \beta \boldsymbol{P}^{\boldsymbol{\psi}}(\mathbf{X}_1, \cdots, \mathbf{X}_p) \\
&= \beta \alpha \boldsymbol{P}(\mathbf{X}_1, \cdots, \mathbf{X}_p) \\
&= \alpha \mathbf{K}(\mathbf{X}_1, \cdots, \mathbf{X}_p)
\end{aligned}$$

leads to equality $\mathbf{K}^{\boldsymbol{\psi}} = \alpha \mathbf{K}$. Therefore, it is possible to conclude that the coordinate of any tensor $\mathbf{K}^{\boldsymbol{\psi}}$ is always α, which is then related only to function $\boldsymbol{\psi}$. Because of this exclusivity, scalar α is defined to be the **determinant of operator $\boldsymbol{\psi}$** and then, from definition (3.75), equality

$$\boldsymbol{P}(\boldsymbol{\psi}(\mathbf{X}_1), \cdots, \boldsymbol{\psi}(\mathbf{X}_p)) = \mathrm{Det}(\boldsymbol{\psi}) \boldsymbol{P}(\mathbf{X}_1, \cdots, \mathbf{X}_p), \tag{3.76}$$

is valid for all $\boldsymbol{P} \in \{\boldsymbol{\mathcal{L}\mathcal{A}}_{\mathbb{F}}^{(p)}(\mathsf{A}^p) \setminus \mathbf{0}\}$ and all $\mathbf{X}_i \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}^{(n)}(\mathsf{U}^{\times m})$. From this equality, given a basis $\mathsf{B} = \{\mathbf{B}_1, \cdots, \mathbf{B}_p\}$ of tensor space $\boldsymbol{\mathcal{L}}_{\mathbb{F}}^{(n)}(\mathsf{U}^{\times m})$ where $p = n^m$, if operator $\boldsymbol{\psi}$ is proper unitary (see property vii below), then

$$\boldsymbol{P}(\boldsymbol{\psi}(\mathbf{B}_1), \cdots, \boldsymbol{\psi}(\mathbf{B}_p)) = \boldsymbol{P}(\mathbf{B}_1, \cdots, \mathbf{B}_p),$$

when we say that $\{\boldsymbol{\psi}(\mathbf{B}_1), \cdots, \boldsymbol{\psi}(\mathbf{B}_p)\}$, which is also a basis according to theorem 11, has the same orientation of B. Now, according to theorem 12, in the particular case where

[9] See also Bishop & Goldberg[6], pp.97–100, and Backus[5], pp.17–21.

operator $\boldsymbol{\psi} \in \mathcal{L}_{\mathbb{F}}(\boldsymbol{\mathcal{L}}_{\mathbb{F}}^{(p)}(\mathrm{U}), \boldsymbol{\mathcal{L}}_{\mathbb{F}}^{(p)}(\mathrm{U}))$ is a 1-cotensor of a second order tensor $\mathbf{T} \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}^{(p)}(\mathrm{U}^2)$, that is, $\boldsymbol{\psi} = \mathbf{T}_1^*$, it is possible to state that $\mathrm{Det}(\mathbf{T}) := \mathrm{Det}(\mathbf{T}_1^*)$ and then

$$\boldsymbol{P}(\mathbf{T}_1^*(x_1^*), \cdots, \mathbf{T}_1^*(x_p^*)) = \mathrm{Det}(\mathbf{T})\boldsymbol{P}(x_1^*, \cdots, x_p^*) \tag{3.77}$$

or

$$\mathbf{P}(\boldsymbol{t}(x_1), \cdots, \boldsymbol{t}(x_p)) = \mathrm{Det}(\mathbf{T})\mathbf{P}(x_1, \cdots, x_p)\,, \tag{3.78}$$

from corollary 12.2, where tensor $\mathbf{P} \in \boldsymbol{\mathcal{L}\mathcal{A}}_{\mathbb{F}}^{(p)}(\mathrm{U}^p)$ and operator $\boldsymbol{t} \in \mathcal{L}_{\mathbb{F}}(\mathrm{U}, \mathrm{U})$ is the representative function of $\mathbf{T}$.

Now, considering arbitrary functions $\boldsymbol{\psi}, \boldsymbol{\varphi} \in \mathcal{L}_{\mathbb{F}}(\boldsymbol{\mathcal{L}}_{\mathbb{F}}^{(n)}(\mathrm{U}^{\times m}), \boldsymbol{\mathcal{L}}_{\mathbb{F}}^{(n)}(\mathrm{U}^{\times m}))$ and tensors $\mathbf{T}, \mathbf{T}_1, \mathbf{T}_2 \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}^{(p)}(\mathrm{U}^2)$, the following properties of determinant are valid:

i. $\mathrm{Det}(\boldsymbol{\imath}) = 1$;

ii. $\mathrm{Det}(\boldsymbol{\psi} \circ \boldsymbol{\varphi}) = \mathrm{Det}(\boldsymbol{\psi})\,\mathrm{Det}(\boldsymbol{\varphi})$;

iii. If $\boldsymbol{\psi}$ is a bijection, then $\mathrm{Det}(\boldsymbol{\psi}^{-1}) = (\mathrm{Det}(\boldsymbol{\psi}))^{-1}$;

iv. If $\mathrm{Det}(\boldsymbol{\psi}) \neq 0$, then $\boldsymbol{\psi}$ is a bijection;

v. If B and C are bases of $\boldsymbol{\mathcal{L}}_{\mathbb{F}}^{(n)}(\mathrm{U}^{\times m})$, then

$$\mathrm{Det}(\boldsymbol{\psi}) = \mathrm{Det}([\boldsymbol{\psi}_{\mathsf{B}}]^{\mathsf{C}})\,; \tag{3.79}$$

vi. If $\boldsymbol{t}$ is the representative function of $\mathbf{T}$, then

$$\mathrm{Det}(\mathbf{T}) = \det(\boldsymbol{t}) = \det([\boldsymbol{t}])\,; \tag{3.80}$$

vii. $\mathrm{Det}(\boldsymbol{\psi}^{\dagger}) = \overline{\mathrm{Det}(\boldsymbol{\psi})}$;

viii. If $\boldsymbol{\psi}$ is unitary, then $\mathrm{Det}(\boldsymbol{\psi}) = \pm 1$;

ix. If $\mathbf{T} = \mathbf{T}_1 \odot_1 \mathbf{T}_2$, then $\mathrm{Det}(\mathbf{T}) = \mathrm{Det}(\mathbf{T}_1)\mathrm{Det}(\mathbf{T}_2)$.

Proof. Verification of item i is a straightforward consequence of considering $\boldsymbol{\psi} = \boldsymbol{\imath}$ in (3.76). The second item is proved through the following equalities:

$$\begin{aligned}\mathrm{Det}(\boldsymbol{\psi} \circ \boldsymbol{\varphi})\boldsymbol{P}(\mathbf{X}_1, \cdots, \mathbf{X}_p) &= \boldsymbol{P}(\boldsymbol{\psi}[\boldsymbol{\varphi}(\mathbf{X}_1)], \cdots, \boldsymbol{\psi}[\boldsymbol{\varphi}(\mathbf{X}_p)]) \\ &= \mathrm{Det}(\boldsymbol{\psi})\boldsymbol{P}(\boldsymbol{\varphi}(\mathbf{X}_1), \cdots, \boldsymbol{\varphi}(\mathbf{X}_p)) \\ &= \mathrm{Det}(\boldsymbol{\psi})\mathrm{Det}(\boldsymbol{\varphi})\boldsymbol{P}(\mathbf{X}_1, \cdots, \mathbf{X}_p)\,.\end{aligned}$$

Equalities $1 = \mathrm{Det}(\boldsymbol{\psi}^{-1} \circ \boldsymbol{\psi}) = \mathrm{Det}(\boldsymbol{\psi}^{-1})\,\mathrm{Det}(\boldsymbol{\psi})$ prove the third item. For the item iv, let's consider $\mathrm{Det}(\boldsymbol{\psi}) \neq 0$ and as equality (3.76) is valid for any $\mathbf{X}_i \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}^{(n)}(\mathrm{A})$, let's choose a linearly independent set $\{\mathbf{X}_1, \cdots, \mathbf{X}_p\}$, which causes $\boldsymbol{P}(\mathbf{X}_1, \cdots, \mathbf{X}_p) \neq 0$ since $\boldsymbol{P}$ is antisymmetric. These conditions imply inequality $\boldsymbol{P}(\boldsymbol{\psi}(\mathbf{X}_1), \cdots, \boldsymbol{\psi}(\mathbf{X}_p)) \neq 0$, from which we conclude that $\{\boldsymbol{\psi}(\mathbf{X}_1), \cdots, \boldsymbol{\psi}(\mathbf{X}_p)\}$ is also linearly independent. Therefore, $\boldsymbol{\psi}$ must be an injection and consequently a bijection since it is an unary operator. Considering the idea in the development that leads to (2.42) in order to verify item v, let's say that $\mathbf{B}_l = \mathbf{f}_{l_1}^{B_1} \otimes \cdots \otimes \mathbf{f}_{l_m}^{B_m}$ and $\mathbf{C}_i = \mathbf{f}_{i_1}^{C_1} \otimes \cdots \otimes \mathbf{f}_{i_m}^{C_m}$ are arbitrary elements of bases B and C of $\boldsymbol{\mathcal{L}}_{\mathbb{F}}^{(n)}(\mathrm{A})$.

Thereby, equalities

$$\boldsymbol{\psi}(\mathbf{B}_l) = \sum_{i=1}^{p}\sum_{j=1}^{p} [\boldsymbol{\psi}_B]^{C}{}_{ij}\,\delta_{lj}\,\mathbf{C}_i = \sum_{i=1}^{p}\sum_{k=1}^{p} [\boldsymbol{\psi}_B]^{C}{}_{il}\,[\boldsymbol{\imath}_C]^{B}_{ki}\,\mathbf{B}_k = \sum_{k=1}^{p} [\boldsymbol{\psi}_B]^{B}{}_{kl}\,\mathbf{B}_k$$

assure that the use of only one basis B to represent $\boldsymbol{\psi}$ is valid for the following results. From the left hand side of (3.76) and considering the previous development, we can write

$$\begin{aligned}
\boldsymbol{P}(\boldsymbol{\psi}(\mathbf{B}_1),\cdots,\boldsymbol{\psi}(\mathbf{B}_p)) &= \sum_{i_1=1}^{p}\cdots\sum_{i_p=1}^{p} [\boldsymbol{\psi}_B]^{B}{}_{i_1 1}\cdots[\boldsymbol{\psi}_B]^{B}{}_{i_p p}\,\boldsymbol{P}(\mathbf{B}_{i_1},\cdots,\mathbf{B}_{i_p})\\
&= \sum_{i_1=1}^{p}\cdots\sum_{i_p=1}^{p} [\boldsymbol{\psi}_B]^{B}{}_{i_1 1}\cdots[\boldsymbol{\psi}_B]^{B}{}_{i_p p}\,\epsilon_{i_1\cdots i_p}\,\boldsymbol{P}(\mathbf{B}_1,\cdots,\mathbf{B}_p)\\
&= \det([\boldsymbol{\psi}_B]^{B})\boldsymbol{P}(\mathbf{B}_1,\cdots,\mathbf{B}_p),
\end{aligned}$$

according to (3.15) and (1.35). Thus, from (3.76), we have $\mathrm{Det}(\boldsymbol{\psi}) = \det([\boldsymbol{\psi}_B]^B) = \mathrm{Det}([\boldsymbol{\psi}_B]^B)$. Now, considering item v and comparing (3.17) with (3.78), proof of item vi is trivial. For the item vii verification, let's consider the previous development and equality (3.67), when we have

$$\boldsymbol{\psi}^{\dagger}(\mathbf{B}_l) = \sum_{k=1}^{p} [\boldsymbol{\psi}^{\dagger}{}_B]^{B}{}_{kl}\,\mathbf{B}_k = \sum_{k=1}^{p} [\boldsymbol{\psi}_B]^{B\dagger}_{kl}\,\mathbf{B}_k = \sum_{k=1}^{p} \overline{[\boldsymbol{\psi}_B]^{B}{}_{lk}}\,\mathbf{B}_k$$

and consequently

$$\begin{aligned}
\boldsymbol{P}(\boldsymbol{\psi}^{\dagger}(\mathbf{B}_1),\cdots,\boldsymbol{\psi}^{\dagger}(\mathbf{B}_p)) &= \sum_{i_1=1}^{p}\cdots\sum_{i_p=1}^{p} \overline{[\boldsymbol{\psi}_B]^{B}{}_{1i_1}\cdots[\boldsymbol{\psi}_B]^{B}{}_{pi_p}}\,\epsilon_{i_1\cdots i_p}\,\boldsymbol{P}(\mathbf{B}_1,\cdots,\mathbf{B}_p)\\
&= \overline{\det([\boldsymbol{\psi}_B]^{B})}\boldsymbol{P}(\mathbf{B}_1,\cdots,\mathbf{B}_p)\\
&= \overline{\mathrm{Det}(\boldsymbol{\psi})}\boldsymbol{P}(\mathbf{B}_1,\cdots,\mathbf{B}_p),
\end{aligned}$$

since an equality similar to (1.35) can be obtained by considering $i_k = k$ on (1.34). Now, equalities $1 = \mathrm{Det}(\boldsymbol{\imath}) = \mathrm{Det}(\boldsymbol{\psi}\circ\boldsymbol{\psi}^{\dagger}) = \mathrm{Det}(\boldsymbol{\psi})\mathrm{Det}(\boldsymbol{\psi}^{\dagger}) = |\mathrm{Det}(\boldsymbol{\psi})|^2$ prove item viii. Property ix is a straightforward consequence of property ii and equality (3.54). □

Theorem 14 – Unitary Tensor Operators and Alternating Tensors

Considering an arbitrary alternating tensor $\mathbf{A}\in\mathcal{LA}_{\mathbb{F}}^{(n)}(\mathrm{U}^n)$ *and an arbitrary unitary tensor operator* $\boldsymbol{\psi}\in\mathcal{N}_{\mathbb{F}}(\mathrm{U}^n)$*, it is valid that* $\mathbf{A} = \mathrm{Det}(\boldsymbol{\psi})\boldsymbol{\psi}(\mathbf{A})$.

Proof. From (3.42), we know that a unitary tensor operator $\boldsymbol{\psi}\in\mathcal{N}_{\mathbb{F}}(\mathrm{U}^n)$ is always related to a unitary operator $\mathbf{g}\in\mathcal{N}_{\mathbb{F}}(\mathrm{U})$. As unitary operators preserve inner product, from (3.15), (3.37) and (3.76),

$$\begin{aligned}
\mathbf{A}(\mathbf{g}(\mathbf{x}_1),\cdots,\mathbf{g}(\mathbf{x}_n)) &= \det(\mathbf{g})\mathbf{A}(\mathbf{x}_1,\cdots,\mathbf{x}_n)\\
\mathbf{A}\cdot(\mathbf{g}(\mathbf{x}_1)^*\otimes\cdots\otimes\mathbf{g}(\mathbf{x}_n)^*) &= \det(\mathbf{g})\mathbf{A}\cdot(\mathbf{x}_1^*\otimes\cdots\otimes\mathbf{x}_n^*)\\
\mathbf{A}\cdot\boldsymbol{\psi}(\mathbf{x}_1^*\otimes\cdots\otimes\mathbf{x}_n^*) &= \det(\mathbf{g})\mathbf{A}\cdot(\mathbf{x}_1^*\otimes\cdots\otimes\mathbf{x}_n^*)\\
\boldsymbol{\psi}\circ\boldsymbol{\psi}^{-1}(\mathbf{A})\cdot\boldsymbol{\psi}(\mathbf{x}_1^*\otimes\cdots\otimes\mathbf{x}_n^*) &= \det(\mathbf{g})\mathbf{A}\cdot(\mathbf{x}_1^*\otimes\cdots\otimes\mathbf{x}_n^*)\\
\boldsymbol{\psi}^{-1}(\mathbf{A})\cdot(\mathbf{x}_1^*\otimes\cdots\otimes\mathbf{x}_n^*) &= \det(\mathbf{g})\mathbf{A}\cdot(\mathbf{x}_1^*\otimes\cdots\otimes\mathbf{x}_n^*)\\
\mathbf{A} &= \det(\mathbf{g})\boldsymbol{\psi}(\mathbf{A}).
\end{aligned}$$

Now, given a unimodular tensor $\mathbf{A}_B$, where B is an orthonormal basis of $\mathrm{U}_{\mathbb{F}}$, we know that tensor $\mathbf{A}_B = \det(\mathbf{g})\boldsymbol{\psi}(\mathbf{A}_B)$. As $\mathbf{A}_B$ is a basis, so is $\mathbf{A}_C := \boldsymbol{\psi}(\mathbf{A}_B)$, according to theorem 11, and then if $\det(\mathbf{g}) = 1$, these bases obviously have the same orientation since $\mathbf{A}_B = \mathbf{A}_C$, that is, $\mathrm{Det}(\boldsymbol{\psi}) = 1$. Conversely, if $\det(\mathbf{g}) = -1$, they don't have the same orientation or $\mathrm{Det}(\boldsymbol{\psi}) = -1$. Therefore, $\det(\mathbf{g}) = \mathrm{Det}(\boldsymbol{\psi})$. □

Now that we know the determinant of linear tensor operators and its properties, let $\mathcal{L}_{\mathbb{F}}^{(n)}(U^{\times m})$ be a Hilbert tensor space and function $\boldsymbol{\varphi}$ be an arbitrary element of a group $\mathcal{G}_{\mathbb{F}}(U^{\times m})$ of invertible continuous linear tensor operators whose domain is $\mathcal{L}_{\mathbb{F}}^{(n)}(U^{\times m})$. A tensor $\mathbf{T} \in \mathcal{L}_{\mathbb{F}}^{(n)}(U^{\times m})$ is said to be **invariant** on group $\mathcal{G}_{\mathbb{F}}(U^{\times m})$ if $\boldsymbol{\varphi}(\mathbf{T}) = \mathbf{T}$. This equality and the linearity of $\boldsymbol{\varphi}$ enable us to affirm that a set V constituted by all tensors of $\mathcal{L}_{\mathbb{F}}^{(n)}(U^{\times m})$ that are invariant on group $\mathcal{G}_{\mathbb{F}}(U^{\times m})$ is a tensor subspace of $\mathcal{L}_{\mathbb{F}}^{(n)}(U^{\times m})$, when V is represented by the invariant tensor space $\mathcal{LV}_{\mathbb{F}}(U^{\times m})$ on group $\mathcal{G}_{\mathbb{F}}(U^{\times m})$. This group is represented by $\mathcal{V}_{\mathbb{F}}(U^{\times m})$ and called the **invariance group** of $\mathcal{LV}_{\mathbb{F}}(U^{\times m})$ if it is constituted by all the invertible continuous linear tensor operators on which every element of $\mathcal{LV}_{\mathbb{F}}(U^{\times m})$ is invariant. Thus, it is evident that operator $\mathbf{i} \in \mathcal{V}_{\mathbb{F}}(U^{\times m})$. The unitary group $\mathcal{N}_{\mathbb{F}}(U^{\times m}) \subseteq \mathcal{V}_{\mathbb{F}}(U^{\times m})$ is the invariance group of the invariant tensor space $\mathcal{LN}_{\mathbb{F}}(N^{\times m})$, whose tensors are said to be **isotropic**[10] relative to $\mathcal{N}_{\mathbb{F}}(U^{\times m})$, and then $\mathcal{LV}_{\mathbb{F}}(U^{\times m}) \subseteq \mathcal{LN}_{\mathbb{F}}(U^{\times m})$. Moreover, the unitary group $\mathcal{N}_{\mathbb{F}}^{+}(U^{\times m}) \subset \mathcal{N}_{\mathbb{F}}(U^{\times m})$, where $\mathrm{Det}(\boldsymbol{\varphi}) = 1$ for all $\boldsymbol{\varphi} \in \mathcal{N}_{\mathbb{F}}^{+}(U^{\times m})$, is the invariance group of $\mathcal{LN}_{\mathbb{F}}^{+}(N^{\times m})$, whose tensors are called **anti isotropic** relative to $\mathcal{N}_{\mathbb{F}}^{+}(U^{\times m})$ and tensor space $\mathcal{LN}_{\mathbb{F}}(U^{\times m}) \subset \mathcal{LN}_{\mathbb{F}}^{+}(U^{\times m})$. Then, we can say that every isotropic tensor is also anti isotropic. Tensors that are not isotropic are called **anisotropic**. From definitions (3.41), in the context of Euclidean spaces, we can say that (anti) isotropic tensors are defined by (proper) orthogonal groups.

Theorem 15 – Coordinate Invariance of Isotropic Tensors

Considering an arbitrary unitary operator $\boldsymbol{\psi} \in \mathcal{N}_{\mathbb{F}}(U^{\times m})$ *where* $\dim(U_{\mathbb{F}})_i = n$, *tensor bases* $B = \{\mathbf{B}_1, \cdots, \mathbf{B}_p\}$ *and* $C = \{\boldsymbol{\psi}(\mathbf{B}_1), \cdots, \boldsymbol{\psi}(\mathbf{B}_p)\}$ *of* $\mathcal{LN}_{\mathbb{F}}(U^{\times m})$, *where* $p = n^m$ *and tensor* $\mathbf{B}_i$ *is defined by orthonormal bases, according to theorem 11, then equalities* $\mathbf{f}_i^B(\mathbf{Q}) = \mathbf{f}_i^C(\mathbf{Q})$ *are valid for arbitrary* $\mathbf{Q} \in \mathcal{LN}_{\mathbb{F}}(U^{\times m})$. *Moreover, if* $\mathbf{f}_i^B(\mathbf{T}) = \mathbf{f}_i^C(\mathbf{T})$ *for a certain tensor* $\mathbf{T} \in \mathcal{L}_{\mathbb{F}}^{(n)}(U^{\times m})$, *then* $\mathbf{T}$ *also belongs to* $\mathcal{LN}_{\mathbb{F}}(U^{\times m})$.

Proof. First, let's verify if set V is indeed a vector space of invariant tensors on $\mathcal{G}_{\mathbb{F}}(U^{\times m})$. Considering operator $\boldsymbol{\varphi} \in \mathcal{G}_{\mathbb{F}}(U^{\times m})$, $\alpha, \beta \in \mathbb{F}$ and $\mathbf{A}, \mathbf{B} \in V$ arbitrary elements of their respective sets, the following equalities $\alpha\mathbf{A} + \beta\mathbf{B} = \alpha\boldsymbol{\varphi}(\mathbf{A}) + \beta\boldsymbol{\varphi}(\mathbf{B}) = \boldsymbol{\varphi}(\alpha\mathbf{A} + \beta\mathbf{B})$ prove that sum and multiplication by scalar of invariant tensors result in an invariant tensor, since $\boldsymbol{\varphi}$ is linear. This condition also justifies that $\mathbf{0} \in V$ because $\boldsymbol{\varphi}(\mathbf{0}) = \mathbf{0}$. The verification of other axioms of vector spaces is trivial. Now, let's prove the theorem above. From theorem 11, we know that set $C = \{\boldsymbol{\psi}(\mathbf{B}_1), \cdots, \boldsymbol{\psi}(\mathbf{B}_p)\}$ is indeed a basis of $\mathcal{LN}_{\mathbb{F}}(U^{\times m})$. Since $\boldsymbol{\psi}(\mathbf{Q}) = \mathbf{Q}$, equalities

$$\boldsymbol{\psi}(\mathbf{Q}) = \sum_{i=1}^{p} \mathbf{f}_i^B(\mathbf{Q})\boldsymbol{\psi}(\mathbf{B}_i) \qquad \text{and} \qquad \mathbf{Q} = \sum_{i=1}^{p} \mathbf{f}_i^C(\mathbf{Q})\boldsymbol{\psi}(\mathbf{B}_i)$$

[10] Appleby et al[3] define an isotropic tensor relative to certain changes of bases as the tensor whose coordinates are preserved under these changes, which is assured in our approach by all unitary operators.

enable us to conclude that coordinates $\mathbf{f}_i^B(\mathbf{Q}) = \mathbf{f}_i^C(\mathbf{Q})$. Considering these same equalities for the case of a tensor $\mathbf{T}$ that has equal coordinates on bases B and C, the conclusion of $\boldsymbol{\psi}(\mathbf{T}) = \mathbf{T}$ is straightforward. $\square$

Admitting a tensor space $\mathcal{L}_{\mathbb{F}}^{(m)}(U^p)$ and a group $\mathcal{N}_{\mathbb{F}}(U^p)$, where p is even, by equality (3.23) we know that the coordinates of identity tensor $\mathbf{I} \in \mathcal{L}_{\mathbb{F}}^{(m)}(U^p)$ are the same for any chosen basis defined from orthonormal bases of $U_{\mathbb{F}}$. Therefore, from theorem above, it is clear that $\mathbf{I}$ is isotropic, that is, $\mathbf{I} \in \mathcal{LN}_{\mathbb{F}}(U^p)$. In this context, again by equality (3.23), it is straightforward to see that the coordinates of an arbitrary tensor belonging to span($\{\mathbf{I}\}$) are the same for any chosen basis of $\mathcal{L}_{\mathbb{F}}^{(m)}(U^p)$ defined from orthonormal bases of $U_{\mathbb{F}}$. Therefore, it is possible to affirm that *any linear combination of an identity tensor is isotropic*. Now, from the context of theorem 14, it is straightforward to conclude that an arbitrary alternating tensor $\mathbf{A} \in \mathcal{LA}_{\mathbb{F}}^{(m)}(U^m)$ is anti isotropic, that is, equality $\mathbf{A} = \boldsymbol{\psi}(\mathbf{A})$ is valid only when $\mathrm{Det}(\boldsymbol{\psi}) = 1$. Moreover, we can say that *any linear combination of alternating tensors is anti isotropic*. Finally, it can be obtained that *every permutation of an (anti) isotropic tensor and a linear combination of such permutations are also (anti) isotropic*. It is important to say also that tensor products preserve (anti) isotropy the following way:

i. If $\mathbf{Q}_1, \mathbf{Q}_2 \in \mathcal{LN}_{\mathbb{F}}(U^m)$ then $\mathbf{Q}_1 \otimes \mathbf{Q}_2 \in \mathcal{LN}_{\mathbb{F}}(U^m \times U^m)$;

ii. If $\mathbf{Q}_1 \in \mathcal{LN}_{\mathbb{F}}(U^m)$ and $\mathbf{Q}_2 \in \mathcal{LN}_{\mathbb{F}}^+(U^m)$ then $\mathbf{Q}_1 \otimes \mathbf{Q}_2, \mathbf{Q}_2 \otimes \mathbf{Q}_1 \in \mathcal{LN}_{\mathbb{F}}^+(U^m \times U^m)$.

Proof. First, let's verify the (anti) isotropy of permutation of tensors. Considering equality (3.42) and the definition of unitary function, in the context where unitary operators $\boldsymbol{\psi}^{-1}$ and $\mathbf{h}$ are related, for an arbitrary isotropic tensor $\mathbf{Q} \in \mathcal{LN}_{\mathbb{F}}(U^m)$, the following development proves that $\mathbf{Q}_\pi = \boldsymbol{\psi}(\mathbf{Q}_\pi)$:

$$\begin{aligned}
\mathbf{Q}_\pi(\mathbf{x}_1, \cdots, \mathbf{x}_m) &= \boldsymbol{\psi}(\mathbf{Q})(\mathbf{x}_{\pi(1)}, \cdots, \mathbf{x}_{\pi(m)}) \\
&= \boldsymbol{\psi}(\mathbf{Q}) \cdot (\mathbf{x}_{\pi(1)} \otimes \cdots \otimes \mathbf{x}_{\pi(m)}) \\
&= \mathbf{Q} \cdot \boldsymbol{\psi}^{-1}(\mathbf{x}_{\pi(1)} \otimes \cdots \otimes \mathbf{x}_{\pi(m)}) \\
&= \mathbf{Q} \cdot (\mathbf{h}(\mathbf{x}_{\pi(1)}) \otimes \cdots \otimes \mathbf{h}(\mathbf{x}_{\pi(m)})) \\
&= \mathbf{Q}_\pi(\mathbf{h}(\mathbf{x}_1), \cdots, \mathbf{h}(\mathbf{x}_m)) \\
&= \mathbf{Q}_\pi \cdot \boldsymbol{\psi}^{-1}(\mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_m) \\
&= \boldsymbol{\psi}(\mathbf{Q}_\pi)(\mathbf{x}_1, \cdots, \mathbf{x}_m).
\end{aligned}$$

By this same process, it is then trivial to prove that a linear combination of transposition of tensors is also (anti) isotropic. Now we verify the properties above. From the context of equality (3.42), development

$$\begin{aligned}
\mathbf{Q}_1 \otimes \mathbf{Q}_2(\mathbf{x}_1, \cdots, \mathbf{x}_m, \mathbf{y}_1, \cdots, \mathbf{y}_m) &= \boldsymbol{\psi}(\mathbf{Q}_1)(\mathbf{x}_1, \cdots, \mathbf{x}_m)\, \boldsymbol{\psi}(\mathbf{Q}_2)(\mathbf{y}_1, \cdots, \mathbf{y}_m) \\
&= \mathbf{Q}_1(\mathbf{g}(\mathbf{x}_1), \cdots, \mathbf{g}(\mathbf{x}_m))\, \mathbf{Q}_2(\mathbf{g}(\mathbf{y}_1), \cdots, \mathbf{g}(\mathbf{y}_m)) \\
&= \mathbf{Q}_1 \otimes \mathbf{Q}_2(\mathbf{g}(\mathbf{x}_1), \cdots, \mathbf{g}(\mathbf{x}_m), \mathbf{g}(\mathbf{y}_1), \cdots, \mathbf{g}(\mathbf{y}_m)) \\
&= \boldsymbol{\phi}(\mathbf{Q}_1 \otimes \mathbf{Q}_2)(\mathbf{x}_1, \cdots, \mathbf{x}_m, \mathbf{y}_1, \cdots, \mathbf{y}_m)
\end{aligned}$$

where $\boldsymbol{\phi} \in \mathcal{N}_{\mathbb{F}}(U^m \times U^m)$, proves the first property. The second property is proved by this same development, which is valid only if $\boldsymbol{\psi} \in \mathcal{U}_{\mathbb{F}}^+(U^m)$ since $\mathbf{Q}_2$ is anti isotropic. $\square$

At this point, we sacrifice once again self-containednes and humbly stand upon the shoulders of a giant: the german mathematician Hermann Weyl (1885–1955). By proving his Theorem 2.9.A, Weyl[58] showed that identity tensor $\mathbf{I} \in \mathcal{LN}_{\mathbb{F}}(\mathcal{U}^{p})$, where p is even, its transpositions and their linear combinations constitute the isotropy group, that is, tensor space

$$\mathcal{LN}_{\mathbb{F}}(\mathcal{U}^{p}) = \operatorname{span}(\{\mathbf{I}\}) \cup \operatorname{span}(\{\mathbf{I}_{\pi}\}) \,. \tag{3.81}$$

For the case of anti isotropic tensors, he arrived at

$$\mathcal{LN}^{+}_{\mathbb{F}}(\mathcal{U}^{p} \times \mathcal{U}^{m}) = \operatorname{span}(\{\mathbf{I} \otimes \mathbf{A}\}) \cup \operatorname{span}(\{(\mathbf{I} \otimes \mathbf{A})_{\pi}\}) \,, \tag{3.82}$$

where $\mathbf{A} \in \mathcal{LA}^{(m)}_{\mathbb{F}}(\mathcal{U}^{m})$. Compiling the results obtained by Backus[5], pp. 80–88, for tensor spaces of order from 0 to 4, we build table 3.1.

Order of $\mathcal{L}_{\mathbb{F}}(\mathcal{U}^{p})$	Value of $\dim(\mathcal{U}_{\mathbb{F}})$	Space $\mathcal{LN}_{\mathbb{F}}(\mathcal{U}^{p})$	Space $\mathcal{LN}^{+}_{\mathbb{F}}(\mathcal{U}^{p})$	Post Conditions
$p = 0$	> 0	$\mathbb{F}$	$\mathbb{F}$	-
$p = 1$	$\geqslant 2$	$\{\mathbf{0}\}$	$\{\mathbf{0}\}$	-
$p = 2$	> 0	$\operatorname{span}(\{\mathbf{I}\})$	$\emptyset$	$\mathbf{I} \in \mathcal{L}_{\mathbb{F}}(\mathcal{U}^{p})$,
	$= 2$		$\operatorname{span}(\{\mathbf{I}, \mathbf{A}\})$	$\forall \mathbf{A} \neq \mathbf{0}$,
	$\geqslant 3$		$\operatorname{span}(\{\mathbf{I}\})$	$\mathbf{A} \in \mathcal{LA}^{(p)}_{\mathbb{F}}(\mathcal{U}^{p})$
$p = 3$	> 0	$\{\mathbf{0}\}$	$\emptyset$	
	$= 3$		$\operatorname{span}(\{\mathbf{A}\})$	
	$\neq 3$		$\{\mathbf{0}\}$	
$p = 4$	> 0	$\operatorname{span}(\{\mathbf{I}, \mathbf{I}_{(2,3)}, \mathbf{I}_{(2,4)}\})$	$\emptyset$	
	$\in \{2, 4\}$		$\operatorname{span}(\{\mathbf{I}, \mathbf{I}_{(2,3)}, \mathbf{I}_{(2,4)}\})$	
	$= 4$		$\operatorname{span}(\{\mathbf{I}, \mathbf{I}_{(2,3)}, \mathbf{I}_{(2,4)}, \mathbf{A}\})$	
	$= 2$	$\operatorname{span}(\{\mathbf{I} \otimes \mathbf{I}, (\mathbf{I} \otimes \mathbf{I})_{(2,3)}, (\mathbf{I} \otimes \mathbf{I})_{(2,4)}\})$	$\operatorname{span}(\{\mathbf{I} \otimes \mathbf{I}, (\mathbf{I} \otimes \mathbf{I})_{(2,3)}, (\mathbf{I} \otimes \mathbf{I})_{(2,4)}, \mathbf{I} \otimes \mathbf{A}, (\mathbf{I} \otimes \mathbf{A})_{(2,3)}, (\mathbf{I} \otimes \mathbf{A})_{(2,4)}, \mathbf{A} \otimes \mathbf{I}, (\mathbf{A} \otimes \mathbf{I})_{(2,3)}, (\mathbf{A} \otimes \mathbf{I})_{(2,4)}\})$	$\mathbf{I} \in \mathcal{L}_{\mathbb{F}}(\mathcal{U}^{2})$, $\forall \mathbf{A} \neq \mathbf{0}$, $\mathbf{A} \in \mathcal{LA}^{(2)}_{\mathbb{F}}(\mathcal{U}^{2})$

Table 3.1 – Anti isotropic and isotropic tensors in $\mathcal{L}_{\mathbb{F}}(\mathcal{U}^{p})$ according to Backus[5].

Chapter

4 Topics on Affine Geometry

Although vector spaces are building blocks of the several important concepts presented so far on the two previous chapters, it is still not possible to recognize shapes in their definer sets. However, the study of mathematical shapes does not belong to Linear Algebra, but to the realms of Geometry, whose non empty sets we shall consider here to be constituted by "shape" elements called **points**, which are geometric objects of primitive notion, devoid of dimensional features in order to morphologically represent physical locations. If a given set of such points is conveniently related to a vector space by a group action, other shapes can be obtained and thereby further morphological concepts can be developed, when we say that a certain geometry is defined. As physical spaces are usually abstracted by geometric sets, we can not prescind from studying at least the basic topics of the so called Affine Geometry, which includes the well known Euclidean Geometry and is sufficiently governed by the concept of parallelism.

4.1 Affine Spaces

Recalling the concept of group action presented on chapter 1, let the set of points $\mathcal{U}$ be the G–set of a vector space $\mathrm{U}_{\mathbb{F}}$ through a simply transitive group action $\oplus$. Given two arbitrary points a and b of this G–set $\mathcal{U}$, now called a **point space**, and a vector $\mathbf{u} \in \mathrm{U}_{\mathbb{F}}$ where point $\oplus(\mathbf{u}, a) = b$ or $\mathbf{u} \oplus a = b$, the axioms of simply transitive group actions can be rewritten the following way:

i. $\mathbf{0} \oplus a = a, \forall\, a \in \mathcal{U}$;

ii. $\mathbf{u} \oplus (\mathbf{v} \oplus a) = (\mathbf{u} + \mathbf{v}) \oplus a, \forall\, a \in \mathcal{U}, \forall\, \mathbf{u}, \mathbf{v} \in \mathrm{U}_{\mathbb{F}}$;

iii. $\exists!\, \mathbf{u} \in \mathrm{U}_{\mathbb{F}}$ such that $\mathbf{u} \oplus a = b$.

In this context, the triple $(\mathsf{U}_{\mathbb{F}}, \mathcal{U}, \oplus)$ is called an affine space, henceforth represented by $\mathcal{U}_{\mathbb{F}}$, which will also be used to refer to the point space $\mathcal{U}$, as we adopted similarly for other cases in order to simplify notation. Moreover, vector space $\mathsf{U}_{\mathbb{F}}$ is usually called the **direction space** of affine space $\mathcal{U}_{\mathbb{F}}$. Therefore, we can say that an affine space is defined by a direction space, a point space and a group action. From axiom iii above, every double of points uniquely identifies a certain vector: if $\mathbf{u} \oplus a = b$, we define that double (a, b) identifies vector $\mathbf{u}$, when this vector is represented by $\overrightarrow{ab}$ and then $\overrightarrow{ab} \oplus a = b$. *The converse is not true: a certain vector can be identified by infinite doubles of points because there is always a* $y = \mathbf{u} \oplus x$ *for all* $x \in \mathcal{U}_{\mathbb{F}}$ *and then* $\overrightarrow{ab} = \overrightarrow{xy}$. Now, from axioms i and ii, considering point $\boldsymbol{v} \oplus a = c$, we conclude that $-\boldsymbol{v} \oplus c = a$, and then vector $\boldsymbol{v} = \overrightarrow{ac} = -\overrightarrow{ca}$. From this characteristic relationship between points and vectors, the classical pictorial representation of a vector identified by a double of points (a, b) as an arrow whose tail "starts" on a and head points to b arises naturally (figure 4.1). Since

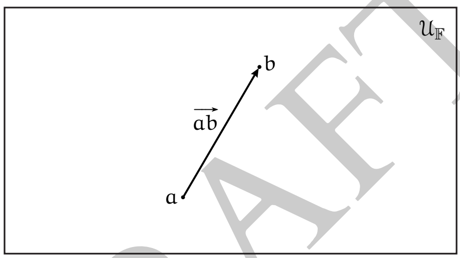

Figure 4.1 – Vector as an arrow identified by points a and b.

an arbitrary vector can be identified by infinite doubles of points, given arbitrary points $x, y, a, b \in \mathcal{U}_{\mathbb{F}}$, the second operand of $\overrightarrow{ab} + \overrightarrow{xy}$ can be described by a vector $\overrightarrow{bc}$, where point $c = \overrightarrow{xy} \oplus b$, and then equalities $(\overrightarrow{ab} + \overrightarrow{bc}) \oplus a = (\overrightarrow{bc} + \overrightarrow{ab}) \oplus a = \overrightarrow{bc} \oplus b = c$ enable us to conclude that if $(\overrightarrow{ab} + \overrightarrow{bc}) \oplus a = c$ then $\overrightarrow{ab} + \overrightarrow{bc} = \overrightarrow{ac}$. From this algebraic property, it is possible to depict graphically a sum of vectors by concatenating each one of the their representative arrows in such a way that a head point to a tail, as shown in the following figure 4.2. It is then straightforward to conclude that a vector $\alpha\overrightarrow{ab}$, where $\alpha \in \mathbb{R}$, is

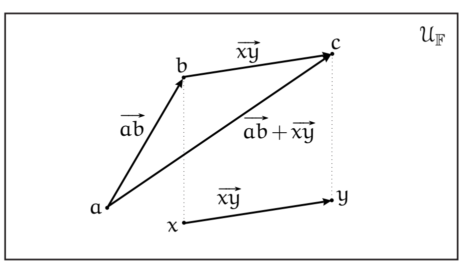

Figure 4.2 – Sum of vectors $\overrightarrow{ab}$ and $\overrightarrow{xy}$ where $\overrightarrow{xy} \oplus b = c$.

represented by a "compressed" $(\alpha \leqslant 1)$ or "expanded" $(\alpha > 1)$ vector $\overrightarrow{ab}$. Moreover, we can say that the coordinates of vector $\alpha\overrightarrow{ab}$ are the coordinates of $\overrightarrow{ab}$ equally multiplied, that is equally "compressed" or "expanded" by the real number α. For the case of a complex affine space $\mathcal{U}_{\mathbb{C}}$, the polar complex coordinates of a vector when multiplied by a complex α are equally multiplied by the real scalar $|\alpha|$ and equally rotated by $\arg(\alpha)$ on the complex plane, according to the basic theory of complex number multiplication. Therefore, in affine geometric representation, a complex vector $\alpha\overrightarrow{ab} \in U_{\mathbb{C}}$ results also in a "compressed" or "expanded" complex vector $\overrightarrow{ab}$.

Considering previous conditions, given an arbitrary point $a \in \mathcal{U}_{\mathbb{F}}$ and a vector subspace $S_{\mathbb{F}} \subset U_{\mathbb{F}}$, the triple $(S_{\mathbb{F}}, \mathcal{S}_a, \oplus)$ is called an **affine subspace** of $\mathcal{U}_{\mathbb{F}}$, represented by $(\mathcal{S}_a)_{\mathbb{F}}$, if point space $\mathcal{S}_a := \{\mathbf{x} \oplus a : \forall \mathbf{x} \in S_{\mathbb{F}}\}$, to which point a, called the **seed pointf** $(\mathcal{S}_a)_{\mathbb{F}}$, also belongs since vector $\mathbf{0} \in S_{\mathbb{F}}$. Note that if a point $b \in (\mathcal{S}_a)_{\mathbb{F}}$ defines the affine subspace $(\mathcal{S}_b)_{\mathbb{F}}$, then

$$(\mathcal{S}_a)_{\mathbb{F}} = \{\mathbf{x} \oplus a : \forall \mathbf{x} \in S_{\mathbb{F}}\} = \{\mathbf{x} + \overrightarrow{ab} \oplus b : \forall \mathbf{x} \in S_{\mathbb{F}}\} = (\mathcal{S}_b)_{\mathbb{F}}. \tag{4.1}$$

In the case of affine space $\mathcal{U}_{\mathbb{F}}$ defined by a m-dimensional vector space, an arbitrary affine subspace $(\mathcal{S}_a)_{\mathbb{F}} \subset \mathcal{U}_{\mathbb{F}}$ becomes another fundamental geometric object called a **hyperplane** when its definer vector subspace $S_{\mathbb{F}}$ is $(m-1)$-dimensional: we specifically call the hyperplane a **line** or a **plane** when $S_{\mathbb{F}}$ is one or two-dimensional respectively. Since it is the vector space which bears a dimensional feature, it is defined that $\dim(\mathcal{U}_{\mathbb{F}}) = \dim(U_{\mathbb{F}}) = m$ and then we use $\mathcal{U}_{\mathbb{F}}^m$ to represent a m-dimensional affine space. Now, from the definition of affine spaces using a group action approach, as we have done, angles and distances cannot be obtained, and then it results that requiring these two concepts means two additional features that turn affine geometry into Euclidean geometry, when we state that Euclidean is a more restricted form of affine geometry. On the other hand, our algebraic definition of affine spaces does not require to explicitly present the so called Playfair's Axiom as a restriction, as an axiom itself to be observed because this restriction becomes a natural property, a theorem. In order to present this very important property, we must first deal with the concept of parallelism using previous definitions. Considering our algebraic approach, two affine subspaces are considered to be **parallel** if the direction space of one is an improper subset of the other[1]. In mathematical terms, we define two affine subspaces $(\mathcal{S}_a)_{\mathbb{F}}^n$ and $(\mathcal{V}_b)_{\mathbb{F}}^r$ of $\mathcal{U}_{\mathbb{F}}^m$ to be parallel, represented by $(\mathcal{S}_a)_{\mathbb{F}}^n \parallel (\mathcal{V}_b)_{\mathbb{F}}^r$, if direction space $S_{\mathbb{F}} \subseteq V_{\mathbb{F}}$ or $V_{\mathbb{F}} \subseteq S_{\mathbb{F}}$. From this definition of parallelism, when $(\mathcal{S}_a)_{\mathbb{F}}^n$ and $(\mathcal{V}_b)_{\mathbb{F}}^r$ are indeed parallel we obtain the following properties:

i. $n = r \iff S_{\mathbb{F}} = V_{\mathbb{F}}$;

ii. $a \in (\mathcal{V}_b)_{\mathbb{F}}^r \wedge S_{\mathbb{F}} \subseteq V_{\mathbb{F}} \iff (\mathcal{S}_a)_{\mathbb{F}}^n \subseteq (\mathcal{V}_b)_{\mathbb{F}}^r$;

iii. $a \notin (\mathcal{V}_b)_{\mathbb{F}}^r \vee b \notin (\mathcal{S}_a)_{\mathbb{F}}^n \iff (\mathcal{S}_a)_{\mathbb{F}}^n \cap (\mathcal{V}_b)_{\mathbb{F}}^r = \emptyset$.

[1] This concept of parallelism admits that a hyperplane (a line or a plane, for example) is parallel to itself.

Figure 4.3 depicts a line and plane parallel to each other for the case of a five-dimensional real affine space. Now, we say that subspaces $(\mathcal{W}_c)^r_{\mathbb{F}}$ and $(\mathcal{S}_a)^n_{\mathbb{F}}$ are **perpendicular**, represented by $(\mathcal{W}_c)^r_{\mathbb{F}} \perp (\mathcal{S}_a)^n_{\mathbb{F}}$, if vectors of $W_{\mathbb{F}}$ and $V_{\mathbb{F}}$ do not have an incidence interrelationship, that is, if $W_{\mathbb{F}} \perp V_{\mathbb{F}}$. Geometrical representation of perpendicularity is more intuitive in the context of affine Euclidean spaces through the concept of angle, which we shall present later in this section.

Proof. For the first item of the previous properties, since $S_{\mathbb{F}} \subseteq V_{\mathbb{F}}$ or $V_{\mathbb{F}} \subseteq S_{\mathbb{F}}$, a direction space is a proper subset of the other only if they have different dimensions, otherwise they are equal. The inverse implication is trivial. For the second item, if $a \in (\mathcal{V}_b)^r_{\mathbb{F}}$ there is a vector $\boldsymbol{w} \in V_{\mathbb{F}}$ where $a = \boldsymbol{w} \oplus b$. Then, from axiom ii on page 103 and since $S_{\mathbb{F}} \subseteq V_{\mathbb{F}}$, we can write $\boldsymbol{x} \oplus a = (\boldsymbol{x} + \boldsymbol{w}) \oplus b$, $\forall \boldsymbol{x} \in S_{\mathbb{F}}$, from which we conclude that $(\mathcal{S}_a)^n_{\mathbb{F}} \subseteq (\mathcal{V}_b)^r_{\mathbb{F}}$. Now, if $(\mathcal{S}_a)^n_{\mathbb{F}} \subseteq (\mathcal{V}_b)^r_{\mathbb{F}}$ is true we always have a vector $\boldsymbol{y} \in V_{\mathbb{F}}$ related to an arbitrary $\boldsymbol{x} \in S_{\mathbb{F}}$ through $\boldsymbol{x} \oplus a = \boldsymbol{y} \oplus b$ or $a = (\boldsymbol{y} - \boldsymbol{x}) \oplus b$, from which we conclude that $a \in (\mathcal{V}_b)^r_{\mathbb{F}}$ and $S_{\mathbb{F}} \subseteq V_{\mathbb{F}}$. For property iii, we prove only for $a \notin (\mathcal{V}_b)^r_{\mathbb{F}}$. In this case, there is no vector in $V_{\mathbb{F}}$ relating the pair of points (b, a). Since $S_{\mathbb{F}} \subseteq V_{\mathbb{F}}$ or $V_{\mathbb{F}} \subseteq S_{\mathbb{F}}$, we can write for all $\boldsymbol{x} \in S_{\mathbb{F}}$ and $\boldsymbol{y} \in V_{\mathbb{F}}$ that

$$\begin{aligned} a &\neq \boldsymbol{y} \oplus b \\ \boldsymbol{x} \oplus a &\neq \boldsymbol{x} \oplus (\boldsymbol{y} \oplus b) \\ \boldsymbol{x} \oplus a &\neq (\boldsymbol{x} + \boldsymbol{y}) \oplus b, \end{aligned}$$

from which we conclude that $(\mathcal{S}_a)^n_{\mathbb{F}} \cap (\mathcal{V}_b)^r_{\mathbb{F}} = \emptyset$. The inverse implication is verified considering $(\mathcal{S}_a)^n_{\mathbb{F}} \cap (\mathcal{V}_b)^r_{\mathbb{F}} = \emptyset$ valid and then, for all $\boldsymbol{x} \in S_{\mathbb{F}}$ and $\boldsymbol{y} \in V_{\mathbb{F}}$,

$$\begin{aligned} \boldsymbol{x} \oplus a &\neq \boldsymbol{y} \oplus b \\ -\boldsymbol{x} \oplus (\boldsymbol{x} \oplus a) &\neq -\boldsymbol{x} \oplus (\boldsymbol{y} \oplus b) \\ a &\neq (\boldsymbol{y} - \boldsymbol{x}) \oplus b, \end{aligned}$$

from which we conclude that $a \notin (\mathcal{V}_b)^r_{\mathbb{F}}$. □

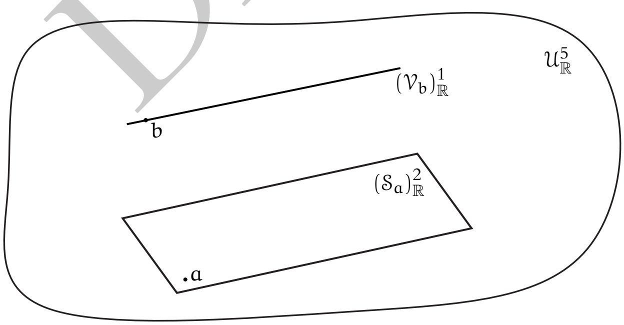

Figure 4.3 – Line and plane parallel to each other.

Theorem 16 – Playfair's "Axiom"

If $\mathcal{U}^{n+1}_{\mathbb{F}}$ is an affine space where $(\mathcal{S}_a)^n_{\mathbb{F}}$ is one of its hyperplanes and $k \notin (\mathcal{S}_a)^n_{\mathbb{F}}$ one of its points, there is a unique hyperplane $(\mathcal{V}_b)^n_{\mathbb{F}} \parallel (\mathcal{S}_a)^n_{\mathbb{F}}$ such that $k \in (\mathcal{V}_b)^n_{\mathbb{F}}$.

Proof. If $(\mathcal{S}_a)^n_{\mathbb{F}}$ exists, so does $(\mathcal{S}_b)^n_{\mathbb{F}}$, where $b \neq a$. Since these subspaces are obviously parallel, the existence of subspace $(\mathcal{V}_b)^n_{\mathbb{F}} = (\mathcal{S}_b)^n_{\mathbb{F}}$ is proved. Uniqueness is verified by the following rationale: since point $k \in (\mathcal{V}_b)^n_{\mathbb{F}}$ then $(\mathcal{V}_b)^n_{\mathbb{F}} = (\mathcal{V}_k)^n_{\mathbb{F}}$ according to (4.1); and, supposing a third hyperplane $(\mathcal{W}_c)^n_{\mathbb{F}} \parallel (\mathcal{S}_a)^n_{\mathbb{F}}$ where $k \in (\mathcal{W}_c)^n_{\mathbb{F}}$, we also have $(\mathcal{W}_c)^n_{\mathbb{F}} = (\mathcal{W}_k)^n_{\mathbb{F}}$. Moreover, from the first property of parallelism, it is clear that $V_{\mathbb{F}} = S_{\mathbb{F}} = W_{\mathbb{F}}$ because the three hyperplanes have obviously the same dimension and then $(\mathcal{W}_k)^n_{\mathbb{F}} = (\mathcal{V}_k)^n_{\mathbb{F}}$. Since $(\mathcal{V}_b)^n_{\mathbb{F}} = (\mathcal{V}_k)^n_{\mathbb{F}}$ and $(\mathcal{W}_c)^n_{\mathbb{F}} = (\mathcal{W}_k)^n_{\mathbb{F}}$, then $(\mathcal{V}_b)^n_{\mathbb{F}} = (\mathcal{W}_c)^n_{\mathbb{F}}$. □

Considering previous conditions, given a point $o \in (\mathcal{S}_a)^n_{\mathbb{F}}$ and a basis $B = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ of $S_{\mathbb{F}}$, where $n < m$, we call double (o, B) an **affine coordinate system** of $(\mathcal{S}_a)^n_{\mathbb{F}}$ because point space $\mathcal{S}_a$ can be obtained from o and B as follows:

$$\mathcal{S}_a = \{\boldsymbol{x} \oplus a : \forall \boldsymbol{x} \in S_{\mathbb{F}}\} = \{(\boldsymbol{x} + \overrightarrow{oa}) \oplus o : \forall \boldsymbol{x} \in \mathrm{span}(B)\}. \tag{4.2}$$

In this context, point o is called the **origin** of the coordinate system and the coordinates of an arbitrary vector $\boldsymbol{v} \in S_{\mathbb{F}}$ on B are considered to be the coordinates of point $v := \boldsymbol{v} \oplus o$ on (o, B), that is, $[v]^B := [\boldsymbol{v}]^B$. Thereby, it is trivial to obtain that the coordinates of point $\boldsymbol{u} \oplus v$ are the coordinates of vector $\boldsymbol{u} + \boldsymbol{v}$ or that $[\boldsymbol{u} \oplus v]^B = [\boldsymbol{u} + \boldsymbol{v}]^B$. Moreover, each line $((\mathcal{B}_i)_o)^1_{\mathbb{F}}$ defined by vector space $(B_i)_{\mathbb{F}} = \mathrm{span}(\{\boldsymbol{v}_i\})$ is called an **axis** of (o, B), which is depicted in figure 4.4. On section 1.6, we said that in a group action a biunivocal

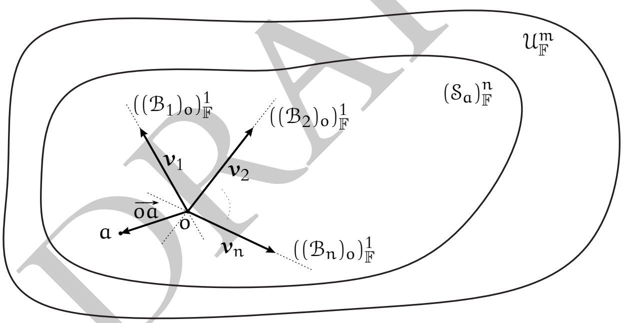

Figure 4.4 – n-dimensional affine coordinate system (o, B) and its axes.

relationship between its definer set and its definer group is possible if an element of the set is fixed. For the case of affine spaces, this fixation is established by a coordinate system, which enables to define a point v from a vector $\boldsymbol{v}$ through an origin o, as we have done. In mathematical terms, considering origin o, the mapping $\upsilon_o : (\mathcal{S}_a)^n_{\mathbb{F}} \mapsto S_{\mathbb{F}}$ is bijective if

$$\upsilon_o(x) = \overrightarrow{ox}, \tag{4.3}$$

where bijection υ_o is said to "vectorize" its argument and then, since the origin is fixed, we can represent $\upsilon_o(v) = \overrightarrow{ov}$ by $\vec{v}$ or $\boldsymbol{v}$. Therefore, it is clear that

$$\upsilon_o^{-1}(\vec{x}) = \vec{x} \oplus o. \tag{4.4}$$

Considering arbitrary $\boldsymbol{u}, \boldsymbol{v} \in S_{\mathbb{F}}$, $a \in (\mathcal{S}_a)^n_{\mathbb{F}}$ and $\alpha \in \mathbb{F}$, there are some interesting properties involving this "vectorization" function, namely

i. $\upsilon_o(\mathbf{u} \oplus a) = \mathbf{u} + \vec{a}$;

ii. $\upsilon_o[(\mathbf{u}+\mathbf{v}) \oplus a] = \mathbf{u}+\mathbf{v}+\vec{a}$;

iii. $\upsilon_o(\alpha\mathbf{u} \oplus a) = \alpha\mathbf{u} + \vec{a}$.

Proof. The first item is proved through the following: if point $c = \mathbf{u} \oplus a$ then, by the sum of vectors, it is true that $\upsilon_o(c) = \overrightarrow{oc} = \overrightarrow{oa} + \mathbf{u} = \vec{a} + \mathbf{u}$. The second item is a straightforward consequence of the first and the third is a straightforward consequence of the second. □

The association $((\mathcal{S}_a)^n_{\mathbb{F}}, o, B)$ of an affine subspace with a coordinate system can be structured like a vector space and then inherits the same classification, outlined in figure 2.3, from its direction space $S_{\mathbb{F}}$ if suitable expressions for metric, norm and inner product are defined. Thereby, for arbitrary points v and u of $(\mathcal{S}_a)^n_{\mathbb{F}}$, the triple $((\mathcal{S}_a)^n_{\mathbb{F}}, o, B)$ is called an affine metric space if $\rho(u,v) := \rho(\vec{u}, \vec{v})$, an affine normed space if a norm $\|\vec{v}\|$ is defined or an affine inner product space if there is the scalar $\vec{v} \cdot \vec{u}$. *Since an affine subspace is itself an affine space, representation of triple* $((\mathcal{S}_a)^n_{\mathbb{F}}, o, B)$ *will be shortened by* $\mathcal{S}^n_{\mathbb{F}}$*, where the seed point* a *and the coordinate system are implicit.* Since we are dealing in this chapter with morphological matters, what are the geometric manifestations of metric, norm and inner product? We already said that metric refers to the concept of distance and thus scalar $\rho(u,v)$ informs the distance between arbitrary points u and v in the context of affine metric spaces. Norm however is not applicable to points, being geometrically represented by the size of a vector, which also depicts the concept of vector intensity: a "bigger" vector is considered to be a more "intense" vector. In the context of affine Banach spaces, where distance

$$\rho(u,v) := \|\vec{u} - \vec{v}\| = \|\overrightarrow{vo} + \vec{u}\| = \|\overrightarrow{vu}\| = \|\overrightarrow{uv}\|, \tag{4.5}$$

norm enables distance measurement. Now, regarding the geometric effects of the inner product, we already said on section 2.1 that this product is closely related to vector incidence, being a scalar measure of this fundamental concept[2]. In the context of affine normed inner product spaces, the incidence interrelationship of non zero vectors $\vec{v}$ and $\vec{u}$ is expressed both by the **projection** $\zeta(\vec{v}, \vec{u})$ of $\vec{v}$ onto the line defined by the set $\operatorname{span}(\{\vec{u}\})$ and the projection $\zeta(\vec{u}, \vec{v})$ of $\vec{u}$ onto the line defined by $\operatorname{span}(\{\vec{v}\})$, where $\zeta : S^2_{\mathbb{F}} \mapsto \mathbb{F}$ whose function rule is

$$\zeta(\vec{x}, \vec{y}) = \frac{\vec{x} \cdot \vec{y}}{\|\vec{y}\|} \tag{4.6}$$

and $S_{\mathbb{F}}$ defines an affine normed inner product space. Considering $\vec{w} := \zeta(\vec{x}, \vec{y})\vec{y}$, we can state that vector $\vec{w} \in \operatorname{span}(\{\vec{y}\})$ where points w and y belong to the line defined by $\operatorname{span}(\{\vec{y}\})$. From the previous rule, if projection is known and not the inner product, we can write that $\vec{x} \cdot \vec{y} = \zeta(\vec{x}, \vec{y})\|\vec{y}\|$, which is usually and imprecisely regarded as the

[2]See page 35.

"geometric definition" of the inner product. It is important to note that if arguments $\vec{x}$ and $\vec{y}$ are orthogonal, then $\zeta(\vec{x},\vec{y})=0$. Moreover, in the context of affine Hilbert spaces, where $\|\vec{x}\|^2=\vec{x}\cdot\vec{x}$, when $\vec{y}=\alpha\vec{x}$ then

$$\zeta(\vec{x},\alpha\vec{x})=\frac{\overline{\alpha}(\vec{x}\cdot\vec{x})}{|\alpha|\|\vec{x}\|}=\frac{\overline{\alpha}\|\vec{x}\|}{|\alpha|},\forall\alpha\in\{\mathbb{F}\setminus 0\}. \tag{4.7}$$

For the case of affine Euclidean spaces, these equalities lead to

$$\zeta(\vec{x},\alpha\vec{x})=\operatorname{sgn}(\alpha)\,\|\vec{x}\|,\ \forall\,\alpha\in\{\mathbb{R}\setminus 0\}. \tag{4.8}$$

Concerning projections on affine Hilbert spaces, there is a property, presented on the following theorem, that will be important for future geometric concepts.

Theorem 17 – Modulus of Projection

Given the arbitrary vectors $\vec{x},\vec{y}\in S_{\mathbb{F}}$, where $S_{\mathbb{F}}$ defines an affine Hilbert space $\mathcal{S}^n_{\mathbb{F}}$, the modulus of projection $\zeta(\vec{x},\vec{y})$ is never greater than the norm of the projected vector. In other words, $|\zeta(\vec{x},\vec{y})|\leqslant\|\vec{x}\|$.

Proof. From the Cauchy-Schwartz Inequality (2.7) and the "geometric definition" of the inner product, the following development proves the theorem.

$$\begin{aligned}|\vec{x}\cdot\vec{y}|^2&\leqslant(\vec{x}\cdot\vec{x})(\vec{y}\cdot\vec{y})\\|\zeta(\vec{x},\vec{y})|^2\|\vec{y}\|^2&\leqslant\|\vec{x}\|^2\|\vec{y}\|^2\\|\zeta(\vec{x},\vec{y})|&\leqslant\|\vec{x}\|.\end{aligned}$$

□

Now, for an affine Euclidean space $\mathcal{S}^n_{\mathbb{R}}$, if θ is the smallest angle defined by vectors $\vec{x},\vec{y}\in S_{\mathbb{R}}$, the linear combination $\vec{z}:=\vartheta\vec{y}/\|\vec{y}\|$ and the scalar $\vartheta\in\mathbb{R}$ are called respectively the **vector and scalar geometric projections** of $\vec{x}$ onto $\vec{y}$ when

$$\vartheta:=\|\vec{x}\|\cos\theta, \tag{4.9}$$

which can be visualized in figure 4.5. In this context, the following corollary establishes a connection between geometric and algebraic projections. From this connection and rule (4.6), the so called "geometric definition" of the Euclidean inner product results the classic expression

$$\vec{x}\cdot\vec{y}=\|\vec{x}\|\|\vec{y}\|\cos\theta. \tag{4.10}$$

From this equality and the scalar geometric projection ϑ of $\vec{x}$ onto $\vec{y}$, then

$$\zeta(\vec{x},\vec{y})=\frac{\vec{x}\cdot\vec{y}}{\|\vec{y}\|}=\frac{\|\vec{x}\|\|\vec{y}\|\cos\theta}{\|\vec{y}\|}=\|\vec{x}\|\cos\theta=\vartheta. \tag{4.11}$$

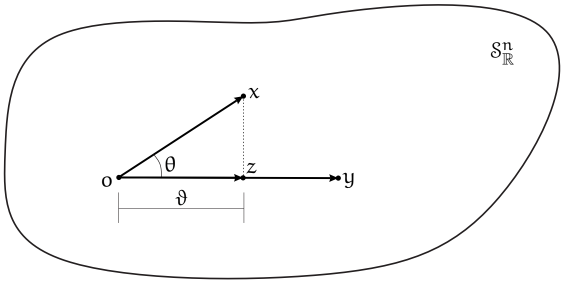

Figure 4.5 – Vector and scalar geometric projections.

From the previous equality and definition (4.9), if there is no geometric projection between $\vec{x}$ and $\vec{y}$, that is, $\vartheta = 0$, we conclude that the angle $\theta = \pi(k+1/2)$, $k = 0, 1, \ldots$ Thereby, if line $(\mathcal{W}_b)^1_{\mathbb{R}}$ and plane $(\mathcal{V}_a)^2_{\mathbb{R}}$, subspaces of affine Euclidean space $\mathcal{S}^5_{\mathbb{R}}$, are perpendicular, we can represent them as depicted on figure 4.6.

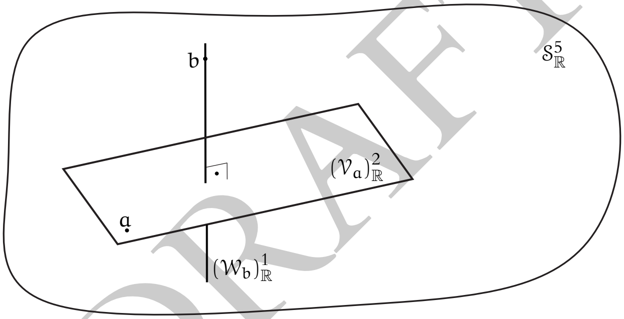

Figure 4.6 – Line and plane perpendicular to each other.

4.2 Affinities

The subject of Geometry is to study not only morphological sets and their elements, as we have done on the previous section, but also the correlation of geometrical objects, that is, functions whose domain and image are morphological sets. Thereby, in the context of Affine Geometry, this section deals exclusively with affinities, here understood as "geometric" functions that correlate points, defined by linear vector functions on Hilbert spaces. In more precise terms, given a Hilbert affine space $\mathcal{U}^m_{\mathbb{F}}$, a bijective mapping $g : \mathcal{U}^m_{\mathbb{F}} \mapsto \mathcal{U}^m_{\mathbb{F}}$ and a vector space $\mathcal{L}_{\mathbb{F}}(\mathsf{U}, \mathsf{U})$, we call $f : \mathcal{U}^m_{\mathbb{F}} \mapsto \mathcal{U}^m_{\mathbb{F}}$ an **affine transformation** and its function an **affinity** when there is a unique linear bijection $\mathbf{f} \in \mathcal{L}_{\mathbb{F}}(\mathsf{U}, \mathsf{U})$ where equality

$$f(\mathbf{x} \oplus a) = \mathbf{f}(\mathbf{x}) \oplus g(a)\,,\ \forall\, a \in \mathcal{U}^m_{\mathbb{F}}\,, \mathbf{x} \in \mathsf{U}_{\mathbb{F}}\,, \tag{4.12}$$

is valid. It is straightforward to note that since $\mathbf{f}$ and g are bijections, so results affinity f. Among other properties, it is possible to obtain that every affinity preserves parallelism and dimensionality; in other words, parallel subspaces on the domain remain parallel on the image and n-dimensional subspaces on the domain, $n \leqslant m$, remain n-dimensional on the image: for example, lines are mapped to lines, planes are mapped to planes and so on. From corollary 12.2 and equalities (3.51), we know that operator $\mathbf{f}$ is the unique representative function of a tensor $\mathbf{T} \in \mathcal{L}_{\mathbb{F}}(U^2)$ in such a way that the previous equality can be rewritten as

$$f(\boldsymbol{x} \oplus a) = (\mathbf{L} \hat{\odot}_1 \boldsymbol{x}^*) \oplus g(a), \; \forall a \in \mathcal{U}_{\mathbb{F}}^m, \boldsymbol{x} \in U_{\mathbb{F}}, \tag{4.13}$$

where $\mathbf{L} := \mathbf{T}^\dagger$ is called the **affinity tensor** of f. Sometimes it is very useful to "vectorize" affinity f in order to work only with vectors. In this sense, vector function $\boldsymbol{f} := \upsilon_o \circ f \circ \upsilon_o^{-1}$ is said to be a **vector affinity** and then, from the definition of the "vectorization" function and its properties, it is straightforward to obtain that

$$\boldsymbol{f}(\mathbf{x} + \vec{a}) = (\mathbf{L} \hat{\odot}_1 \mathbf{x}^*) + \overrightarrow{g(a)}. \tag{4.14}$$

Proof. Let's prove that an affinity preserves parallelism and dimension of subspaces. Considering $(\mathcal{S}_a)_{\mathbb{F}}^n$ and $(\mathcal{V}_b)_{\mathbb{F}}^r$ parallel subspaces of $\mathcal{U}_{\mathbb{F}}^m$, where $S_{\mathbb{F}} \subseteq V_{\mathbb{F}}$, we can say that if $\{\mathbf{u}_1, \cdots, \mathbf{u}_n, \cdots, \mathbf{u}_r, \cdots, \mathbf{u}_m\}$ is a basis of $U_{\mathbb{F}}$, so are $\{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ and $\{\mathbf{u}_1, \cdots, \mathbf{u}_n, \cdots, \mathbf{u}_r\}$ bases of $S_{\mathbb{F}}$ and $V_{\mathbb{F}}$ respectively. From definition (4.12), function $\mathbf{f}$ is the sole responsible for preserving parallelism or not. Thereby, since $\mathbf{f}$ is a bijective linear unary operator, we already know that $\{\mathbf{f}(\mathbf{u}_1), \cdots, \mathbf{f}(\mathbf{u}_n)\}$ and $\{\mathbf{f}(\mathbf{u}_1), \cdots, \mathbf{f}(\mathbf{u}_n), \cdots, \mathbf{f}(\mathbf{u}_r)\}$ are bases of $S_{\mathbb{F}}$ and $V_{\mathbb{F}}$ (see p. 43) and then $S_{\mathbb{F}} \subseteq V_{\mathbb{F}}$ is preserved. From these same arguments, it is straightforward to conclude that $\mathbf{f}$ does not change vector space dimensions. □

Considering previous conditions, an affinity f is said to be centered at point a, represented by f_a, when g is the identity function, that is, when point $f_a(\mathbf{x} \oplus a) = \mathbf{f}(\mathbf{x}) \oplus a$. From this equality, when $\mathbf{f}(\mathbf{x}) = \mathbf{u} + \mathbf{x}$, we have $f_a(\mathbf{x} \oplus a) = \mathbf{u} \oplus (\mathbf{x} \oplus a)$ and then affinity f_a is called a **translation**, represented by $t_{\mathbf{u}}$, described by the rule

$$t_{\mathbf{u}}(x) = \mathbf{u} \oplus x \tag{4.15}$$

and depicted in figure 4.7, where $t_{\mathbf{u}}$ preserves the shape of the subspace $\mathcal{S}_{\mathbb{F}}^n$. Given an arbitrary vector $\boldsymbol{v} \in U_{\mathbb{F}}$, from equality $t_{\boldsymbol{v}} \circ t_{\mathbf{u}}(x) = (\boldsymbol{v} + \mathbf{u}) \oplus x$, it is straightforward to conclude that $t_{\boldsymbol{v}} \circ t_{\mathbf{u}} = t_{\boldsymbol{v}+\mathbf{u}}$. Moreover, if $\boldsymbol{v} = -\mathbf{u}$ we can affirm that $t_{-\mathbf{u}} = t_{\mathbf{u}}^{-1}$ and translation t_0 is the identity function. Previous equality $t_{\boldsymbol{v}} \circ t_{\mathbf{u}} = t_{\boldsymbol{v}+\mathbf{u}}$ also allows us to conclude that the set T of all translations on $\mathcal{U}_{\mathbb{F}}^m$ defines an abelian group $(T, \circ)$ because the composition of translations clearly observes the axioms of

i. associativity, where $t_{\mathbf{u}} \circ (t_{\boldsymbol{v}} \circ t_{\boldsymbol{w}}) = (t_{\mathbf{u}} \circ t_{\boldsymbol{v}}) \circ t_{\boldsymbol{w}}, \forall \mathbf{u}, \boldsymbol{v}, \boldsymbol{w} \in U_{\mathbb{F}}$,

ii. identity element, where $t_{\mathbf{u}} \circ t_0 = t_0 \circ t_{\mathbf{u}} = t_{\mathbf{u}}, \forall \mathbf{u} \in U_{\mathbb{F}}$,

iii. inverse element, where $t_{\mathbf{u}} \circ t_{-\mathbf{u}} = t_{-\mathbf{u}} \circ t_{\mathbf{u}} = t_0$, $\forall \mathbf{u} \in \{U_{\mathbb{F}} \setminus 0\}$, and of

iv. commutativity, where $t_{\mathbf{u}} \circ t_{\boldsymbol{v}} = t_{\boldsymbol{v}} \circ t_{\mathbf{u}}, \forall \mathbf{u}, \boldsymbol{v} \in U_{\mathbb{F}}$.

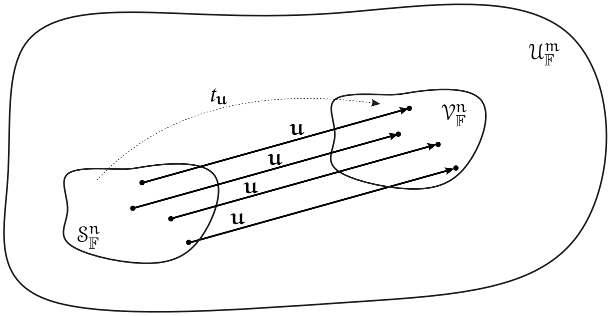

Figure 4.7 – Hyperplane $\mathcal{S}_{\mathbb{F}}^n$ is mapped to $\mathcal{V}_{\mathbb{F}}^n$ by translation $t_{\mathbf{u}}$.

According to (4.14), the correspondent vector affinity of translation $t_{\mathbf{u}}$ is called a **vector translation**, being described by equality

$$\boldsymbol{t}_{\mathbf{u}}(\overrightarrow{x}) = \mathbf{u} + \overrightarrow{x}\,, \tag{4.16}$$

from which it is clear that vector function $\boldsymbol{t}_{\mathbf{u}}$ is linear, that is, $\boldsymbol{t}_{\mathbf{u}} \in \mathcal{L}_{\mathbb{F}}(\mathrm{U},\mathrm{U})$.

Still considering previous conditions, the affinity in $s : \mathcal{U}_{\mathbb{F}}^m \mapsto \mathcal{U}_{\mathbb{F}}^m$ is called a **dilation** and $\boldsymbol{s} := \upsilon_o \circ h \circ \upsilon_o^{-1}$ a **vector dilation** if bijection $\mathbf{h} \in \mathcal{L}_{\mathbb{F}}(\mathrm{U},\mathrm{U})$ in equality

$$s(\mathbf{x} \oplus \mathrm{a}) = \mathbf{s}(\mathbf{x}) \oplus g(\mathrm{a}) \tag{4.17}$$

or in equality

$$\boldsymbol{s}(\mathbf{x} + \overrightarrow{\mathrm{a}}) = \mathbf{s}(\mathbf{x}) + \overrightarrow{g(\mathrm{a})} \tag{4.18}$$

is a positive-definite Hermitian operator, usually called a **stretch operator**, representing, according to (3.56) and (3.58), a positive-definite tensor $\mathbf{S} = \mathbf{S}^{\dagger}$, called **stretch tensor**, which is the affinity tensor of s. In this context and by considering the conditions of the bases in (3.70), from which equality $[\mathbf{S}]^{\mathrm{B}} = [\mathbf{s}_{\mathrm{B}_2}]^{\mathrm{B}_1}$ can be concluded, we can write that the eigenvalues

$$\lambda(\mathbf{S}) := \lambda(\mathbf{s}) = \lambda_1, \cdots, \lambda_m\,. \tag{4.19}$$

We already know from section 2.5 that these eigenvalues, called **stretch coefficients**, are all positive real scalars. Thereby, if coefficient $\lambda_i > 1$ or $\lambda_i \leqslant 1$, it is called an **expansion** or a **contraction** respectively and in the case of proportional stretch, where $\lambda_1 = \cdots = \lambda_m$, function $\mathbf{s}$ (or tensor $\mathbf{S}$) is equally called an expansion or contraction operator(tensor). It is important to note that, from corollary 12.3, in the specific context of Euclidean spaces, where $\mathbf{s}^{\dagger} = \mathbf{s}^{\mathrm{T}}$, the stretch tensor $\mathbf{S}$ is symmetric. As any other affinity, a dilation s can be centered at point a, represented by s_{a}, where $s_{\mathrm{a}}(\mathbf{u} \oplus \mathrm{a}) = \mathbf{s}(\mathbf{u}) \oplus \mathrm{a}$. Figure 4.8 shows different types of dilations. Given two arbitrary dilations s_1 and s_2, equalities

$$s_1 \circ s_2(\mathbf{x} \oplus \mathrm{a}) = s_1(\mathbf{s}_2(\mathbf{x}) \oplus g_2(\mathrm{a})) = \mathbf{s}_1 \circ \mathbf{s}_2(\mathbf{x}) \oplus g_1 \circ g_2(\mathrm{a})$$

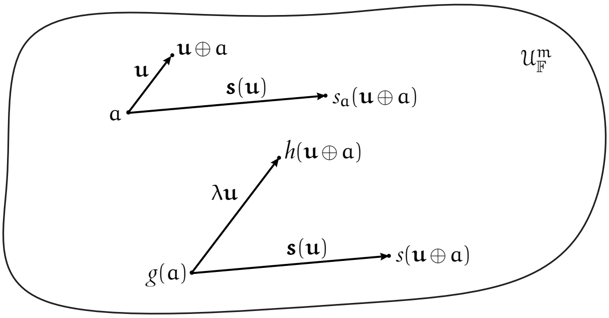

Figure 4.8 – Affinities s, s_a and h are ordinary, centered and proportional dilations.

prove that $s_1 \circ s_2$ is also a dilation and that $s_1^{-1}(\mathbf{x}\oplus a) = \mathbf{s}_1^{-1}(\mathbf{x}) \oplus g_1^{-1}(a)$. Considering the identity dilation i, similarly to the case of translations, it is straightforward to verify that dilations also obey the axioms of associativity, identity element, inverse element and commutativity on the operation of composition. Therefore, the set $\mathcal{S}$ of all dilations on $\mathcal{U}^m_{\mathbb{F}}$ defines the abelian group $(\mathcal{S}, \circ)$.

Still considering previous conditions, an affinity r is called **isometric** when its correspondent linear operator $\mathbf{r}$ is isometric, that is, $\mathbf{r} \in \mathcal{I}_{\mathbb{F}}(\mathrm{U})$. In the context of Euclidean spaces, we already know that an isometric operator is also an orthogonal operator[3], that is, the isometric group $\mathcal{I}_{\mathbb{R}}(\mathrm{U})$ equals the orthogonal group $\mathcal{O}_{\mathbb{R}}(\mathrm{U})$, whose elements preserve inner product and have determinant ± 1. The centered isometric affinity r_a, described by $r_a(\mathbf{x}\oplus a) = \mathbf{r}(\mathbf{x}) \oplus a$, is called a **rotation** if isometric operator $\mathbf{r}$ is proper orthogonal, or $\mathbf{r} \in \mathcal{O}^+{}_{\mathbb{R}}(\mathrm{U})$; otherwise, it is called a **rotoreflection**. Considering that $\mathbf{r}$ represents $\mathbf{T}$, the affinity tensor $\mathbf{R} := \mathbf{T}^{\mathrm{T}}$ of a rotation is called a **rotation tensor** and, similarly, the affinity tensor $\bar{\mathbf{R}}$ of a rotoreflection is called a **rotoreflection tensor**.

Theorem 18 – Rotodilation Function Decompositions

Let a linear bijection $\mathbf{f}$ be polar decomposed by orthogonal operator $\mathbf{r} \in \mathcal{L}_{\mathbb{R}}(\mathrm{U},\mathrm{U})$ and positive-definite symmetric operators $\mathbf{s}_1, \mathbf{s}_2 \in \mathcal{L}_{\mathbb{R}}(\mathrm{U},\mathrm{U})$ where $\mathbf{f} = \mathbf{r} \circ \mathbf{s}_1^{1/2} = \mathbf{s}_2^{1/2} \circ \mathbf{r}$ and $\mathrm{U}_{\mathbb{R}}$ is an m-dimensional Euclidean space. Given an affine transformation $f : \mathcal{U}^m_{\mathbb{R}} \mapsto \mathcal{U}^m_{\mathbb{R}}$ described by $f(\mathbf{x}\oplus a) = \mathbf{f}(\mathbf{x}) \oplus g(a)$, affinity

$$f = r_{g(a)} \circ s_1 = s_2 \circ r_{g(a)} \,,$$

where $r_{g(a)}, s_1, s_2$ are respectively a rotation (rotoreflection) defined by orthogonal operator $\mathbf{r}$ and dilations defined by stretch operators $\mathbf{s}_1^{1/2}$ and $\mathbf{s}_2^{1/2}$.

[3]See p. 51.

Proof. By the right polar decomposition, the development

$$f(\mathbf{x}\oplus\mathbf{a}) = \mathbf{f}(\mathbf{x})\oplus g(\mathbf{a}) = \mathbf{r}\circ\mathbf{s}_1^{1/2}(\mathbf{x})\oplus g(\mathbf{a}) = r_{g(\mathbf{a})}[\mathbf{s}_1^{1/2}(\mathbf{x})\oplus g(\mathbf{a})] = r_{g(\mathbf{a})}\circ s_1(\mathbf{x}\oplus\mathbf{a})$$

proves the theorem for s_1. Proof for s_2 is similar. □

Corollary 18.1 – Rotodilation Tensor Decompositions

If $\mathbf{L}\in\mathcal{L}_{\mathbb{R}}(\mathrm{U}^2)$ is an affinity tensor then $\mathbf{L}=\mathbf{R}\odot_1\mathbf{S}_1=\mathbf{S}_2\odot_1\mathbf{R}$, where $\mathbf{R}$ is a rotation (rotoreflection) tensor and $\mathbf{S}_1$, $\mathbf{S}_2$ are stretch tensors with the same stretch coefficients.

Proof. According to (4.13), development

$$\begin{aligned} f(\mathbf{x}\oplus\mathbf{a}) &= r_{g(\mathbf{a})}\circ s_1(\mathbf{x}\oplus\mathbf{a}) \\ &= r_{g(\mathbf{a})}[(\mathbf{S}_1\hat{\odot}_1\mathbf{x}^*)\oplus g(\mathbf{a})] \\ &= \mathbf{R}\hat{\odot}_1(\mathbf{S}_1\odot_1\mathbf{x}^*)\oplus g(\mathbf{a}) \\ &= [(\mathbf{R}\odot_1\mathbf{S}_1)\hat{\odot}_1\mathbf{x}^*]\oplus g(\mathbf{a}) \end{aligned}$$

proves the theorem for $\mathbf{S}_1$. Proof for $\mathbf{S}_2$ is similar. Considering a tensor basis B, from the rotodilation decomposition we just proved, equality $[\mathbf{R}]^{\mathrm{B}}[\mathbf{S}_1]^{\mathrm{B}}=[\mathbf{S}_2]^{\mathrm{B}}[\mathbf{R}]^{\mathrm{B}}$ is valid, from which we conclude that matrices $[\mathbf{S}_i]^{\mathrm{B}}$ are similar through $\mathbf{R}$ and therefore isospectral, according to theorem 2. □

Now, recalling the concept of unimodular tensors presented in section 3.2, let $(\mathrm{U}_{\mathbb{R}},\mathbf{A}_{\mathrm{B}})$ be an m-dimensional oriented Euclidean space where $\mathbf{A}_{\mathrm{B}}\in\mathcal{LA}_{\mathbb{R}}^{(m)}(\mathrm{U}^m)$ and $\mathrm{B}=\{\hat{\mathbf{u}}_1,\cdots,\hat{\mathbf{u}}_m\}$. Considering r_a a rotation and $\bar{r}_a$ a rotoreflection, since every pair of Euclidean orthonormal basis can be related through an orthogonal operator[4], we conclude that rotation $\mathrm{C}=\{\mathbf{r}(\hat{\mathbf{u}}_1),\cdots,\mathbf{r}(\hat{\mathbf{u}}_m)\}$ and rotoreflection $\overline{\mathrm{C}}=\{\bar{\mathbf{r}}(\hat{\mathbf{u}}_1),\cdots,\bar{\mathbf{r}}(\hat{\mathbf{u}}_m)\}$ of basis B are positively and negatively oriented bases respectively because equalities $\mathbf{A}_{\mathrm{C}}(\hat{\mathbf{u}}_1,\cdots,\hat{\mathbf{u}}_m)=-\mathbf{A}_{\overline{\mathrm{C}}}(\hat{\mathbf{u}}_1,\cdots,\hat{\mathbf{u}}_m)=1$ are valid according to definition (3.16) and equality (3.18). In order to geometrically represent in a simple way basis rotation and rotoreflection, we specify the affine Euclidean space in question to be two-dimensional, that is, $m=2$, and rotoreflection $\bar{\mathbf{r}}=-\mathbf{r}$. Thereby, for the basis vectors $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$ the smallest angle θ between them leads to equalities $0=\hat{\mathbf{u}}_1\cdot\hat{\mathbf{u}}_2=\|\hat{\mathbf{u}}_1\|\|\hat{\mathbf{u}}_2\|\cos\theta=\cos\theta$, and since orthogonal functions preserve inner product, then θ is also the smallest angle between $\mathbf{r}(\hat{\mathbf{u}}_1)$ and $\mathbf{r}(\hat{\mathbf{u}}_2)$, or between $\bar{\mathbf{r}}(\hat{\mathbf{u}}_1)$ and $\bar{\mathbf{r}}(\hat{\mathbf{u}}_2)$, because the following equalities are all valid: $\hat{\mathbf{u}}_1\cdot\hat{\mathbf{u}}_2=\mathbf{r}(\hat{\mathbf{u}}_1)\cdot\mathbf{r}(\hat{\mathbf{u}}_2)=\bar{\mathbf{r}}(\hat{\mathbf{u}}_1)\cdot\bar{\mathbf{r}}(\hat{\mathbf{u}}_2)=0$. From these angle and size preserving features, we can represent an anticlockwise acute angle ϕ rotation and rotoreflection of basis B according to figure 4.9. We can verify that the function on the right of this figure is indeed a rotation and on the left a rotoreflection by proving that the rotated and reflected bases are positively and negatively oriented respectively in relation to $(\mathrm{U}_{\mathbb{R}},\mathbf{A}_{\mathrm{B}})$, that is, by calculating $\mathbf{A}_{\mathrm{C}}(\hat{\mathbf{u}}_1,\hat{\mathbf{u}}_2)$ and $\mathbf{A}_{\overline{\mathrm{C}}}(\hat{\mathbf{u}}_1,\hat{\mathbf{u}}_2)$. From definition (3.16) and

[4]See p. 50.

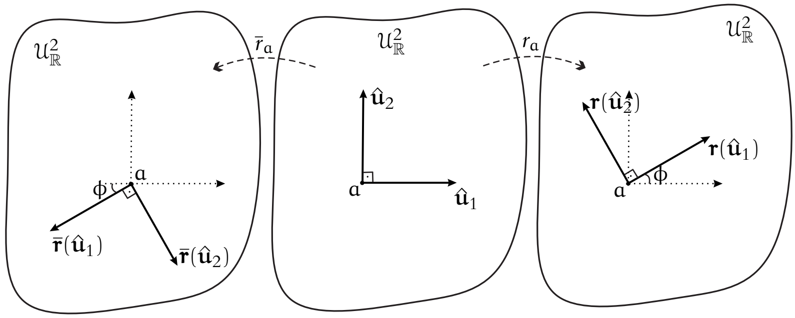

Figure 4.9 – Two-dimensional rotoreflexion and rotation.

equality (4.10), development

$$\begin{aligned}
\mathbf{A}_C(\hat{\mathbf{u}}_1,\hat{\mathbf{u}}_2) &= \sum_{i=1}^{2}\sum_{j=1}^{2} \mathbf{f}_i^C(\hat{\mathbf{u}}_1)\mathbf{f}_j^C(\hat{\mathbf{u}}_2)\epsilon_{ij} \\
&= (\hat{\mathbf{u}}_1\cdot\mathbf{r}(\hat{\mathbf{u}}_1))(\hat{\mathbf{u}}_2\cdot\mathbf{r}(\hat{\mathbf{u}}_2)) - (\hat{\mathbf{u}}_1\cdot\mathbf{r}(\hat{\mathbf{u}}_2))(\hat{\mathbf{u}}_2\cdot\mathbf{r}(\hat{\mathbf{u}}_1)) \\
&= (\cos\phi\cos\phi) - (-\sin\phi\sin\phi) = 1
\end{aligned}$$

proves that C is positively oriented and, similarly, $\overline{C}$ results negatively oriented.

Still considering the oriented Euclidean space $(U_{\mathbb{R}},\mathbf{A}_B)$ where $B=\{\hat{\mathbf{u}}_1,\cdots,\hat{\mathbf{u}}_m\}$, the **simple bivector** z defined by vectors $\mathbf{x},\mathbf{y}\in U_{\mathbb{R}}$ is here understood as the mathematical object, endowed with magnitude and orientation, used to express rotational physical quantities such as angular momentum and torque. In geometrical terms, the magnitude of this simple bivector z is defined to be the area of the parallelogram constructed from line segments $\overline{a\mathbf{x}\oplus a}$ and $\overline{a\mathbf{y}\oplus a}$ while its orientation is related to $\mathbf{A}_B$ and to rotating $\overline{a\mathbf{x}\oplus a}$ to $\overline{a\mathbf{y}\oplus a}$, which is the inverse of rotating $\overline{a\mathbf{y}\oplus a}$ to $\overline{a\mathbf{x}\oplus a}$ for the case of the simple bivector $-z$. It is important to say that this new entity is here loosely defined because it is a particular case of a more general concept called multivector, which is out of the scope of this elementary study. Nevertheless, to our good fortune, this simple bivector can be represented as a vector if $\dim(U_{\mathbb{R}})=3$, when it is called an **axial vector**. In this context, considering an axial vector $\mathbf{z}$ defined by the simple bivector z, where $\{\mathbf{x},\mathbf{y}\}$ is linearly independent, line $(\{\mathbf{z}\}_a)^1_{\mathbb{R}}$ is defined to be perpendicular to the plane $(\{\mathbf{x},\mathbf{y}\}_a)^2_{\mathbb{R}}$ and vector $\overrightarrow{a\mathbf{z}\oplus a}$ must be oriented in such a way that basis $\{\mathbf{z},\mathbf{x},\mathbf{y}\}$ is positively oriented, that is, $\mathbf{A}_B(\mathbf{z},\mathbf{x},\mathbf{y})>0$. There is a classic mnemonic rule for obtaining the direction of $\overrightarrow{a\mathbf{z}\oplus a}$ that works as follows: considering my right hand, I must point my index and middle fingers to the same directions of $\overrightarrow{\hat{\mathbf{u}}_1\oplus o}$ and $\overrightarrow{\hat{\mathbf{u}}_2\oplus o}$ respectively, and then if my thumb points to the same direction of $\overrightarrow{\hat{\mathbf{u}}_3\oplus o}$, then space $(U_{\mathbb{R}},\mathbf{A}_B)$ is oriented according to the **right-hand rule**; on the other hand, literally, space $(U_{\mathbb{R}},\mathbf{A}_B)$ is oriented according to the **left-hand rule**. Once a hand is found, if I point my index and middle fingers to the same directions of $\overrightarrow{a\mathbf{x}\oplus a}$ and $\overrightarrow{a\mathbf{y}\oplus a}$ respec-

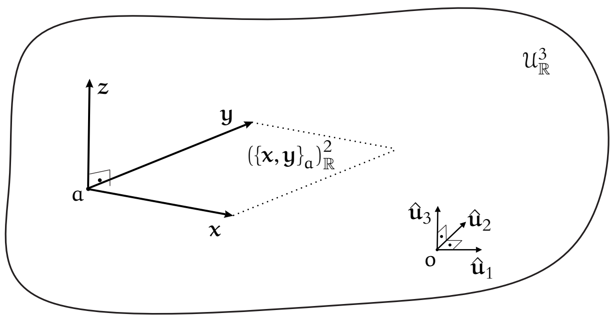

Figure 4.10 – Axial vector z defined by x, y and oriented Euclidean space $(\mathrm{U}_{\mathbb{R}}, \mathbf{A}_{\mathrm{B}})$, where basis $\mathrm{B} = \{\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2, \hat{\mathbf{u}}_3\}$.

tively, then the direction of my thumb defines the direction of $\overrightarrow{az \oplus a}$. For practical purposes, a three-dimensional Euclidean affine space is usually oriented according to the right-hand rule. An axial vector is always the result of a **cross product** of two vectors, that is, $z = x \times y$ in the present case, where

$$x \times y := \mathbf{A}_{\mathrm{B}} \hat{\odot}_2 (x^* \otimes y^*). \tag{4.20}$$

For arbitrary vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{k} \in \mathrm{U}_{\mathbb{R}}$ and scalars $\alpha, \beta, \theta \in \mathbb{R}$, where θ is the smallest angle defined by vectors $\vec{u}, \vec{v}$, the cross product has the following properties.

i. Sum: $\mathbf{u} \times \mathbf{v} + \mathbf{w} \times \mathbf{k}$ is an axial vector;

ii. Mixed products: $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \mathbf{A}_{\mathrm{B}}(\mathbf{u}, \mathbf{v}, \mathbf{w})$;

iii. Anticommutativity: $\mathbf{u} \times \mathbf{v} = -(\mathbf{v} \times \mathbf{u})$;

iv. Self cross product: $\mathbf{u} \times \mathbf{u} = \mathbf{0}$;

v. Distributivity over addition: $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = (\mathbf{u} \times \mathbf{v}) + (\mathbf{u} \times \mathbf{w})$;

vi. Scalar multiplication: $\alpha\beta(\mathbf{u} \times \mathbf{v}) = (\alpha\mathbf{u}) \times (\beta\mathbf{v}) = (\beta\mathbf{u}) \times (\alpha\mathbf{v})$;

vii. Positive orientation: $\mathbf{A}_{\mathrm{B}}(\mathbf{u} \times \mathbf{v}, \mathbf{u}, \mathbf{v}) > 0$;

viii. Orthogonality: $(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{u} = (\mathbf{u} \times \mathbf{v}) \cdot \mathbf{v} = 0$;

ix. Triple cross product: $\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{w}$;

x. Area of parallelogram: $\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \|\mathbf{v}\| \sin\theta$.

Proof. Considering $x^* \otimes y^* = \mathbf{u}^* \otimes \mathbf{v}^* + \mathbf{w}^* \otimes \mathbf{k}^*$, we prove item i through the following equalities:

$$\mathbf{u} \times \mathbf{v} + \mathbf{w} \times \mathbf{k} = \mathbf{A}_{\mathrm{B}} \hat{\odot}_2 (\mathbf{u}^* \otimes \mathbf{v}^*) + \mathbf{A}_{\mathrm{B}} \hat{\odot}_2 (\mathbf{w}^* \otimes \mathbf{k}^*) = \mathbf{A}_{\mathrm{B}} \hat{\odot}_2 (x^* \otimes y^*) = x \times y\,.$$

Item ii is verified by

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \mathbf{u} \cdot [\mathbf{A}_{\mathrm{B}} \hat{\odot}_2 (\mathbf{v}^* \otimes \mathbf{w}^*)] = \mathbf{A}_{\mathrm{B}} \odot_3 (\mathbf{u}^* \otimes \mathbf{v}^* \otimes \mathbf{w}^*) = \mathbf{A}_{\mathrm{B}}(\mathbf{u}, \mathbf{v}, \mathbf{w})\,.$$

Given an arbitrary vector $x \in \mathrm{U}_{\mathbb{R}}$ and $\mathbf{A}_{\mathrm{B}} = \mathbf{a}^* \otimes \mathbf{b}^* \otimes \mathbf{c}^*$, item iii is verified by the following equalities:

$$x \cdot (\mathbf{u} \times \mathbf{v}) = (\mathbf{b} \cdot \mathbf{u})(\mathbf{c} \cdot \mathbf{v})(x \cdot \mathbf{a}) = \mathbf{A}_{\mathrm{B}}(x, \mathbf{u}, \mathbf{v}) = -\mathbf{A}_{\mathrm{B}}(x, \mathbf{v}, \mathbf{u}) = -x \cdot (\mathbf{v} \times \mathbf{u})\,.$$

Item iii is a straightforward corollary of item iii. We prove item v similarly:

$$\mathbf{x}\cdot[\mathbf{u}\times(\mathbf{v}+\mathbf{w})]=\mathbf{A}_B(\mathbf{x},\mathbf{u},\mathbf{v}+\mathbf{w})=\mathbf{A}_B(\mathbf{x},\mathbf{v},\mathbf{u})+\mathbf{A}_B(\mathbf{x},\mathbf{v},\mathbf{w})=\mathbf{x}\cdot[(\mathbf{v}\times\mathbf{u})+(\mathbf{v}\times\mathbf{w})].$$

By this same procedure, proof of property vi is trivial. Now, property vii is verified by expression

$$\mathbf{A}_B(\mathbf{u}\times\mathbf{v},\mathbf{u},\mathbf{v})=\mathbf{A}_B((\mathbf{b}\cdot\mathbf{u})(\mathbf{c}\cdot\mathbf{v})\mathbf{a},\mathbf{u},\mathbf{v})=(\mathbf{a}\cdot\mathbf{a})^2(\mathbf{b}\cdot\mathbf{u})^2(\mathbf{c}\cdot\mathbf{v})^2>0$$

and, since $\mathbf{A}_B$ is antisymmetric, equalities $(\mathbf{u}\times\mathbf{v})\cdot\mathbf{u}=\mathbf{A}_B(\mathbf{u},\mathbf{u},\mathbf{v})=0$ prove item viii. Considering an arbitrary orthonormal basis $X=\{\hat{\mathbf{x}}_1,\hat{\mathbf{x}}_2,\hat{\mathbf{x}}_3\}$, equality (3.19) and identity (1.36), development

$$\begin{aligned}
\mathbf{u}\times(\mathbf{v}\times\mathbf{w})&=\mathbf{A}_B\hat{\odot}_2[\mathbf{u}^*\otimes\mathbf{A}_B\hat{\odot}_2(\mathbf{v}^*\otimes\mathbf{w}^*)]\\
&=\sum_{i,j,k=1}^{m}\epsilon_{ijk}\hat{\mathbf{x}}_i^*\otimes\hat{\mathbf{x}}_j^*\otimes\hat{\mathbf{x}}_k^*\hat{\odot}_2[\sum_{j=1}^{m}\mathbf{f}_j^X(\mathbf{u})\hat{\mathbf{x}}_j\otimes\sum_{k,r,s=1}^{m}\epsilon_{krs}\mathbf{f}_r^X(\mathbf{v})\mathbf{f}_s^X(\mathbf{w})\hat{\mathbf{x}}_k]\\
&=\sum_{i,j,k=1}^{m}\sum_{r,s=1}^{m}-\epsilon_{kji}\epsilon_{krs}\mathbf{f}_j^X(\mathbf{u})\mathbf{f}_r^X(\mathbf{v})\mathbf{f}_s^X(\mathbf{w})\hat{\mathbf{x}}_i\\
&=\sum_{i,j,k,r,s=1}^{m}(\delta_{js}\delta_{ir}-\delta_{jr}\delta_{is})\mathbf{f}_j^X(\mathbf{u})\mathbf{f}_r^X(\mathbf{v})\mathbf{f}_s^X(\mathbf{w})\hat{\mathbf{x}}_i\\
&=\sum_{i,j,k,r,s=1}^{m}\mathbf{f}_s^X(\mathbf{u})\mathbf{f}_i^X(\mathbf{v})\mathbf{f}_s^X(\mathbf{w})\hat{\mathbf{x}}_i-\mathbf{f}_r^X(\mathbf{u})\mathbf{f}_r^X(\mathbf{v})\mathbf{f}_i^X(\mathbf{w})\hat{\mathbf{x}}_i\\
&=(\mathbf{u}\cdot\mathbf{w})\mathbf{v}-(\mathbf{u}\cdot\mathbf{v})\mathbf{w}
\end{aligned}$$

proves property ix. Now, in order to prove the last item, equality $\mathbf{u}\cdot\mathbf{A}_B\hat{\odot}_2(\mathbf{v}^*\otimes\mathbf{w}^*)=\mathbf{A}_B(\mathbf{u},\mathbf{v},\mathbf{w})$, property x and equality (4.10) are required. Then,

$$\begin{aligned}
\|\mathbf{u}\times\mathbf{v}\|^2&=(\mathbf{u}\times\mathbf{v})\cdot(\mathbf{u}\times\mathbf{v})\\
&=\mathbf{A}_B(\mathbf{u}\times\mathbf{v},\mathbf{u},\mathbf{v})\\
&=-\mathbf{A}_B(\mathbf{v},\mathbf{u},\mathbf{u}\times\mathbf{v})\\
&=-\mathbf{v}\cdot(\mathbf{u}\times(\mathbf{u}\times\mathbf{v}))\\
&=\mathbf{v}\cdot[(\mathbf{u}\cdot\mathbf{u})\mathbf{v}-(\mathbf{u}\cdot\mathbf{v})\mathbf{u}]\\
&=\|\mathbf{u}\|^2\|\mathbf{v}\|^2-(\mathbf{u}\cdot\mathbf{v})^2=\|\mathbf{u}\|^2\|\mathbf{v}\|^2-\|\mathbf{u}\|^2\|\mathbf{v}\|^2\cos^2\theta=\|\mathbf{u}\|^2\|\mathbf{v}\|^2\sin^2\theta.
\end{aligned}$$

□

Still considering the conditions of the previous paragraph, since basis B is positively oriented, it is obvious that $\mathbf{A}_B(\hat{\mathbf{u}}_i,\hat{\mathbf{u}}_j,\hat{\mathbf{u}}_k)=\epsilon_{ijk}$, from which the second property above enables us to easily conclude equality $(\hat{\mathbf{u}}_i-\hat{\mathbf{u}}_j\times\hat{\mathbf{u}}_k)\epsilon_{ijk}=\mathbf{0}$. Thereby, an arbitrary vector $\mathbf{v}\in U_{\mathbb{R}}$ can be decomposed as

$$\mathbf{v}=\mathbf{f}_1^B(\mathbf{v})(\hat{\mathbf{u}}_2\times\hat{\mathbf{u}}_3)+\mathbf{f}_2^B(\mathbf{v})(\hat{\mathbf{u}}_3\times\hat{\mathbf{u}}_1)+\mathbf{f}_3^B(\mathbf{v})(\hat{\mathbf{u}}_1\times\hat{\mathbf{u}}_2).$$

Taking into account properties i. and vi., we conclude that all the vectors of $U_{\mathbb{R}}$ are axial. This fact is an implicit precondition for the following theorem.

Theorem 19 – Antisymmetric Tensors and Axial Vectors

Given a three dimensional oriented Euclidean space $(U_{\mathbb{R}},\mathbf{A}_B)$*, the function in mapping* $\boldsymbol{\psi}:\mathcal{LA}_{\mathbb{R}}(U^2)\mapsto U_{\mathbb{R}}$ *is a linear bijection if it is described by* $\boldsymbol{\psi}(\mathbf{x}^*\otimes\mathbf{y}^*)=\mathbf{A}_B\odot_2(\mathbf{x}^*\otimes\mathbf{y}^*)$.

Proof. The proof of the linearity of $\boldsymbol{\psi}$ is trivial. Considering an arbitrary vector $\boldsymbol{v} \in \mathsf{U}_{\mathbb{R}}$, if $\boldsymbol{v} = \mathbf{u} \times \boldsymbol{w}$, from property iii. above, the following development simultaneously proves that $\mathbf{u}^* \otimes \boldsymbol{w}^*$ is antisymmetric and $\boldsymbol{\psi}$ is a surjection:

$$\begin{aligned}
\mathbf{u} \times \boldsymbol{w} &= -(\boldsymbol{w} \times \mathbf{u}) \\
\mathbf{A}_{\mathrm{B}} \hat{\odot}_2 (\mathbf{u}^* \otimes \boldsymbol{w}^*) &= -\mathbf{A}_{\mathrm{B}} \hat{\odot}_2 (\boldsymbol{w}^* \otimes \mathbf{u}^*) \\
\boldsymbol{\psi}(\mathbf{u}^* \otimes \boldsymbol{w}^*) &= \boldsymbol{\psi}[-(\boldsymbol{w}^* \otimes \mathbf{u}^*)].
\end{aligned}$$

Supposing that $\mathbf{u}_1^* \otimes \boldsymbol{w}_1^*$ and $\mathbf{u}_2^* \otimes \boldsymbol{w}_2^*$ are different tensors related to the same vector through $\boldsymbol{\psi}$, the following equalities contradict this supposition since $\mathbf{A}_{\mathrm{B}}$ is not zero, proving that $\boldsymbol{\psi}$ is an injection.

$$0 = \boldsymbol{\psi}(\mathbf{u}_1^* \otimes \boldsymbol{w}_1^*) - \boldsymbol{\psi}(\mathbf{u}_2^* \otimes \boldsymbol{w}_2^*) = \mathbf{A}_{\mathrm{B}} \hat{\odot}_2 [(\mathbf{u}_2^* \otimes \boldsymbol{w}_1^*) - (\mathbf{u}_2^* \otimes \boldsymbol{w}_2^*)].$$

$\square$

Considering property x., the geometrical definition of inner product (4.10) and arbitrary vectors $\mathbf{x}, \mathbf{y}, \boldsymbol{z} \in \mathsf{U}$, if θ_1 is the smallest angle defined by $\overline{a\mathbf{x} \oplus a}$ and $\overline{a(\mathbf{y} \times \boldsymbol{z}) \oplus a}$, and θ_2 the smallest angle defined by defined by $\overline{a\mathbf{y} \oplus a}$ and $\overline{a\boldsymbol{z} \oplus a}$, scalar

$$\mathbf{A}_{\mathrm{B}}(\mathbf{x}, \mathbf{y}, \boldsymbol{z}) = \mathbf{x} \cdot (\mathbf{y} \times \boldsymbol{z}) = \|\mathbf{x}\|(\|\mathbf{y}\|\|\boldsymbol{z}\| \sin\theta_2)\cos\theta_1 = \underbrace{\|\mathbf{y}\|\|\boldsymbol{z}\|\sin\theta_2}_{\alpha}\underbrace{\|\mathbf{x}\|\cos\theta_1}_{\beta}, \tag{4.21}$$

where α is the area of the parallelogram defined by $\overline{a\mathbf{y} \oplus a}$ and $\overline{a\boldsymbol{z} \oplus a}$, as we already know. Moreover, it is possible to conclude that from vector $\mathbf{h} = \beta(\mathbf{y} \times \boldsymbol{z})/\|\mathbf{y} \times \boldsymbol{z}\|$, the line segments $\overline{a\mathbf{x} \oplus a}$, $\overline{a\mathbf{y} \oplus a}$ and $\overline{a\boldsymbol{z} \oplus a}$ define a parallelepiped of height $\overline{a\mathbf{h} \oplus a}$ and volume $\alpha\beta$, represented in figure 4.11. It is interesting to note that if $\{\mathbf{x}, \mathbf{y}, \boldsymbol{z}\}$ is negatively oriented then volume $\alpha\beta$ is negative since θ_1 is obtuse. Considering

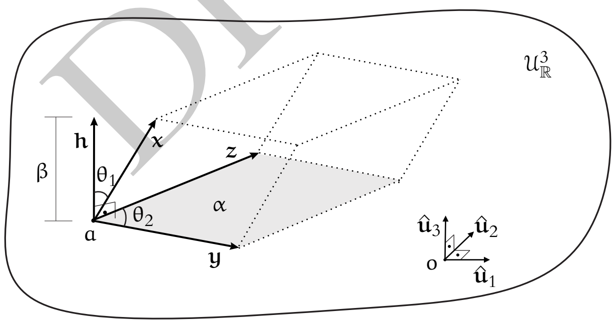

Figure 4.11 – Parallelogram of volume $\alpha\beta$ defined by $\mathbf{x}, \mathbf{y}, \boldsymbol{z}$.

definition (3.78) in the present context of a three-dimensional Euclidean space and $\{\mathbf{x}, \mathbf{y}, \boldsymbol{z}\} \subset \mathsf{U}$ an arbitrary linearly independent set, for a second order tensor $\mathbf{T} \in \mathcal{L}_{\mathbb{R}}(\mathsf{U}^2)$ equalities

$$\mathrm{Det}(\mathbf{T}) = \frac{\mathbf{A}_{\mathrm{B}}(\boldsymbol{t}(\mathbf{x}), \boldsymbol{t}(\mathbf{y}), \boldsymbol{t}(\boldsymbol{z}))}{\mathbf{A}_{\mathrm{B}}(\mathbf{x}, \mathbf{y}, \boldsymbol{z})} = \frac{\boldsymbol{t}(\mathbf{x}) \cdot [\boldsymbol{t}(\mathbf{y}) \times \boldsymbol{t}(\boldsymbol{z})]}{\mathbf{x} \cdot (\mathbf{y} \times \boldsymbol{z})} \tag{4.22}$$

show that $\mathrm{Det}(\mathbf{T})$ measures the change of volume of a parallelogram defined by $\mathbf{x}, \mathbf{y}, \boldsymbol{z}$ and "modified" by $\mathbf{T}$. Considering the orthonormal vectors of basis B and (3.52), we conclude from the previous equality that

$$\mathrm{Det}(\mathsf{T}) = \boldsymbol{t}(\hat{\mathbf{u}}_1) \cdot [\boldsymbol{t}(\hat{\mathbf{u}}_2) \times \boldsymbol{t}(\hat{\mathbf{u}}_3)] = \mathsf{T}(\boldsymbol{t}(\hat{\mathbf{u}}_2) \times \boldsymbol{t}(\hat{\mathbf{u}}_3), \hat{\mathbf{u}}_1)\,. \tag{4.23}$$

Chapter

5 Basic Tensor Calculus

Analysis is the branch of Mathematics that studies the behaviour of functions by using the idea of limit, considered fundamental, and its subordinate concepts. When such concepts involve differentiation and integration, as well as their related definitions, Analysis is commonly known as Calculus. Intuitively speaking, up to this point, we have dealt with functions in a kind of "static" approach: on previous chapters, the sole definition of a mapping, with its associated function explicitly described or not, was sufficient for developing the concepts presented. Now, we are also interested in a "dynamic" approach for functions, where it is important not only to specify them but also to describe how their values evolve on their images, or how they behave, through derivatives and integrals. In this context, considering the study of tensors started on chapter 3, the following pages introduce the Calculus of tensor functions, usually called Tensor Calculus. Since a tensor function is the most general type of function presented so far, we do not deviate from our main objective of being as abstract as possible on behalf of mathematical beauty.

5.1 Differentiation

In the context of Banach tensor spaces, which are normed and complete, to differentiate a tensor function $\boldsymbol{\psi}$ is to obtain its derivative, here understood as a linear tensor function that measures qualitatively and quantitatively **local sensitivity**, that is, the sensitivity of $\boldsymbol{\psi}$ on an element of the domain. Intuitively speaking, sensitivity refers to the intrinsic "susceptibility" of a function, or how it would "react", by varying its values, to an *eventual small change* on its argument. In the specific case of derivatives, this function "reaction" is *linearly estimated* on an element of its domain. However, not all tensor functions can be differentiable: the following mathematical conditions and definitions will describe which functions admit differentiation. Considering $\mathcal{L}_{\mathbb{F}}(\mathsf{U}^{\times m})$ and $\mathcal{L}_{\mathbb{F}}(\mathsf{V}^{\times s})$ Banach tensor spaces, *both defined by the same field* $\mathbb{F}$, let $\mathfrak{U} \subset \mathcal{L}_{\mathbb{F}}(\mathsf{U}^{\times m})$ be an open set and element $\mathbf{X}_0 \in \mathfrak{U}$ an arbitrary tensor at which local sensitivity is measured. The functions in $\boldsymbol{\psi} : \mathfrak{U} \mapsto \mathcal{L}_{\mathbb{F}}(\mathsf{V}^{\times s})$ and $\boldsymbol{\kappa} : \mathfrak{U} \mapsto \mathcal{L}_{\mathbb{F}}(\mathsf{V}^{\times s})$ are said to be **tangent** to each other at $\mathbf{X}_0$ if

$$\lim_{\mathbf{X}\to\mathbf{0}} \frac{\boldsymbol{\psi}(\mathbf{X}_0+\mathbf{X}) - \boldsymbol{\kappa}(\mathbf{X}_0+\mathbf{X})}{\|\mathbf{X}\|} = \mathbf{0}, \tag{5.1}$$

In other words, as variable $\mathbf{X}$ tends to the zero tensor, the numerator tends to the zero tensor faster than $\|\mathbf{X}\|$ tends to zero and then, on the limit, equality $\boldsymbol{\psi}(\mathbf{X}_0) = \boldsymbol{\kappa}(\mathbf{X}_0)$ must hold. From this definition, it is possible to obtain that if two functions are tangent to $\boldsymbol{\psi}$ at $\mathbf{X}_0$, they are tangent to each other at this tensor.

Proof. Since the sum of limits is the limit of sums, the previous statement can be verified by considering functions $\boldsymbol{\kappa}_1$ and $\boldsymbol{\kappa}_2$ tangent to $\boldsymbol{\psi}$ at $\mathbf{X}_0$ and subtracting both limit expressions described by (5.1) :

$$\begin{aligned}\mathbf{0} &= \lim_{\mathbf{X}\to\mathbf{0}} \frac{\boldsymbol{\psi}(\mathbf{X}_0+\mathbf{X}) - \boldsymbol{\kappa}_1(\mathbf{X}_0+\mathbf{X}) - \boldsymbol{\psi}(\mathbf{X}_0+\mathbf{X}) + \boldsymbol{\kappa}_2(\mathbf{X}_0+\mathbf{X})}{\|\mathbf{X}\|} \\ &= \lim_{\mathbf{X}\to\mathbf{0}} \frac{\boldsymbol{\kappa}_2(\mathbf{X}_0+\mathbf{X}) - \boldsymbol{\kappa}_1(\mathbf{X}_0+\mathbf{X})}{\|\mathbf{X}\|}.\end{aligned}$$

□

Proceeding with the above conditions, if there is one and only one tensor function $\boldsymbol{\psi}_d$, tangent to $\boldsymbol{\psi}$ at $\mathbf{X}_0$, described by the rule

$$\boldsymbol{\psi}_d(\mathbf{X}) = \boldsymbol{\psi}(\mathbf{X}_0) + [\boldsymbol{\psi}'(\mathbf{X}_0)](\mathbf{X} - \mathbf{X}_0), \tag{5.2}$$

where function $\boldsymbol{\psi}'(\mathbf{X}_0) \in \mathcal{L}_{\mathbb{F}}(\mathfrak{U}, \mathcal{L}_{\mathbb{F}}(\mathsf{V}^{\times s}))$ is bounded[1], then $\boldsymbol{\psi}$ is called **Fréchet differentiable**, **totally differentiable** or simply **differentiable** at $\mathbf{X}_0$, while the linear function $\boldsymbol{\psi}'(\mathbf{X}_0)$ is called the **derivative** of $\boldsymbol{\psi}$ at $\mathbf{X}_0$, whose value at $\mathbf{X}_0$ is described by tensor

$$\mathrm{D}\boldsymbol{\psi}(\mathbf{X}_0) := [\boldsymbol{\psi}'(\mathbf{X}_0)](\mathbf{X}_0), \tag{5.3}$$

from which we conclude that tensor function $\mathrm{D}\boldsymbol{\psi} \in \mathcal{L}_{\mathbb{F}}(\mathfrak{U}, \mathcal{L}_{\mathbb{F}}(\mathsf{V}^{\times s}))$. From the preceding argumentation to equality (2.23), linearity and boundedness imply continuity, and then $\boldsymbol{\psi}'(\mathbf{X}_0)$ belongs to function space $\mathcal{C}\mathcal{L}_{\mathbb{F}}(\mathfrak{U}, \mathcal{L}_{\mathbb{F}}(\mathsf{V}^{\times s}))$, which we specify to be also a Banach space of continuous linear functions. Moreover, the function in $\boldsymbol{\psi}' : \mathfrak{U} \mapsto \mathcal{C}\mathcal{L}_{\mathbb{F}}(\mathfrak{U}, \mathcal{L}_{\mathbb{F}}(\mathsf{V}^{\times s}))$ is said to be the **derivative function** or simply the derivative of $\boldsymbol{\psi}$. Rewriting expression (5.1) for $\boldsymbol{\psi}$ and $\boldsymbol{\psi}_d$ tangent at $\mathbf{X}_0$, the result is

$$\lim_{\mathbf{X}\to\mathbf{0}} \frac{\boldsymbol{\psi}(\mathbf{X}_0+\mathbf{X}) - \boldsymbol{\psi}(\mathbf{X}_0) - [\boldsymbol{\psi}'(\mathbf{X}_0)](\mathbf{X})}{\|\mathbf{X}\|} = \mathbf{0}, \tag{5.4}$$

from which we can conclude that the numerator tends to the zero tensor faster than the denominator tends to zero, otherwise there would be no definite limit. If numerator tends to the zero tensor, the term $[\boldsymbol{\psi}'(\mathbf{X}_0)](\mathbf{X})$ results an approximation for the

[1] The concept of boundedness defined in (2.23) is also valid for tensor functions.

difference $\boldsymbol{\psi}(\mathbf{X}_0+\mathbf{X})-\boldsymbol{\psi}(\mathbf{X}_0)$, and that is why the derivative is said to be a local linear approximation of a function and $[\boldsymbol{\psi}'(\mathbf{X}_0)](\mathbf{X})$ is called the **differential** of $\boldsymbol{\psi}$ at $\mathbf{X}_0$. Moreover, since the previous equality is valid for $\mathbf{X}$ tending to the zero tensor through an arbitrary "path", let's choose a specific direction by considering an arbitrary tensor $\mathbf{H}\in\mathfrak{U}$ in such a way that $\mathbf{X}=\alpha\mathbf{H}$, where $\alpha\in\mathbb{R}$ is the variable, and then

$$\lim_{\alpha\to 0}\frac{\boldsymbol{\psi}(\mathbf{X}_0+\alpha\mathbf{H})-\boldsymbol{\psi}(\mathbf{X}_0)-[\boldsymbol{\psi}'(\mathbf{X}_0)](\alpha\mathbf{H})}{|\alpha|}=\mathbf{0}.$$

Multiplying both sides by $\operatorname{sgn}(\alpha)$, we have

$$\lim_{\alpha\to 0}\frac{\boldsymbol{\psi}(\mathbf{X}_0+\alpha\mathbf{H})-\boldsymbol{\psi}(\mathbf{X}_0)}{\alpha}-\lim_{\alpha\to 0}\frac{\alpha[\boldsymbol{\psi}'(\mathbf{X}_0)](\mathbf{H})}{\alpha}=\operatorname{sgn}(\alpha)\mathbf{0},$$

which results

$$[\boldsymbol{\psi}'(\mathbf{X}_0)](\mathbf{H})=\lim_{\alpha\to 0}\frac{\boldsymbol{\psi}(\mathbf{X}_0+\alpha\mathbf{H})-\boldsymbol{\psi}(\mathbf{X}_0)}{\alpha}, \tag{5.5}$$

where $[\boldsymbol{\psi}'(\mathbf{X}_0)](\mathbf{H})\in\mathcal{L}_{\mathbb{F}}(V^{\times s})$ is called the **directional derivative** of $\boldsymbol{\psi}$ along $\mathbf{H}$ at $\mathbf{X}_0$. From this definition, when there is a directional derivative of $\boldsymbol{\psi}$ for any $\mathbf{X}_0\in\mathfrak{U}$, the function described by

$$[\boldsymbol{\psi}'(\mathbf{X})](\mathbf{H})=\lim_{\alpha\to 0}\frac{\boldsymbol{\psi}(\mathbf{X}+\alpha\mathbf{H})-\boldsymbol{\psi}(\mathbf{X})}{\alpha} \tag{5.6}$$

is said to be the directional derivative[2] of $\boldsymbol{\psi}$ along $\mathbf{H}$.

Proof. The above theory would be useless if there were no tangent function $\boldsymbol{\delta}_d$ for any given function $\boldsymbol{\delta}$. Considering the rules $\boldsymbol{\delta}(\mathbf{X})=\alpha\mathbf{X}$ and $\boldsymbol{\delta}_d(\mathbf{X})=\boldsymbol{\delta}(\mathbf{X}_0)+\alpha(\mathbf{X}-\mathbf{X}_0)=\alpha\mathbf{X}$, function $\boldsymbol{\delta}_d$ tangent to $\boldsymbol{\delta}$ on $\mathbf{X}_0$ can be easily verified from (5.4). Now, let's prove that there is only one function $\boldsymbol{\psi}_d$ tangent to $\boldsymbol{\psi}$ by supposing functions $\boldsymbol{\psi}_{d1}$ and $\boldsymbol{\psi}_{d2}$ tangent to $\boldsymbol{\psi}$. One equality (5.4) can be obtained for each tangent function and subtracting these equalities results

$$\lim_{\mathbf{X}\to\mathbf{0}}\frac{[\boldsymbol{\psi}_2'(\mathbf{X}_0)](\mathbf{X})-[\boldsymbol{\psi}_1'(\mathbf{X}_0)](\mathbf{X})}{\|\mathbf{X}\|}=\lim_{\mathbf{X}\to\mathbf{0}}\frac{\{[\boldsymbol{\psi}_2'-\boldsymbol{\psi}_1'](\mathbf{X}_0)\}(\mathbf{X})}{\|\mathbf{X}\|}=\mathbf{0}.$$

Since these equalities are valid for any $\mathbf{X}\in\mathfrak{U}$, let's admit $\mathbf{X}=\alpha\mathbf{H}$, where $\mathbf{H}\in\mathfrak{U}$ is a non zero given

[2] It is important to say a few words about a less restricted or weaker type of derivative than the Fréchet derivative just presented. It is not defined from the concept of tangency but from a directional derivative in which the derivative involved is not necessarily bounded and linear. Considering the above conditions, the tensor function in $\boldsymbol{\varphi}:\mathfrak{U}\mapsto\mathcal{L}_{\mathbb{F}}(V^{\times s})$ is said to be **Gâteaux differentiable** at $\mathbf{X}_0$ if there exists a mapping $\boldsymbol{\varphi}_G'(\mathbf{X}_0):\mathfrak{U}\mapsto\mathcal{L}_{\mathbb{F}}(V^{\times s})$, not necessarily bounded and linear, where

$$[\boldsymbol{\varphi}_G'(\mathbf{X}_0)](\mathbf{X})=\lim_{\alpha\to 0}\frac{\boldsymbol{\varphi}(\mathbf{X}_0+\alpha\mathbf{X})-\boldsymbol{\varphi}(\mathbf{X}_0)}{\alpha}.$$

The tensor $[\boldsymbol{\varphi}_G'(\mathbf{X}_0)](\mathbf{X})$ is called the **Gâteaux differential** of $\boldsymbol{\varphi}$ at $\mathbf{X}_0$ and $\boldsymbol{\varphi}_G'$ is the **Gâteaux derivative** of $\boldsymbol{\varphi}$. Every Fréchet differentiable function is also Gâteaux differentiable, but the converse is obviously not true. What is less obvious is the following: even when Gâteaux derivative $\boldsymbol{\varphi}_G'(\mathbf{X}_0)$ is bounded and linear, Gâteaux differentiability, which means differentiability along all directions at a point, does not ensure Fréchet differentiability. This present study will require only the stronger restrictions of the Fréchet differentiation. For further developments of the Gâteaux derivative, see Wouk[59], chapter 12.

tensor and α is a positive real number. Thereby, previous equalities lead to

$$\lim_{\alpha\to 0}\frac{\{[\boldsymbol{\psi}_2'-\boldsymbol{\psi}_1'](\mathbf{X}_0)\}(\alpha\mathbf{H})}{\|\alpha\mathbf{H}\|}=\frac{\{[\boldsymbol{\psi}_2'-\boldsymbol{\psi}_1'](\mathbf{X}_0)\}(\mathbf{H})}{\|\mathbf{H}\|}=\mathbf{0}$$

which are valid for any chosen $\mathbf{X}_0$ and then $\boldsymbol{\psi}_2'=\boldsymbol{\psi}_1'$. □

From previous conditions, it is possible to obtain the following important properties for derivatives by considering the function in mapping $\boldsymbol{\psi}:\mathcal{U}\mapsto\mathcal{L}_{\mathbb{F}}(V^{\times s})$ differentiable in its domain, an arbitrary tensor $\mathbf{H}\in\mathcal{U}$ and a scalar $\beta\in\mathbb{F}$.

i. If $\boldsymbol{\psi}(\mathbf{X})=\mathbf{H}$ then $\boldsymbol{\psi}'=\mathbf{0}$;

ii. If $\boldsymbol{\psi}(\mathbf{X})=\beta\mathbf{X}$, tensor $[\boldsymbol{\psi}'(\mathbf{X})](\mathbf{H})=\beta\mathbf{H}$ and therefore $\boldsymbol{\psi}'(\mathbf{X})=\boldsymbol{\psi}=\boldsymbol{\imath}$ for $\beta=1$;

iii. Sum rule: if $\boldsymbol{\psi}=\boldsymbol{\psi}_1+\boldsymbol{\psi}_2$ then $\boldsymbol{\psi}'=\boldsymbol{\psi}_1'+\boldsymbol{\psi}_2'$;

iv. **Chain rule**: if $\boldsymbol{\psi}=\boldsymbol{\psi}_1\circ\boldsymbol{\psi}_2$, where $\boldsymbol{\psi}_1:\mathcal{W}\mapsto\mathcal{L}_{\mathbb{F}}(V^{\times s})$, $\boldsymbol{\psi}_2:\mathcal{U}\mapsto\mathcal{W}$ and $\mathcal{W}\subset\mathcal{L}_{\mathbb{F}}(W^{\times r})$ is open, then

$$\boldsymbol{\psi}'(\mathbf{X})=[\boldsymbol{\psi}_1'\circ\boldsymbol{\psi}_2(\mathbf{X})]\circ\boldsymbol{\psi}_2'(\mathbf{X})\,; \tag{5.7}$$

v. **Product rule**: if $\boldsymbol{\psi}(\mathbf{X})=\boldsymbol{\psi}_1(\mathbf{X})\odot_q\boldsymbol{\psi}_2(\mathbf{X})$, where mappings $\boldsymbol{\psi}_1:\mathcal{U}\mapsto\mathcal{L}_{\mathbb{F}}(V^{\times r}\times W^{\times q})$ and $\boldsymbol{\psi}_2:\mathcal{U}\mapsto\mathcal{L}_{\mathbb{F}}(W^{\times q}\times V^{\times p})$ are defined for $s=r+p$, then

$$[\boldsymbol{\psi}'(\mathbf{X})](\mathbf{H})=\boldsymbol{\psi}_1(\mathbf{X})\odot_q[\boldsymbol{\psi}_2'(\mathbf{X})](\mathbf{H})+[\boldsymbol{\psi}_1'(\mathbf{X})](\mathbf{H})\odot_q\boldsymbol{\psi}_2(\mathbf{X})\,; \tag{5.8}$$

vi. **Leibniz's rule**: if function in $\boldsymbol{\psi}:\mathcal{U}\mapsto\mathbb{F}$ is described by $\boldsymbol{\psi}(\mathbf{X})=\boldsymbol{\psi}_1(\mathbf{X})\cdot\boldsymbol{\psi}_2(\mathbf{X})$, where $\boldsymbol{\psi}_1:\mathcal{U}\mapsto\mathcal{L}_{\mathbb{F}}(V^{\times s})$ and $\boldsymbol{\psi}_2:\mathcal{U}\mapsto\mathcal{L}_{\mathbb{F}}(V^{\times s})$, then

$$[\boldsymbol{\psi}'(\mathbf{X})](\mathbf{H})=\boldsymbol{\psi}_1(\mathbf{X})\cdot[\boldsymbol{\psi}_2'(\mathbf{X})](\mathbf{H})+[\boldsymbol{\psi}_1'(\mathbf{X})](\mathbf{H})\cdot\boldsymbol{\psi}_2(\mathbf{X})\,; \tag{5.9}$$

vii. If $\boldsymbol{\psi}(\mathbf{X})=\boldsymbol{\psi}_1(\mathbf{X})\otimes\boldsymbol{\psi}_2(\mathbf{X})$, where $\boldsymbol{\psi}_1:\mathcal{U}\mapsto\mathcal{L}_{\mathbb{F}}(V^{\times s-r})$ and $\boldsymbol{\psi}_2:\mathcal{U}\mapsto\mathcal{L}_{\mathbb{F}}(V^{\times r})$, then

$$[\boldsymbol{\psi}'(\mathbf{X})](\mathbf{H})=\boldsymbol{\psi}_1(\mathbf{X})\otimes[\boldsymbol{\psi}_2'(\mathbf{X})](\mathbf{H})+[\boldsymbol{\psi}_1'(\mathbf{X})](\mathbf{H})\otimes\boldsymbol{\psi}_2(\mathbf{X})\,; \tag{5.10}$$

viii. If $\boldsymbol{\psi}(\mathbf{X})=\mathrm{Det}(\mathbf{X})$, where $\boldsymbol{\psi}:\mathcal{U}\mapsto\mathbb{F}$ and domain $\mathcal{U}\subset\mathcal{L}_{\mathbb{F}}(U^2)$ is constituted by tensors whose representative functions are invertible, then

$$[\boldsymbol{\psi}'(\mathbf{X})](\mathbf{H})=\mathrm{Det}(\mathbf{X})\,\mathrm{tr}\,(\mathbf{X}^{-1}\odot_1\mathbf{H})\,; \tag{5.11}$$

Proof. Verification of items i and ii are straightforward results after specifying their respective rules of $\boldsymbol{\psi}$ on definition (5.6). For the case of item iii, the use of (5.6) leads to

$$[\boldsymbol{\psi}'(\mathbf{X})](\mathbf{H})=[\boldsymbol{\psi}_1'(\mathbf{X})](\mathbf{H})+[\boldsymbol{\psi}_2'(\mathbf{X})](\mathbf{H})=[\boldsymbol{\psi}_1'(\mathbf{X})+\boldsymbol{\psi}_2'(\mathbf{X})](\mathbf{H})=[[\boldsymbol{\psi}_1'+\boldsymbol{\psi}_2'](\mathbf{X})](\mathbf{H})\,.$$

Now, in order to prove the chain rule, we adapt the strategy of ZEIDLER[60], p. 248, as follows. Considering the concept of derivative as an approximation, from (5.6) it is correct to say that there is a residue $r_1(\beta\mathbf{Z})$ that tends to the zero tensor when non-zero $\beta\in\mathbb{F}$ tends to zero, where

$$\beta\,[\boldsymbol{\psi}_1'(\mathbf{Y})](\mathbf{Z})=\boldsymbol{\psi}_1(\mathbf{Y}+\beta\mathbf{Z})-\boldsymbol{\psi}_1(\mathbf{Y})+r_1(\beta\mathbf{Z})$$

and $\mathbf{Y},\mathbf{Z}\in\mathcal{L}_{\mathbb{F}}(W^{\times r})$ are arbitrary. For the case of $\boldsymbol{\psi}_2$, we write

$$\boldsymbol{\psi}_2(\mathbf{X})=\boldsymbol{\psi}_2(\mathbf{X}+\alpha\mathbf{H})-\alpha[\boldsymbol{\psi}_2'(\mathbf{X})](\mathbf{H})+r_2(\alpha\mathbf{H}).$$

Specifying $\mathbf{Y}=\boldsymbol{\psi}_2(\mathbf{X})$ and $\mathbf{Z}=\alpha/\beta\,[\boldsymbol{\psi}_2'(\mathbf{X})](\mathbf{H})$ on the first equality, then

$$\alpha[\boldsymbol{\psi}_1'\circ\boldsymbol{\psi}_2(\mathbf{X})]([\boldsymbol{\psi}_2'(\mathbf{X})](\mathbf{H}))=\boldsymbol{\psi}_1(\boldsymbol{\psi}_2(\mathbf{X})+\alpha[\boldsymbol{\psi}_2'(\mathbf{X})](\mathbf{H}))-\boldsymbol{\psi}_1(\mathbf{Y})+r_1(\alpha[\boldsymbol{\psi}_2'(\mathbf{X})](\mathbf{H}))$$
$$\alpha[[\boldsymbol{\psi}_1'\circ\boldsymbol{\psi}_2(\mathbf{X})]\circ\boldsymbol{\psi}_2'(\mathbf{X})](\mathbf{H})=\boldsymbol{\psi}_1(\boldsymbol{\psi}_2(\mathbf{X}+\alpha\mathbf{H})+r_2(\alpha\mathbf{H}))-\boldsymbol{\psi}_1\circ\boldsymbol{\psi}_2(\mathbf{X})+r_1(\alpha[\boldsymbol{\psi}_2'(\mathbf{X})](\mathbf{H}))$$

Dividing both sides by α and taking them to the limit of α tending to zero, item iv is verified. The product rule is proved by the following development: adding and subtracting the right hand side of

$$[\boldsymbol{\psi}'(\mathbf{X})](\mathbf{H})=\lim_{\alpha\to 0}1/\alpha[\boldsymbol{\psi}_1(\mathbf{X}+\alpha\mathbf{H})\odot_q\boldsymbol{\psi}_2(\mathbf{X}+\alpha\mathbf{H})-\boldsymbol{\psi}_1(\mathbf{X})\odot_q\boldsymbol{\psi}_2(\mathbf{X})]$$

by $\boldsymbol{\psi}_1(\mathbf{X}+\alpha\mathbf{H})\odot_q\boldsymbol{\psi}_2(\mathbf{X})$, we arrive at

$$\begin{aligned}[\boldsymbol{\psi}'(\mathbf{X})](\mathbf{H})&=\lim_{\alpha\to 0}1/\alpha\{\boldsymbol{\psi}_1(\mathbf{X}+\alpha\mathbf{H})\odot_q[\boldsymbol{\psi}_2(\mathbf{X}+\alpha\mathbf{H})-\boldsymbol{\psi}_2(\mathbf{X})]+\\&\qquad+[\boldsymbol{\psi}_1(\mathbf{X}+\alpha\mathbf{H})-\boldsymbol{\psi}_1(\mathbf{X})]\odot_q\boldsymbol{\psi}_2(\mathbf{X})\}\\&=\boldsymbol{\psi}_1(\mathbf{X})\odot_q\lim_{\alpha\to 0}1/\alpha[\boldsymbol{\psi}_2(\mathbf{X}+\alpha\mathbf{H})-\boldsymbol{\psi}_2(\mathbf{X})]+\\&\qquad+\lim_{\alpha\to 0}1/\alpha[\boldsymbol{\psi}_1(\mathbf{X}+\alpha\mathbf{H})-\boldsymbol{\psi}_1(\mathbf{X})]\odot_q\boldsymbol{\psi}_2(\mathbf{X}).\end{aligned}$$

Since item vi is a corollary of the product rule, its verification is trivial. Item vii is verified by a similar procedure we used for the product rule. In order to prove viii, we consider the conditions of (5.6) and recall equalities (3.69) and (3.80). Thereby, from (1.47) and (1.48), we develop the following:

$$\begin{aligned}\mathrm{Det}(\mathbf{X}^{-1}\odot_1\mathbf{H}+\alpha^{-1}\mathbf{I})&=(-1)^n(-\alpha)^{-n}+a_1(-\alpha)^{1-n}+a_2(-\alpha)^{2-n}+\cdots+a_{n-1}(-\alpha)+a_n\\&=\alpha^{-n}+a_1(-1)^{1-n}\alpha^{1-n}+a_2(-1)^{2-n}\alpha^{2-n}+\cdots-a_{n-1}\alpha+a_n.\end{aligned}$$

Considering this last equality and (1.49), the following development proves item viii.

$$\begin{aligned}[\boldsymbol{\psi}'(\mathbf{X})](\mathbf{H})&=\lim_{\alpha\to 0}\frac{\mathrm{Det}(\mathbf{X}+\alpha\mathbf{H})-\mathrm{Det}(\mathbf{X})}{\alpha}\\&=\lim_{\alpha\to 0}\frac{\mathrm{Det}(\mathbf{X})\alpha^n\mathrm{Det}(\mathbf{X}^{-1}\odot_1\mathbf{H}+\alpha^{-1}\mathbf{I})-\mathrm{Det}(\mathbf{X})}{\alpha}\\&=\lim_{\alpha\to 0}\frac{\mathrm{Det}(\mathbf{X})[1+a_1(-1)^{1-n}\alpha^1+a_2(-1)^{2-n}\alpha^2+\cdots-a_{n-1}\alpha^{n+1}+a_n\alpha^n-1]}{\alpha}\\&=\lim_{\alpha\to 0}\mathrm{Det}(\mathbf{X})[a_1(-1)^{1-n}+a_2(-1)^{2-n}\alpha+\cdots-a_{n-1}\alpha^n+a_n\alpha^{n-1}]\\&=\mathrm{Det}(\mathbf{X})a_1(-1)^{1-n}=\mathrm{Det}(\mathbf{X})\mathrm{tr}\,(\mathbf{X}^{-1}\odot_1\mathbf{H}).\end{aligned}$$

□

Considering a mapping $\boldsymbol{\varphi}:\mathcal{U}_1\times\cdots\times\mathcal{U}_q\mapsto\mathcal{L}_{\mathbb{F}}(V^{\times s})$, where $\mathcal{U}_i\subset\mathcal{L}_{\mathbb{F}}(U^{\times m})$ is open and the values of $\boldsymbol{\varphi}$ are represented by the multivariate notation $\boldsymbol{\varphi}(\mathbf{X}_1,\cdots,\mathbf{X}_q)$, if there is a tangent function $\boldsymbol{\varphi}_d$ to $\boldsymbol{\varphi}$ at $(\mathbf{Y}_1,\cdots,\mathbf{Y}_q)\in\mathcal{U}_1\times\cdots\times\mathcal{U}_q$ where

$$\boldsymbol{\varphi}_d(\mathbf{X}_1,\cdots,\mathbf{X}_q)=\boldsymbol{\varphi}(\mathbf{Y}_1,\cdots,\mathbf{Y}_q)+[\boldsymbol{\varphi}'(\mathbf{Y}_1,\cdots,\mathbf{Y}_q)](\mathbf{X}_1-\mathbf{Y}_1,\cdots,\mathbf{X}_q-\mathbf{Y}_q)$$

and $\boldsymbol{\varphi}'(\mathbf{Y}_1,\cdots,\mathbf{Y}_q)$ is multilinear, then $\boldsymbol{\varphi}$ is said to be totally differentiable and, by a similar development to the case of univariate tensor functions, we arrive at the multivariate definition for directional derivatives:

$$[\boldsymbol{\varphi}'(\mathbf{X}_1,\cdots,\mathbf{X}_q)](\mathbf{H}_1,\cdots,\mathbf{H}_q)=\lim_{\alpha\to 0}\frac{\boldsymbol{\varphi}(\mathbf{X}_1+\alpha\mathbf{H}_1,\cdots,\mathbf{X}_q+\alpha\mathbf{H}_q)-\boldsymbol{\varphi}(\mathbf{X}_1,\cdots,\mathbf{X}_q)}{\alpha},\quad(5.12)$$

where $\mathbf{H}_i \in \mathcal{U}_i$ is an arbitrary direction. When the strategy is to study each variable $\mathbf{X}_i$ independently, *caeteris paribus*, development from the concept of tangency is the same as the univariate case, but now the derivative of $\boldsymbol{\varphi}$ is called **partial derivative** of $\boldsymbol{\varphi}$ with respect to $\mathbf{X}_r$, $1 \leqslant r \leqslant q$, represented by $\partial_{\mathbf{X}_r}\boldsymbol{\varphi}$, where

$$[\partial_{\mathbf{X}_r}\boldsymbol{\varphi}(\mathbf{X}_1,\cdots,\mathbf{X}_q)](\mathbf{H}_r) = \lim_{\alpha\to 0}\frac{\boldsymbol{\varphi}(\mathbf{X}_1,\cdots,\mathbf{X}_r+\alpha\mathbf{H}_r,\cdots,\mathbf{X}_q)-\boldsymbol{\varphi}(\mathbf{X}_1,\cdots,\mathbf{X}_q)}{\alpha}. \tag{5.13}$$

In the particular case of $\mathbf{H}_r = \mathbf{X}_r$, we have

$$D_{\mathbf{X}_r}\boldsymbol{\varphi}(\mathbf{X}_1,\cdots,\mathbf{X}_q) := [\partial_{\mathbf{X}_r}\boldsymbol{\varphi}(\mathbf{X}_1,\cdots,\mathbf{X}_q)](\mathbf{X}_r). \tag{5.14}$$

From the conditions presented so far, the property

$$[\boldsymbol{\varphi}'(\mathbf{X}_1,\cdots,\mathbf{X}_q)](\mathbf{H}_1,\cdots,\mathbf{H}_q) = \sum_{i=1}^{q}[\partial_{\mathbf{X}_i}\boldsymbol{\varphi}(\mathbf{X}_1,\cdots,\mathbf{X}_q)](\mathbf{H}_i) \tag{5.15}$$

relates total and partial derivatives of multivariate tensor functions and the following equality is commonly known as the **Chain Rule for Partial Derivatives:**

$$\begin{aligned}[\boldsymbol{\varphi}'(\boldsymbol{\varphi}_1(\mathbf{Y}_1),\cdots,\boldsymbol{\varphi}_q(\mathbf{Y}_q))](\mathbf{H}_1,\cdots,\mathbf{H}_q) &= \\ &\sum_{i=1}^{q}[\partial_{\mathbf{X}_i}\boldsymbol{\varphi}(\boldsymbol{\varphi}_1(\mathbf{Y}_1),\cdots,\boldsymbol{\varphi}_q(\mathbf{Y}_q))\circ\boldsymbol{\varphi}_i'(\mathbf{Y}_i)](\mathbf{H}_i),\end{aligned} \tag{5.16}$$

where the function in $\boldsymbol{\varphi}_i : \mathcal{W}_i \mapsto \mathcal{U}_i$, $\mathcal{W}_i \subset \boldsymbol{\mathcal{L}}_{\mathbb{F}}(W^{\times r})$, must obviously be differentiable.

Proof. Let's prove these two previous properties. From (5.12) and (5.13) it is clear that, for each direction $\mathbf{H}_i$, we have q equalities of the form

$$[\boldsymbol{\varphi}'(\mathbf{X}_1,\cdots,\mathbf{X}_q)](\mathbf{0},\cdots,\mathbf{H}_i,\cdots,\mathbf{0}) = [\partial_{\mathbf{X}_i}\boldsymbol{\varphi}(\mathbf{X}_1,\cdots,\mathbf{X}_q)](\mathbf{H}_i).$$

By summing up these q equalities, we prove the first property because function $\boldsymbol{\varphi}'(\mathbf{X}_1,\cdots,\mathbf{X}_q)$ is multilinear. Applying the chain rule for deriving $\boldsymbol{\varphi}$ with respect to argument $\boldsymbol{\varphi}_i(\mathbf{Y}_i)$, q other equalities are also obtained; and since the derivatives involved are linear and multilinear, the second property is verified by also summing up these equalities. □

It is interesting to note that derivatives may be differentiable themselves, and also derivatives of derivatives, and so on. Considering previous conditions, the tensor function $\boldsymbol{\psi}$ is defined to be the zero order derivative of itself, function $\boldsymbol{\psi}'$ is the order one derivative of $\boldsymbol{\psi}$, $\boldsymbol{\psi}''$ the order two and $\boldsymbol{\psi}^{(k)}$ the order $k \geqslant 0$. Thereby, from definition (5.2), a tensor function $\boldsymbol{\psi}$ is said to be differentiable of order two at $\mathbf{X}_0$ if it is differentiable at $\mathbf{X}_0$ and there is a unique function $\boldsymbol{\psi_d}^{(1)}$ tangent to the function in $\boldsymbol{\psi}' : \mathcal{U} \mapsto \mathcal{CL}_{\mathbb{F}}(\mathcal{U},\boldsymbol{\mathcal{L}}_{\mathbb{F}}(V^{\times s}))$ at $\mathbf{X}_0$ where

$$\boldsymbol{\psi_d}^{(1)}(\mathbf{X}) = \boldsymbol{\psi}'(\mathbf{X}_0) + [\boldsymbol{\psi}''(\mathbf{X}_0)](\mathbf{X}-\mathbf{X}_0).$$

Since function $\boldsymbol{\psi}'(\mathbf{X}_0) \in \mathcal{CL}_{\mathbb{F}}(\mathfrak{U}, \boldsymbol{\mathcal{L}}_{\mathbb{F}}(V^{\times s}))$ then $\boldsymbol{\psi}'' : \mathfrak{U} \mapsto \mathcal{CL}_{\mathbb{F}}(\mathfrak{U}, \mathcal{CL}_{\mathbb{F}}(\mathfrak{U}, \boldsymbol{\mathcal{L}}_{\mathbb{F}}(V^{\times s})))$. Similarly, if $\boldsymbol{\psi}$ is differentiable of order three at $\mathbf{X}_0$, then it is differentiable of order two at $\mathbf{X}_0$ and there is a unique $\boldsymbol{\psi_d}^{(2)}$ tangent to $\boldsymbol{\psi}''$ at $\mathbf{X}_0$ where

$$\boldsymbol{\psi_d}^{(2)}(\mathbf{X}) = \boldsymbol{\psi}''(\mathbf{X}_0) + [\boldsymbol{\psi}'''(\mathbf{X}_0)](\mathbf{X} - \mathbf{X}_0)$$

and $\boldsymbol{\psi}''' : \mathfrak{U} \mapsto \mathcal{CL}_{\mathbb{F}}(\mathfrak{U}, \mathcal{CL}_{\mathbb{F}}(\mathfrak{U}, \mathcal{CL}_{\mathbb{F}}(\mathfrak{U}, \boldsymbol{\mathcal{L}}_{\mathbb{F}}(V^{\times s}))))$. Generically, tensor function $\boldsymbol{\psi}$ is said to be differentiable of order $k+1$ at $\mathbf{X}_0$ if it is differentiable of order k at $\mathbf{X}_0$ and there is a unique function $\boldsymbol{\psi_d}^{(k)}$ tangent to function $\boldsymbol{\psi}^{(k)}$ at $\mathbf{X}_0$ described by

$$\boldsymbol{\psi_d}^{(k)}(\mathbf{X}) = \boldsymbol{\psi}^{(k)}(\mathbf{X}_0) + [\boldsymbol{\psi}^{(k+1)}(\mathbf{X}_0)](\mathbf{X} - \mathbf{X}_0), \tag{5.17}$$

where

$$\boldsymbol{\psi}^{(k+1)} : \mathfrak{U} \mapsto \mathcal{CL}_{\mathbb{F}}(\mathfrak{U}_{k+1}, \mathcal{CL}_{\mathbb{F}}(\mathfrak{U}_k, \cdots \mathcal{CL}_{\mathbb{F}}(\mathfrak{U}_1, \boldsymbol{\mathcal{L}}_{\mathbb{F}}(V^{\times s}))))$$

and $\mathfrak{U}_i = \mathfrak{U}$. By a similar development we performed for one order derivative, rule (5.17) leads to equality

$$[\boldsymbol{\psi}^{(k+1)}(\mathbf{X})](\mathbf{H}) = \lim_{\alpha \to 0} \frac{\boldsymbol{\psi}^{(k)}(\mathbf{X} + \alpha \mathbf{H}) - \boldsymbol{\psi}^{(k)}(\mathbf{X})}{\alpha}, \tag{5.18}$$

which is the $k+1$ order directional derivative of $\boldsymbol{\psi}$ along $\mathbf{H} \in \mathfrak{U}$. From previous equality, if $k > 0$ the directional derivative of order $k+1$ along a given tensor is always a function and then by choosing arbitrary tensors $\mathbf{H}_1, \cdots, \mathbf{H}_{k+1} \in \mathfrak{U}$, we conclude that

$$[[[\boldsymbol{\psi}^{(k+1)}(\mathbf{X})](\mathbf{H}_{k+1})](\mathbf{H}_k) \cdots](\mathbf{H}_1) \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(V^{\times s}). \tag{5.19}$$

In this sense, when $\mathbf{H}_i = \mathbf{X}$, tensor

$$D^{(k+1)}\boldsymbol{\psi}(\mathbf{X}) := [[[\boldsymbol{\psi}^{(k+1)}(\mathbf{X})](\mathbf{X})](\mathbf{X}) \cdots](\mathbf{X}). \tag{5.20}$$

Considering the conditions of (5.14), if multivariate function $\boldsymbol{\varphi}$ is differentiable of order q, we specify

$$\partial_{\mathbf{X}_q, \cdots, \mathbf{X}_1} \boldsymbol{\varphi}(\mathbf{X}_1, \cdots, \mathbf{X}_q) := \partial_{\mathbf{X}_q} \cdots \partial_{\mathbf{X}_1} \boldsymbol{\varphi}(\mathbf{X}_1, \cdots, \mathbf{X}_q) \tag{5.21}$$

and

$$D_{\mathbf{X}_q, \cdots, \mathbf{X}_1} \boldsymbol{\varphi}(\mathbf{X}_1, \cdots, \mathbf{X}_q) := D_{\mathbf{X}_q} \cdots D_{\mathbf{X}_1} \boldsymbol{\varphi}(\mathbf{X}_1, \cdots, \mathbf{X}_q). \tag{5.22}$$

Therefore,

$$D_{\mathbf{X}_q, \cdots, \mathbf{X}_1} \boldsymbol{\varphi}(\mathbf{X}_1, \cdots, \mathbf{X}_q) = [[[\partial_{\mathbf{X}_q, \cdots, \mathbf{X}_1} \boldsymbol{\varphi}(\mathbf{X}_1, \cdots, \mathbf{X}_q)](\mathbf{X}_1)](\mathbf{X}_2) \cdots](\mathbf{X}_q). \tag{5.23}$$

Now, let's restrict our attention to continuous tensor functions, that is, to elements of function space $\mathcal{C}_{\mathbb{F}}(\mathfrak{U}, \boldsymbol{\mathcal{L}}_{\mathbb{F}}(V^{\times s}))$ in this present case, when they are said to be of **differentiability class** $\mathcal{C}^0$. Function $\boldsymbol{\psi}$ is of class $\mathcal{C}^1$ or **continuously differentiable** if it is of class $\mathcal{C}^0$ and differentiable of order one, with derivative $\boldsymbol{\psi}'$ also of differentiability class $\mathcal{C}^0$. In general terms, $\boldsymbol{\psi}$ is of class $\mathcal{C}^{k+1}$ if it is of class $\mathcal{C}^k$ and differentiable of order $k+1$, with $\boldsymbol{\psi}^{(k+1)}$ of class $\mathcal{C}^k$. Stating that function $\boldsymbol{\psi}$ is of class $\mathcal{C}^{k+1}$ is equal to stating that $\boldsymbol{\psi}$ is **smooth** of order $k+1$ and for the case where $\boldsymbol{\psi}$ is smooth of any order, we say that it is of class $\mathcal{C}^\infty$ or simply call it smooth. A differentiable tensor function is called a **diffeomorphism** if it is a bijection and its inverse is also differentiable. When a diffeomorphism is smooth of order $k+1$, it is called a $\mathcal{C}^{k+1}$-diffeomorphism. For the case of the multivariate tensor function in $\boldsymbol{\varphi} : \mathfrak{U}_1 \times \cdots \times \mathfrak{U}_q \mapsto \boldsymbol{\mathcal{L}}_{\mathbb{F}}(V^{\times s})$, a given level of smoothness must be presented on each open set $\mathfrak{U}_i$. In other words, $\boldsymbol{\varphi}$ is said to be continuous if it is continuous on each $\mathfrak{U}_i$, being classified as $\mathcal{C}^0$. Thereby, it is of class $\mathcal{C}^1$ or continuously differentiable if it is $\mathcal{C}^0$ and all of its q partial derivatives are continuous; it is $\mathcal{C}^2$ if it is a $\mathcal{C}^1$ function and its q^2 second order partial derivatives are continuous; and so on.

5.2 The Gradient

The concept of gradient we shall detail in this section is restricted to the context of Hilbert tensor spaces because it relies on the Riesz-Fréchet Representation of Tensors, described on theorem 12. Similarly to what we already presented in this chapter, an open set $\mathfrak{U} \subset \boldsymbol{\mathcal{L}}_{\mathbb{F}}(U^{\times m})$ and a differentiable function in $\boldsymbol{\psi} : \mathfrak{U} \mapsto \boldsymbol{\mathcal{L}}_{\mathbb{F}}(V^{\times s})$ are also considered, but $\boldsymbol{\mathcal{L}}_{\mathbb{F}}(U^{\times m})$ and $\boldsymbol{\mathcal{L}}_{\mathbb{F}}(V^{\times s})$ are Hilbert tensor spaces here and $m+s>0$. Considering the conditions of the aforementioned theorem, there is a unique tensor $\nabla\boldsymbol{\psi}(\mathbf{X}_0) \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(U^{\times m} \times V^{\times s})$, called the **gradient** of $\boldsymbol{\psi}$ at $\mathbf{X}_0$, which the derivative $\boldsymbol{\psi}'(\mathbf{X}_0) \in \mathcal{C}\mathcal{L}_{\mathbb{F}}(\mathfrak{U}, \boldsymbol{\mathcal{L}}_{\mathbb{F}}(V^{\times s}))$ is the m-cotensor of and then

$$[\boldsymbol{\psi}'(\mathbf{X}_0)](\mathbf{X}) = \mathbf{X} \odot_m \nabla\boldsymbol{\psi}(\mathbf{X}_0)\,. \tag{5.24}$$

Moreover, the tensor function in $\nabla\boldsymbol{\psi} : \mathfrak{U} \mapsto \boldsymbol{\mathcal{L}}_{\mathbb{F}}(U^{\times m} \times V^{\times s})$, called the gradient of $\boldsymbol{\psi}$, assigns to a tensor of order m a tensor of order $m+s$. If our interest is to inspect function values only at $\mathbf{X}_0 = \mathbf{X}$, previous equality results

$$D\boldsymbol{\psi}(\mathbf{X}) = \mathbf{X} \odot_m \nabla\boldsymbol{\psi}(\mathbf{X})\,. \tag{5.25}$$

If tensor function $\boldsymbol{\psi}$ is a differentiable unary tensor operator on $\mathfrak{U}$, that is tensor space $\boldsymbol{\mathcal{L}}_{\mathbb{F}}(U^{\times m}) = \boldsymbol{\mathcal{L}}_{\mathbb{F}}(V^{\times s}) = \boldsymbol{\mathcal{L}}_{\mathbb{F}}(U^m)$, its gradient assigns to a m-th order tensor a $2m$-th order tensor. Considering the conditions of expression (5.13), which defines partial derivatives, we write

$$[\partial_{\mathbf{X}_r}\boldsymbol{\varphi}(\mathbf{X}_1, \cdots, \mathbf{X}_q)](\mathbf{H}_r) = \mathbf{H}_r \odot_m \nabla_{\mathbf{X}_r}\boldsymbol{\varphi}(\mathbf{X}_1, \cdots, \mathbf{X}_q)\,, \tag{5.26}$$

where $\nabla_{\mathbf{X}_r}\boldsymbol{\varphi}$ is the gradient of $\boldsymbol{\varphi}$ with respect to $\mathbf{X}_r$.

For the particular case of $m = s = 1$, tensor spaces $\boldsymbol{\mathcal{L}}_{\mathbb{F}}(U) = U^*_{\mathbb{F}}$ and $\boldsymbol{\mathcal{L}}_{\mathbb{F}}(V) = V^*_{\mathbb{F}}$, gradient $\nabla\boldsymbol{\psi}(\mathbf{x}^*_0) \in \boldsymbol{\mathcal{L}}_{\mathbb{R}}(U \times V)$ and derivative $\boldsymbol{\psi}'(\mathbf{x}^*_0) \in \mathcal{CL}_{\mathbb{F}}(\boldsymbol{\mathfrak{U}}, V^*_{\mathbb{F}})$, expression (5.24) turns out to be

$$[\boldsymbol{\psi}'(\mathbf{x}^*_0)](\mathbf{x}^*) = \mathbf{x}^* \odot_1 \nabla\boldsymbol{\psi}(\mathbf{x}^*_0)\,. \tag{5.27}$$

Now, from theorem 6, we present vector functions

$$\hat{\boldsymbol{\psi}} := \Phi^{-1} \circ \boldsymbol{\psi} \circ \Phi\,, \quad \hat{\boldsymbol{\psi}}'(\mathbf{x}_0) := \Phi^{-1} \circ \boldsymbol{\psi}'(\mathbf{x}^*_0) \circ \Phi \quad \text{and} \quad \nabla\hat{\boldsymbol{\psi}} := \nabla\boldsymbol{\psi} \circ \Phi\,, \tag{5.28}$$

from which we can rewrite equality (5.27) as

$$[\hat{\boldsymbol{\psi}}'(\mathbf{x}_0)](\mathbf{x}) = \nabla\hat{\boldsymbol{\psi}}(\mathbf{x}_0)^\dagger \hat{\odot}_1 \mathbf{x}^*\,. \tag{5.29}$$

Comparing this equality with (3.53), we can state that $\hat{\boldsymbol{\psi}}'(\mathbf{x}_0)$ is the representative function of second order tensor $\nabla\hat{\boldsymbol{\psi}}(\mathbf{x}_0)$, called the gradient of $\hat{\boldsymbol{\psi}}$ at $\mathbf{x}_0$. Now, considering equalities (3.52) and second order tensor $\nabla\hat{\boldsymbol{\psi}}(\mathbf{x}_0) = \nabla\boldsymbol{\psi}(\mathbf{x}^*_0)$, we also have for the present case that

$$[\nabla\hat{\boldsymbol{\psi}}(\mathbf{x}_0)](\mathbf{x}, \mathbf{y}) = \mathbf{y} \cdot [\hat{\boldsymbol{\psi}}'(\mathbf{x}_0)](\mathbf{x}^c)\,, \forall \mathbf{x}, \mathbf{y} \in \mathcal{U}\,, \tag{5.30}$$

where $\mathcal{U}$ is constituted by the corresponding vectors to the elements of $\boldsymbol{\mathfrak{U}}$. Considering rule (5.6), from definitions (5.28), it is possible to develop the following for an arbitrary direction $\mathbf{h}^* \in \boldsymbol{\mathfrak{U}}$:

$$\begin{aligned}
[\hat{\boldsymbol{\psi}}'(\mathbf{x})](\mathbf{h}) &= \Phi^{-1} \circ [\boldsymbol{\psi}'(\mathbf{x}^*)](\mathbf{h}^*) \\
&= \lim_{\alpha \to 0} \frac{\Phi^{-1} \circ \boldsymbol{\psi}(\mathbf{x}^* + \alpha\mathbf{h}^*) - \Phi^{-1} \circ \boldsymbol{\psi}(\mathbf{x}^*)}{\alpha} \\
&= \lim_{\alpha \to 0} \frac{\Phi^{-1} \circ \boldsymbol{\psi} \circ \Phi(\mathbf{x} + \alpha\mathbf{h}) - \Phi^{-1} \circ \boldsymbol{\psi} \circ \Phi(\mathbf{x})}{\alpha} \\
&= \lim_{\alpha \to 0} \frac{\hat{\boldsymbol{\psi}}(\mathbf{x} + \alpha\mathbf{h}) - \hat{\boldsymbol{\psi}}(\mathbf{x})}{\alpha}\,,
\end{aligned} \tag{5.31}$$

which results the rule of the directional derivative of the vector function in $\hat{\boldsymbol{\psi}} : \mathcal{U} \mapsto V_{\mathbb{F}}$ along $\mathbf{h} \in \mathcal{U}$. From this result, *the derivative* $\hat{\boldsymbol{\psi}}'(\mathbf{x})$*, as a linear vector function, inherits all the previous and forthcoming definitions of the linear tensor function* $\boldsymbol{\psi}'(\mathbf{x}^*)$. For example, from (5.25) and (5.27), we can write that vector

$$D\hat{\boldsymbol{\psi}}(\mathbf{x}) := [\hat{\boldsymbol{\psi}}'(\mathbf{x})](\mathbf{x}) = \nabla\hat{\boldsymbol{\psi}}(\mathbf{x})^\dagger \hat{\odot}_1 \mathbf{x}^*\,. \tag{5.32}$$

If the vector function $\hat{\boldsymbol{\psi}}$ is scalar valued, that is, $V_{\mathbb{F}} = \mathbb{F}$, so is $[\hat{\boldsymbol{\psi}}'(\mathbf{x}_0)](\mathbf{x})$. Therefore, rewriting equality (5.29) as $[\hat{\boldsymbol{\psi}}'(\mathbf{x}_0)](\mathbf{x}) = \mathbf{x}^* \hat{\odot}_1 \nabla\hat{\boldsymbol{\psi}}(\mathbf{x}_0)$, we can conclude that $\nabla\hat{\boldsymbol{\psi}}(\mathbf{x}_0)$ is a covector. The correspondent vector to this covector is here defined as $\mathrm{grad}\hat{\boldsymbol{\psi}}(\mathbf{x}_0)$, that is, vector

$$\mathrm{grad}\hat{\boldsymbol{\psi}}(\mathbf{x}_0) := \Phi^{-1} \circ \nabla\hat{\boldsymbol{\psi}}(\mathbf{x}_0)\,, \tag{5.33}$$

where Φ is defined in theorem 6. Therefore, from (5.27),

$$[\hat{\boldsymbol{\psi}}'(\mathbf{x}_0)](\mathbf{x}) = \mathbf{x}\cdot \operatorname{grad}\hat{\boldsymbol{\psi}}(\mathbf{x}_0) = \overline{\operatorname{grad}\hat{\boldsymbol{\psi}}(\mathbf{x}_0)\cdot \mathbf{x}} \tag{5.34}$$

Since these equalities are valid for all $\mathbf{x}\in U_{\mathbb{F}}$, let $B=\{\hat{\mathbf{u}}_1,\cdots,\hat{\mathbf{u}}_n\}$ be an arbitrary orthonormal basis of $U_{\mathbb{F}}$, from which the n scalars

$$\overline{[\hat{\boldsymbol{\psi}}'(\mathbf{x}_0)]}(\hat{\mathbf{u}}_i) = \operatorname{grad}\hat{\boldsymbol{\psi}}(\mathbf{x}_0)\cdot\hat{\mathbf{u}}_i = \mathbf{f}_i^B[\operatorname{grad}\hat{\boldsymbol{\psi}}(\mathbf{x}_0)]$$

obviously constitute the coordinates of vector $\operatorname{grad}\hat{\boldsymbol{\psi}}(\mathbf{x}_0)$. Therefore,

$$\operatorname{grad}\hat{\boldsymbol{\psi}}(\mathbf{x}_0) = \sum_{i=1}^{n}\overline{[\hat{\boldsymbol{\psi}}'(\mathbf{x}_0)]}(\hat{\mathbf{u}}_i)\hat{\mathbf{u}}_i\,. \tag{5.35}$$

Now, in the case of a vector valued scalar function $\hat{\boldsymbol{\psi}}$, when $U_{\mathbb{F}}=\mathbb{F}$, from mapping $\hat{\boldsymbol{\psi}}:\mathcal{U}\mapsto V_{\mathbb{F}}$, we can write that $[\hat{\boldsymbol{\psi}}'(x)](\beta)=\beta[\hat{\boldsymbol{\psi}}'(x)](1)$ for all $\beta\in\mathbb{F}$, since the derivative is linear. In this context, our interest is to describe the directional derivative

$$\frac{d\hat{\boldsymbol{\psi}}}{dx}(x) := [\hat{\boldsymbol{\psi}}'(x)](1) = \lim_{\alpha\to 0}\frac{\hat{\boldsymbol{\psi}}(x+\alpha)-\hat{\boldsymbol{\psi}}(x)}{\alpha}\,. \tag{5.36}$$

Thereby, following the idea of (5.27), we conclude that vector

$$\frac{d\hat{\boldsymbol{\psi}}}{dx}(x) = \operatorname{grad}\hat{\boldsymbol{\psi}}(x)\,. \tag{5.37}$$

Moreover, if $\hat{\boldsymbol{\psi}}$ is differentiable of order $k\geqslant 1$, we adopt

$$\frac{d^k\hat{\boldsymbol{\psi}}}{dx^k}(x) := [\hat{\boldsymbol{\psi}}^{(k)}(x)](1)\,. \tag{5.38}$$

Moreover, given an arbitrary real number x_0 in a context where $U_{\mathbb{F}}=V_{\mathbb{F}}=\mathbb{R}$, real function $\hat{\psi}$ is real valued and then we have the classical geometric interpretation of real number $[d\hat{\psi}/dx](x_0)=[\hat{\psi}'(x_0)](1)$ as the slope of the tangent line $\hat{\psi}_d$ to the curve $\hat{\psi}$ at x_0. Still in this context and again following the idea of (5.27), scalar

$$\frac{d\hat{\psi}}{dx}(x) = \operatorname{grad}\hat{\psi}(x)\,. \tag{5.39}$$

Now returning to the conditions of definition (5.24), from the directional derivative of $\boldsymbol{\psi}$ along $\mathbf{H}\in\boldsymbol{\mathcal{U}}$ and rule (5.6), it is easy to obtain that

$$\mathbf{H}\odot_m\nabla\boldsymbol{\psi}(\mathbf{X}) = \lim_{\alpha\to 0}\frac{\boldsymbol{\psi}(\mathbf{X}+\alpha\mathbf{H})-\boldsymbol{\psi}(\mathbf{X})}{\alpha}\,. \tag{5.40}$$

For the following important items, the properties of derivatives and their definer conditions presented on previous section were considered; the functions in gradients are obviously differentiable on their domains.

i. If $\boldsymbol{\psi} = \boldsymbol{\psi}_1 + \boldsymbol{\psi}_2$ then $\nabla\boldsymbol{\psi} = \nabla\boldsymbol{\psi}_1 + \nabla\boldsymbol{\psi}_2$;

ii. If $\boldsymbol{\psi} = \mathbf{i}$ then $\nabla\boldsymbol{\psi}(\mathbf{X}) = \nabla\mathbf{i}(\mathbf{X}) = \mathbf{I}$ for all $\mathbf{X} \in \boldsymbol{\mathcal{U}}$;

iii. If $\boldsymbol{\psi} = \boldsymbol{\psi}_1 \circ \boldsymbol{\psi}_2$, where $\boldsymbol{\psi}_1 : \boldsymbol{\mathcal{L}}_{\mathbb{F}}(W^{\times r}) \mapsto \boldsymbol{\mathcal{L}}_{\mathbb{F}}(V^{\times s})$ and $\boldsymbol{\psi}_2 : \boldsymbol{\mathcal{U}} \mapsto \boldsymbol{\mathcal{L}}_{\mathbb{F}}(W^{\times r})$, then

$$\nabla\boldsymbol{\psi}(\mathbf{X}) = \nabla\boldsymbol{\psi}_2(\mathbf{X}) \odot_r \nabla\boldsymbol{\psi}_1 \circ \boldsymbol{\psi}_2(\mathbf{X})\,; \tag{5.41}$$

iv. If $\boldsymbol{\psi}$ is a diffeomorphism then $\nabla\boldsymbol{\psi}(\mathbf{X}) \neq \mathbf{0}$ because

$$\mathbf{I} = \nabla\boldsymbol{\psi}(\mathbf{X}) \odot_s \nabla\boldsymbol{\psi}^{-1} \circ \boldsymbol{\psi}(\mathbf{X})\,, \tag{5.42}$$

where $\mathbf{I} \in \boldsymbol{\mathcal{L}}_{\mathbb{F}}(V^{\times s} \times V^{\times s})$;

v. If $\boldsymbol{\psi}(\mathbf{X}) = \boldsymbol{\psi}_1(\mathbf{X}) \cdot \boldsymbol{\psi}_2(\mathbf{X})$, where $\boldsymbol{\psi}_1 : \boldsymbol{\mathcal{U}} \mapsto \boldsymbol{\mathcal{L}}_{\mathbb{R}}(V^{\times s})$, $\boldsymbol{\psi}_2 : \boldsymbol{\mathcal{U}} \mapsto \boldsymbol{\mathcal{L}}_{\mathbb{R}}(V^{\times s})$ and domain $\boldsymbol{\mathcal{U}} \subset \boldsymbol{\mathcal{L}}_{\mathbb{R}}(U^{\times m})$, then

$$\nabla\boldsymbol{\psi}(\mathbf{X}) = \nabla\boldsymbol{\psi}_1(\mathbf{X}) \odot_s \boldsymbol{\psi}_2(\mathbf{X}) + \nabla\boldsymbol{\psi}_2(\mathbf{X}) \odot_s \boldsymbol{\psi}_1(\mathbf{X})\,; \tag{5.43}$$

vi. If $\nabla\boldsymbol{\psi}$ is a constant function of value $\mathbf{T} \in \boldsymbol{\mathcal{L}}_{\mathbb{R}}(U^{\times m} \times V^{\times s})$, that is, $\nabla\boldsymbol{\psi}(\mathbf{X}) = \mathbf{T}$ for all $\mathbf{X} \in \boldsymbol{\mathcal{U}}$, then

$$\boldsymbol{\psi}(\mathbf{H}_1) - \boldsymbol{\psi}(\mathbf{H}_2) = (\mathbf{H}_1 - \mathbf{H}_2) \odot_m \mathbf{T}\,; \tag{5.44}$$

vii. If $\hat{\boldsymbol{\psi}}$ is a diffeomorphism then

$$\nabla\hat{\boldsymbol{\psi}}^{-1} \circ \hat{\boldsymbol{\psi}}(\mathbf{x}) = \nabla\hat{\boldsymbol{\psi}}(\mathbf{x})^{-1}; \tag{5.45}$$

viii. If $\hat{\boldsymbol{\psi}}(\mathbf{x}) = \hat{\boldsymbol{\psi}}_1(\mathbf{x}) \times \hat{\boldsymbol{\psi}}_2(\mathbf{x})$, where mappings $\hat{\boldsymbol{\psi}}_1 : \mathcal{U} \mapsto U_{\mathbb{R}}$ and $\hat{\boldsymbol{\psi}}_2 : \mathcal{U} \mapsto U_{\mathbb{R}}$ are properly defined to enabling cross products,

$$\hat{\boldsymbol{\psi}}'(\mathbf{x})(\mathbf{h}) = \hat{\boldsymbol{\psi}}_1(\mathbf{x}) \times \hat{\boldsymbol{\psi}}'_{2\mathbf{x}}(\mathbf{h}) + \hat{\boldsymbol{\psi}}'_{1\mathbf{x}}(\mathbf{h}) \times \hat{\boldsymbol{\psi}}_2(\mathbf{x})\,, \forall \mathbf{h} \in \mathcal{U}\,, \tag{5.46}$$

and then

$$D\hat{\boldsymbol{\psi}}(\mathbf{x}) = \hat{\boldsymbol{\psi}}_1(\mathbf{x}) \times D\hat{\boldsymbol{\psi}}_2(\mathbf{x}) + D\hat{\boldsymbol{\psi}}_1(\mathbf{x}) \times \hat{\boldsymbol{\psi}}_2(\mathbf{x})\,. \tag{5.47}$$

ix. If the function in $\hat{\boldsymbol{\varphi}} : \mathcal{U}_1 \times \mathcal{U}_2 \mapsto V_{\mathbb{R}}$, where $\mathcal{U}_i \subset U_{\mathbb{R}}$, is $\mathcal{C}^2$ then

$$\nabla_{\mathbf{x}_1} D_{\mathbf{x}_2} \hat{\boldsymbol{\varphi}}(\mathbf{x}_1, \mathbf{x}_2) = D_{\mathbf{x}_2} \nabla_{\mathbf{x}_1} \hat{\boldsymbol{\varphi}}(\mathbf{x}_1, \mathbf{x}_2)\,. \tag{5.48}$$

x. **Mean Value Theorem**: given the closed interval $[\alpha, \beta] \subset \mathbb{R}$ where $\alpha < \beta$, if the function in $\hat{\psi} : [\alpha, \beta] \mapsto \mathbb{R}$ is continuous on its domain and differentiable on open interval (α, β), then there is always a scalar $\kappa \in [\alpha, \beta]$ where

$$\operatorname{grad}\hat{\psi}(\kappa) = \frac{\hat{\psi}(\beta) - \hat{\psi}(\alpha)}{\beta - \alpha}\,. \tag{5.49}$$

xi. **Mean Value Theorem for Real Valued Vector Functions:** considering the mapping $\hat{\psi} : \mathcal{U} \mapsto \mathbb{R}$, where $\mathcal{U} \subset \mathrm{U}_{\mathbb{R}}$ is open, and an Euclidean affine space $\mathcal{S}^n_{\mathbb{R}}$ defined by that $\mathrm{U}_{\mathbb{R}}$, given arbitrary vectors $\vec{a}, \vec{b} \in \mathcal{U}$, if every point x in the line segment $\overline{ab} \subset \mathcal{S}^n_{\mathbb{R}}$ defines a vector $\vec{x} \in \mathcal{U}$, there is always a point $c \in \overline{ab}$ where

$$(\vec{b} - \vec{a}) \cdot \operatorname{grad} \hat{\psi}(\vec{c}) = \hat{\psi}(\vec{b}) - \hat{\psi}(\vec{a}). \tag{5.50}$$

Proof. The first item is verified through the following development:

$$\begin{aligned}
[[\boldsymbol{\psi}'](\mathbf{X})(\mathbf{H}) &= \{[\boldsymbol{\psi}_1' + \boldsymbol{\psi}_2'](\mathbf{X})\}(\mathbf{H}) \\
\mathbf{H} \odot_m \nabla\boldsymbol{\psi}(\mathbf{X}) &= \boldsymbol{\psi}_1'(\mathbf{X})(\mathbf{H}) + \boldsymbol{\psi}_2'(\mathbf{X})(\mathbf{H}) \\
&= \mathbf{H} \odot_m \nabla\boldsymbol{\psi}_1(\mathbf{X}) + \mathbf{H} \odot_m \nabla\boldsymbol{\psi}_2(\mathbf{X}) \\
&= \mathbf{H} \odot_m [\nabla\boldsymbol{\psi}_1 + \nabla\boldsymbol{\psi}_2](\mathbf{X}).
\end{aligned}$$

Property ii is verified by the second property of directional derivatives, from which the non-zero tensor $[\mathbf{i}'(\mathbf{X})](\mathbf{H}) = \mathbf{H} = \mathbf{H} \odot_m \nabla\mathbf{i}(\mathbf{X})$, together with equality (3.43). Proof of item iii is obtained from the chain rule, where $[\boldsymbol{\psi}'(\mathbf{X})](\mathbf{H}) = \{[\boldsymbol{\psi}_1' \circ \boldsymbol{\psi}_2(\mathbf{X})] \circ \boldsymbol{\psi}_2'(\mathbf{X})\}(\mathbf{H})$. We can develop the right hand side of this equality using equalities (3.44) because $\boldsymbol{\psi}_1' \circ \boldsymbol{\psi}_2(\mathbf{X})$ and $\boldsymbol{\psi}_2'(\mathbf{X})$ are the r-cotensor and m-cotensor of $\nabla\boldsymbol{\psi}_1 \circ \boldsymbol{\psi}_2(\mathbf{X})$ and $\nabla\boldsymbol{\psi}_2(\mathbf{X})$ respectively. Item iii is proved by

$$\begin{aligned}
[\boldsymbol{\psi}'(\mathbf{X})](\mathbf{H}) &= \{[\boldsymbol{\psi}_1' \circ \boldsymbol{\psi}_2(\mathbf{X})] \circ \boldsymbol{\psi}_2'(\mathbf{X})\}(\mathbf{H}) \\
\mathbf{H} \odot_m \nabla\boldsymbol{\psi}(\mathbf{X}) &= \mathbf{H} \odot_m [\nabla\boldsymbol{\psi}_2(\mathbf{X}) \odot_r \nabla\boldsymbol{\psi}_1 \circ \boldsymbol{\psi}_2(\mathbf{X})].
\end{aligned}$$

Now, applying property iii on $\mathbf{i} = \boldsymbol{\psi}^{-1} \circ \boldsymbol{\psi}$ and considering (3.43), proof of (5.42) is straightforward and the statement of item iv is obvious. From the Leibniz's rule we have

$$\begin{aligned}
[\boldsymbol{\psi}'(\mathbf{X})](\mathbf{H}) &= [\boldsymbol{\psi}_1'(\mathbf{X})](\mathbf{H}) \cdot \boldsymbol{\psi}_2(\mathbf{X}) + [\boldsymbol{\psi}_2'(\mathbf{X})](\mathbf{H}) \cdot \boldsymbol{\psi}_1(\mathbf{X}) \\
\mathbf{H} \odot_m [\nabla\boldsymbol{\psi}](\mathbf{X}) &= \mathbf{H} \odot_m [\nabla\boldsymbol{\psi}_1](\mathbf{X}) \odot_s \boldsymbol{\psi}_2(\mathbf{X}) + \mathbf{H} \odot_m [\nabla\boldsymbol{\psi}_2](\mathbf{X}) \odot_s \boldsymbol{\psi}_1(\mathbf{X}) \\
&= \mathbf{H} \odot_m \{[\nabla\boldsymbol{\psi}_1](\mathbf{X}) \odot_s \boldsymbol{\psi}_2(\mathbf{X}) + [\nabla\boldsymbol{\psi}_2](\mathbf{X}) \odot_s \boldsymbol{\psi}_1(\mathbf{X})\},
\end{aligned}$$

which proves item v. For item vi, equality $\boldsymbol{\psi}(\mathbf{X}) = \mathbf{X} \odot_m \mathbf{T} + \mathbf{C}$, labeled by (a) and where $\mathbf{C} \in \boldsymbol{\mathcal{L}}_{\mathbb{R}}(V^{\times s})$ is arbitrary, is a rule from which $\nabla\boldsymbol{\psi}(\mathbf{X}) = \mathbf{T}$ when we consider the product rule of directional derivatives, property ii and equality (3.43) in the context of real fields. The subtraction of equalities (a) defined by $\mathbf{X} = \mathbf{H}_1$ and $\mathbf{X} = \mathbf{H}_2$ proves the item. Item vii is a straighforward consequence of (5.42) and (3.57). Considering $(\mathrm{U}_{\mathbb{R}}, \mathbf{A}_B)$ a three dimensional oriented Euclidean space, we prove the equalities of item viii first by the development

$$\begin{aligned}
\hat{\boldsymbol{\psi}}'(\mathbf{x})(\mathbf{h}) &= \lim_{\alpha \to 0} 1/\alpha \{\mathbf{A}_B \odot_2 [\hat{\boldsymbol{\psi}}_1(\mathbf{x} + \alpha\mathbf{h}) \otimes \hat{\boldsymbol{\psi}}_2(\mathbf{x} + \alpha\mathbf{h})] - \mathbf{A}_B \odot_2 [\hat{\boldsymbol{\psi}}_1(\mathbf{x}) \otimes \hat{\boldsymbol{\psi}}_2(\mathbf{x})]\} \\
&= \mathbf{A}_B \odot_2 \lim_{\alpha \to 0} 1/\alpha \{\hat{\boldsymbol{\psi}}_1(\mathbf{x} + \alpha\mathbf{h}) \otimes \hat{\boldsymbol{\psi}}_2(\mathbf{x} + \alpha\mathbf{h}) - \hat{\boldsymbol{\psi}}_1(\mathbf{x}) \otimes \hat{\boldsymbol{\psi}}_2(\mathbf{x})\}
\end{aligned}$$

and then by adding and subtracting the limit on the right hand side by $\mathbf{A}_B \odot_2 [\hat{\boldsymbol{\psi}}_1(\mathbf{x} + \alpha\mathbf{h}) \otimes \hat{\boldsymbol{\psi}}_2(\mathbf{x})]$ we obtain

$$\hat{\boldsymbol{\psi}}'(\mathbf{x})(\mathbf{h}) = \mathbf{A}_B \odot_2 [\hat{\boldsymbol{\psi}}_1(\mathbf{x}) \otimes \hat{\boldsymbol{\psi}}_{2\mathbf{x}}'(\mathbf{h})] + \mathbf{A}_B \odot_2 \hat{\boldsymbol{\psi}}_{1\mathbf{x}}'(\mathbf{h}) \otimes \hat{\boldsymbol{\psi}}_2(\mathbf{x})],$$

from which we easily verify the first equality of viii. By definition (5.32), proof of (5.47) is straightforward. Now, we prove item ix. From (5.8), (5.14) and (5.32), we can write that

$$D_{\mathbf{x}_1} D_{\mathbf{x}_2} \hat{\boldsymbol{\varphi}}(\mathbf{x}_1, \mathbf{x}_2) = \mathbf{x}_1^* \hat{\odot}_1 \nabla_{\mathbf{x}_1} D_{\mathbf{x}_2} \hat{\boldsymbol{\varphi}}(\mathbf{x}_1, \mathbf{x}_2)$$

and

$$D_{\mathbf{x}_2} D_{\mathbf{x}_1} \hat{\boldsymbol{\varphi}}(\mathbf{x}_1, \mathbf{x}_2) = \mathbf{x}_1^* \hat{\odot}_1 D_{\mathbf{x}_2} \nabla_{\mathbf{x}_1} \hat{\boldsymbol{\varphi}}(\mathbf{x}_1, \mathbf{x}_2).$$

The item is proved by both of these equalities and the so called Theorem of Schwarz or Clairaut's theorem, which states that second derivatives of $\mathcal{C}^2$ functions are symmetric or commute (see LANG[35], Theorem 7.3), that is, $D_{x_1} D_{x_2} \hat{\boldsymbol{\varphi}}(x_1, x_2) = D_{x_2} D_{x_1} \hat{\boldsymbol{\varphi}}(x_1, x_2)$. Now we prove item x by considering a function rule $g(x) = \hat{\psi}(x) - \alpha x$, where α is a real constant and g is obviously continuous on interval $[a, b]$. The condition where $g(a) = g(b)$ implies that $\alpha = (\hat{\psi}(b) - \hat{\psi}(a))/(b - a)$ and that there is a scalar $c \in (a, b)$ where $\operatorname{grad} g(c) = 0$, according to Rolle's Theorem (see Spivak[51], p.193). But since $\operatorname{grad} g(c) = \operatorname{grad} \hat{\psi}(c) - \alpha = 0$, replacing α by its value results (5.49). In order to prove item xi, let the function in $f : [0,1] \mapsto \mathcal{U}$ be described by the rule $f(t) = t\vec{b} + (1-t)\vec{a}$. Therefore, for an arbitrary $t \in [0,1]$, there is a point $x \in \overline{ab}$ where $\hat{\psi}(\vec{x}) = \hat{\psi} \circ f(t)$. Since the function $\hat{\phi} := \hat{\psi} \circ f$ observes the conditions of property x, there is always a real number $\kappa \in [0,1]$ where $\hat{\phi}(1) - \hat{\phi}(0) = \operatorname{grad} \hat{\phi}(\kappa)$. From this equality, considering (5.41) and $\vec{c} = f(\kappa)$, the following equalities prove the property.

$$\hat{\psi}(\vec{b}) - \hat{\psi}(\vec{a}) = \operatorname{grad} \hat{\phi}(\kappa) = \operatorname{grad} f(\kappa) \cdot \operatorname{grad} \hat{\psi} \circ f(\kappa) = (\vec{b} - \vec{a}) \cdot \operatorname{grad} \hat{\psi}(\vec{c}).$$

□

Still in the context of Hilbert spaces, we define from the concept of gradient three important and very useful tensor functions, namely divergence, Laplacian and curl. Considering the mapping $\boldsymbol{\varphi} : \mathfrak{U} \mapsto \mathcal{L}_{\mathbb{F}}(U^{\times m} \times V^{\times s})$ where function $\boldsymbol{\varphi}$ is differentiable on its domain, the function in $\operatorname{div} \boldsymbol{\varphi} : \mathfrak{U} \mapsto \mathcal{L}_{\mathbb{F}}(V^{\times s})$ is called the **divergence** of $\boldsymbol{\varphi}$ if its rule is described by

$$\operatorname{div} \boldsymbol{\varphi}(\mathbf{X}) = \mathbf{I} \odot_{2m} \nabla \boldsymbol{\varphi}(\mathbf{X}), \tag{5.51}$$

where identity tensor $\mathbf{I} \in \mathcal{L}_{\mathbb{F}}(U^{\times m} \times U^{\times m})$ and $\nabla \boldsymbol{\varphi}(\mathbf{X}) \in \mathcal{L}_{\mathbb{F}}(U^{\times m} \times U^{\times m} \times V^{\times s})$. While the gradient always assigns to a tensor $\mathbf{X}$ a tensor of higher order, the divergence assigns to $\mathbf{X}$ a tensor of lower order only when $s < m$. From this concept of divergence and from mapping $\nabla \boldsymbol{\psi} : \mathfrak{U} \mapsto \mathcal{L}_{\mathbb{F}}(U^{\times m} \times V^{\times s})$, it is valid and convenient to define a function

$$\Delta \boldsymbol{\psi} = \operatorname{div} \nabla \boldsymbol{\psi}, \tag{5.52}$$

called the **Laplacian** of $\boldsymbol{\psi}$, where $\Delta \boldsymbol{\psi}(\mathbf{X}) \in \mathcal{L}_{\mathbb{F}}(V^{\times s})$. If $\boldsymbol{\varphi}$ is a unary tensor operator on $\mathfrak{U}$, that is $\boldsymbol{\varphi} : \mathfrak{U} \mapsto \mathcal{L}_{\mathbb{F}}(U^m)$, then its divergence is a scalar function for all $m > 0$.

Now, let $(U_{\mathbb{R}}, \mathbf{A}_B)$ be a three dimensional oriented Euclidean space, $V_{\mathbb{R}}$ another Euclidean space and $\boldsymbol{\varphi} : \mathfrak{U} \mapsto \mathcal{L}_{\mathbb{R}}(U \times V)$ a mapping where $\mathfrak{U}$ is an open subset of $U^*_{\mathbb{R}}$. In this context, considering an open set $\mathfrak{U} \subset U^*_{\mathbb{R}}$, it is possible to define a linear mapping $\operatorname{curl} \boldsymbol{\varphi} : \mathfrak{U} \mapsto \mathcal{L}_{\mathbb{R}}(U \times V)$, whose function is called the **curl** of $\boldsymbol{\varphi}$, with rule

$$\operatorname{curl} \boldsymbol{\varphi}(\boldsymbol{x}^*) = \mathbf{A}_B \odot_2 \nabla \boldsymbol{\varphi}(\boldsymbol{x}^*), \tag{5.53}$$

where tensor $\nabla \boldsymbol{\varphi}(\boldsymbol{x}^*) \in \mathcal{L}_{\mathbb{R}}(U^2 \times V)$. In the specific case of $\boldsymbol{\varphi} : \mathfrak{U} \mapsto U^*_{\mathbb{R}}$, we can define a vector valued vector function $\hat{\boldsymbol{\varphi}} = \Phi^{-1} \circ \boldsymbol{\varphi} \circ \Phi$, from which $\nabla \times \hat{\boldsymbol{\varphi}} := \Phi^{-1} \circ \operatorname{curl} \boldsymbol{\varphi} \circ \Phi$. From this definition, it is clear that axial vector

$$\nabla \times \hat{\boldsymbol{\varphi}}(\boldsymbol{x}) = \mathbf{A}_B \hat{\odot}_2 \nabla \hat{\boldsymbol{\varphi}}(\boldsymbol{x}). \tag{5.54}$$

Similarly to the case of gradients, there are some notable properties of divergence and curl, presented as follows, where the appropriate previous conditions are implicit.

i. If $\boldsymbol{\varphi} = \boldsymbol{\varphi}_1 + \boldsymbol{\varphi}_2$ then $\operatorname{div} \boldsymbol{\varphi} = \operatorname{div} \boldsymbol{\varphi}_1 + \operatorname{div} \boldsymbol{\varphi}_2$ and $\operatorname{curl} \boldsymbol{\varphi} = \operatorname{curl} \boldsymbol{\varphi}_1 + \operatorname{curl} \boldsymbol{\varphi}_2$;

ii. If $\boldsymbol{\varphi} = \boldsymbol{\varphi}_1 \circ \boldsymbol{\varphi}_2$, where mapping $\boldsymbol{\varphi}_1 : \mathcal{L}_{\mathbb{F}}(W^{\times r}) \mapsto \mathcal{L}_{\mathbb{F}}(U^{\times m} \times V^{\times s})$ and mapping $\boldsymbol{\varphi}_2 : \mathcal{U} \mapsto \mathcal{L}_{\mathbb{F}}(W^{\times r})$ are defined, then

$$\operatorname{div} \boldsymbol{\varphi}(\mathbf{X}) = \operatorname{div} \boldsymbol{\varphi}_2(\mathbf{X}) \odot_r \nabla \boldsymbol{\varphi}_1 \circ \boldsymbol{\varphi}_2(\mathbf{X}), \tag{5.55}$$

iii. If $\boldsymbol{\varphi} = \boldsymbol{\varphi}_1 \circ \boldsymbol{\varphi}_2$, where mappings $\boldsymbol{\varphi}_1 : \mathcal{L}_{\mathbb{F}}(U \times W) \mapsto \mathcal{L}_{\mathbb{F}}(U \times V)$ and $\boldsymbol{\varphi}_2 : \mathcal{U} \mapsto \mathcal{L}_{\mathbb{F}}(U \times W)$ are defined, then

$$\operatorname{curl} \boldsymbol{\varphi}(\boldsymbol{x}^*) = \operatorname{curl} \boldsymbol{\varphi}_2(\boldsymbol{x}^*) \odot_2 \nabla \boldsymbol{\varphi}_1 \circ \boldsymbol{\varphi}_2(\boldsymbol{x}^*); \tag{5.56}$$

iv. If $\boldsymbol{\varphi}(\mathbf{X}) = \boldsymbol{\varphi}_1(\mathbf{X}) \cdot \boldsymbol{\varphi}_2(\mathbf{X})$, where mapping $\boldsymbol{\varphi}_1 : \mathcal{U} \mapsto \mathcal{L}_{\mathbb{F}}(U^{\times m} \times V^{\times s})$ and mapping $\boldsymbol{\varphi}_2 : \mathcal{U} \mapsto \mathcal{L}_{\mathbb{F}}(U^{\times m} \times V^{\times s})$ are defined, then

$$\operatorname{div} \boldsymbol{\varphi}(\mathbf{X}) = \operatorname{div} \boldsymbol{\varphi}_1(\mathbf{X}) \odot_s \boldsymbol{\varphi}_2(\mathbf{X}) + \operatorname{div} \boldsymbol{\varphi}_2(\mathbf{X}) \odot_s \boldsymbol{\varphi}_1(\mathbf{X}), \tag{5.57}$$

and for $m = s = 1$,

$$\operatorname{curl} \boldsymbol{\varphi}(\boldsymbol{x}^*) = \operatorname{curl} \boldsymbol{\varphi}_1(\boldsymbol{x}^*) \odot_1 \boldsymbol{\varphi}_2(\boldsymbol{x}^*) + \operatorname{curl} \boldsymbol{\varphi}_2(\boldsymbol{x}^*) \odot_1 \boldsymbol{\varphi}_1(\boldsymbol{x}^*); \tag{5.58}$$

v. If $\boldsymbol{\varphi} : \mathcal{U} \mapsto U^*_{\mathbb{R}}$, where $\mathcal{U} \subset U^*_{\mathbb{R}}$, then $\nabla \times \nabla \hat{\boldsymbol{\varphi}}(\boldsymbol{x}) = \mathbf{0}$.

Proof. Considering definitions (5.51) and (5.53), since partial inner products are left (right) distributive, proof of item i is trivial. Pre multiplying properties (5.41), (5.43) and (5.47) adequately by $\mathbf{I}$ or $\mathbf{A}_B$, proofs of items i, ii, iii and iv are straightforward. For item v, since we already know that $\boldsymbol{u} \cdot \mathbf{A}_B \hat{\odot}_2 \boldsymbol{v}^* \otimes \boldsymbol{w}^* = \mathbf{A}_B(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w})$, if we consider tensor $\nabla \hat{\boldsymbol{\varphi}}(\boldsymbol{x}) = \boldsymbol{v}^* \otimes \boldsymbol{w}^*$, then $\boldsymbol{u} \cdot [\nabla \times \nabla \hat{\boldsymbol{\varphi}}(\boldsymbol{x})] = \mathbf{A}_B(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w})$ for all $\boldsymbol{u} \in \mathcal{U}$. Thereby, choosing vector $\boldsymbol{u} = \boldsymbol{v}$, we have $\boldsymbol{v} \cdot [\nabla \times \nabla \hat{\boldsymbol{\varphi}}(\boldsymbol{x})] = \mathbf{A}_B(\boldsymbol{v}, \boldsymbol{v}, \boldsymbol{w}) = 0$ because $\mathbf{A}_B$ is antisymmetric, and since $\boldsymbol{v} \neq \mathbf{0}$, we prove the item. □

Recalling the concept of higher order differentiation presented on the previous section and the conditions of theorem 12, we can state that if $\boldsymbol{\psi}$ is $k+1$ differentiable at $\mathbf{X}_0$ then there is a unique tensor $\nabla \boldsymbol{\psi}^{(k)}(\mathbf{X}_0) \in \mathcal{L}_{\mathbb{F}}((U^{\times m})^{k+1} \times V^{\times s})$ which the derivative

$$\boldsymbol{\psi}^{(k+1)}(\mathbf{X}_0) \in \mathcal{CL}_{\mathbb{F}}(\mathcal{U}_{k+1}, \mathcal{CL}_{\mathbb{F}}(\mathcal{U}_k, \cdots \mathcal{CL}_{\mathbb{F}}(\mathcal{U}_1, \mathcal{L}_{\mathbb{F}}(V^{\times s}))))$$

is the m-cotensor of and each $\mathcal{U}_i = \mathcal{U}$ defines a Hilbert space of continuous functions. Moreover, choosing arbitrary directions $\mathbf{H}_1, \cdots, \mathbf{H}_{k+1} \in \mathcal{U}$ and specifying $r = m \cdot (k+1)$, statement (5.19) can be rewritten as

$$\mathbf{H}_1 \otimes \cdots \otimes \mathbf{H}_{k+1} \odot_r \nabla \boldsymbol{\psi}^{(k)}(\mathbf{X}) \in \mathcal{L}_{\mathbb{F}}(V^{\times s}). \tag{5.59}$$

5.3 Integration

The integration approach we shall deal with in this section relies fundamentally on Lebesgue's theory, which is based on assigning a real number to a set, whatever measure this scalar expresses. The most commom measure related to set is the geometric concept of size, that is, length, area, volume or **hypervolume** of respectively one, two, three or n dimensional Hilbert spaces. In generic terms, a measure must obey certain rules that require and lead to new important definitions. In this context, from specific scalar functions with domains called measurable, integrable tensor functions can be defined.[3]

From the concept of class and its dependent definitions presented on section 1.2, the double $(M,\mathfrak{M})$ is said to be a **measurable space** when $\mathfrak{M}$ is a σ–ring and a class of the set M. If this set defines a group or a field we also qualify them "measurable". Now, considering the mapping $\mu:\mathfrak{M}\mapsto\mathbb{R}^+$ and an arbitrary disjoint countable class $\{A_1,\cdots,A_n\}\subseteq\mathfrak{M}$, set function μ is called additive if

$$\mu(\bigcup_{i=1}^{n} A_i)=\sum_{i=1}^{n}\mu(A_i)\,. \tag{5.60}$$

This equality implies that $\mu(\emptyset)=\mu(\emptyset\cup\emptyset)=2\mu(\emptyset)$, from which we conclude that $\mu(\emptyset)=0$. Another important conclusion is that, given $A\in\mathfrak{M}$, a set

$$B\subseteq A \implies \mu(B)\leqslant\mu(A)\,. \tag{5.61}$$

Proof. This property is easily verified by considering the obvious equality $A=B\cup(A/B)$ and then from development $\mu(B)+\mu(A)=\mu(A\cup B)=\mu(B)+\mu(A/B)+\mu(B)$, since μ is additive. □

Moreover, μ is called a **measure** if

$$\mu(\bigcup_{i=1}^{\infty} B_i)=\sum_{i=1}^{\infty}\mu(B_i)\,, \tag{5.62}$$

where $\{B_1,B_2,\cdots\}\subseteq\mathfrak{M}$ is also an arbitrary disjoint countable class. Now, we choose to restrict our study to measures where

$$\mu(A)=0 \iff A=\emptyset\,, \tag{5.63}$$

a condition that avoids terms like "almost all" or "almost everywhere" in forthcoming definitions. In this context, a double constituted by a measurable space and a measure, that is $((M,\mathfrak{M}),\mu)$, is called a **measure space**, which will hereafter be represented by the triple $(M,\mathfrak{M},\mu)$.

Given a measurable field $(\mathbb{R},\mathfrak{F})$ and an arbitrary element $B\in\mathfrak{F}$, the function in $s:M\mapsto\mathbb{R}$ is called **measurable** if the preimage $R_B^{-1}\in\mathfrak{M}$. Now, considering an arbitrary set $A\in\mathfrak{M}$, a class $\mathfrak{A}=\{A_1,\cdots,A_n\}\subseteq\mathfrak{M}$ is called a **partition** of A if it is disjoint and $A=\bigcup_{i=1}^{n}A_i$. Given a sequence $\alpha_1,\cdots,\alpha_n\in\mathbb{R}$, the measurable function in $s_{\mathfrak{A}}:M\mapsto\mathbb{R}$ is called **step** or **simple** with respect to $\mathfrak{A}$ if its rule is described by

$$s_{\mathfrak{A}}(x)=\sum_{i=1}^{n}\alpha_i\mathbf{1}_{A_i}(x)\,, \tag{5.64}$$

[3]The concise theory presented in this section is based mainly on the texts of HALMOS[25] and LANG[35].

where the function in $\mathbf{1}_{A_i} : A_i \mapsto \mathbb{R}$ is commonly known as the **indicator** or **characteristic** function of A_i, defined by

$$\mathbf{1}_{A_i}(x) = \begin{cases} 0 & \text{if } x \notin A_i \\ 1 & \text{if } x \in A_i \end{cases}.$$

The step function $s_{\mathfrak{A}}$ is defined to be **integrable** on set $E \in \mathfrak{M}$ when for every $\alpha_i \neq 0$, scalar $\mu(E \cap A_i)$ is finite. If this is the case, the real number

$$\int_E s_{\mathfrak{A}} := \sum_{i=1}^{n} \alpha_i \mu(E \cap A_i) \tag{5.65}$$

is called the **integral** of $s_{\mathfrak{A}}$ on E. Now, if $\mathfrak{B} = \{B_1, \cdots, B_m\} \subseteq \mathfrak{M}$ is another partition of A, we conclude that there is no element of A which does not belong to some $A_i \cap B_j$ and then it is obvious that class

$$\mathfrak{C} = \{A_1 \cap B_1, A_1 \cap B_2, \cdots, A_1 \cap B_m, A_2 \cap B_1, \cdots, A_n \cap B_m\}$$

is disjoint and also a partition of A. Moreover, since $A = \bigcup_{i=1}^{n} A_i = \bigcup_{j=1}^{m} B_i$, it is obvious that $A_i = \bigcup_{j=1}^{m} A_i \cap B_j$ and $B_j = \bigcup_{i=1}^{n} B_j \cap A_i$. From previous definition, we can write

$$\int_E s_{\mathfrak{C}} = \sum_{i=1}^{n} \sum_{j=1}^{m} \gamma_{ij} \mu(E \cap A_i \cap B_j),$$

which means that

$$s_{\mathfrak{C}}(x) = \sum_{i=1}^{n} \sum_{j=1}^{m} \gamma_{ij} \mathbf{1}_{A_i \cap B_j}(x) = \sum_{i=1}^{n} \underbrace{\sum_{j=1}^{m} \gamma_{ij} \mathbf{1}_{A_i \cap B_j}(x)}_{\alpha_i \mathbf{1}_{A_i}(x)} = \sum_{j=1}^{m} \underbrace{\sum_{i=1}^{n} \gamma_{ij} \mathbf{1}_{B_i \cap A_j}(x)}_{\beta_j \mathbf{1}_{B_j}(x)}.$$

Thereby, from these equalities, scalar

$$\begin{aligned}
\int_E s_{\mathfrak{A}} &= \sum_{i=1}^{n} \alpha_i \mu(E \cap A_i) \\
&= \sum_{i=1}^{n} \alpha_i \mu(E \cap \bigcup_{j=1}^{m} A_i \cap B_j) \\
&= \sum_{j=1}^{m} \sum_{i=1}^{n} \alpha_i \mu(E \cap A_i \cap B_j) \\
&= \sum_{j=1}^{m} \beta_j \mu(E \cap \bigcup_{i=1}^{n} B_j \cap A_i) \\
&= \sum_{j=1}^{m} \beta_j \mu(E \cap B_j)
\end{aligned}$$

$$= \int_E s_{\mathfrak{B}} \,,$$

since intersection is distributive on union and μ is additive according to (5.60). This equality enables us to replace the subscript of the partition of A in the step function for the set A itself, that is, $s_A : M \mapsto \mathbb{R}$, since the integral of s_A is the same for any partition of $A \subseteq M$. For practical purposes, we impose that the step function s_A is specified so as

$$E \cap A \neq \emptyset, \tag{5.66}$$

which means that there exists at least one element X in the partition of A where $\mu(E \cap X) \neq 0$. Now, from definition (5.65), given a step function described by the rule $s_E(x) = \sum_{i=1}^{m} \mathbf{1}_{E_i}(x)$ where $\{E_1, \cdots, E_m\} \subseteq \mathfrak{M}$ is a partition of E, since μ is additive, it is straightforward to conclude that

$$\int_E s_E = \mu(E) \,. \tag{5.67}$$

In the set of all measurable functions on $(M, \mathfrak{M}, \mu)$, let (S, ρ) be a normed metric space of integrable step functions on E with norm[4] and metric defined by

$$\|s\|_1 = \int_E |s| \qquad \text{and} \qquad \rho(s_1, s_2) = \|s_1 - s_2\|_1 \,,$$

where $|s|(x) := |s(x)|$ and $s, s_1, s_2 \in S$. In this context, if a Cauchy sequence of integrable step functions $s_1, s_2, \cdots \in S$ exists and converges to the measurable functional in $f : M \mapsto \mathbb{R}$, we say that f is integrable or Labesgue integrable on E and scalar

$$\int_E f := \lim_{i \to \infty} \int_E s_i \tag{5.68}$$

is the integral or Labesgue integral of f on E. It is important to say that if $s_1, s_2, \cdots \in S$ and $t_1, t_2, \cdots \in S$ are convergent Chauchy sequences of integrable step functions, so is the sequence $s_1 + t_1, s_2 + t_2, \cdots$. Moreover, for arbitrary $\alpha, \beta \in \mathbb{R}$, if the function in $g : M \mapsto \mathbb{R}$ is measurable,

$$s_1, s_2, \cdots \to f \;\wedge\; t_1, t_2, \cdots \to g \implies \alpha s_1 + \beta t_1, \alpha s_2 + \beta t_2, \cdots \to \alpha f + \beta g \,. \tag{5.69}$$

Proof. Since scalar multiplication does not affect the convergence of Cauchy sequences $s_1, s_2, \cdots$ and $t_1, t_2, \cdots$, we will verify only the effect of the sum. First we prove that $h_1 + r_1, h_2 + r_2, \cdots$, where $h_i = \alpha s_i$ and $r_i = \beta t_i$, is also a Cauchy sequence. If both expressions $\lim_{\min(i,j) \to \infty} \rho(h_i, h_j) = 0$ and $\lim_{\min(i,j) \to \infty} \rho(r_i, r_j) = 0$ are valid, then, by summing these two equalities, we obtain $\lim_{\min(i,j) \to \infty} \rho(h_i, h_j) + \rho(r_i, r_j) = 0$. From the property of triangular inequality of norms, since

$$\rho(h_i, h_j) + \rho(r_i, r_j) = \|h_i - h_j\| + \|r_i - r_j\| \geqslant \|h_i - h_j + r_i - r_j\| \,,$$

[4]For a generic function f, scalar $\|f\|_1$ is actually the *seminorm* of f, but because of our restriction (5.66) on a step function s, statement $\|s\|_1 = 0 \iff s = 0$ is valid.

it is obvious that $\lim_{\min(i,j)\to\infty} \|(h_i + r_i) - (h_j + r_j)\| = 0$, which proves the sum to be a Cauchy sequence. The convergence to $\alpha f + \beta g$ can be proved by this same procedure. □

Given the above conditions, we present some fundamental properties of integrals:

i. If f is integrable and nonnegative, then its integral on E is also nonnegative;

ii. If f is integrable on every element of disjoint class $\{E_1, \cdots, E_m\} \subseteq \mathfrak{M}$, then

$$\int_{\bigcup_{i=1}^{m} E_i} f = \sum_{i=1}^{m} \int_{E_i} f; \tag{5.70}$$

iii. If f is integrable on $E \in \mathfrak{M}$ and $\mu(E) = 0$, then

$$\int_E f = 0; \tag{5.71}$$

iv. If f and g are integrable on $E \in \mathfrak{M}$ and scalars $\alpha, \beta \in \mathbb{R}$ are arbitrary, then

$$\int_E (\alpha f + \beta g) = \alpha \int_E f + \beta \int_E g. \tag{5.72}$$

Proof. In order to prove item i, let's specify a sequence of integrable step functions $s_1, s_2, \cdots$ where

$$s_k(x) = \sum_{i=1}^{n} (f(x) + k^{-1}) \mathbf{1}_{A_i^{(k)}}(x)$$

is defined from partitioning $A^{(k)} \subset M$. This is obviously a Cauchy sequence that converges to f as $k \to \infty$ and since the coefficients of s_k are nonnegative, then

$$\int_E f = \lim_{k\to\infty} \sum_{i=1}^{n} (f(x) + k^{-1}) \mu(E \cap A_i) \geqslant 0.$$

Considering equality (5.62) and definition (5.68), we prove item ii by

$$\begin{aligned}
\int_{\bigcup_{i=1}^{m} E_i} f &= \lim_{k\to\infty} \int_{\bigcup_{i=1}^{m} E_i} s_k \\
&= \lim_{k\to\infty} \sum_{j=1}^{n} \alpha_j^{(k)} \mu(A_j^{(k)} \cap \bigcup_{i=1}^{m} E_i) \\
&= \sum_{i=1}^{m} \lim_{k\to\infty} \sum_{j=1}^{n} \alpha_j^{(k)} \mu(E_i \cap A_j^{(k)}) \\
&= \sum_{i=1}^{m} \lim_{k\to\infty} \int_{E_i} s_k = \sum_{i=1}^{m} \int_{E_i} f.
\end{aligned}$$

Now, since μ is nonnegative and additive, that is $\mu(X \cup Y) = 0 \iff \mu(X) = \mu(Y) = 0$, and partitioned set A on (5.65) is arbitrary, then choosing E to be partitioned we prove item iii by

$$\int_E f = \lim_{k\to\infty} \sum_{i=1}^{n} \alpha_i^{(k)} \mu(E \cap E_i^{(k)}) = \lim_{k\to\infty} \sum_{i=1}^{n} \alpha_i^{(k)} \underbrace{\mu(E_i^{(k)})}_{0} = 0$$

For item iv, let's specify two partitions $\{A_1^{(k)}, \cdots, A_n^{(k)}\}$ and $\{B_1^{(k)}, \cdots, B_m^{(k)}\}$ of a set $A^{(k)} \subset M$ which define two step functions $s_k = \sum_{i=1}^n \alpha_i^{(k)} \mathbf{1}_{A_i^{(k)}}$ and $t_k = \sum_{j=1}^m \beta_j^{(k)} \mathbf{1}_{B_j^{(k)}}$ in such a way that

$$\begin{aligned}
\alpha \int_E f + \beta \int_E g &= \lim_{k\to\infty} \alpha \int_E s_k + \lim_{k\to\infty} \beta \int_E t_k \\
&= \lim_{k\to\infty} \sum_{i=1}^n \alpha\alpha_i^{(k)} \mu(E \cap A_i^{(k)}) + \lim_{k\to\infty} \sum_{j=1}^m \beta\beta_j^{(k)} \mu(E \cap B_j^{(k)}) \\
&= \lim_{k\to\infty} \sum_{i=1}^n \sum_{j=1}^m \alpha\alpha_i^{(k)} \mu(E \cap A_i^{(k)} \cap B_j^{(k)}) + \sum_{i=1}^n \sum_{j=1}^m \beta\beta_j^{(k)} \mu(E \cap A_i^{(k)} \cap B_j^{(k)}) \\
&= \lim_{k\to\infty} \sum_{i=1}^n \sum_{j=1}^m (\alpha\alpha_i^{(k)} + \beta\beta_j^{(k)}) \mu(E \cap A_i^{(k)} \cap B_j^{(k)}). \qquad \text{(a)}
\end{aligned}$$

Now, let's verify that a sum of two step functions is also a step function:

$$\begin{aligned}
\alpha s_k + \beta t_k &= \sum_{i=1}^n \alpha\alpha_i^{(k)} \mathbf{1}_{A_i^{(k)}} + \sum_{j=1}^m \beta\beta_j^{(k)} \mathbf{1}_{B_j^{(k)}} \\
&= \sum_{i=1}^n \sum_{j=1}^m \alpha\alpha_i^{(k)} \mathbf{1}_{A_i^{(k)} \cap B_j^{(k)}} + \sum_{i=1}^n \sum_{j=1}^m \beta\beta_j^{(k)} \mathbf{1}_{A_i^{(k)} \cap B_j^{(k)}} \\
&= \sum_{i=1}^n \sum_{j=1}^m (\alpha\alpha_i^{(k)} + \beta\beta_j^{(k)}) \mathbf{1}_{A_i^{(k)} \cap B_j^{(k)}}.
\end{aligned}$$

Thereby, from property (5.69), we have

$$\int_E \alpha f + \beta g = \lim_{k\to\infty} \int_E \alpha s_k + \beta t_k = \lim_{k\to\infty} \sum_{i=1}^n \sum_{j=1}^m (\alpha\alpha_i^{(k)} + \beta\beta_j^{(k)}) \mu(E \cap A_i^{(k)} \cap B_j^{(k)}),$$

which is equal to result (a). □

Now, our goal is to integrate a broader class of functions other than functionals, namely, tensor and vector functions. Considering $\mathcal{L}_{\mathbb{R}}(U^{\times m})$ and $\mathcal{L}_{\mathbb{R}}(V^{\times s})$ Euclidean tensor spaces, let $\boldsymbol{\psi} : \mathcal{U} \mapsto \mathcal{L}_{\mathbb{R}}(V^{\times s})$ be a tensor mapping, where $\mathcal{U} \subseteq \mathcal{L}_{\mathbb{R}}(U^{\times m})$, and $f : \mathcal{U} \mapsto \mathbb{R}$ a mapping whose functional is described by the rule $f = \mathbf{T}^{\circledast} \circ \boldsymbol{\psi}$, such that cotensor $\mathbf{T}^{\circledast}$ is the representative tensor functional of an arbitrary tensor $\mathbf{T} \in \mathcal{L}_{\mathbb{R}}(V^{\times s})$. We define that function $\boldsymbol{\psi}$ is integrable on $\mathcal{U}$ if and only if functional f is integrable on $\mathcal{U}$. If this is the case, by specifying a rule

$$\mathbf{h}(\mathbf{X}) = \int_{\mathcal{U}} \mathbf{X}^{\circledast} \circ \boldsymbol{\psi}, \qquad (5.73)$$

where tensor functional $\mathbf{h} \in \mathcal{L}_{\mathbb{R}}^*(V^{\times s})$ is linear because of property iv above, there exists a unique tensor $\mathbf{H} \in \mathcal{L}_{\mathbb{R}}(V^{\times s})$ which $\mathbf{h}$ is the cotensor of, according to corollary 12.1. In this context, tensor

$$\int_{\mathcal{U}} \boldsymbol{\psi} := \mathbf{H} \qquad (5.74)$$

is called the integral of tensor function $\boldsymbol{\psi}$ on $\mathcal{U}$. Therefore, since $\mathbf{h}(\mathbf{T}) = \mathbf{T} \cdot \mathbf{H}$, we conclude from (5.73) that scalar

$$\mathsf{T}\cdot\int_{\mathfrak{U}}\boldsymbol{\psi}=\int_{\mathfrak{U}}\mathsf{T}^{\circledast}\circ\boldsymbol{\psi}. \tag{5.75}$$

Similarly as above, for the case of the vector valued vector function $\hat{\boldsymbol{\psi}}=\Phi^{-1}\circ\boldsymbol{\psi}\circ\Phi$, as defined in (5.28), we also specify a vector functional $f=\boldsymbol{\nu}^*\circ\hat{\boldsymbol{\psi}}$ and a rule

$$\mathbf{h}^*(\mathbf{x})=\int_{\mathfrak{U}}\mathbf{x}^*\circ\hat{\boldsymbol{\psi}}, \tag{5.76}$$

where $\boldsymbol{\nu}^*,\mathbf{h}^*\in V^*_{\mathbb{R}}$, from which $\hat{\boldsymbol{\psi}}$ is said to be integrable on $\mathfrak{U}$. Thereby, vector

$$\int_{\mathfrak{U}}\hat{\boldsymbol{\psi}}:=\mathbf{h} \tag{5.77}$$

is called the integral of $\hat{\boldsymbol{\psi}}$ on $\mathfrak{U}$ and, since $\mathbf{h}^*(\boldsymbol{\nu})=\boldsymbol{\nu}\cdot\mathbf{h}$,

$$\boldsymbol{\nu}\cdot\int_{\mathfrak{U}}\hat{\boldsymbol{\psi}}=\int_{\mathfrak{U}}\boldsymbol{\nu}^*\circ\hat{\boldsymbol{\psi}}. \tag{5.78}$$

5.4 Integral Identities

In rough geometrical terms, domains whose boundary shapes show a specific form of regularity – where sharp edges and different types of self intersection are not allowed, as shown in figure 5.1 – enable identities that equalize integrations on surfaces and interiors, which are fundamental for Mathematical Physics. First of all, we recall the biunivocal correspondence between an affine space and a vector space established by defining a coordinate system, as shown in section 4.1, in order to henceforth make geometrical statements and assumptions on finite dimensional Euclidean spaces. In this context and from the concept of smoothness presented on section 5.1, a set of strong mathematical requirements that avoid the irregularities cited above is given by introductory texts on Differential Geometry, such as Do Carmo[17], where surfaces called smooth are defined. However, for the sake of generality, we choose to study a less restrictive regularity condition, which defines the so called Lipschitz surface[5]; that also does not present the above irregularities and is less complex to work with than smooth surfaces. We shall adopt this concept by adapting the accessible approach of Adams & Fournier[1], which is more than sufficient for our purposes[6].

Let $U_{\mathbb{R}}$ be a n-dimensional Euclidean space, $n\geqslant 2$, and $\mathfrak{U}$ one of its proper subsets, which we assume to be a domain of a certain function. Considering an arbitrary non zero scalar $\alpha\in\mathbb{R}^+$, a disjoint cover $\mathfrak{C}=\{\widehat{C}_1,\cdots,\widehat{C}_r\}$ of boundary $\partial\mathfrak{U}$ and a class $\{A_1,\cdots,A_r\}$ where each set $A_i\subset C_i$ defines a $(n-1)$-dimensional Euclidean subspace of $U_{\mathbb{R}}$ from the first $n-1$ dimensions of $U_{\mathbb{R}}$, the subset $\mathfrak{U}$ is called a **Lipschitz domain** and its boundary a **Lipschitz surface** if both of the following conditions are observed.

[5]See Ciarlet[12], p.32.

[6]For a more rigorous treatment, see Agranovich[2], chapter 3.

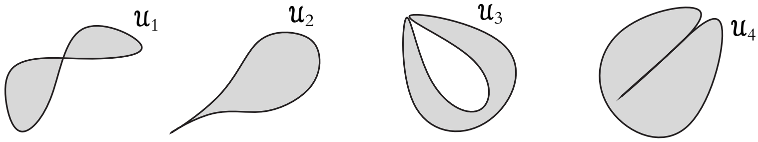

Figure 5.1 – Examples of two dimensional irregular domains.

i. Every element $\widehat{C}_j$ of $\mathfrak{C}$ must have a specific shape such that arbitrary vectors

$$\mathbf{u},\mathbf{v} \in \{\mathbf{x} \in \mathfrak{U} : \min_{\mathbf{y}\in\partial\mathfrak{U}} \rho(\mathbf{x},\mathbf{y}) < \alpha\},$$

where distance $\rho(\mathbf{u},\mathbf{v}) < \alpha$, are also elements of the subset

$$\{\mathbf{x} \in \widehat{C}_j : \min_{\mathbf{y}\in\partial C_j} \rho(\mathbf{x},\mathbf{y}) > \alpha\};$$

ii. Every set $\mathfrak{U} \cap \widehat{C}_j$ must be on the same "side" of boundary $\partial\mathfrak{U}$, whose points coordinates must constitute the graph of a Lipschitz continuous function. In other words, there must exist a set of Lipschitz continuous functionals $\{\mathbf{g}_1,\cdots,\mathbf{g}_r\}$ where $\mathbf{g}_j : A_j \mapsto \mathbb{R}$ in such a way that for an orthonormal basis $B_j = \{\hat{\mathbf{v}}_1^{(j)},\cdots,\hat{\mathbf{v}}_n^{(j)}\}$ of $\widehat{C}_j$, the nonnegative real number

$$|\mathbf{f}_n^{B_j}(\mathbf{x})| < |\mathbf{g}_j[\sum_{i=1}^{n-1} \mathbf{f}_i^{B_j}(\mathbf{x})\hat{\mathbf{v}}_i^{(j)}]|, \forall \mathbf{x} \in \mathfrak{U} \cap \widehat{C}_j.$$

Considering $\mathfrak{U} \subset U_{\mathbb{R}}$ a Lipschitz domain, we additionally require orientation of the Lipschitz surface $\partial\mathfrak{U}$ on each of its points. In order to adequately accomplish this, let $\mathcal{W}$ be a $(n-1)$-dimensional subspace of $U_{\mathbb{R}}$ and the function in $\mathbf{g} : \mathcal{W} \mapsto \mathbb{R}$ be Lipschitz continuous whose graph describes $\partial\mathfrak{U}$ in such a way that

$$\mathbf{f}_n^B(\mathbf{v}) = \mathbf{g}(\sum_{i=1}^{n-1} \mathbf{f}_i^B(\mathbf{v})\hat{\mathbf{u}}_i), \forall \mathbf{v} \in \partial\mathfrak{U},$$

where $B = \{\hat{\mathbf{u}}_1,\cdots,\hat{\mathbf{u}}_n\}$ is an arbitrary orthonormal basis of $U_{\mathbb{R}}$. Thereby, the function in mapping $\mathbf{s} : U_{\mathbb{R}} \mapsto \mathbb{R}$ where

$$\mathbf{s}(\mathbf{v}) = \mathbf{f}_n^B(\mathbf{v}) - \mathbf{g}(\sum_{i=1}^{n-1} \mathbf{f}_i^B(\mathbf{v})\hat{\mathbf{u}}_i) = 0 \quad \text{and} \quad \operatorname{grad}\mathbf{s}(\mathbf{v}) \neq 0, \forall \mathbf{v} \in \partial\mathfrak{U}, \tag{5.79}$$

is called the **implicit function** of $\partial\mathfrak{U}$ and since it is constant, equality $[\mathbf{s}'(\mathbf{v})](\mathbf{h}) = 0$ is valid along any direction $\mathbf{h} \in \partial\mathfrak{U}$. Therefore, from (5.34), we can say that $\mathbf{h}\cdot\operatorname{grad}\mathbf{s}(\mathbf{v}) = 0$ for all $\mathbf{h} \in \{\{\mathbf{v}\}_{\mathbf{v}} \cap \partial\mathfrak{U}\}$, that is, vector $\operatorname{grad}\mathbf{s}(\mathbf{v})$ is orthogonal to surface $\partial\mathfrak{U}$ at arbitrary $\mathbf{v} \in \partial\mathfrak{U}$. From this conclusion, surface $\partial\mathfrak{U}$ is called **orientable** if it is possible to define a mapping $\mathbf{n} : \partial\mathfrak{U} \mapsto \mathsf{U}_{\mathbb{R}}$ whose function, called the **unit normal function** on $\partial\mathfrak{U}$, is described by

$$\mathbf{n}(x) = \frac{\operatorname{grad} \mathbf{s}(x)}{\|\operatorname{grad} \mathbf{s}(x)\|}. \tag{5.80}$$

Additionally to $\mathbf{n}$, it is true that an implicit function $-\mathbf{s}$ leads to a unit normal function $-\mathbf{n}$ and when one of these directions $\mathbf{n}$ or $-\mathbf{n}$ is specified as positive, we say that $\partial\mathfrak{U}$ is **oriented**. For the purposes of this study, we choose $\mathbf{n}$ to orientate the surfaces. In the case of three-dimensional Euclidean spaces, the values of function $\mathbf{n}$ are outward unit vectors with respect to the surface and the values of $-\mathbf{n}$ are inward unit vectors.

From what we have studied so far, it is finally possible to attain the objective of this section which is to present a fundamental identity of elementary Calculus, called the **Gauss-Ostrogradsky** or the **Divergence** theorem. We shall not deal with its most general form, the notorious Generalized Stokes' theorem, because it would demand us to cover an extensive material on Exterior Calculus, which is out of the scope of this work[7]. In general terms, the following theorem relates projections of function values on unit normals along the domain boundary (flux) with the divergence of the function on the domain interior.

Theorem 20 – Gauss-Ostrogradsky

Considering $\mathsf{U}_{\mathbb{R}}$ and $\mathcal{L}_{\mathbb{R}}(\mathsf{U} \times \mathsf{V}^{\times s})$ finite dimensional Euclidean spaces, $\boldsymbol{\varphi} : \mathfrak{U} \mapsto \mathcal{L}_{\mathbb{R}}(\mathsf{U} \times \mathsf{V}^{\times s})$ a continuous mapping where $\mathfrak{U} \subset \mathsf{U}_{\mathbb{R}}$ is a Lipschitz domain, $\boldsymbol{\varphi}$ continuously differentiable on interior $\hat{\mathfrak{U}}$, surface $\partial\mathfrak{U}$ oriented by a unit normal vector function $\mathbf{n}$ and $\boldsymbol{\varphi}_{\odot} : \mathfrak{U} \mapsto \mathcal{L}_{\mathbb{R}}(\mathsf{U} \times \mathsf{V}^{\times s})$ a continuous mapping where $\boldsymbol{\varphi}_{\odot}(x) = \mathbf{n}(x) \odot_1 \boldsymbol{\varphi}(x)$,

$$\int_{\hat{\mathfrak{U}}} \operatorname{div} \boldsymbol{\varphi} = \int_{\partial\mathfrak{U}} \boldsymbol{\varphi}_{\odot}\,.$$

Proof. Instead of developing here this very long demonstration, we invite the interested reader to see the elegant proof presented by FIGUEIREDO[20] in the context of n-dimensional Euclidean spaces. □

Corollary 20.1 – Gauss-Ostrogradsky for Gradients

For a continuous mapping $\boldsymbol{\varphi}_{\otimes} : \mathfrak{U} \mapsto \mathcal{L}_{\mathbb{R}}(\mathsf{U} \times \mathsf{V}^{\times s})$ where $\boldsymbol{\varphi}_{\otimes}(x) = \mathbf{n}(x) \otimes \boldsymbol{\varphi}(x)$,

$$\int_{\hat{\mathfrak{U}}} \nabla \boldsymbol{\varphi} = \int_{\partial\mathfrak{U}} \boldsymbol{\varphi}_{\otimes}\,.$$

Proof. Considering identity tensor $\mathbf{I} \in \mathcal{L}_{\mathbb{R}}(\mathsf{U}^2)$ and property (5.75), we can write that

$$\mathbf{I} \odot_2 \int_{\hat{\mathfrak{U}}} \nabla \boldsymbol{\varphi} = \mathbf{I} \odot_2 \int_{\partial\mathfrak{U}} \boldsymbol{\varphi}_{\otimes}$$

[7] For an introduction to Exterior Calculus, see FLANDERS[21], LOOMIS & STERNBERG[37] and SPIVAK[50].

$$\int_{\hat{\mathfrak{U}}} \mathbf{I}^{\circledast} \circ \nabla \boldsymbol{\varphi} = \int_{\partial \mathfrak{U}} \mathbf{I}^{\circledast} \circ \boldsymbol{\varphi}_{\otimes} .$$

But from definition (3.46), we have equalities $\mathbf{I}^{\circledast} \circ \nabla \boldsymbol{\varphi}(\mathbf{x}) = \mathbf{I} \cdot \nabla \boldsymbol{\varphi}(\mathbf{x}) = \mathbf{I} \odot_2 \nabla \boldsymbol{\varphi}(\mathbf{x}) = \operatorname{div} \boldsymbol{\varphi}(\mathbf{x})$ and from (3.20) we have $\mathbf{I}^{\circledast} \circ \boldsymbol{\varphi}_{\otimes}(\mathbf{x}) = \mathbf{I} \cdot [\mathbf{n}(\mathbf{x}) \otimes \boldsymbol{\varphi}(\mathbf{x})] = \mathbf{n}(\mathbf{x}) \odot_1 \boldsymbol{\varphi}(\mathbf{x}) = \boldsymbol{\varphi}_{\odot}(\mathbf{x})$. From theses results, we arrive at the previous theorem. □

Corollary 20.2 – Gauss-Ostrogradsky for Curls

In the case of continuous mappings $\boldsymbol{\varphi} : \mathfrak{U} \mapsto \mathrm{U}_{\mathbb{R}}$ and $\boldsymbol{\varphi}_{\times} : \mathfrak{U} \mapsto \mathrm{U}_{\mathbb{R}}$ where $(\mathrm{U}_{\mathbb{R}}, \mathbf{A}_{\mathrm{B}})$ is an three dimensional oriented Euclidean space and $\boldsymbol{\varphi}_{\times}(\mathbf{x}) = \mathbf{n}(\mathbf{x}) \times \boldsymbol{\varphi}(\mathbf{x})$,

$$\int_{\hat{\mathfrak{U}}} \operatorname{curl} \boldsymbol{\varphi} = \int_{\partial \mathfrak{U}} \boldsymbol{\varphi}_{\times} .$$

Proof. Considering (5.75), (4.20), (3.46), we have $\mathbf{n}(\mathbf{x}) \times \boldsymbol{\varphi}(\mathbf{x}) = \mathbf{A}_{\mathrm{B}} \odot_2 [\mathbf{n}(\mathbf{x}) \otimes \boldsymbol{\varphi}(\mathbf{x})] = \mathbf{A}_{\mathrm{B}}^{\circledast} \circ \boldsymbol{\varphi}_{\otimes}(\mathbf{x})$. Thereby, we arrive at corollary 20.1 by

$$\begin{aligned}
\int_{\hat{\mathfrak{U}}} \operatorname{curl} \boldsymbol{\varphi} &= \int_{\partial \mathfrak{U}} \boldsymbol{\varphi}_{\times} \\
\int_{\hat{\mathfrak{U}}} \mathbf{A}_{\mathrm{B}} \odot_2 \nabla \boldsymbol{\varphi} &= \int_{\partial \mathfrak{U}} \mathbf{A}_{\mathrm{B}}^{\circledast} \circ \boldsymbol{\varphi}_{\otimes} \\
\mathbf{A}_{\mathrm{B}} \odot_2 \int_{\hat{\mathfrak{U}}} \nabla \boldsymbol{\varphi} &= \mathbf{A}_{\mathrm{B}} \odot_2 \int_{\partial \mathfrak{U}} \boldsymbol{\varphi}_{\otimes} .
\end{aligned}$$

□

For the next theorem, we first need to define a function that assigns a tangent vector to each of the elements of a curve[8]. Given an open interval $(\alpha_1, \alpha_2) \subset \mathbb{R}$ and a n–dimensional Euclidean space $\mathrm{U}_{\mathbb{R}}$, if the function in $r : (\alpha_1, \alpha_2) \mapsto \mathrm{U}_{\mathbb{R}}$ is continuously differentiable on its domain and its image is a curve $\mathfrak{R}$, then r is called the parametric function of $\mathfrak{R}$. In this context, we are interested in parametric functions that are injections, when their generated curves are classified as **simple**. From the terms of definition (5.24), we conclude mapping $\nabla r : (\alpha_1, \alpha_2) \mapsto \mathrm{U}_{\mathbb{R}}$, where $\nabla r(x)$ is a n–dimensional vector, and then specify a mapping $\boldsymbol{\tau} : \mathfrak{R} \mapsto \mathrm{U}_{\mathbb{R}}$ such that $\mathfrak{R}$ is a simple curve and $\boldsymbol{\tau}$ is called the **unit tangent vector function** on $\mathfrak{R}$ if it is described by

$$\boldsymbol{\tau}(\mathbf{x}) = \frac{\nabla r \circ r^{-1}(\mathbf{x})}{\|\nabla r \circ r^{-1}(\mathbf{x})\|} . \tag{5.81}$$

There are Lipschitz surfaces in three dimensional Euclidean spaces for which it is possible to define a boundary, like elliptic paraboloids for example. In general terms, the intersection of a bounded set A with an open ball centered at an arbitrary element of boundary ∂A is not a subset of the interior of A. Considering this idea on a three dimensional Euclidean space $\mathrm{U}_{\mathbb{R}}$, the boundary of a surface $\mathfrak{S} \subset \mathrm{U}_{\mathbb{R}}$ is a curve $\partial \mathfrak{S} \subset \mathfrak{S}$ constituted by points $\mathbf{x} \in \partial \mathfrak{S}$ such that $B_{\mathbf{x},r} \cap \mathfrak{S} \not\subset \widehat{\mathfrak{S}}$ for all $r \in \mathbb{R}$. For three dimensional Lipschitz surfaces with boundaries that are simple curves, we present the following

[8] For an introductory approach of the theory of curves, see Do Carmo[17] or Kreyszig[33].

classical theorem that establishes a relationship between an integral on a surface and an integral on its boundary.

Theorem 21 – Kelvin-Stokes

Let $(\mathsf{U}_{\mathbb{R}}, \mathbf{A}_{\mathrm{B}})$ *be a three dimensional oriented Euclidean space and* $\boldsymbol{\varphi} : \mathcal{S} \mapsto \boldsymbol{\mathcal{L}}_{\mathbb{R}}(\mathsf{U} \times \mathsf{V})$ *a mapping where* $\boldsymbol{\varphi}$ *is continuously differentiable on Lipschitz surface* $\widehat{\mathcal{S}} \subset \mathsf{U}_{\mathbb{R}}$*, which is oriented by an outward unit normal vector function* $\mathbf{n}$*. Considering* $\boldsymbol{\tau}$ *a unit tangent vector function on boundary* $\partial\mathcal{S} \subset \mathcal{S}$ *such that* $\boldsymbol{\varphi}_{\partial\mathcal{S}} : \partial\mathcal{S} \mapsto \mathsf{V}_{\mathbb{R}}$ *and* $\boldsymbol{\varphi}_{\widehat{\mathcal{S}}} : \widehat{\mathcal{S}} \mapsto \mathsf{V}_{\mathbb{R}}$ *are continuous mappings where* $\boldsymbol{\varphi}_{\partial\mathcal{S}}(\boldsymbol{x}) = \boldsymbol{\tau}(\boldsymbol{x}) \odot_1 \boldsymbol{\varphi}(\boldsymbol{x})$ *and* $\boldsymbol{\varphi}_{\widehat{\mathcal{S}}}(\boldsymbol{x}) = \mathbf{n}(\boldsymbol{x}) \odot_1 \operatorname{curl} \boldsymbol{\varphi}(\boldsymbol{x})$*, then*

$$\int_{\widehat{\mathcal{S}}} \boldsymbol{\varphi}_{\widehat{\mathcal{S}}} = \int_{\partial\mathcal{S}} \boldsymbol{\varphi}_{\partial\mathcal{S}} \,.$$

Proof. We recommend the reader to see the proof presented by BACKUS[5], p.131. □

Bibliography

[1] Adams, R.; Fournier J. **Sobolev Spaces.** 2 ed. Kidlington: Academic Press, 2003.

[2] Agranovich, M. S. **Sobolev Spaces, Their Generalizations, and Elliptic Problems in Smooth and Lipschitz Domains.** 1 ed. London: Springer-Verlag, 2015.

[3] Appleby, P. G. et al *On the Classification of Isotropic Tensors* in **Glasgow Mathematical Journal**, Vol. 29, No. 2. Cambridge: Cambridge University Press, 1987, pp. 185–196.

[4] Artzy, R. **Linear Geometry.** 1 ed. Reading: Addison-Wesley Publishing Company, 1965.

[5] Backus, G. **Continuum Mechanics.** Available at `<http://samizdat.mines.edu/backus/>`, Accessed on feb 2004.

[6] Bishop, R. L.; Goldberg, S. I. **Tensor Analysis on Manifolds.** 1 ed. Mineola: Dover Publications, 1980.

[7] Bowen, R. M.; Wang, C. -C. **Introduction to Vectors and Tensors.** 1 ed. Mineola: Dover Publications, 2008.

[8] Cameron, P. J. **Sets, Logic and Categories.** 1 ed. London: Springer-Verlag, 1999.

[9] Cervantes, M. **Don Quijote De La Mancha** 1 ed. Madrid: Castalia, 2010.

[10] Chadwick, P. **Continuum Mechanics: Concise Theory and Problems.** 2 ed. Mineola: Dover Publications, 1999.

[11] Christoffel, E. B. *Ueber die Transformation der homogenen Differentialausdrücke zweiten Grades* in **Journal für die Reine und Angewandte Mathematik.** Vol. 70, 1869, pp. 46–70.

[12] Ciarlet, P. G. **Mathematical Elasticity – Volume 1: Three Dimensional Elasticity.** 1. ed. Amsterdam: Elsevier Science Publishers BV, 1988.

[13] Connel, E. H. **Elements of Abstract and Linear Algebra.** Available at `<http://www.math.miami.edu/~ec/book/>`, Accessed on feb 2004.

[14] Conrad, F. **Tensor Products.** Available at `<http://www.math.uconn.edu/~kconrad/blurbs/linmultialg/tensorprod.pdf>`, Accessed on sept 2017.

[15] Derbyshire, J. **Unknown Quantity: a Real and Imaginary History of Algebra.** 1 ed. Washington,D.C.: Joseph Henry Press, 2006.

[16] Dieudonné, J. **Foundations Of Modern Analysis.** 1 ed. New York: Academic Press, 1969.

[17] Do Carmo, M. P. **Differential Geometry of Curves and Surfaces.** 2 ed. Mineola: Dover Publications, 2016.

[18] Dodson, C.T.J.; Poston, T. **Tensor Geomery: The Geometric Viewpoint and its Uses.** 1 ed. Berlin: Springer-Verlag Inc., 1991.

[19] Einstein, A. *Zur allgemeinen Relativitätstheorie* in **Preussische Akademie der Wissenschaften, Sitzungsberichte.** Berlim, 1915, pp. 778–786, 799–801.

[20] Figueiredo, D. G. *A Simplified Proof of the Divergence Theorem* in **The American Mathematical Monthly**, Vol. 71, No. 6, 1964, pp. 619–622.

[21] Flanders, H. **Differential Forms With the Applications to the Physical Sciences.** 1 ed. Mineola: Dover Publications, 1989.

[22] Gibbs, J. W. **Elements of Vector Analysis Arranged for the Use of Students in Physics.** 1. ed. New Haven: Tuttle, Morehouse & Taylor, 1881.

[23] Gurtin, M. E. **An Introduction to Continuum Mechanics.** 1. ed. San Diego: Academic Press, 1981.

[24] Gurtin, M. E. et al. **The Mechanics and Thermodynamics of Continua.** 1. ed. Cambridge: Cambridge University Press, 2010.

[25] Halmos, P. R. **Measure Theory.** 1 ed. New York: Springer-Verlag Inc., 1974.

[26] Halmos, P. R. **Naive Set Theory.** 1 ed. New York: Springer-Verlag Inc., 1974.

[27] Hamilton, W. R. *On Quaternions; or on a new System of Imaginaries in Algebra* in **The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science: Volume XXIX.** 1846, pp. 26–31.

[28] Hermann, R. **Ricci and Levi-Civita's Tensor Analysis Paper.** 1 ed. Brookline: Math Sci Press, 1975.

[29] Hermes, H.; Markwald, W. *Foundations of Mathematics* in **Fundamentals of Mathematics: Volume I.** 1 ed. Massachusetts: MIT Press, 1987, pp. 1–86.

[30] Johnson, C. R. *Positive Definite Matrices* in **The American Mathematical Monthly** Vol. 77, No. 3. New York: Mathematical Assossiation of America, 1970, pp. 259–264.

[31] Knuth, D. E. **The Art of Computer Programming, Volume 1.** 1 ed. Boston: Addison-Wesley, 1997.

[32] Kreyszig, E. **Introductory Functional Analysis With Applications.** 1 ed. Boston: John Wiley & Sons, 1978.

[33] Kreyszig, E. **Differential Geometry.** 1 ed. Mineola: Dover Publications, 1991.

[34] Lanczos, C. **Space Through The Ages: The Evolution of Geometrical Ideas From Pythagoras to Hilbert and Einstein.** 1 ed. London: Academic Press, 1970.

[35] Lang, S. **Real and Functional Analysis.** 3 ed. Berlin: Springer-Verlag, 1993.

[36] Levi-Civita, T. **The Absolute Differential Calculus (Calculus of Tensors).** 1 ed. Mineola: Dover Publications, 2013.

[37] Loomis, L. H.; Sternberg, S. **Advanced Calculus.** 1 ed. Hackensack: World Scientific Publishing Co., 2014.

[38] Luque, J.; Thibon, J. **Hyperdeterminantal Calculations of Selberg's And Aomoto's Integrals.** Available at <`https://arxiv.org/pdf/math/0607410.pdf`>, Accessed on july 2018.

[39] Malament, D. B. **Notes on Geometry and Spacetime.** Available at <`http://www.lps.uci.edu/home/fac-staff/faculty/malament/geometryspacetimedocs/`>, Accessed on aug 2006.

[40] Marcus, M. **Finite Dimensional Multilinear Algebra, Part I.** 1 ed. New York: Marcel Dekker Inc., 1973.

[41] Milne J. S. **Group Theory.** Available at <http://www.jmilne.org/math/CourseNotes/>, Accessed on feb 2004.

[42] Munroe, M. E. **Measure and Integration** 1 ed. Reading: Addison-Wesley Publishing Company, 1971.

[43] Ogden, R.W. **Non-Linear Elastic Deformations.** 1 ed. Mineola: Dover Publications, 1997.

[44] Ricci, G.; Levi-Civita, T. *Méthodes de Calcul Différentiel Absolu et Leurs Applications* in **Mathematische Annalen.** Vol 54, No. 1-2. Berlin: Springer, 1900, pp. 125-201.

[45] Segel, L. **Mathematics Applied to Continuum Mechanics.** 1 ed. New York: Macmillan Publishing Co., 1977.

[46] Shen, A.; Vereshchagin, N. K. **Basic Set Theory** 1 ed. Providence: American Mathematical Society, 2002.

[47] Snapper, E.; Troyer, R. J. **Metric Affine Geometry.** 1 ed. New York: Academic Press, 1971.

[48] Strang, G. **Linear Algebra and Its Applications.** 4 ed. Pacific Grove: Brooks-Cole Publishing, 2006.

[49] Soutas-Little, R. W. **Elasticity.** 1 ed. Mineola: Dover Publications, 1999.

[50] Spivak, M. **Calculus on Manifolds: a Modern Approach to Classical Theorems of Advanced Calculus** 1 ed. Princeton: Perseus Books Publishing, 1965.

[51] Spivak, M. **Calculus** 4 ed. Huston: Publish or Perish Inc, 2008.

[52] Truesdell, C. A.; Noll, W. **The Non-Linear Field Theories of Mechanics.** 1 ed. Berlin: Springer-Verlag Inc., 1992.

[53] Vaisman, I. **Foundations of Three Dimensional Euclidean Geometry.** New York: Marcel Dekker Inc., 1980.

[54] Vinberg, E. B. **A Course in Algebra.** Rhode Island: American Mathematical Society, 2003.

[55] Voigt, W. **Die Fundamentalen Physikalischen Eigenschaften der Krystalle in Elementarer Darstellung.** Leipzig: Verlag Von Veit & Comp., 1898.

[56] Waerden, B. **A History of Algebra: From al-Khwārizmī to Emmy Noether** 1 ed. Berlin: Springer-Verlag Inc., 1985.

[57] Waleffe, F. **Tensor Product and Tensors.** Available at `<http://www.math.wisc.edu/~milewski/>`, Accessed on october 2004.

[58] Weyl, H. **The Classical Groups: Their Invariants and Representations.** 2 ed. Princeton: Princeton University Press, 1953.

[59] Wouk, A. **A Course of Applied Functional Analysis.** 1 ed. New York: John Wiley & Sons Inc., 1979.

[60] Zeidler, E. **Applied Functional Analysis: Main Principles and Their Applications.** 1 ed. New York: Springer-Verlag Inc., 1995.

Part II

Elementary Continuum Mechanics

Chapter

6

Early Continuum Mechanics: a Biobibliographical Account

This chapter deals with some early historical aspects of Continuum Mechanics in order to situate the reader on the subject, which is a fundamental part of the science of Mechanics as a whole. By history of Continuum Mechanics we mean the direct contributions to the study of motion, strength and deformation of non rigid bodies; in other words, contributions to the mechanics of deformable bodies. Here, it is sufficient to rely on intuitive notions of all these concepts, which will be rigorously defined in the following chapters. Since our description is focused on non rigid bodies, works that contributed chiefly to the mechanics of material points will not be covered; otherwise, our history would be very, very long.

6.1 Leonardo and Galileo

An adequate historical account of the main contributors to the subject of Continuum Mechanics must start with the painter Leonardo da Vinci (1452–1519). Born in the city of Vinci, a comune of Florence, in the italian region of Tuscany, da Vinci became an apprentice, at the age of fourteen, in Andrea del Verrocchio's workshop, the most famous florentine artist of the time. Already informally educated on Latin, Geometry and Mathematics, at the workshop, he was taught, besides artistic abilities, a wide range of technical skills, including engineering, architecture, metallurgy and chemistry. Among his notable works, he left not only famous paintings like *Mona Lisa* and *The Last Supper*, but also notebooks, whose descriptive annotations and sketches cover a great variety of themes, from prosaic issues of his everyday life to exercises on Mathematics, architectural drawings, studies on painting and sculpture, engineering, cartography, astronomy, optics, botany, human anatomy, motions of fluids and solids, machine design and so on. Although it is not our purpose to criticize or profoundly analyze da Vinci's work, we must inform that in his notebooks there can be found the earliest known records on the mechanics of deformable bodies: a study on tensile test in notebook Codex Atlanticus f.222r; studies on beams and columns in notebooks Codices Atlanticus f.562r, 908r, Forster I f.88v, 89r, Madrid I f.84v, 85v, 135v, 136r, 139r, 177v and Paris Manuscript A f.45v. On figure 6.1, some of these records are

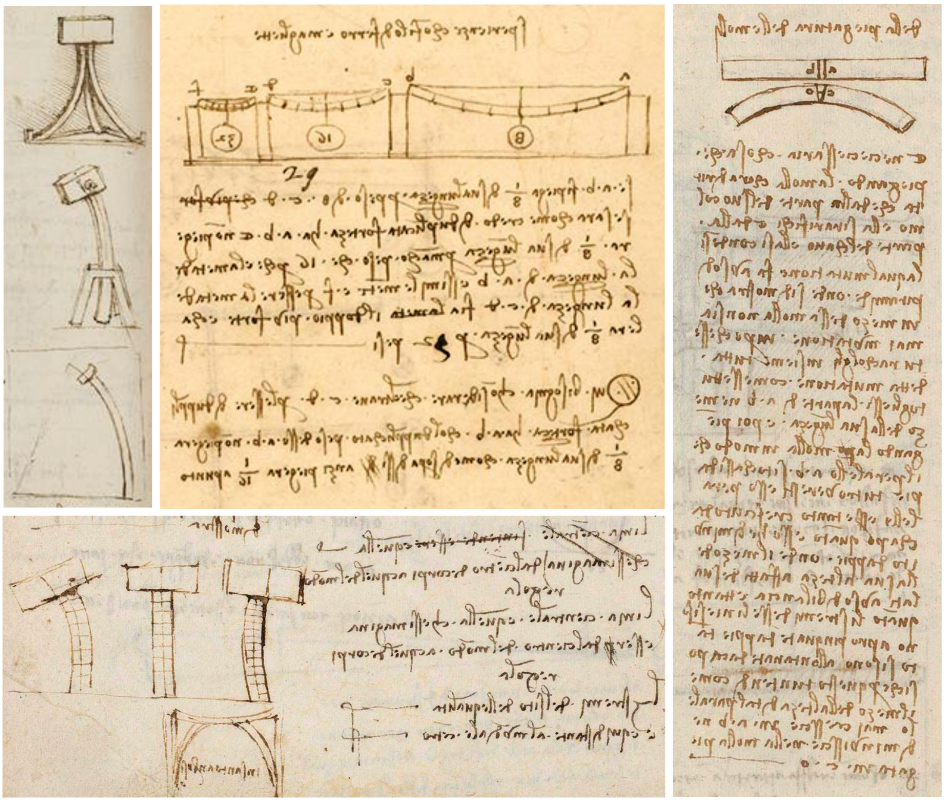

Figure 6.1 – Paris A f.45v, Atlanticus f.908r, Madrid I f.84v (right), 177v (bottom).

shown: sketches Paris A f.45v and Madrid I f.177v are studies on buckling of columns and the others are studies on bending of beams. In particular, the text on Madrid I f.84v, translated by Zammattio[71], reads: *"If a straight spring is bent, it is necessary that its convex part become thinner and its concave part, thicker. This modification is pyramidal, and consequently, there will never be a change in the middle of the spring. You shall discover, if you consider all of the aforementioned modifications, that by taking part 'ab' in the middle of its length and then bending the spring in a way that the two parallel lines, 'a' and 'b' touch a the bottom, the distance between the parallel lines has grown as much at the top as it has diminished at the bottom. Therefore, the center of its height has become much like a balance for the sides.*

And the ends of those lines draw as close at the bottom as much as they draw away at the top. From this you will understand why the center of the height of the parallels never increases in 'ab' nor diminishes in the bent spring at 'co'." Moreover, the striking sketch on Codex Atlanticus f.222r, entitled *Testing The Strength of Iron wires of Various Lengths*, shown in figure 6.2, is the first known record on strength of materials. The drawing on the

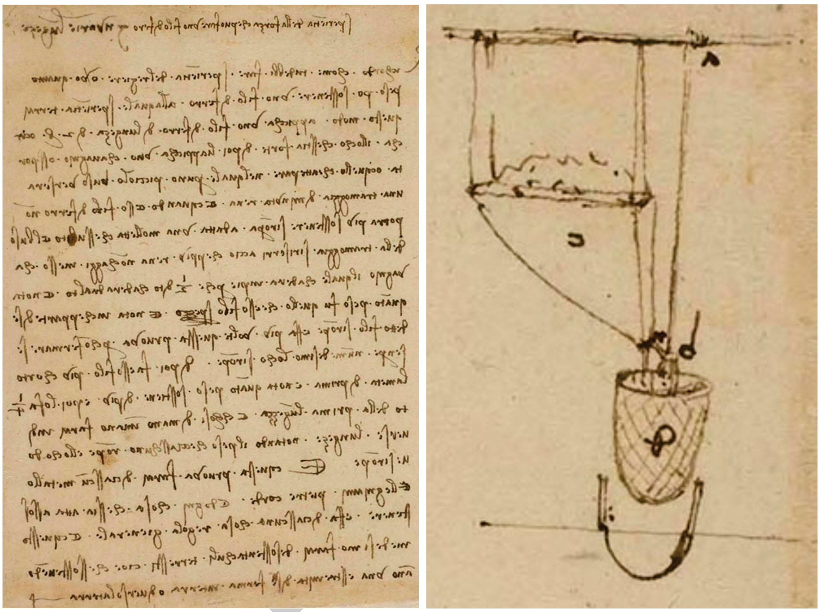

Figure 6.2 – Codex Atlanticus f.222r.

figure represents a test scheme of an iron wire with length 'ab' and a given thickness. The wire suspends an initially empty basket 'q' which is slowly loaded with sand by a hopper 'c'. When the wire breaks, a spring closes the hopper and the basket falls a short distance into a hole, so as to not drop the sand. The sand in the basket is then weighted to obtain the strength of the wire. Still on this figure, an excerpt of the text reads: *"The object of this test is to find the load an iron wire can carry. Attach an iron wire 2 braccia long to something which will firmly support it, then attach a basket or similar container to the wire and feed into the basket some fine sand through a small hole placed at the end of the hopper. A spring is fixed so that it will close the hole as soon as the wire breaks. The basket is not upset while falling, since it falls through a very short distance. The weight of sand and the location of the fracture of the wire are to be recorded. The test is repeated several times to check the results. Then a wire of 1/2 the previous length is tested and the additional weight it carries is recorded; then a wire of 1/4 length is tested and so forth, noting the ultimate strength and the location of the fracture.*[1]*"* Concerning the subject of Mechanics, which is our interest

[1] Lund & Byrne[36], p.3.

here, some scholars on the field contest the alleged scientific value of da Vinci's works: the most aggressive criticism is given by Truesdell[65], which doubts, in his typical verbosity, whether da Vinci's sketches and annotations on engineering are really his creation or merely reproduce common technical knowledge of his time. Moreover, Truesdell argues that da Vinci's proposed physical laws, all of them linear, conceived intuitively from simple rules of three, are mostly wrong and the right ones just happened to be correct, since the laws in Physics are either linear or nonlinear. According to Dugas[20], da Vinci *"...cuts the figure of a gifted amateur... He tackled all kinds of problem, often with more faith than success. Frequently he returned to the same problem by very different paths, and did not scruple to contradict himself."* Therefore, from today's perspective, it is not possible to state that da Vinci's work on Mechanics directly anticipated relevant definitions or concepts on the field, despite his undeniable outstanding efforts as a curious and creative individual. But even if da Vinci's contributions were scientifically substantial, they would not serve as reference for the study of subsequent professional scholars because none of his notebooks was published in his time or shortly after his death.

Figure 6.3 – Presumed self-portrait by Leonardo da Vinci (Uffizi Gallery).

Unlike Leonardo da Vinci, his fellow countryman Galileo Galilei (1564–1642) received a formal education and, at the age of sixteen, enrolled the university of his hometown, Pisa, to study medicine. Among other disciplines, the course required knowledge on philosophy, provided by professors Francesco Buonamici and Girolamo Borro, as well as Mathematics and Astronomy, given by Father Filippo Fantoni, a monk. Both Buonamici and Borro were strict aristotelians, which meant that their teachings conflicted with the Christian dogmas of creationism, afterlife, immortality of the soul and others. Some biographers believe that this teachings intensified the impetuous spirit of Galileo, which would cause him trouble with fellow catholic academicians and also with the Inquisition years later. At a certain moment in 1583, Galileo attended lectures on Euclid's *Elements* given by the former mathematician Ostilio Ricci, instructor at the Medici court. Amazed by the performance of the student, Ricci managed to convince Vincenzo Galilei, Galileo's father, to allow his son to switch from Medicine, which would assure a promising career, to Mathematics and Natural Philosophy, areas that captivated the young Galileo during his studies at the faculty. In 1585, following a not unusual practice among noble youngsters of his time, Galileu dropped out of the university, without a degree, in order to get a job: at Florence and Siena, he started private teaching his most interested subjects as a preparation for an eventual post of Mathematics professor on some important university. Meanwhile he kept attending the lectures of Ricci and also studying the work of other eminent mathematicians such as Giovanni Battista Benedetti, Christoph Clavius and Guidobaldo del Monte. In 1588, Galileo was invited to lecture at the Florentine Academy about the location and dimensions of the hell in Dante's *Inferno*[2]. The favorable impressions that these lectures caused on the tuscan nobility and the quitting of Father Fantoni, as well as recommendations from professor Clavius and other mathematicians, enable Galileo to get a chair on Mathematics at the University of Pisa in 1589. After three years of professorship in Pisa, Galileo was convinced by the fellow Medicine professor Girolamo Mercuriale to apply for a vacant chair on Mathematics at the University of Padua, which paid three times more than Pisa. Starting in December 1592, he worked at Padua for eighteen years and then returned to Florence as a mathematician of the Medici court.

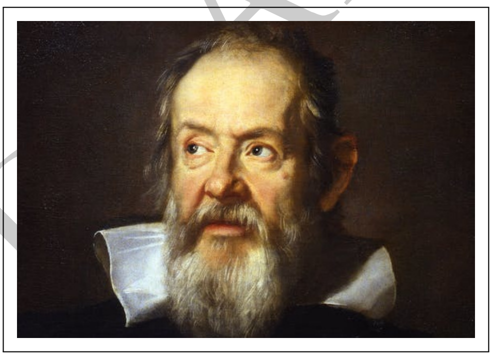

Figure 6.4 – Portrait of Galileo Galilei by Justus Sustermans (Uffizi Gallery).

During his Padua years, which he considered the happiest of his life, Galileo developed his most notable works, particularly on Kinematics and Astronomy. He also endeavored to develop studies on strength of materials, which would be published in *Dialogs Concerning Two New Sciences* forty years later. From this book, whose dialogs between characters Salviati, Sagredo and Simplicio occur during four days, it is noteworthy for our purposes to present the attempt of Galileo to calculate the strength of cantilever beams. On the second day, Salviati proposes a solution to the problem depicted on the right in figure 6.5. He says to his interlocutors that "*the* momento *of*

[2] See Wallace[68].

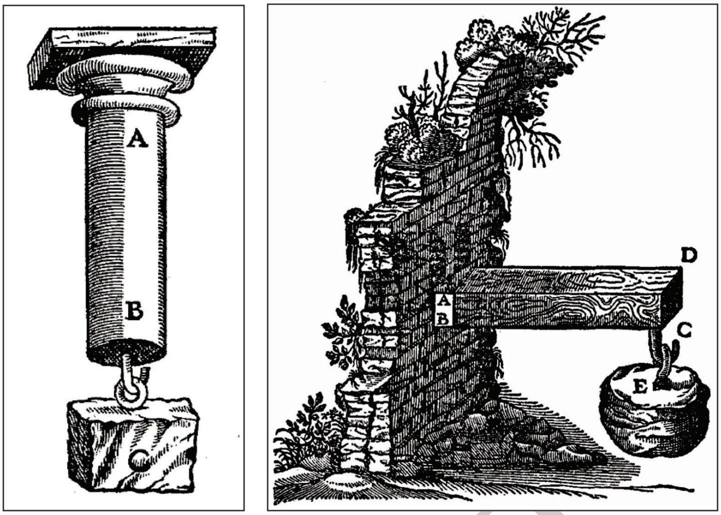

Figure 6.5 – Beams in tension and bending (Galileo[23]).

the force applied at C bears to the momento *of the resistance, found in the thickness of the prism, i. e., in the attachment of the base BA to its contiguous parts, the same ratio which the length CB bears to half the length BA; if we now define absolute resistance to fracture as that offered to a longitudinal pull* – [drawing on the left in previous figure]... *then it follows that the absolute resistance of the prism BD is to the breaking load placed at the end of the lever BC in the same radio as the length BC is to the half of AB, in the case of a prism, or the semidiameter in the case of a cylinder.*[3]" In other words, if we call h the thickness of the prism, $\tau_u.h.AB$ his "absolute resistance to fracture" and E_u his "breaking load", then Salviati states that

$$\frac{\tau_u.h.AB}{E_u} = \frac{BC}{AB/2} \quad \text{or} \quad E_u = \frac{\tau_u.h.AB^2}{2.BC}\,.$$

But if Simplicio and Sagredo were more argumentative and cautious, they would verify this categoric statement and would see that Salviati's proposal is dangerously wrong for the case of beams made of steel: the ultimate load for rectangular cross sectional cantilever steel beams on pure bending is actually three times less[4] than E_u. When Salviati specifies a *momento* from a load with an arm $AB/2$ that balances the *momento* of the lever BC caused by E, he inadvertently defines that the tensile load distribution on section AB is uniform and that there is a longitudinal compressive load concentrated at fulcrum B in order to balance this tensile load distribution. The reasons that made Galileo arrive at this conclusion are subject of debate: Higdon et al[29] speculate that he might have observed the failure of bending beams made of stone. On brittle

[3] Galileo[23], p.115.
[4] See Popov[51], p.286.

materials like this, the shape of the fracture caused by axial loads and pure bending are very similar, which probably induced him to wrongly consider an axial resistance on the bending beam. Since Galileo was the first to formally address the problem of the strength of a cantilever beam subjected to pure bending, it is commonly known as **Galileo's Problem**[5].

6.2 Mariotte

The French scholar Edmé Mariotte (?–1684) also contributed to the mechanics of deformable bodies, but he is best known for his work on the properties of air, entitled *Discourse on the Nature of Air*, published in 1676, in which he proposed the notorious inverse relation between volume and pressure: today, we call it the Boyle–Mariotte Law for gases. The early life of Mariotte is completely unknown and the first registered document of his existence is a letter he sent from Dijon in 1668 to the dutch physicist Christiaan Huygens, where he announced the discovery of the blind spot in human eye. It is also unclear how the Paris Academy of Sciences became acquainted with Mariotte's works on plant physiology, but the excellent impression that he caused on the academicians when invited to go to Paris in order to present his theories and experiments soon enabled his engagement at the Academy on 27 July 1667, as a *physicien*. Following a usual practice of the scholars of his time, Mariotte wrote articles on a great variety of subjects: Mathematics, Mechanics, Astronomy, medicine, hydrology, musical theory, plant biology, among others. His body of published work is extensive and the most important texts are the following: four essays, gathered under the title *Essays on Pshyics*, of which the already mentioned *Discourse on the Nature of Air* is part; a *Treatise on the Motion of Water and Other Fluid Bodies*, published unfinished and postumously in 1686, where Mariotte studies natural springs, artificial fountains and the flow of water through pipes; and a *Treatise on the Collision or Shock of Bodies*, first published in 1673, which covers the topic of elastic and inelastic collisions and shows Mariotte's abilities as a gifted experimenter. It is important to account that Mariotte's work heavily relied on the subjects commonly studied and produced by others at the Paris Academy and, apart from the volume–pressure relation, there is no relevant discoveries attributed to him. However, since rigorous and exhaustive experimentation is the fundamental basis of his efforts, some authors recognize him as the man who introduced the experimental physics into France. In corollary 6 of his third axiom of motion, Newton[46] cites Mariotte's book on collisions: "*But Wren additionally proved the truth of these rules before the Royal Society by means of an experiment with pendulums, which the eminent Mariotte soon after thought worthy to be made the subject of a whole book.*"

In part V of his *Treatise on the Motion of Water and Other Fluid Bodies*, called *On Water Motion and Pipe Strength*, Mariotte undertakes the study of Galileo's Problem in

[5]See Benvenuto[4], p.177.

order to correctly design the dimensions of water pipes because he verified experimentally that Galileo's proposition of an axial strength on the bending beam was incorrect for the case of iron and wood. On the discourse II of this same part V, he starts by specifying that the axial strength of a beam will be measured not by a maximum load (absolute resistance to fracture), but by a maximum extension: the beam performs a certain extension in order to sustain a certain load; there is then a maximum extension beyond which the beam breaks. This is the first known quantitative consideration of deformation in the study of the resistance of deformable bodies. Still in part V, discourse II, Mariotte also describes in literal form a load-extension relation: "*... if a solid of wood needs to extend two lines to break, and a weight of 500 pounds make this extension, a weight of 125 pounds make it extend about half a line, 250 pounds, about one line, etc. Thus, each extension will balance with a certain weight.*[6]". The drawing on the left in figure 6.6 is a model conceived by Mariotte to study the behavior of the material fibers of a cantilever beam using identical chords DI, GL and HM subjected to tension by a lever N with a given fulcrum C. The rectangle ACQP represents the part of the cantilever,

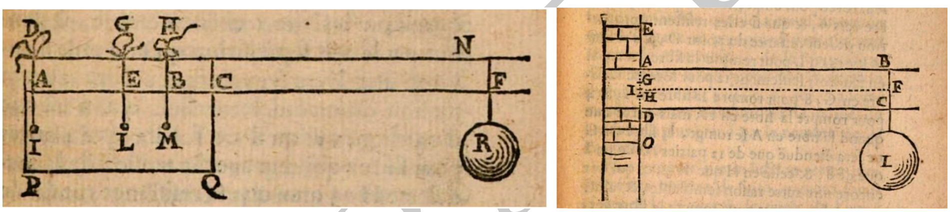

Figure 6.6 – Extension of chords and Galileo's Problem (Mariotte[38], pp.353-355).

shown on the right in this same figure, that is inside the wall and AC = 2.EC = 4.BC. Mariotte then measured the extensions ϵ_{DI}, ϵ_{GL} and ϵ_{HM} of the chords until rupture for different applied loads R and observed the relation $\epsilon_{DI} = 2.\epsilon_{GL} = 4.\epsilon_{HM}$, which is the same for the loads on the chords, assuming his load-extension law. For an infinite number of chords, the load distribution on face AC decreases linearly to zero from A to C and therefore the resultant load acts at a distance AC/3 from A. Now considering the drawing on the right in figure 6.6, from his extensive experiments, Mariotte came to the conclusion that the beam fibers on the face AD under the middle point I are in compression and the fibers over I are subjected to tension. These fibers in tension behave just like the chords of his model, that is, their extension decreases linearly to zero, from A to I, and the resultant load[7] is at AI/3 from A. For the compressed fibers, Mariotte considered the same triangular distribution from his model of chords, applied to compressive loads, and thus the resultant compressive load is at DI/3 from D. He then considers that "*these extensions and these compressions will share the force of the weight* L*: adding one third of the thickness* IA *to the third of the thickness* ID*, the whole will be equal*

[6] Benvenuto[4], p.266.
[7] Points G and H do not refer to AI/3 and DI/3 but to another development by Mariotte[38].

to one third of the whole thickness AD*; from which the same thing will follow as if all the parts extended.*[8]" Since Mariotte did not express this vague reasoning in a literal or mathematical expression for the breaking load, it is at least improper to attribute to him the incorrect triangular distribution of loads on face AD, as is usually done. Despite the imprecision of Mariotte's proposition, we dare to interpret it as follows: the resistance of the extended fibers corresponds to half of the total resistance of the beam. In this context, isolating the upper part AI, we must consider a compressive longitudinal load concentrated on I in order to balance the tensile loads on face AI. Thereby, let the beam be a prism, just like the case of Galileo's Problem, h its thickness, $\tau_u.h.AD$ is Galileo's absolute resistance to fracture, L_u the breaking load. The balance of *momenta* relative to fulcrum I is

$$\frac{L_u}{2}.DC = \frac{\tau_u.h.AD}{2}\frac{2.AI}{3} \quad \text{or} \quad L_u = \frac{\tau_u.h.AD^2}{3.DC},$$

which results in a breaking load two times greater than the ultimate load for rectangular cross sectional cantilever steel beams subjected to pure bending.

6.3 Hooke

Regarding his load-extension law, described previously, it is improbable that Mariotte knew the *Lectures* de Potentia Restitutiva *Or of Spring*, written by the member of the Royal Society of London Robert Hooke (1635–1703) and published in 1678, eight years before Mariotte's treatise on fluids. In this book, Hooke's approach on material deformation is broader than his French contemporary's: a linear load-extension law is presented as part of a general constitutive property of "springing bodies", as classified by him to express the behavior of deformed bodies that recover their initial shape after load removal, a feature currently known as elasticity. But before exploring this work, let's learn a little about its author's life. Hooke was born at the village of Freshwater, a peninsula on the west of the English Island of Wight, on July 18 and baptized eight days later by his own father, John Hooke, minister of that remote Anglican parish. Because of a weak constitution and recurrent illnesses until the age of seven, the family doubted Hooke would survive beyond childhood. After this period, but still suffering from frequent headaches, the boy did not show interest on the religious studies oriented by his father and preferred to spend most of hist time building mechanical toys. After his father's death in 1648, when he inherited a small sum of money, there are no records of Hooke's life until the age of twenty, when he went to live at the house of Richard Busby, the headmaster of the Westminster school. There, he became proficient on Latin and Greek, as well as Hebrew and other oriental languages. Under Busby orientation, he mastered Euclid's Elements by himself, "*and thence proceeded orderly from that sure Basis to the other parts of the Mathematicks, and after to the*

[8] Benvenuto[4], p.268.

application thereof to Mechanicks, his first and last Mistress.[9]" In 1653, he became a student of the Christ-Church, a constituent college of the University of Oxford, and received the degree of Master of Arts in 1662 or 1663. During this period at the college, he first worked as assistant to the medicine professor Dr. Thomas Willis and then to the eminent natural philosopher Dr. Robert Boyle from 1655 to 1662. In order to drive a pneumatic engine designed by Boyle, which enabled the publication of the famous Boyle's Law on gases (the same as Mariotte's) in 1662, Hooke contrived and perfected an air pump. In this same period, he attempted designs of structures attached to the human body in order to enable flying, but concluded that it was impossible because human muscles were not strong enough for the task. He took active part on the meetings with the group of scholars that would soon found the Royal Society of London in 1660. Already member of the Royal Society and famous for his incomparable skills as a mechanical experimenter, Hooke was nominated the curator of experiments of the Society in November 1661 and could take up his post with the blessings of Boyle, who released him. On the beginning, Hooke was in charge of providing every weekly meeting of the Royal Society with some new experiment or practical presentation on Mechanics, his expertise. However, since specialization at that time was not a virtue, Hooke published in 1665 the first relevant book on biological microscopic observations called *Micrographia*, to which he built his own microscope and where the term "cell" was coined to refer to small biological structures. The following year, after the great fire of London, Hooke was appointed one of the surveyors to rebuild the city and was heavily demanded as an architect, which assured him a substantial financial income for the following ten years. During this period, the most productive of his career, Hooke kept working at the Society, giving lectures and publishing papers: the aforementioned *Lectures* de Potentia Restitutiva *Or of Spring* is one of them.

Figure 6.7 – Presumed portrait of Robert Hooke by Mary Beale.

In a very pragmatic style, Hooke does not waste time with preambles and at the very beginning of these lectures on springs, he starts describing the famous load-extension law that would bare his name until today: "*About two years since I printed*

[9]HOOKE[33], p.iii.

this Theory in an Anagram at the end of my Book of the Descriptions of Helioscopes, viz, ceiiinosssttuu, id est, Ut tensio sic vis*; That is, the Power [force] of any Spring is in the same proportion with the [ex]Tension thereof: that is, if one power [force] stretch or bend it one space, two will bend it two, and three will bend it three, and so forward... And this is the Rule or Law of Nature, upon which all manner of Restituent or Springing motion does proceed... It is very evident that the Rule or Law of Nature in every springing body is, that the force or power thereof to restore itself to its natural position is always proportionate to the Distance or space it is removed therefrom...*[10]" In mathematical words, for each pair of applied load F_i and extension x_i, value $k = F_i/x_i$ is a constant that is constitutive of the spring. Thereby, **Hooke's Law** can be described as the following: in other to attain an arbitrary extension x of a spring with constant k, the applied force needs to be

$$F = kx, \tag{6.1}$$

and thus the restitutive force of the spring is obviously $-kx$. It is common to assume that when the spring is stretched, x is positive; when compressed, it is negative. In the cited paragraph, Hooke also states that the load-extension relation of every springing body (elastic beam) is always linear, but this is incorrect since not all elastic bodies behave linearly. Therefore, in the context of elasticity, a body is called **Hookean** when it observes Hooke's Law, that is, when it deforms linearly. Figure 6.8 illustrates the two springs built by Hooke to obtain the results for his Law.

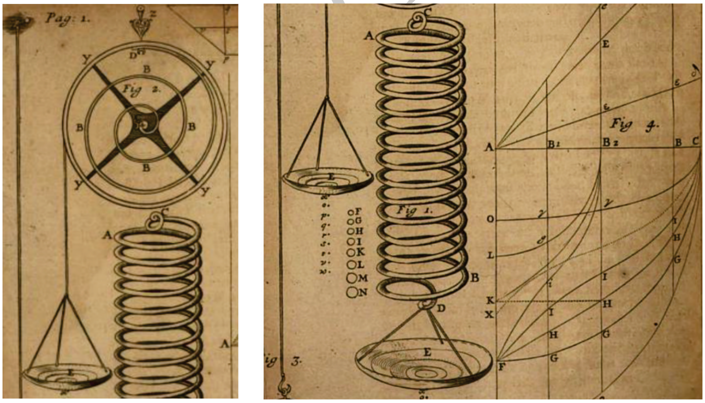

Figure 6.8 – Springs built by Hooke[32] for his Law.

6.4 The Bernoullis

Concerning Galileo's Problem, Hooke did not apply his theory of springing bodies to solve it, but only to describe a "compound way of springing[11]", resulting from the

[10] Hooke[32], pp.1,2,4.
[11] Hooke[32], p. 15.

curvature of the deformed cantilever, where its inner elastic fibers are compressed and the external are stretched. Speaking of curvature, the mathematical description of the deflection or curvature of bending beams was first addressed by the swiss mathematician Jakob Bernoulli (1654–1705) on an article[12] in 1694, a study he called the problem of the *curvatura laminae elasticae* (elastic band curvature), or simply ***elastica***. Born in the city of Basel, Jakob was the eldest son of pharmacist Nikolaus Bernoulli and Margaretha Schönauer, daughter of a banker. Nikolaus's father, also Jakob Bernoulli, was a merchant and emigrated from Amsterdam to Basel, where he became a swiss citizen through marriage. Jakob received his master of arts in philosophy in 1671 and, compelled by his father, graduated in theology in 1676, both at the University of Basel; but, to his parents dismay, during these university years, Jacob became strongly interested in Mathematics and astronomy. In the same year of 1676, working as a tutor of nobles, he traveled to Geneva and then to France, where he lived for two years and could study the works of Descartes and followers. From 1681 to 1682, Jakob resumed his travels, but this time to make contact with scholars and mathematicians in Holland and England, where he met Robert Boyle and Robert Hooke. Back to Basel, during 1683, he started giving lectures on Mechanics and also publishing articles on Geometry and Algebra in the prominent scientific periodicals *Journal des Sçavans* and the *Acta Eruditorum*. In 1684, Jakob married Judith Stupanus, daughter of a pharmacist, with whom he had two children. He kept working and publishing on Mathematics, as well as exchanging an intense scientific epistolary correspondence with acquaintances he made during his travels, including Leibniz, Huygens and others. In this period, his brother Johann Bernoulli (1667–1748) was studying medicine at the university, also compelled by Nikolaus, but Johann secretly started studying Mathematics by himself, under the orientation of Jakob. After the appointment of Jakob as professor of Mathematics at the University of Basel in 1687, both brothers started studying the differential calculus of Leibniz and adopted his differential notation for derivatives, that is, they became "Leibnizians", as opposed to british "Newtonians", who worked with Newton's fluxions and fluents, concepts that embodied the infinitely small number o. In a paper published in the 1690 edition of *Acta Eruditorum*, concerning the solutions of the problem of the catenary curve proposed by Huygens and Leibniz, Jakob coined the term "integral" in the sense we currently use. Among his works, the main contributions deal with problems and propositions on Calculus, Probability and Series. According to Maugin[42], Jakob "*may have been less creative than John [Johann], but nonetheless played an essential role in the dissemination of integral calculus ..., in the establishment of the theory of probabilities, and in solving critical problems in mechanics (the isochronous curve, the* elastica*)*."

From 1690 to 1695, Jakob decided to study the so called flexible lines; mainly the catenary, the isochronous curve and the *elastica*. Our interest here is his contribution

[12] See Truesdell[64], p. 89.

Figure 6.9 – Jakob Bernoulli.

on the deflection and curvature of the *elastica*, which is an elastic band deformed by its own weight or by a weight or couple applied to one or both of its endings. Jakob had referred to the problem of the *elastica* with an ending fixed and the other deflected by a load, proposed by Leibniz in private letters, in an article published 1691, announcing that he would present a solution the following year[13]; but it took three years until *Curvatura Laminae Elasticae, &c.* was finally submitted to *Acta Eruditorum*. Jakob starts this article by saying that "*After a silence of three years I keep my word; but in such a way as right richly to compensate for that delay, which else the reader might have borne with annoyance, since I exhibit the curvature of springs not in one way only (as I had promised in the beginning) but generally for any hypothesis on the elongations; which, unless I err, I am the first to achieve, after the problem was attempted in vain by many.*[14]" In order to understand what he did, we reproduce on figure 6.10 one of the diagrams presented in the article. Fixed to the support SRYX, he draws a beam SVAR with uniform rectangular section, the inner fiber of which is the *elastica* to be studied. The beam sustains a weight Z, whose chord is perpendicular to the tangent αu on A. The curve AC represents the load-extension relation at an arbitrary point y, where the load, on the abscissa, increases from A to B and the extension increases from B to C. And that's it: with the help of the shape and proportions of the geometry presented, the reader is supposed to understand and accept that the curvature $\kappa := 1/Qn$ at point Q is directly proportional to distance QP. The year following the publication of this article, Jakob submitted another to the same periodical, making things clearer: "*I consider a lever with fulcrum* Q*, in which the thickness* Qy *of the band [beam] forms the shorter arm, the part of the curve* AQ *the longer. Since* Qy *and the attached weight* Z *remain the same, it is clear that the force [load] stretching the filament [fiber] at* y *... is proportional to the segment* QP.[15]" In other words, he assumes that the moment QP.Z is balanced by Qy.F, where F represents the load that stretches y an extension e, measured relative to initial length of the *elastica* AR. Since Qy and Z

[13]See Truesdell[64], p. 88.
[14]Jakob Bernoulli[5], pp. 262-263, translated from the Latin by Truesdell[64].
[15]Jakob Bernoulli[6], p. 538, translated from the Latin by Truesdell[64].

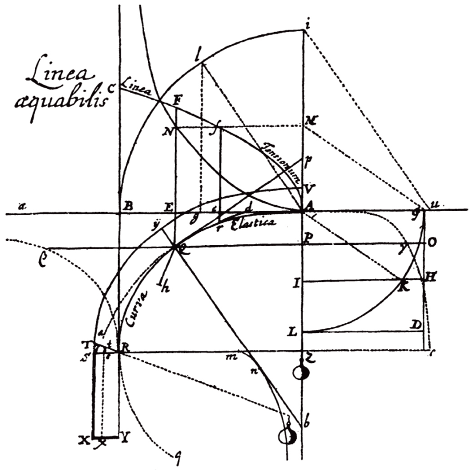

Figure 6.10 – Study on the *curva elastica* by Jakob Bernoulli[5].

are constants, F is directly proportional to QP. Moreover, since F is directly proportional to e, according to his load-extension relation, then e is directly proportional to QP; but since e is inversely proportional to Qn, then k is directly proportional to QP. Mathematically, equality $Qy.F = QP.Z$ implies that $QP \propto F \propto e \propto 1/Qn$. Then, Jakob proposed that

$$\kappa = \frac{e}{Qy}$$

and, knowing the formula for the curvature of a plane curve, called by him the "golden theorem", he arrived at an equation of the deflection[16]. However, for some strange reason, Jakob incorrectly considered only the moment $Qy.F$ to balance $QP.Z$ and not the moment correspondent to the whole distribution of forces in section Qy, a distribution he studied because lines ST and st are drawn to represent the elongations of two different points on the cross section RS.

With fundamental studies on differential equations, Calculus of Variations and other problems of Mathematics, the "more creative" brother Johann is considered to be the most brilliant mathematician of the Bernoulli dynasty on this field[17]. From his early studies on medicine, Johann defended in 1694 one the first known doctoral thesis on biomechanics, where he described the motion of muscles. It would be incorrect to say that he did not specifically contribute to the subject of deformable bodies: besides introducing the concept of what today is known as the principle of virtual work in a letter to the French mathematician Pierre Varignon (1654–1722) in February 26th, 1715, Johann is the father and master of Daniel Bernoulli (1700–1782), who, among other notable works, substantially developed and generalized the ideas of his

[16] See Truesdell[64], p. 93.
[17] See Maugin[42], p. 8.

uncle Jakob on the *elastica*. The second son of Johann Bernoulli and Dorothea Falkner, Daniel was born in the dutch city of Groningen, when his father held a post of Mathematics professor at the university. At the age of thirteen, with the family already living at Basel, Daniel was sent to the university to study philosophy and logic. During this time, he also studied Mathematics privatley, oriented by his father and his elder brother Nikolaus. However, when Daniel finished his studies in 1716, Johann determined he would be a merchant and sent him to learn the profession. Daniel showed enough evidences that he had no talent for commercial business and finally expressed his wishes to study Mathematics, but Johann strongly disagreed, arguing that Mathematics did not bring financial comfort. A compromise solution was found and Daniel ended up enrolling at the University of Basel to study medicine, while keeping his private studies in Mathematics and Physics. As a result, and just like his father Johann, Daniel defended a thesis on biomechanics, more specifically on the mechanics of breathing, when he completed the medicine course in 1721. In order to follow an academic career at Basel, Daniel applied firstly for a chair on anatomy and botany and secondly for a vacant chair on logic, but both chairs were defined by lottery, which he lost out. Disappointed by this bad luck, Daniel went to Venice in order to practice medicine, but did not abandon Mathematics, and published a book entitled *Exercitationes Quaedam Mathematicae* in 1724, with the help of German mathematician Christian Goldbach (1690–1764). The work presents four "exercises": a probabilistic description of a game; a discussion on the flow of water in a hole, which was incorrect; a study on differential equations; and a study on geometry of circular figures. His medical studies on blood flow greatly influenced his second exercise, which would stimulate him to keep studying the motion of fluids. Although *Exercitationes* did not make any new or relevant contribution, it enabled the name of Daniel Bernoulli to be known in scientific circles. When his father Johann dropped an offer for a chair at the Imperial Academy of Sciences and Arts in St. Petersburg on behalf of him, the name of Daniel was immediately accepted. He assumed the chair of Mathematics at the Academy in 1725, together with his elder brother Nikolaus, a condition previously imposed by Daniel to accept the arrangement of his father. Eight months after their arrival at St. Petersburg Nikolaus died, a personal tragedy that affected Daniel in such a way that made him write to his father communicating his intentions to return to Basel. In order to alleviate Daniel sadness and also to help the young Leonhard Euler (1707–1783), Johann arranged a way to make his outstanding pupil be accepted to work with his son at the Imperial Academy. Euler arrived at St. Petersburg in 1727 and until Daniel return to Basel in 1733, they collaborated intensely and proficuously. In this period, Daniel and Euler produced fundamental contributions to the vibration theory, probability and political economy, as well as the mechanics of elastic and flexible bodies. Still in St. Petersburg, Daniel developed and finished his most notable work, which he would publish only in 1738, called *Hydrodynamica*. In this book, the conservation of energy, conceived by his father, where an isolated body in motion balances kinetic and potential energies, is the base for a particular version applied to fluids, the famous Bernoulli's principle, which states, in current terminology, that an isolated fluid in incompressible flow (reversible and adiabatic process) balances kinetic energy and static pressure. Moreover, according to Mikhailov[45], in this work Daniel also "*develops the first model of the kinetic theory of gases, approaches the principle of conservation of energy, establishes a foundation for the analysis of efficiency of machines, and he develops a theory of hydroreactive (water-jet) ship propulsion, including a solution of the first problem of motion of a variable-mass system.*" From 1731 on, Daniel began applying for posts at the University of Basel, but he only got the chair of botany, which was a sufficient reason for him to leave St. Petersburg in 1733. Already at home, Daniel submitted a study on astronomy to the Grand Prize of the Paris Academy in 1734; a prize to which Johann had also submitted a similar work. The fact that father and son were declared joint winners led to strong quarrels between them and the relationship, which had always been full of tension and envy, suffered a definitive break up. After this event, Daniel never again developed himself any relevant mathematical work, although he kept an active correspondence and collaboration with Euler, that stayed in St. Petersburg.

Figure 6.11 – Daniel Bernoulli.

6.5 Euler

In the Bernoulli-Euler epistolary collaboration on the study flexible lines, including statical and vibrational behaviors, Daniel acted mainly as an instigator, proposing problems and suggesting directions to follow, while Euler conceived brilliant solutions, definitions, methods and generalizations; in other words, Daniel oriented and Euler made it happen, beautifully. But before we describe their great contribution to the mathematical description of elastic bands, which is our concern here, we request the reader to indulgently follow us now through some words about Euler the man. First-born child of the calvinist pastor Paulus Euler and Margaretha Brucker, daughter of a hospital vicar, the swiss mathematician Leonhard Euler was born in Basel, from where the family moved when he was around one year old due to the appointment of his father to be the minister of a parish in the nearby town of Riehen. Paulus had enrolled the University of Basel back in 1685 and chose protestant theology to be his field of study. During the initial semesters, in October of 1688, to be precise, he attended a course on geometry and algebra given by the renowned mathematician Jakob Bernoulli, whose family Paulus became acquainted with. This course on Mathematics, whose final work was a thesis on Algebra he had to defend, probably inspired Paulus to tutor the young Euler on the first elementary algebraic concepts, and then *Algebra*, written by the German mathematician Christoff Rudolff (1499–1545), was adopted as a textbook. Shortly after, the boy, who was eight years old, was sent to live in Basel with his maternal grandmother Maria Magdalena in order to begin his Latin studies. Since the education at the gymnasium was poor, Euler also had private complementary lessons with the young theologian Johannes Burckhardt (1691–1743), who was a great enthusiastic of mathematical matters and, most probably, heavily influenced his pupil: in a letter, Daniel Bernoulli referred to Burckhardt as being the math teacher of Euler. Following the tradition of that period, Euler was supposed to trace the same steps of his father and then, at the age of thirteen, he enrolled the University of Basel to course the philosophical faculty, a kind of technical high school, as a prerequisite to study theology which would allow him to become a minister. The early semesters at the university included a variety of compulsory disciplines, including Mathematics, and Euler ended up attending the classes of the eminent professor Johann Bernoulli on arithmetic and geometry. His exceptional mathematical skills and the help of fellow classmate Johann II, Johann's youngest son, enabled Euler to attend advanced lectures, including the famous Saturday afternoons *privatissima*, which constituted a kind of mentorship for outstanding students on higher mathematics and Physics given by the always severe Johann Bernoulli. In 1723, Euler acquired the *magister* degree, which meant he had successfully finished the philosophical faculty and, to fulfill the wishes of his father, he immediately enrolled the theological faculty; but never abandoned his math studies. It is believed that Euler's performance and achievements amazed Johann Bernoulli in such a way that both master, almost sixty years old, and his eighteen-year-old pupil traveled to Riehen in order to convince Paulus Euler, Johann's old college fellow, that it was God's will that his son be a mathematician rather than a calvinist minister of some rural parish. According to biographer Calinger[10], "*Euler, like Kepler, Newton, and the Bernoullis, assumed that intimate connections existed between religion and the mathematical sciences. From his college days on, both reason and religious faith inspired Euler's research.*" Paulus apparently accepted the arguments of professor Johann and from then on Euler could spend even more time in his mathematical studies and production. His first work was published in the 1726 edition of *Acta Eruditorum*, entitled *Construction Of Isochronal Curves In Any Forms Of Resistant Media*, in which Euler tries to obtain a friction function on a generic curve so as this curve behaves isochronally; but he did not succeed. In this same year, he signed up for a contest offered by the Paris Academy on the best way to set up a mast on a ship, when he shared the second place with other competitor. By this time, as we already know, Daniel Bernoulli was working at St. Petersburg, when his father Johann managed to send a letter recommending Euler to the president of the Imperial Academy Laurentius Blumentrost (1692–1755). In September of 1726, Leonhard received a letter from Daniel informing that there was a vacant post of adjunct professor of physiology at the Academy, available to him until June 1727, since there were no chairs available in Mathematics; but Euler decided to accept it only if his efforts to get a chair at the University of Basel were not successful. In the spring 1727, the twenty-year-old Euler disputed a vacant Physics chair at Basel with the thesis *Dissertation On The Theory Of Sound*, but the selection committee did not accept him for the final dispute due to his young age. Shortly after this bad result, which many scholars consider the most fortunate for the development of Physics and Mathematics, Euler enrolled a medicine course on physiology at Basel in order to gain proficiency for his professorship in Russia, but only three days later, on April 5, 1727, he left Basel on the way to St. Petersburg. The causes of this sudden departure are unknown, maybe some quarrel between Euler an his family: after this departure, Leonhard would never return to Basel again. He arrived at St. Petersburg on May 24, the same day the unexpected death of empress Catherine I was reported to the public. In Euler's own words, "*my salary was 300 Rbl. along with free lodging, firewood, and light, and since my inclinations were directed solely and exclusively toward mathematical studies, I was appointed as an adjunct of higher mathematics, and the suggestion to occupy myself with medicine was dropped altogether.*[18]" Under Blumentrost, the Imperial Academy had expended substantial efforts and money to recruit foreign renowned math scholars: Georg Bilfinger (1693–1750) and Christian Goldbach (1690–1764) from German, Jakob Hermann (1678–133) from Switzerland, Joseph-Nicholas Delisle (1688–1768) from France, among others. Because of the highly favorable and generous work conditions the Academy offered to its professors, with laboratories, very few students and an ample library, the place was considered to be, in scientific circles, the paradise of scholars. There, Euler began his longtime collaboration with Daniel Bernoulli: some biographers suspect that Daniel's most notable work *Hydrodynamica* has a direct not acknowledged contribution of Euler because he was also studying the subject at the time. In 1731, with the departure of professor Bilfinger, his chair on Physics was given to Euler and two years later, soon after Daniel's return to Basel, Leonhard finally got the post of professor on Mathematics. Three years later, he married to Katharina Gsell (1707–1773), born in Amsterdam and daughter of the Russian court painter Georg Gsell and his first wife. With Katharina, Euler had thirteen children, but only five survived childhood. In 1736, he published his first masterpiece, a two-volume book entitled *Mechanica Or The Exposition Of The Analysis Of The Science of Motion*, where the

[18] Fellmann[22], p. 6.

mathematical tools of Calculus were systematically applied for the first time to describe the theory and study the problems of Mechanics. Moreover, Euler produced proficuously on Mathematics and completed in 1738 an outstanding work on hydrodynamics that would later be published as the two volume book entitled *Scientia Navalis*, where he, among other achievements, "*defines via the (directionally independent) fluid pressure an ideal fluid, which later served Cauchy as a model for the definition of the stress tensor.*[19]" Since the death of empress Catherina I, political instability in Russia dangerously increased and the old favorable conditions to foreign scholars at the Imperial Academy deteriorated rapidly. Turmoils among the populace were not uncommon and the result was fire in houses and buildings. When, on February 15, 1741, Euler received an offer by the king Frederick II of Prussia for a post at the Berlin Academy, his fearful wife Katharina started to press him to accept the post and leave St. Petersburg as soon as possible. The Eulers – husband, wife and two sons – arrived at Berlin on July 25, after a three week journey. In his Berlin years, Euler produced around 380 articles and two other masterpieces: *Introduction to the Analysis of the Infinite*, a two-volume book published in 1748, where the branch of functional analysis is created, and *Foundations of differential calculus*, published in 1755, which laid the bases for the modern approach of Calculus. From 1760 to 1762, Euler exchanged letters with German princess Friederike Charlotte on various subjects like philosophy, mechanics, optics, astronomy and theology, which were later compiled in what would become his most popular work *Letters to a German Princess*, published in 1768. Euler had never lost contact with his influential Russian friends and when king Frederick appointed D'Alembert to the presidency of the Berlin Academy instead of him, the recurrent offerings to return to the Imperial Academy became attractive. Although the political situation in Russia had greatly improved, Euler still imposed extravagant conditions to return to St. Petersburg, which were promptly accepted by empress Catherina II. With fifty nine years old, on July 28, 1766, Leonhard Euler and his "entourage" of eighteen people arrived at St. Petersburg, where he would live until the end of his life. During this period, his eyesight problems worsened: with the right vision completely lost due to an unknown infection he had in the first Russian period, now a cataract operation impaired much of his left vision, a limitation that required him to dictate all his written work to scribblers. But this condition did not prevent him from being productive: in his second Russian years, Euler made available more than 300 works, including his most popular math work *Elements of Algebra*, a two-volume textbook published in 1770. In 1773, he lost his 66 year-old wife Katharina and three years later married to her younger half-sister Salome Abigail Gsell (1723-1794). Euler would ask her, around 5 pm on September 18, 1783, from his seat, when playing happily with a grandson and talking to a close friend, if he had already taken two cups of tea. Salome said he had taken just one and served him another. Taking the tea with one hand, a pipe he was smoking slipped

[19] Fellmann[22], p. 45.

from the other and he bent forward to try to pick it up, but couldn't reach it. After recomposing himself, he felt dizzy, grabbed his head with the two hands and fainted, victim of a stroke. Around 11 pm on that same day, Leonhard Euler expired.

Figure 6.12 – Portrait of Leonhard Euler by Jakob Handmann.

Euler's main contribution to the study of deformable bodies is presented on Appendix I, entitled *De Curvis Elasticis*[20], of his book *Methodus Inveniendi Lineas Curvas Maximi Minimive Proprietate Gaudentes*, published in 1744. This appendix is the result of a suggestion made by Daniel Bernoulli to Euler in the letter of October 20, 1742, in which Bernoulli proposes the problem of finding the curvature of the *elastica*, submitted to couples at both fixed endings, through the minimization of a potential. Daniel details his proposition the following way: *"for a naturally straight elastic band I express the potential live force [elastic potential] of the curved band by* $\int ds/r^2$*... Since no one has perfected the isoperimetric method [calculus of variations] as much as you, you will easily solve this problem of rendering* $\int ds/r^2$ *a minimum."* Easily or not, Euler did solve the problem and, with the method he created, curvatures for other forms and loading conditions of the *elastica* could be found as presented in figure 6.13, extracted from *De Curvis Elasticis*, which is considered to be the first mathematical treatise on elasticity. Concerning the problem proposed by Bernoulli for bands of uniform elasticity and cross section, represented on drawing 1 of tabula III, the elastic potential $\int ds/r^2$ is described by $\int Z dx$ on rectangular coordinates, where $Z = q^2/\sqrt{(1+p^2)^5}$, $p = dy/dx$ and $q = dp/dx$. Through some development he did not present, Euler stated that finding the minimum of $\int Z dx$ is the same as solving the equation

$$\frac{d^2Q}{dx^2} - \frac{dP}{dx} + \alpha \frac{d}{dx}\left(\frac{p}{\sqrt{1+p^2}}\right) = 0,$$

where $Q = dZ/dq$, $P = dZ/dp$ and α is a constant. Integrating this differential equation, we can write that

$$\frac{dQ}{dx} - P + \frac{\alpha p}{\sqrt{1+p^2}} = \beta \quad \text{or} \quad \frac{dQ}{dp} q - \frac{dZ}{dp} + \frac{\alpha p}{\sqrt{1+p^2}} = \beta$$

[20] See Oldfather et al[49].

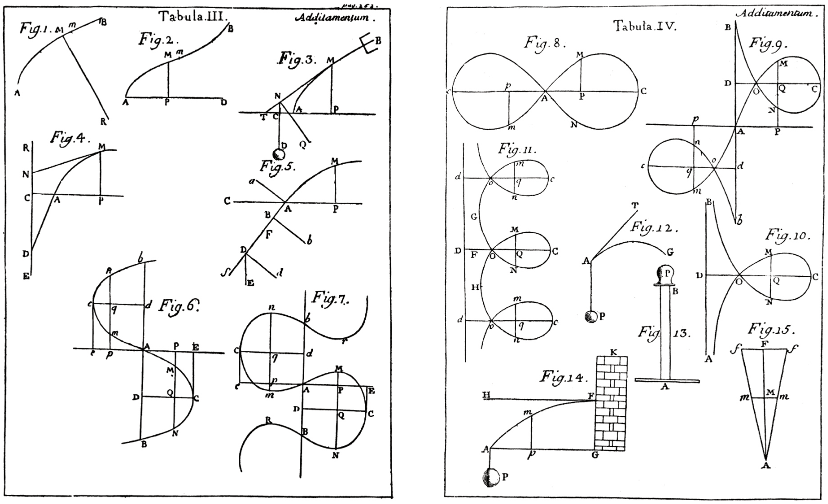

Figure 6.13 – Different configurations of the elastic band on Euler's *De Curvis Elasticis*.

because $dQ/dx = dQ/dp.dp/dx$. Integrating again and isolating q, we arrive at

$$q = \frac{1}{Q}(\beta p + \gamma + Z - \alpha\sqrt{1+p^2}).$$

Since $q = dp/dx = dp/dy.dy/dx = p \cdot dp/dy$, then

$$\frac{dp}{dx} = (1+p^2)^{\frac{5}{4}}\sqrt{\alpha\sqrt{1+p^2}+\beta p+\gamma} \quad \text{and} \quad \frac{dp}{dy} = \frac{1}{p}\frac{dp}{dx}.$$

At this point, since $(du/dv)^{-1} = dv/du$, Euler made use of the so called mathematical sagacity to construct the equality

$$\beta\frac{dx}{dp} - \gamma\frac{dy}{dp} = \frac{(\beta-\gamma p)}{(1+p^2)^{\frac{5}{4}}\sqrt{\alpha\sqrt{1+p^2}+\beta p+\gamma}}$$

whose integration leads to

$$\beta x - \gamma y + \delta = \frac{2\sqrt{\alpha\sqrt{1+p^2}+\beta p+\gamma}}{\sqrt[4]{1+p^2}}. \tag{6.2}$$

Then he proceeded a change of variables by rotating and translating the orthogonal axes in such a way that

$$X = \frac{\beta x - \gamma y + \delta}{\sqrt{\beta^2+\gamma^2}} \quad \text{and} \quad Y = \frac{\gamma x + \beta y + \delta}{\sqrt{\beta^2+\gamma^2}}.$$

From these equalities and considering $\tilde{p} = dY/dX$, it is possible to write that

$$p = \frac{d}{dx}\left(\frac{Y\sqrt{\beta^2+\gamma^2} - \gamma x + \delta}{\beta}\right)$$

$$
\begin{aligned}
&= \frac{1}{\beta}\left(\frac{dY}{dX}\frac{dX}{dx}\sqrt{\beta^2+\gamma^2}-\gamma\right) \\
&= \frac{1}{\beta}[\tilde{p}(\beta-\gamma p)-\gamma] \\
&= \frac{\beta\tilde{p}-\gamma}{\beta+\tilde{p}\gamma}
\end{aligned}
$$

and also that $1+p^2=(\beta^2+\gamma^2)(1+\tilde{p}^2)/(\beta+\gamma\tilde{p})^2$. Substituting all these values on (6.2), Euler obtained that

$$
X\sqrt{\beta^2+\gamma^2} = \frac{2\sqrt{\alpha\sqrt{1+\tilde{p}^2}+\tilde{p}\sqrt{\beta^2+\gamma^2}}}{\sqrt[4]{1+\tilde{p}^2}}
$$

and considered $\beta_1=\sqrt{\beta^2+\gamma^2}$, which led to $X\beta_1\sqrt[4]{1+\tilde{p}^2}=2\sqrt{\alpha\sqrt{1+\tilde{p}^2}+\tilde{p}\beta_1}$. Generically, this last result can be rewritten as

$$
x\beta\sqrt[4]{1+p^2}=2\sqrt{\alpha\sqrt{1+p^2}+p\beta}.
$$

From this equation, through algebraic manipulations, he arrived at

$$
p=\frac{n^2x^2-ma^2}{\sqrt{n^2a^4-(n^2x^2-ma^2)^2}}
$$

where $m=\alpha a^2/4$ and $n=\beta a^2/4$. Making a new change of variables where $n=k$, $x=\tilde{x}+c/2k$ and $m=a^{-2}(c^2/4-bk)$ and developing the previous equation, he finally obtained the deflection equation

$$
\frac{dy}{dx}=\frac{(b+cx+kx^2)}{\sqrt{a^4-(b+cx+kx^2)^2}} \tag{6.3}
$$

of the *elastica* with each of its fixed endings submitted to a given couple; confirming the result already obtained by Jakob Bernoulli in 1695 through an approach which Euler called *a priori* method (causes) instead of Daniel's method of maxima and minima (effects). In order to prove the generality of his result, Euler solves through an *a priori* method the configuration of the *elastica* presented on drawing 3 of tabula III, that is, with an ending fixed and the other attached to a rigid staff subjected to a load P. In this context, from a paper published by Daniel Bernoulli in 1728, in which he stated that the moment $Px=EI\kappa$, where x is the distance of the section from the applied load, E is the Young's modulus, I a geometrical quantity (known today as the moment of inertia) and curvature

$$
\kappa=\frac{d^2y/dx^2}{[1+(dy/dx)^2]^{3/2}},
$$

Euler obtained the equation

$$
\frac{dy}{dx}=\frac{-P(b+cx+x^2/2)}{\sqrt{E^2I^2-P^2(b+cx+x^2/2)^2}},
$$

which is compatible to (6.3) and valid for large elastic deflections. When small deflections are considered, the curvature $\kappa \approx d^2y/dx^2$ and then the deflection equation results $d^2y/dx^2 = Px/EI$. In tabulae III and IV, Euler analyzes other forms of the *elastica* and also the strength of a column to buckling: considering drawing 13, he states that a column with uniform cross section will not bend if $P \leqslant EI(\pi/h)^2$, where h is its height.

6.6 Cauchy

In 1817, the French civil engineer Augustin-Louis Cauchy (1789–1857) was in charge of the disciplines of Mechanics and Analysis at the École Polytechnique in Paris and, from Euler's concept of hydrostatic pressure p in a perfect fluid, he taught his students that the property of p being the same in all directions was a consequence of its perpendicular action on the surfaces involved[21]. At Faculté des Sciences, also in Paris, where he lectured on Mechanics and Mathematics by the year of 1821, he began extending his ideas on hydrostatics to any deformable body. Already established as a renowned mathematician, Cauchy was frequently considered by some fellows in French scientific circles to be the special recipient of their work: in a letter dated July 24, 1821, he acknowledged the mathematician Marie-Sophie Germain (1776–1831) for having sent him *Research on the Theory of Elastic Plates*, published that same year, a work in which she solved a mathematical problem concerning vibration of elastic surfaces and that had already made her the winner of a competition organized by the Paris Academy of Sciences in 1816. Inspired by this work and the ideas of Augustin-Jean Fresnel (1788–1827) and Claude-Louis Navier (1785–1836), which were also working on mathematical models for the mechanical behavior of deformable bodies, Cauchy managed to anticipate these two renowned scholars and published in 1823 an abstract entitled *Researches On The Equilibrium And Internal Motion Of Solid Bodies Or Fluids, Elastic Or Non-elastic* on the January edition of the *Bulletin of the Société Philomatique*, which was submitted on September 30 of the previous year. Read out loud by Cauchy himself on a meeting of the Société Philomatique in February 22, 1823, the abstract delineated the concept of stress, which he adapted from Euler's hydrostatic pressure, and its implications. The great relevance of this abstract as a document attesting the birth of the modern approach to Continuum Mechanics, which is the last period of our brief history, is the reason why we allow ourselves to reproduce it completely, as follows, in our English translation of the original text by Cauchy[11].

[21] See Belhoste[3], p. 92.

Researches On The Equilibrium And Internal Motion Of Solid Bodies Or Fluids, Elastic Or Non-elastic; by Mr. Aug. CAUCHY. *Translated by Mr. R.D.* ALGARTE.

These researches were undertaken on the occasion of a Memoir published by Mr. Navier, on August 14, 1820. The author, in order to establish the equation of equilibrium of an elastic plane, considered two types of forces produced, one by expansion or contraction, the other by flexion of the same plane. Moreover, he assumed, in his calculations, both one and the other perpendicular to the lines or faces against which they are exerted. It seemed to me that these two types of forces could be reduced to one, which must constantly be called tension or pressure, and which was of the same nature of the hydrostatic pressure exerted by a resting fluid against the surface of a solid body. But this new pressure did not always remain perpendicular to the faces submitted to it, nor the same in all directions at a given point. By developing this idea, I soon arrived at the following conclusions.

If in an elastic or non-elastic solid body, a small element of its volume, defined by arbitrary faces, is made rigid and invariable, this small element will experience on its different faces, and at any point of each of them, a certain pressure or tension. This pressure or tension will be similar to the pressure that a fluid exerts against an element of the surface of a solid body, with the only difference that the pressure exerted by a fluid at rest against the surface of a solid body is directed perpendicular to this surface, from outside to inside, and independent, at each point, of the inclination of the surface with respect to the coordinate planes, while the pressure or tension exerted at a given point of a solid body through which a very small surface element passes can be directed perpendicularly or obliquely to this surface, sometimes from outside to inside, if there is contraction, sometimes from inside to outside, if there is expansion, and may depend on the inclination of the surface with respect to the planes in question. In addition, the pressure or tension exerted against each plane can be deduced very easily, both in magnitude and in direction, from the pressures or tensions exerted against three given rectangular planes. I was at this point when Mr. Fresnel, coming to talk to me about the work he was developing on light, and of which he had presented only a part at the institute, told me that, for his part, he had obtained, by the laws according to which elasticity varies in the many directions that emanate from a single point, a theorem analogous to mine. However, at that time, the theorem in question was far from being sufficient for me to attain the objective I had proposed, namely, to obtain the general equations of equilibrium and internal motion of a body; and only recently have I succeeded in establishing new principles that suitably led me to this result and that I will make known.

From the theorem stated above, it follows that the pressure or tension at each point is equivalent to one divided by the vector radius of an ellipsoid. To the three axes of this ellipsoid, three pressures or tensions are related, which we will call *principal*, and we can prove[1] that each of them is perpendicular to the plane against which it is exerted. Amongst these principal pressures or tensions, there are the maximum pressure or tension and the minimum pressure or tension. The other pressures or tensions are distributed symmetrically around the three axes. Moreover, the pressure or tension normal to each plane, that is to say, the component, perpendicular to a plane, of the pressure or tension exerted against this plane is reciprocally proportional to the square of the vector radius of a second ellipsoid. Sometimes this second ellipsoid is replaced by two hyperboloids, one with a single sheet, the other with two sheets, which have the same center, the same axes, and touch each other to infinity defining a conical surface of second degree, whose edges indicate the directions to which the normal pressure or tension decreases to zero.

That being said, if we consider a solid body of variable shape and subjected to any accelerating forces, in order to establish the equilibrium equations of this solid body, it will suffice to write that there is a balance between the motive forces which stress an infinitely small element in the direction of the coordinate axes and the orthogonal components of the external pressures or tensions which act on the faces of this element. We will thus obtain three equilibrium equations which include, as a special case, those of the equilibrium of fluids. But, in the general case, these equations contain six unknown functions of the coordinates x, y, z. It remains to determine the values of these six unknowns; but the solution of this last problem varies according to the nature of the body and its more or less perfect elasticity. Let us now explain how we manage to solve it for elastic bodies.

When an elastic body is in equilibrium by virtue of any accelerating forces, each molecule must be assumed to be displaced from the position it occupied when the body was in its natural state. By virtue of the displacements of this kind, there are, around each point, different contractions or expansions in different directions. Now, it is clear that each expansion produces a tension, and each contraction a pressure. Moreover, I have demonstrated that the various contractions or expansions around a point, decreased or increased by unity, become equal, except for the sign, to the vector radii of an ellipsoid. I call *principal contractions or expansions* those which take place along the axes of this ellipsoid, around which all the others are symmetrically distributed. That being said, it is clear that in an elastic solid, if the tensions or pressures depend only on the contractions or expansions, the principal tensions or pressures will have the same direction of the principal contractions or expansions. Moreover, it is natural to suppose, at least when the displacements of the molecules are very small, that the principal pressures or tensions are respectively proportional to the principal contractions or

[1]Our remark here agrees with the latest research of Mr. Fresnel. (See *Bulletin*, May 1822)

expansions. By admitting this principle, we immediately arrive at the equations of the equilibrium of an elastic body. In the case of very small displacements, the component, perpendicular to a plane, of the pressure or tension exerted against this plane, always preserves the same relation with the contraction or expansion which takes place in the direction of this component, and the equilibrium formulas are reduced to four partial differential equations, one of which separately determines the contraction or expansion of the volume, while each of the others serves to fix the displacement parallel to one of the coordinate axes.

Once the equilibrium equations of an elastic body are obtained, it is easy to deduce the equations of motion by ordinary methods. The latter are still four in number, and each of them is a linear partial differential equation with an appended variable term. They are integrated by the methods exposed in our previous Memoir. One of these equations contains only the unknown which represents the volume contraction or expansion. In the particular case where the accelerating force becomes constant and preserves the same direction, this equation is reduced to that which describes the propagation of sound in air, with the only difference that the constant which it contains, instead of depending on the height of the supposedly homogeneous atmosphere, depends on the linear expansion or contraction of a body under a given pressure. We must conclude that the speed of sound in an elastic solid is constant, as in the air, but it varies from one body to another depending on the matter of which the body is composed. This constancy is even more remarkable since the displacements of successive molecules considered in fluids and elastic solids follow different laws.

My Memoir ends by obtaining the equations of the internal motion of solid bodies entirely devoid of elasticity. In order to achieve this, it suffices to suppose that in these bodies the pressures or tensions around a moving point no longer depend on total contractions or expansions which correspond to the absolute displacements relative to the initial positions of the molecules, but only, during any period of time, on very small contractions or expansions which correspond to the respective displacements of the different points during a very short period of time. We then find that the contraction of volume is determined by an equation similar to that of heat, which establishes a remarkable analogy between the propagation of calorie and the propagation of vibrations in a body entirely devoid of elasticity.

In another Memoir, I will perform the application of the formulas that I obtained to the theory of elastic plates and laminae.

Since there are very few reliable sources of Cauchy's biography, we shall mostly rely on the definitive account of BELHOSTE[3]. Born in Paris in August 21, 1789, one month after the Storming of the Bastille, which marked the effective breakout of the French Revolution, Cauchy was the eldest of the six children of Louis-François Cauchy, a highly ranked senior official of the French Senate, and Marie-Madeleine Desestre, a middle-class Parisienne whose most of the relatives, including her father and brother, occupied lower posts in the Parisian government. At the time of Cauchy's birth, Louis-François had a post at the police of Paris and was closely related to Lieutenant Géneral Louis Thiroux de Crosne, his benefactor. When Thiroux, returning from England, was arrested and beheaded in April 28, 1794, by the revolutionary authorities, Louis-François hasted to flee from Paris with his wife and two sons, the five-year-old Augustin-Louis and the baby Alexandre-Laurent, to live in the commune of Arcueil, in a country house they had bought years before. Living there, the eldest son not only contracted smallpox but also received the first lessons on elementary education from his father, who had been a brilliant Law student at the University of Paris. After the execution of Robespierre in July 27 and the fall of his followers, the widespread turmoil ceased: arrests and executions decreased substantially; the overall conditions improved, which led to the end of the so called Reign of Terror. Since there were no more reasons to fear political persecution, the Cauchys returned to Paris that same year, and Louis-François assumed an administrative post at the Ministry of Interior in Autuum. A fervent supporter of the ideas, plans and political movements that conduced Napoleon Bonaparte to power as the First Council of France in November 1799, following the coup d'etat of 18 Brumaire, Louis-François was rewarded in January 1800 with the position of Secretary-General at the newly created Senate, a highly prestigious post, inferior only to that of Chancellor, which was occupied at that time by senator Pierre-Simon Laplace (1749-1827), the famous polymath. As the Secretary-General of the Senate, Louis-François had obviously to be in daily contact with many senators, an interaction that eventually ended in friendship, as was the case with Chancellor Laplace and senator Joseph-Louis Lagrange (1736-1813). Already informed by Louis-François about the skills and interests in Mathematics of the young Augustin-Louis, the renowned Italian-French mathematician, senator Lagrange, is said to have given the following piece of advice to his friend Louis-François: "*Do not allow him even to open a Mathematics book nor write a single number before he has completed his studies in literature*". And so it was done: under Lagrange influence and conduction, Augustin-Louis entered the École Centrale du Panthéon in 1802 to course Latin and humanities. However, during this period, he performed brilliantly and won many competitions in ancient languages, graduating two years later with distinction. Despite this success in humanities, Augustin-Louis ended up breaking the family tradition of studying Law in order to satisfy his early interests in Mathematics and then decided to enroll the École Polytechnique the following year of 1805 to become an engineer of the French public service. But this surprising decision did not exasperate Louis-François, who provided all the necessary material support and also took advantage of his great prestige to enable the success of his son's career. Examined by the renowned mathematician Jean-Baptiste Biot (1774-1862), among 293 applicants, Augustin-Louis ranked the second place out of 125 approved, and chose to specialize in highways and bridges, the area in the public service where he would work after finishing the course. Among the disciplines at École Polytechnique, he attended classes on Algebra by Jean-Guillaume Garnier (1766-1840), Calculus by Sylvestre Lacroix (1765-1843), Geometry by Gaspard Monge (1746-1818) and Mechanics by Gaspard de Prony (1755-1839); with the tutorship of André-Marie Ampère (1775-1836). Augustin-Louis performed accordingly on these fundamental disciplines, which constituted the first two years of the course. After this period, he entered the École des Ponts et Chaussées at the end of 1807, in order to specialize in highways and bridges, as he had chosen. In the spring of 1808, he joined the team of the Ourcq Canal project, under the supervision of Pierre-Simon Girard (1765-1836), as a practical discipline of the course. In January 1810, he graduated with distinction as a junior engineer and, in February, was sent to work in the construction of Port Napoleon, in Cherbourg. The job was very demanding but Augustin-Louis kept developing his studies on pure Mathematics, particularly Geometry. In this period, he submitted *Studies on polyhedra* in 1811 to the *Journal de L'École Polytechnique*, an article that was highly praised by the judging commission and that brought him a good recognition in Paris. However, Augustin-Louis suddenly found himself drained by the obligations of a job he clearly hated and also very anxious by the opportunities he might be losing of pursuing an academic career on pure science in the capital. Endowed with the typical psychological rigidity that stems from a strong religious education, he could not manage his frustrations and got deeply depressed. In September 1812, when Marie-Madeleine went to visit Augustin-Louis in Cherbourg, she saw her son in a very bad condition and quickly took him back to Paris, definitely. After recovering from his illness, he kept studying Mathematics and finished two papers: one of them on the theory of combinations and the other on the theory of determinants, both submitted to the *Journal de L'École Polytechnique* and accepted. In order not to be sent back to Cherbourg, Augustin-Louis resorted to his father's prestige and influent friends to stay in Paris, when he was appointed to a post of engineer at the Ourcq Canal, a place he already known. While working there, his paper *On Definite Integrals*, submitted to the École Polytechnique in August, 1814, was highly praised and in December he was elected member of the Société Philomatique de Paris. On November, 1815, favorable political conditions led the Governor of the École Polytechnique to appoint Augustin-Louis assistant professor of Analysis, replacing the sickly professor Louis Poinsot (1777-1859). In this same month, he presented a study that would make him internationally famous, *Proof of Fermat's General Theorem on Polygonal Numbers*, published as an appendix of a Legendre's paper in 1816. This sudden notoriety enabled him to apply for a membership at the famous Royal Academy of Sciences and in March, 1816, Louis XVIII surprisingly appointed the twenty-seven-year-old Cauchy a member of the academy, replacing Gaspard Monge, expelled for political reasons. The reactionary political trend in France in that period and the great influence of his highly conservative father would once more favor Augustin-Louis' academic pretensions: in June he was appointed for full professorship in Analysis and Mechanics at the École Polytechnique. He also started teaching as an eventual substitute professor at the Collège de France and at the Faculté des Sciences. In April 4, 1818, Cauchy married to Aloïse de Bure, member of a traditional family of publishers and booksellers; a condition that later on proved to be quite convenient because it enabled him to publish most of his works independently, with the help of his father-in-law, on a periodical called *Exercices de Mathématiques*, dedicated only to his works. He and Aloïse had two daughters, Marie Françoise Alicia, born in 1819, and Marie Mathilde, born in 1823. In 1821, he published the first part of *Cours D'Analyse*[22], called *Analyse Algébrique*: a seminal textbook that, for the first time, presented infinitesimal calculus in a rigorous mathematical approach. It was in this context that Cauchy started developing and presenting his theories on Fluid and Solid Mechanics, which led to the submission of the aforementioned abstract on the equilibrium of deformable bodies in September, 1822. After the publication of this abstract in January, 1823, Navier correctly claimed that his August 1820 paper, cited by Cauchy in the abstract, had still not been published and waited for evaluation by the commission of Société Philomatique together with another paper of his, *On the Laws of Equilibrium and Motion of Elastic Solid Bodies*, submitted in May, 1821, that covered the same subject of Cauchy's abstract. Rumors of plagiarism began to be heard and then Augustin-Louis decided not to publish his achievements. This whole uncomfortable affair was the main reason for him to start publishing the mathematical development of his equilibrium equations only four years later, after Navier published his own paper on the January 1827 edition of the *Mémoires de L'Académie des Sciences de L'Institut de France*. During this delay, Cauchy made improvements on the theory and on its mathematical developments, which resulted in four papers published on his personal periodical *Exercices de Mathématiques* in the following chronological order.

1. March, 1827, pp. 42–57: *Sur La Pression ou Tension Dans Les Corps Solides*[23] (On the Pressure or Tension in Solid Bodies);
2. April, 1827, pp. 60–69: *Sur La Condensation Et La Dilatation Des Corps Solides* (On the Contraction and Expansion of Solid Bodies);
3. April, 1827, pp. 108–111: *Sur Les Relations Qui Existent, Dans L'etat D'equilibre D'un Corps Solide Ou Fluicie, Entre les Pressions Ou Tensions Et Les Forces Accel-*

[22]For an English translation, see Bradley & Sandifer[8].

[23]See Cauchy[12]. In Appendix A, p. 207, a newer version of this paper was translated by me.

eratrices (On The Relations That Exist, in the State of Equilibrium of a Solid or Fluid Body, Between the Pressures or Tensions and the Accelerating Forces);

4. September, 1828, pp. 160–187: *Sur Les Equations Qui Expriment Les Conditions D'equilibre Ou Les Lois Du Mouvement interieur D'un Corps Solide, Elastique Ou Non Elastique* (On the Equations That Express the Conditions of Equilibrium or the Laws of Interior Motion of a Solid Body, Elastic or Non-elastic).

The first of these important works became a classic on the emerging field of Continuum Mechanics, where Cauchy developed the subject presented on the abstract of 1823, and the others were improvements of the topics presented in this first one. Meanwhile, he also published in the March 1826 edition of his personal periodical the article *A New Genre of Calculus Analog to Infinitesimal Calculus*, where he presented the decomposition of rational fractions by means of the calculus of residues. A complex function of a complex variable was defined for the first time in the textbook *Lessons on Differential Calculus*, which he published in 1829, a year that closed a cycle of the greatest productivity in Cauchy's career; the political turmoil of the following period would definitely misguide him. Since the fall of Napoleon in 1814, France had been watching a strong political conflict between the supporters of the constitutional monarchy, led by the Bourbon monarch Charles X, and the bourgeois liberal movement. After the elections of 1830, when the liberals took the majority in the Chamber, the monarchy tried to suspend the freedom of the press and to dissolute the parliament. As a result, the so called Second French Revolution broke out in July 26, 1830, which led to the overthrow of Charles X by his cousin Louis Philippe, from the House of Orléans. Enraged by this profound political transformation that ended with the fall of the regime the Cauchys had passionately supported for so long and also by the oaths that he, as a public server, would have to make for the new king, the anti-liberal and impulsive Augustin-Louis Cauchy, falsely claiming health issues, left all his job positions and his country to the Italian city of Turin in the beginning of September 1830, a trip he imagined would not be so long – since his wife and daughters stayed in Paris – as it really turned out to be. Still abroad in November of that same year, the claim of health problems was no longer convincing and then he lost the position at Faculté des Sciences; in January 1831, he was fired from École Polytechnique and also from the engineering post at Ponts et Chaussées in March, but he managed to keep his membership at the Royal Academy. In Turin, the King Charles Albert of Sardinia created a post of Physics professor at the University specially for the famous French mathematician, a position Cauchy held from 1832 to 1833. In August 1833, he left Turin to Prague in order to be the tutor of the thirteen-year-old Duke of Bordeaux, the exiled grandson of Charles X. One year later, his wife and daughters finally came to live with him after four years apart. The unsuccessful and troublesome tutorship lasted until 1838, a period after which he received the title of baron from Charles X.

Cauchy returned to Paris in the autumn of 1838 and resumed his activities only at the Academy since he still refused to swear loyalty to the king, a mandatory requirement for French professors at that time. In 1843, he applied for a vacant chair of Mathematics at the Collège de France, but was not approved mainly for political reasons, since he was engaged in reestablishing the prestige of catholic university education in Paris, a movement considered by the academic community to be contrary to the Enlightenment ideals. With the fall of King Louis Philippe, the rise of Louis Napoleon Bonaparte in 1848 as the president of the French republic and the abolition of the oath requirement for public servers in 1849, Cauchy got a post of professor of mathematical astronomy at Faculté de Sciences, where he worked until the end of his life. In May 12, 1857, under his doctor's advice, Cauchy left Paris to his country house in Sceaux, in order to spend the whole summer, as a treatment for his unknown illness, but in May 23, at 4 a.m., "*he met death with such a calm that made us ashamed of our unhappiness*"[24], wrote his daughter Alicia.

Figure 6.14 – Daguerreotype of Augustin-Louis Cauchy taken by Charles Reutlinger.

Leonhard Euler starts the first volume of his fundamental *Scientia Navalis* with a lemma stating that *"The pressure which the water exerts upon a submerged body in its several points is normal to the surface of the body; and the force which an arbitrary element of the submerged surface sustains is equal to the weight of a right aqueous cylinder whose base is equal to the element of surface itself, and whose altitude is equal to the depth of the element below the upper surface of the water."*[25] The property of normal punctual compressions on surfaces submerged in water at rest was already known in 1738 and, according to Truesdell[63], it is difficult to trace its origins. It was also known the idea of considering the weight of the water sustained by a submerged surface as analog to it sustaining a column of water, which was first presented by the Belgian engineer Simon Stevin (?–1620) on his work *Hypomnemata mathematica: De statica*, published in 1605.

[24]See Belhoste[3], p.240.
[25]Euler[21], p. 1, translated from the Latin by Truesdell[63].

Today, these properties of water are assigned to a group of fluids called perfect, which have the feature of being "slippery", not capable of resisting tangential (shear) forces. In addition to the clarity and simplicity of a mathematical text, what is brilliant in Euler's statement is his new concept of pressure (punctual compression) as a measure of force per unit of area on each point of the submerged surface. Inspired by this concept of pressure and its corollaries, Cauchy came up with the idea that an element in the interior of a solid body has its surfaces also subjected to pressures, but, unlike the case of a perfect fluid at rest, a pressure on an arbitrary point of a solid surface may not be normal and may not be positive (compressive). In his *Sur La Pression ou Tension Dans Les Corps Solides*[26], from a small volume of solid limited by infinite surfaces with boundaries, Cauchy considered a pressure or tension (negative pressure) p_i acting on the points of an arbitrary surface s_i, defined $(\cos\lambda_i, \cos\mu_i, \cos\nu_i)$ as the cosine directions of p_i relative to rectangular coordinate axes (x, y, z) and took plane xy as a reference to describe the components of forces

$$\iint p_i \cos\lambda_i \sec\gamma_i \, dx dy\,, \quad \iint p_i \cos\mu_i \sec\gamma_i \, dx dy \quad \text{and} \quad \iint p_i \cos\nu_i \sec\gamma_i \, dx dy\,,$$

where γ_i defines the cosine direction relative to z of the unit normal to $\sec\gamma_i \, dx dy$, which is an infinitesimal element of s_i. When the volume of the small element tends to zero, higher order terms become negligible, and the above integrals are respectively simplified to

$$p_i dA_i \cos\lambda_i\,, \quad p_i dA_i \cos\mu_i \quad \text{and} \quad p_i dA_i \cos\nu_i\,,$$

where dA_i is the infinitesimal area of s_i. Concerning the infinite surfaces of the volume element in equilibrium, the balance of all the forces involved results

$$\sum_{i=1}^{\infty} p_i dA_i \cos\lambda_i = 0\,, \quad \sum_{i=1}^{\infty} p_i dA_i \cos\mu_i = 0 \quad \text{and} \quad \sum_{i=1}^{\infty} p_i dA_i \cos\nu_i = 0\,.$$

Cauchy applies this same procedure for linear momenta relative to the coordinate axes and obtains their correspondent balance expressions. Then, he considers a small volume element with the shape of a right prism whose bases s_1 and s_2 are parallel to the plane xy, both having area A. Making this prism infinitesimal by decreasing its height faster than the sides of its bases, this height can be neglected and therefore what remains of the element is an infinitesimal plane of area dA with faces labeled ds_1 and ds_2. In this context, the previous balance of forces can be described by

$$\sum_{i=1}^{2} p_i dA_i \cos\lambda_i = 0\,, \quad \sum_{i=1}^{2} p_i dA_i \cos\mu_i = 0 \quad \text{and} \quad \sum_{i=1}^{2} p_i dA_i \cos\nu_i = 0\,,$$

[26] See also appendix A.

from which equalities $p_1 = p_2$, $\cos\lambda_1 = -\cos\lambda_2$, $\cos\mu_1 = -\cos\mu_2$ and $\cos\nu_1 = -\cos\nu_2$ can be concluded. Although simple, this result is fundamental in Cauchy's mechanical theory because it allows us to assert that *an arbitrary infinitesimal plane, to which a point in equilibrium belongs, defines on it a unique pair of opposing tensions or pressures with the same intensity*. Making a small parallelepiped in equilibrium an infinitesimal volume element, a consequence of the previous statement is that, among the six pressures or tensions on the element faces, only three may be distinct, namely p_1, p_2 and p_3. In this context, Cauchy then obtained from the balance of linear momenta that the tangential pressures or tensions $p_1\cos\mu_1 = p_2\cos\lambda_2$, $p_1\cos\nu_1 = p_3\cos\lambda_3$ and $p_2\cos\nu_2 = p_3\cos\mu_3$.

Considering the set of all deformable bodies as constituted by solids and fluids, the feature of an arbitrary plane defining a pair of opposing pressures or tensions on a point in equilibrium becomes a generalization of the particular case of perfect fluids at rest, where an arbitrary plane, to which a point of depth h belongs, defines on it a pair of opposing equal pressures normal to this plane (hydrostatic). Moreover, if we represent tensions or pressures by geometric vectors and consider the general case of deformable bodies, the set of all the planes to which a point in equilibrium belongs defines on it a set of vectors that may not constitute a sphere, as does the set of all hydrostatic pressures on a point inside a perfect fluid at rest, since the intensity of an hydrostatic pressure depends only on the depth h, according to Euler's lemma. If the geometric shape of the set of tension or pressure vectors on a point in equilibrium may not be a sphere, which shape would it be? In order to answer this question, Cauchy elected his famous tetrahedron as a small volume element, defined pressure p on a point of its basis s as unknown and $(\cos\alpha, \cos\beta, \cos\gamma)$ as the cosine directions of the unit normal to s. From the previous expressions of balance of forces and after lengthy algebraic manipulations, he arrived at the following equation

$$\begin{aligned} p^2 = &(A\cos\alpha + F\cos\beta + E\cos\gamma)^2 + \\ &(F\cos\alpha + B\cos\beta + D\cos\gamma)^2 + \\ &(E\cos\alpha + D\cos\beta + C\cos\gamma)^2 \end{aligned}$$

and the projection of tension or pressure p on the unit normal to s

$$\begin{aligned} p\cos\delta = {} & A\cos^2\alpha + B\cos^2\beta + C\cos^2\gamma + \\ & + 2D\cos\beta\cos\gamma + 2E\cos\gamma\cos\alpha + 2F\cos\alpha\cos\beta \end{aligned}$$

where coefficients $A = p_1\cos\lambda_1$, $B = p_2\cos\mu_2$, $C = p_3\cos\nu_3$, $D = p_2\cos\nu_2$, $E = p_3\cos\lambda_3$ and $F = p_1\cos\mu_1$. Relative to the point on which it acts on s, the vector defined by $p\cos\delta$ has coordinates $x = p\cos\delta\cos\alpha$, $y = p\cos\delta\cos\beta$ and $z = p\cos\delta\cos\gamma$. Considering these coordinates, the previous equality divided by $(p\cos\delta)^2$ results the polynomial

$$Ax^2 + By^2 + Cz^2 + 2Dyz + 2Ezx + 2Fxy - \frac{1}{p\cos\delta} = 0,$$

which describes a quadratic surface. Based on this polynomial, Cauchy then found out that there are three mutually perpendicular bases $s_{(1)}, s_{(2)}, s_{(3)}$ of the tetrahedron to which pressures or tensions $p_{(1)}, p_{(2)}, p_{(3)}$ are respectively normal (hydrostatic) and that two of them are the maximum and the minimum values of all the pressures or tensions on the point under consideration. Because of these features, he called $p_{(1)}, p_{(2)}, p_{(3)}$ the principal pressures or tensions on this point. Considering a new tetrahedron in a rectangular coordinate system defined by the intersection of $s_{(1)}$, $s_{(2)}$ and $s_{(3)}$, the balance of forces becomes rather simple for a basis s where the principal pressures or tensions appear simultaneously on the faces of the tetrahedron:

$$
\begin{aligned}
p\cos\delta\cos\alpha dA - p_{(1)}dA\cos\alpha &= 0,\\
p\cos\delta\cos\beta dA - p_{(2)}dA\cos\beta &= 0,\\
p\cos\delta\cos\gamma dA - p_{(2)}dA\cos\gamma &= 0.
\end{aligned}
$$

Defining $X = p\cos\delta\cos\alpha$, $Y = p\cos\delta\cos\beta$ and $Z = p\cos\delta\cos\gamma$ and by the fundamental trigonometric property of cosine directions, namely $\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$, we easily arrive at

$$
\frac{X^2}{p_{(1)}^2} + \frac{Y^2}{p_{(2)}^2} + \frac{Z^2}{p_{(3)}^2} = 1,
$$

from which it is clear that the principal pressures or tensions define the magnitutes of the axes of an ellipsoid, which is the set defined by all the tension or pressure vectors on the point under consideration. Since the ellipsoid is the geometric generalization of the sphere, pressures or tensions on a point of a perfect fluid at rest can be considered a particular case of Cauchy's theory.

And finally our brief history comes to an end. For the purposes of this book, we decided not to be exhaustive and chose to present only short passages about the life and the work of the outstanding initial contributors on Continuum Mechanics; unfortunately, other important scholars had to be left out. For the reader or student interested in deepening their knowledge on the historical matters of this field, we strongly recommend the monumental accounts of Maugin[41] [42] [43].

Chapter

7

Continuum Kinematics

In practice, this chapter introduces our particular approach of the subject of Continuum Mechanics, which will be presented, here and henceforth, by using the mathematical tools and concepts covered in the initial chapters. Since we didactically decided to study Mathematics first, the development of the physical concepts here will assume that the reader is fluent on the topics of the first part of this book. The present introduction deals with the basic fundamental concepts commonly used in the literature of Continuum Mechanics to describe motion: position, deformation, strain, displacement, velocity and acceleration. This description, Kinematics, is here based mainly on time dependent bivariate and time independent univariate functions of positions or material particles – Eulerian or Lagrangian descriptions respectively.

7.1 Continuum Mechanics: a Mathematical Model

In classical terms, Mechanics is the physical science of **motion** concerned with its descriptive and predictive aspects. The descriptive study of motion, or Kinematics, deals mainly with its geometric measurements while the predictive study, Dynamics, considers its causes. For both of these aspects, the notion of **time** is fundamental, enabling the very idea of motion and most of its related concepts, particularly in Kinematics. For the purposes of Dynamics, **force** is the main concept, being the motive of geometric changes in time, a causality relation, called the **Laws of Motion**, which aims to be as generic as possible and independent of material specificities. But when the influence of these specificities can not be disregarded for the mechanical phenomena under study, particular **constitutive** relations have to be considered.

In a restricted and objective way, Kinematics adopts the common subject-object philosophical approach, where the subject is merely an **observer** and the object the thing observed. In the act of kinematically describing a mechanical event, the observer makes use of two fundamental concepts: position and time. In Newtonian or Classical Mechanics, which is our concern here, **space**, mathematically modeled as a three dimensional Euclidean space, is the context or the framework that enables the subject to perform its positional identifications, when either a certain region of this space or a specific physical entity are the observed objects. The biggest physical space, which includes the whole universe, is defined to be **absolute**, that is, it has a fixed size, it is completely at rest and not affected by its elements, being a totally independent entity. Any observer attached to the absolute space is also called absolute.

When performing a certain measurement on the object under observation, the observer notes that changes may occur: given a point in space or a material element of a physical entity, there may be a set of different measurement values on this point or element that develop one after another, sequentially. Time is the concept that enables the observer to describe these changes by identifying which measurement value follows the other in such a way as to build a set of successive measurement values for every observed point or element. There is always a specific instance or value of time, called **instant**, that can label a certain measurement value; in other words, time is infinite and infinitely divisible[1], or an infinite unlimited complete set, conveniently modeled as a real field. When a limited portion of this set is selected to describe a certain mechanical event, we call this portion a **period**. If the observer measures the positions of the material elements of a limited physical entity, a biunivocal relation of these elements with space points in a given instant is called a **configuration** or **deformation** of this physical entity and the set of these space points is said to be a **shape** of this entity. Moreover, a set of distinct shapes on a certain period of time is defined to be the image of a motion of the physical entity. Still in the context of Newtonian Mechanics, time is considered to be unstoppable, always progressing indefinitely, and the rate of this progress is the same for any observer; in short, time is also considered absolute. Therefore, we can say that every observer's clock is synchronized with a unique reference clock. As a consequence of the ever increasing Newtonian time, an observer describing a certain mechanical event can never label two distinct measurement values relative to the same point of an object by the same instant: simultaneity here makes sense only for different points. Moreover, an observer may eventually become an ob-

[1]On this subject, we cannot help but quoting the great writer Leo Tolstoy: *"For human reason, absolute continuity of movement is incomprehensible. Man begins to understand the laws of any kind of movement only when he examines the arbitrarily chosen units of that movement. But at the same time it is from this arbitrary division of continuous movement into discrete units that the greater part of human errors proceeds... A new branch of mathematics, having attained to the art of dealing with infinitesimal quantities in other, more complex problems of movement as well, now gives answers to questions that used to seem insoluble. This new branch of mathematics, unknown to the ancients, in examining questions of movement, allows for infinitesimal quantities, that is, such as restore the main condition of movement (absolute continuity), and thereby corrects the inevitable error that human reason cannot help committing when it examines discrete units of movement instead of continuous movement."* (Tolstoy[62], p. 955)

ject when another observer can describe its motion. That said, an observer at rest or moving with constant velocity relative to an absolute observer is called **inertial**: it is only from the point of view of an inertial observer that Newton's Axiom of Inertia is valid.

Following Hermann Weyl, "*space and time are commonly regarded as the forms of existence of the real world,* ***matter*** *as its substance.*[2]" It is observed that what constitutes matter, the specific material it is made of, influences its mechanical behavior. In this sense, a mechanical description concerned with material peculiarities is called constitutive. Another topic on the study of matter is its structure, that is, its building blocks, commonly called particles, and the way they are organized. For some reason, the popularity of this topic promoted the preconception, even in academic circles, that a physical theory which seeks to describe the overall behavior of matter becomes more reliable, or is much closer to reality, when it considers structural phenomena. In our opinion, this idea tries to attach what it is not attachable. A mechanical theory, expressed as a mathematical model, is good if its results are sufficiently close to experimental data, when available, and if it describes an ample set of physical phenomena, regardless whether structural variables are considered or not. In studying the mechanical behavior of matter, the approach called Continuum Mechanics, which models a limited portion of matter, or body, as a continuum, presents both of these virtues, notwithstanding it totally disregards structural phenomena.

For the purposes of this book, matter is considered to be impenetrable, that is, a point in space can never be occupied by two different particles at a particular instant of time. In the context of Continuum Mechanics, where the physical space is a three-dimensional Euclidean affine space, time is a real field and body is a continuum, we conclude from the impenetrability of matter that a deformation of a given body is always a bijection – labeled by a real scalar representing an instant of time – that maps the body to a shape, which is a bounded subset of the three-dimensional Euclidean affine space. In this sense, the body is here considered to be a formless portion of matter, related through deformation to an Euclidean geometric shape in a certain instant of time. Considering the biunivocal property of the deformation and since every portion of matter has mass, we can define on an arbitrary shape of a given body a mass density functional, whose volume integral results the mass of the body relative to the shape in question.

7.2 Deformation

Concerning the subject of this chapter, the intuitive concepts and definitions introduced in the previous section will now be detailed from the mathematical material presented in the first part of this book. In our study of Mechanics, the absolute space

[2]See Weyl[70], p.1.

is a three-dimensional Euclidean affine space $\mathcal{U}^3_{\mathbb{R}}$ whose affine coordinate system is the absolute observer. A complete metric space (B,ρ) is called a body B if it is connected and totally bounded; in other words, if it is a continuum. During a motion of this body, its shape at instant t is $\mathcal{B}_t := (\mathcal{B}^3_{\mathbb{R}})_t$, which is a subset of the absolute universe $\mathcal{U}^3_{\mathbb{R}}$, defined by an oriented Euclidean vector space $(\mathsf{U}_{\mathbb{R}},\mathbf{A}_{\mathsf{O}})$, where O is its natural basis. Formally, a **motion** of $\mathfrak{B}$ is the function in the surjective mapping

$$\chi : \mathfrak{B} \times \mathbb{T} \mapsto \mathcal{S}, \tag{7.1}$$

where $\mathbb{T} \subseteq \mathbb{R}$ is the set constituted by instants of time and image

$$\mathcal{S} = \bigcup_{t\in\mathbb{T}} \mathcal{B}_t\,. \tag{7.2}$$

Since time is an independent variable, for most of the forthcoming concepts to be presented in this section, it is convenient to fix an arbitrary instant $t \in \mathbb{T}$ and define a bijective mapping $\chi_t : \mathfrak{B} \mapsto \mathcal{B}_t$ where $\chi_t(x) = \chi(x,t)$. Among other reasons, since surface integrals will be a fundamental tool in this study, we shall deal only with motions that define shapes bounded by Lipschitz surfaces. In order to identify the particles of body $\mathfrak{B}$ by geometric points, we choose a reference instant of time $\bar{t}$ in such a way that an arbitrary particle $x \in \mathfrak{B}$ is uniquely identified by the point $\bar{\chi}(x) := \chi_{\bar{t}}(x)$, from which set $\overline{\mathcal{B}} := \{\bar{\chi}(x) : x \in \mathfrak{B}\}$ is called a reference shape. It is important to note that the reference instant is not necessarily the initial instant t_0 of the period $[t_0, t_f]$ under study, that is, $t_f \geqslant t_0 \geqslant \bar{t}$; a condition that enables body $\mathfrak{B}$ to have three notable shapes: $\overline{\mathcal{B}}$, $\mathcal{B}_{t_0}$ and $\mathcal{B}_{t_f}$. The one-to-one identification of particles by points allows us to call the reference shape also a body, but now, a body with affine Euclidean features. In order to make shapes viable for mathematical calculations, we still specify that

$$\overline{\boldsymbol{\mathcal{B}}} := \{\upsilon_o(x) : \forall x \in \overline{\mathcal{B}}\} \qquad \text{and} \qquad \boldsymbol{\mathcal{B}}_t := \{\upsilon_o(x) : \forall x \in \mathcal{B}_t\}, \tag{7.3}$$

where υ_o "vectorizes" the points of $\overline{\mathcal{B}}$ and $\mathcal{B}_t$ relative to an origin $o \in \mathcal{U}^3_{\mathbb{R}}$, according to definition (4.3). Since υ_o is a bijection, the sets of vectors $\overline{\boldsymbol{\mathcal{B}}}$ and $\boldsymbol{\mathcal{B}}_t$ will also be called body and shape respectively. In this vector framework where $\boldsymbol{\mathcal{S}}$ is the vector image of the motion, we impose that the vector motion $\boldsymbol{\chi}$ in mapping

$$\boldsymbol{\chi} : \overline{\boldsymbol{\mathcal{B}}} \times \mathbb{T} \mapsto \boldsymbol{\mathcal{S}} \tag{7.4}$$

is a $\mathcal{C}^3$ surjection that observes $\boldsymbol{\chi}(\bar{\mathbf{x}},\bar{t}) = \bar{\mathbf{x}}$. By fixing time, we can define the bijective mapping $\boldsymbol{\chi}_t : \overline{\boldsymbol{\mathcal{B}}} \mapsto \boldsymbol{\mathcal{B}}_t$ where function is called the deformation or configuration of $\overline{\boldsymbol{\mathcal{B}}}$ at t if it obeys equality

$$\boldsymbol{\chi}_t(\bar{\mathbf{x}}) = \boldsymbol{\chi}(\bar{\mathbf{x}},t)\,, \forall \bar{\mathbf{x}} \in \overline{\boldsymbol{\mathcal{B}}}, \tag{7.5}$$

from which we conclude that the univariate motion $\boldsymbol{\chi}_t$ observes restriction $\boldsymbol{\chi}_{\bar{t}}(\bar{\mathbf{x}}) = \bar{\mathbf{x}}$ and results a $\mathcal{C}^3$-diffeomorphism due to the impenetrability of matter. Given an arbitrary point $\bar{\mathbf{u}} \in \overline{\boldsymbol{\mathcal{B}}}$ and conveniently chosen non zero vectors $\bar{\mathbf{v}}_i \in \overline{\boldsymbol{\mathcal{B}}}$, from which

$\overline{\mathbf{u}}_i = \overline{\boldsymbol{v}}_i - \overline{\mathbf{u}}$ constitute a linearly independent set $\{\overline{\mathbf{u}}_1, \overline{\mathbf{u}}_2, \overline{\mathbf{u}}_3\}$, since body points cannot collapse during deformation, according to (4.21) volume

$$\mathbf{A}_O[\boldsymbol{\chi}_t(\overline{\mathbf{u}}_1), \boldsymbol{\chi}_t(\overline{\mathbf{u}}_2), \boldsymbol{\chi}_t(\overline{\mathbf{u}}_3)] \neq 0. \tag{7.6}$$

The continuity and bijectivity of $\boldsymbol{\chi}_t$ respectively ensures that Lipschitz surfaces are preserved[3] and no distinct elements of the body become indistinct or collapsed in the shape; while the differentiability of class two provides the minimum level of regularity required by future concepts. The entities and their relationships presented so far are depicted in figure 7.1. From this figure, it is possible to conclude that if both

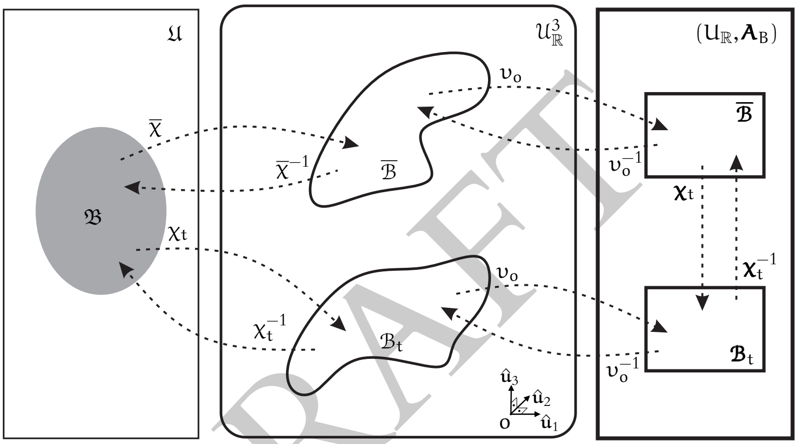

Figure 7.1 – Body shapes, vectorizations and deformations.

deformations $\overline{\chi}$ and χ_t are specified, a deformation $\boldsymbol{\chi}_t$ can be defined by

$$\boldsymbol{\chi}_t = \upsilon_o \circ \chi_t \circ \overline{\chi}^{-1} \circ \upsilon_o^{-1}. \tag{7.7}$$

Moreover, concerning the elements of the domains represented in the figure, we say that a material particle $x \in \mathfrak{B}$, represented by the **material point** $\overline{x} = \overline{\chi}(x)$ or by the **material vector** $\overline{\mathbf{x}} := \upsilon_o \circ \overline{\chi}(x)$, occupies the **spatial place** $x := \chi_t(x)$ or $\mathbf{x} = \upsilon_o \circ \chi_t(x)$. In order to simplify our study, since deformation was defined as a vector function, we shall mostly use vectors to denote body and shape elements. In this context, when a vector function is defined on a material domain $\overline{\boldsymbol{\mathcal{B}}}$ or on a spatial domain $\boldsymbol{\mathcal{B}}_t$, it is called **Lagrangian** or **Eulerian** respectively.

Considering the previous regularity conditions and an open set $\mathbf{W} \subset \overline{\boldsymbol{\mathcal{B}}}$, the derivative of $\boldsymbol{\chi}_t$ at an arbitrary particle $\overline{\mathbf{u}} \in \mathbf{W}$ is the vector function $(\boldsymbol{\chi}_t)'_{\overline{\mathbf{u}}}$ such that, from equality (5.29),

$$(\boldsymbol{\chi}_t)'_{\overline{\mathbf{u}}}(\mathbf{x}) = \nabla\boldsymbol{\chi}_t(\overline{\mathbf{u}})^{\mathrm{T}} \hat{\odot}_1 \mathbf{x}^*, \tag{7.8}$$

[3]See Ciarlet[16], theorem 1.2-8, p.16.

where the function in $\nabla\boldsymbol{\chi}_{\mathrm{t}} : \mathcal{W} \mapsto \mathcal{L}_{\mathbb{R}}(\mathrm{U}^2)$, called **deformation gradient**, has fundamental importance in Continuum Kinematics because it bases the concept of strain. The second order tensor $\nabla\boldsymbol{\chi}_{\mathrm{t}}(\overline{\mathbf{u}}) = \nabla_{\overline{\mathbf{u}}}\boldsymbol{\chi}(\overline{\mathbf{u}}, \mathrm{t})$ is said to be the deformation gradient at $\overline{\mathbf{u}}$ and, since deformation $\boldsymbol{\chi}_{\mathrm{t}}$ is a diffeomorphism, we can recall property (5.45), from which tensor

$$\nabla\boldsymbol{\chi}_{\mathrm{t}}(\overline{\mathbf{u}})^{-1} = \nabla\boldsymbol{\chi}_{\mathrm{t}}^{-1}(\mathbf{u}). \tag{7.9}$$

Now, at reference instant $\mathrm{t} = \bar{\mathrm{t}}$, we know that $\boldsymbol{\chi}_{\bar{\mathrm{t}}}(\overline{\mathbf{x}}) = \overline{\mathbf{x}} = \mathbf{i}_{\overline{\mathcal{B}}}(\overline{\mathbf{x}})$ and therefore, from a basic property of directional derivatives, equality $\boldsymbol{\chi}'_{\bar{\mathrm{t}}\overline{\mathbf{u}}} = \mathbf{i}_{\overline{\mathcal{B}}}$ holamads. But since $\boldsymbol{\chi}'_{\bar{\mathrm{t}}\overline{\mathbf{u}}}$ is the representative function of second order Euclidean tensor $\nabla\boldsymbol{\chi}_{\bar{\mathrm{t}}}(\overline{\mathbf{u}})$, from (4.22) we conclude that the infinitesimal or local volume change $\mathrm{Det}[\nabla\boldsymbol{\chi}_{\bar{\mathrm{t}}}(\overline{\mathbf{u}})] = 1$. Moreover, since deformation $\boldsymbol{\chi}_{\mathrm{t}}$ is a diffeomorphism, the property relative to equality (5.42) is valid and then $\mathrm{Det}[\nabla\boldsymbol{\chi}_{\mathrm{t}}(\overline{\mathbf{u}})] \neq 0$ because $\nabla\boldsymbol{\chi}_{\mathrm{t}}(\overline{\mathbf{u}}) \neq \mathbf{0}$. Therefore, we can state that the local volume change

$$\mathrm{Det}[\nabla\boldsymbol{\chi}_{\mathrm{t}}(\overline{\mathbf{u}})] > 0 \tag{7.10}$$

because motion $\boldsymbol{\chi}$ and its derivative are continuous: during a motion where $\bar{\mathrm{t}} = \mathrm{t}_0$, the local volume change $\mathrm{Det}[\nabla\boldsymbol{\chi}_{\mathrm{t}}(\overline{\mathbf{u}})]$, which is one at $\mathrm{t} = \mathrm{t}_0$ and different from zero otherwise, cannot "jump" zero and be negative. From (4.22) and this previous inequality, we can conclude that vector function $\boldsymbol{\chi}'_{\mathrm{t}\overline{\mathbf{u}}}$ is orientation-preserving[4], a consequence that is used by some authors as a local necessary condition for the vector function $\boldsymbol{\chi}_{\mathrm{t}}$ to be considered a deformation. A deformation $\boldsymbol{\chi}_{\mathrm{t}}$ where the local volume change

$$\mathrm{Det}[\nabla\boldsymbol{\chi}_{\mathrm{t}}(\overline{\mathbf{u}})] = 1 \tag{7.11}$$

is said to be **isochoric** on $\overline{\mathbf{u}}$ or just isochoric when this equality is valid for all $\overline{\mathbf{u}} \in \overline{\mathcal{B}}$.

Recalling affinities and their related concepts presented in section 4.2, a motion $\boldsymbol{\varphi}$ is said to be **affine** if it is described by the rule

$$\boldsymbol{\varphi}(\overline{\mathbf{x}}, \mathrm{t}) = [\nabla\boldsymbol{\varphi}_{\mathrm{t}}(\overline{\mathbf{x}})^{\mathrm{T}} \,\hat{\odot}_1\, \overline{\mathbf{x}}^*] + \mathbf{c}, \tag{7.12}$$

where second order tensor

$$\mathsf{F}_{\overline{\mathbf{u}}} := \nabla\boldsymbol{\varphi}_{\mathrm{t}}(\overline{\mathbf{u}})^{\mathrm{T}} \tag{7.13}$$

is called the **affine deformation gradient** of vector affinity $\boldsymbol{\varphi}_{\mathrm{t}}$ at $\overline{\mathbf{u}}$ and $\mathbf{c} \in \mathrm{U}_{\mathbb{R}}$ is a constant vector. Therefore, if $\mathsf{F}_{\overline{\mathbf{x}}}$ is the identity tensor $\mathbf{I} \in \mathcal{L}_{\mathbb{R}}(\mathrm{U}^2)$ or a stretch tensor for all $\overline{\mathbf{x}} \in \overline{\mathcal{B}}$, then affine deformation $\boldsymbol{\varphi}_{\mathrm{t}}$ is a vector translation or a centered vector dilation respectively, according to (4.16) and (4.18). It is important to observe that an

[4] About this topic, professor Michael Spivak wrote: "*The non-singular linear maps* $\mathsf{f} : \mathsf{V} \mapsto \mathsf{V}$ *from a finite dimensional vector space to itself fall into two groups, those with* $\det \mathsf{f} > 0$ *and those with* $\det \mathsf{f} < 0$*; linear transformations in the first group are called* ***orientation preserving*** *and the others are called* ***orientation reversing****...There is no way to pass continuously between these two groups...*" (SPIVAK[58], p. 84)

affine deformation can be decomposed in a rotodilation, according to theorem 18 and corollary 18.1, in such a way that

$$\boldsymbol{\varphi}_t = \boldsymbol{s}_2 \circ \boldsymbol{r}_c = \boldsymbol{r}_c \circ \boldsymbol{s}_1 \tag{7.14}$$

and

$$\mathbf{F}_{\overline{u}} = \mathbf{V}_{\overline{u}} \odot_1 \mathbf{R}_{\overline{u}} = \mathbf{R}_{\overline{u}} \odot_1 \mathbf{U}_{\overline{u}}\,, \forall \overline{\mathbf{u}} \in \overline{\mathcal{B}}, \tag{7.15}$$

where $\mathbf{V}_{\overline{u}}$ and $\mathbf{U}_{\overline{u}}$, usually called **left** and **right stretch tensors**, are the affinity tensors of $\boldsymbol{s}_2$ and $\boldsymbol{s}_1$, both having the same stretch coefficients $\lambda_1, \lambda_2, \lambda_3$. Since the local volume change is always positive, as we already know, $\mathbf{R}_{\overline{u}}$ can never be a rotoreflection tensor. Moreover, from the polar decomposition (7.14), we can conclude that stretch operators $\mathbf{s}_1^{1/2}$ is Lagrangian and $\mathbf{s}_2^{-1/2}$ is Eulerian, representing $\mathbf{V}_{\overline{u}}^{-1}$, whose stretch coefficients are $\lambda_1^{-1}, \lambda_2^{-1}, \lambda_3^{-1}$. The stretch coefficients $\lambda_1, \lambda_2, \lambda_3$ and $\lambda_1^{-1}, \lambda_2^{-1}, \lambda_3^{-1}$ are also said to be the **principal stretches**, at instant t, on material point $\overline{\mathbf{u}}$ and on spatial point $\mathbf{u}$ respectively. In practice, it is convenient not to deal with square root operators $\mathbf{s}_1^{1/2}$ and $\mathbf{s}_2^{1/2}$, which represent the above stretch tensors and are troublesome to compute, but instead to specify

$$\mathbf{B}_{\overline{u}} := \mathbf{V}_{\overline{u}}^2 \qquad \text{and} \qquad \mathbf{C}_{\overline{u}} := \mathbf{U}_{\overline{u}}^2, \tag{7.16}$$

known as the **left** and **right Cauchy-Green deformation tensors**, represented by operators $\mathbf{s}_2$ and $\mathbf{s}_1$ respectively, according to (3.54), whose stretch coefficients are $\lambda_1^2, \lambda_2^2, \lambda_3^2$. Since $\mathbf{s}_2^{-1}$ and $\mathbf{s}_1$ are respectively Eulerian and Lagrangian, their represented tensors $\mathbf{B}_{\overline{u}}^{-1}$ and $\mathbf{C}_{\overline{u}}$ are also called Eulerian and Lagrangian. In terms of the affine deformation gradient,

$$\mathbf{B}_{\overline{u}} = \mathbf{F}_{\overline{u}} \odot_1 \mathbf{F}_{\overline{u}}^T \qquad \text{and} \qquad \mathbf{C}_{\overline{u}} = \mathbf{F}_{\overline{u}}^T \odot_1 \mathbf{F}_{\overline{u}}. \tag{7.17}$$

Proof. Firstly, the following development

$$\begin{aligned} 0 &< \mathrm{Det}[\nabla \boldsymbol{\varphi}_t(\overline{\mathbf{u}})] \\ &< \mathrm{Det}(\mathbf{F}_{\overline{u}}^T) \\ &< \mathrm{Det}(\mathbf{R}_{\overline{u}}^{-1} \odot_1 \mathbf{V}_{\overline{u}}) \\ &< [\mathrm{Det}(\mathbf{R}_{\overline{u}})]^{-1} \mathrm{Det}(\mathbf{V}_{\overline{u}}) \end{aligned}$$

proves that $\mathbf{R}_{\overline{u}}$ is not a rotoreflection tensor, that is $\mathrm{Det}(\mathbf{R}_{\overline{u}}) > 0$, because $\mathrm{Det}(\mathbf{V}_{\overline{u}}) > 0$ as a consequence of $\mathbf{V}_{\overline{u}}$ being symmetric positive-definite. Now, we'll verify one of the properties (7.17).

$$\begin{aligned} \mathbf{B}_{\overline{u}} &= \mathbf{F}_{\overline{u}} \odot_1 \mathbf{F}_{\overline{u}}^T \\ &= \mathbf{V}_{\overline{u}} \odot_1 \mathbf{R}_{\overline{u}} \odot_1 (\mathbf{V}_{\overline{u}} \odot_1 \mathbf{R}_{\overline{u}})^T \\ &= \mathbf{V}_{\overline{u}} \odot_1 \mathbf{R}_{\overline{u}} \odot_1 \mathbf{R}_{\overline{u}}^T \odot_1 \mathbf{V}_{\overline{u}}^T \\ &= \mathbf{V}_{\overline{u}} \odot_1 \mathbf{R}_{\overline{u}} \odot_1 \mathbf{R}_{\overline{u}}^{-1} \odot_1 \mathbf{V}_{\overline{u}} \\ &= \mathbf{V}_{\overline{u}} \odot_1 \mathbf{V}_{\overline{u}}. \end{aligned}$$

□

When the affine deformation gradient is constant on $\overline{\mathfrak{B}}$, the deformation is called **homogeneous**, represented by $\tilde{\boldsymbol{\varphi}}_t$. In other words, tensor $\mathsf{F}_{\overline{u}} = \mathsf{F}$ for all $\overline{u} \in \overline{\mathfrak{B}}$ and then

$$\tilde{\boldsymbol{\varphi}}_t(\overline{x}) = (\mathsf{F} \hat{\odot}_1 \overline{x}^*) + \mathbf{c}. \tag{7.18}$$

It is interesting to note that if $\mathsf{F} = \mathbf{I}$, deformation $\tilde{\boldsymbol{\varphi}}_t$ results a vector translation. In the rotodilation decomposition of the homogeneous affinity tensor F, the subscripts identifying point dependence can be removed, resulting

$$\mathsf{F} = \mathsf{V} \odot_1 \mathsf{R} = \mathsf{R} \odot_1 \mathsf{U}, \qquad \mathsf{B} = \mathsf{F} \odot_1 \mathsf{F}^{\mathrm{T}} \quad \text{and} \quad \mathsf{C} = \mathsf{F}^{\mathrm{T}} \odot_1 \mathsf{F}. \tag{7.19}$$

Example 7.1. Before examples of notable homogeneous deformations are presented, we specify eight body points that constitute the vertices of a cube $\overline{\mathfrak{C}}$: an arbitrarily chosen $\overline{u}$, three orthonormal vectors $\overline{u}_i = \hat{e}_i + \overline{u}$, $i = 1,2,3$, that constitute basis U, where $\hat{e}_i$ constitute the natural basis O; points $\overline{u}_4 = \hat{e}_1 + \hat{e}_2 + \overline{u}$, $\overline{u}_5 = \hat{e}_1 + \hat{e}_3 + \overline{u}$, $\overline{u}_6 = \hat{e}_2 + \hat{e}_3 + \overline{u}$ and $\overline{u}_7 = \hat{e}_1 + \hat{e}_2 + \hat{e}_3 + \overline{u}$. If we consider $\mathbf{f}$ the representative function of deformation gradient F, rule (7.18) can be rewritten as $\tilde{\boldsymbol{\varphi}}_t(\overline{x}) = \mathbf{f}(\overline{x}) + \mathbf{c}$. An homogeneous deformation where $\mathsf{F} = \mathsf{R}$ is called **rigid** and then $\mathbf{f}$ results proper orthogonal. Figure 7.2 depicts a rigid deformation of $\overline{\mathfrak{C}}$ where it translates and rotates a counterclockwise angle θ around the axis defined by $\overline{u}$ and $\overline{u}_3$. In matrix terms, the rule of this deformation is

$$[\tilde{\boldsymbol{\varphi}}_t(\overline{x})]^{\mathsf{U}} = \underbrace{\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{[\mathbf{f}_{\mathsf{U}}]^{\mathsf{U}}} [\overline{x}]^{\mathsf{U}} + [\mathbf{c}]^{\mathsf{U}}.$$

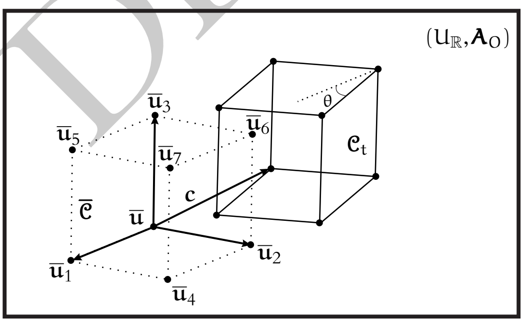

Figure 7.2 – Rigid deformation.

In the case of $\mathsf{R} = \mathbf{I}$, $\mathbf{c} = \mathbf{0}$ and $\mathsf{U} = \mathsf{V} = \sum_{i=1}^{3} \lambda_i \overline{u}_i \otimes \overline{u}_i$, where λ_i is a stretch coefficient, then $\tilde{\boldsymbol{\varphi}}_t$ is called a **pure stretch**. Figure 7.3 shows $\overline{\mathfrak{C}}$ subjected to a pure stretch described by

$$[\tilde{\boldsymbol{\varphi}}_t(\overline{x})]^{\mathsf{U}} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} [\overline{x}]^{\mathsf{U}}.$$

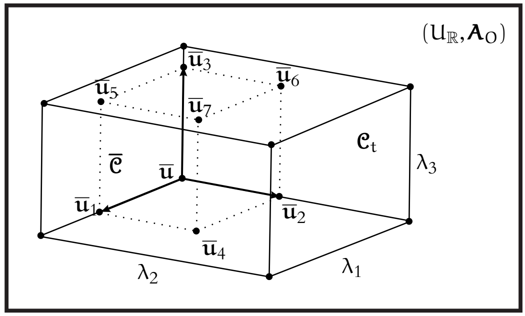

Figure 7.3 – Pure stretch.

We recall that when stretch coefficients are identical, pure stretch is said to be proportional. Now, if $\mathbf{c} = \mathbf{0}$ and deformation gradient $\mathsf{F} = \mathbf{I} + \gamma \overline{u}_i \otimes \overline{u}_j$, where $i \neq j$, then $\tilde{\boldsymbol{\varphi}}_t$ is called a **simple shear.** Considering $i = 2$ and $j = 3$, we have $\overline{\mathfrak{C}}$ subjected to a simple shear represented on figure 7.4 and described by

$$[\tilde{\boldsymbol{\varphi}}_t(\overline{x})]^U = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \gamma \\ 0 & 0 & 1 \end{bmatrix} [\overline{x}]^U .$$

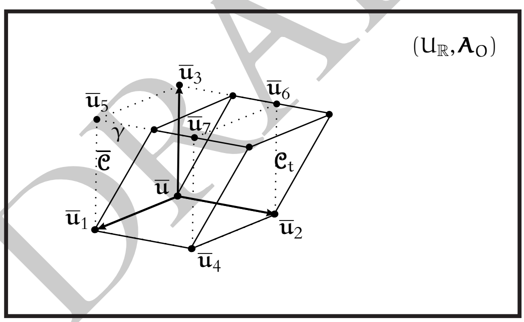

Figure 7.4 – Simple shear.

The quantity or level of local deformation can be obtained by measuring the change in an infinitesimal length of the body. This local length change, called **strain,** can be measured from the stretch tensors related to an arbitrary material point $\overline{u}$. Thereby, in the context of the normalized eigenbases X and Y of the stretch operators that represent $\mathbf{U}_{\overline{u}}$ and $\mathbf{V}_{\overline{u}}$ respectively, most of the strain measures evaluate the deviation of a given stretch from the stretch of the rigid deformation where $\lambda_1 = \lambda_2 = \lambda_3 = 1$. The simplest type of this deviation is an algebraic difference called **extension**, whose Lagrangian and Eulerian forms are respectively described by $\delta_i = \lambda_i - 1$ and $\Delta_i = 1 - \lambda_i^{-1}$, which are zero for rigid deformations. But since principal stretches λ_i are related to annoying square root operators, convenience led to the definition of extensions described by

$$\tilde{\delta}_i = \lambda_i^2 - 1 \qquad \text{and} \qquad \tilde{\Delta}_i = 1 - \lambda_i^{-2}, \tag{7.20}$$

which are valid because $\tilde{\delta}_i = \tilde{\Delta}_i = 0$ in the context of rigid deformations. Such extensions can respectively be written in matrix form as $[\mathbf{U}_{\overline{u}}^2]^X - I$ and $I - [\mathbf{V}_{\overline{u}}^{-2}]^Y$, which are matrix representations of the classical definitions

$$2\mathbf{E}_{\overline{u}} = \mathbf{C}_{\overline{u}} - \mathbf{I} \qquad \text{and} \qquad 2\boldsymbol{e}_{\overline{u}} = \mathbf{I} - \mathbf{B}_{\overline{u}}^{-1}, \tag{7.21}$$

where Lagrangian tensor $\mathbf{E}_{\overline{u}}$ is called the **Green-St Venant** strain tensor and Eulerian tensor $\boldsymbol{e}_{\overline{u}}$ the **Almansi-Hamel** strain tensor. From both of these tensors, DOYLE & ERICKSEN[18] and SETH[54] independently proposed a generalization that include most of the different strain measures available. In more complicated terms, they specified the extension

$$\delta_i^{(k)} := \begin{cases} (\lambda_i^k - 1)/k & \text{if } k \neq 0 \\ \ln \lambda_i & \text{if } k = 0 \end{cases}, \forall\, k \in \mathbb{R}, \tag{7.22}$$

usually called the **Doyle-Ericksen extension**, which is valid because $\delta_i^{(k)} = 0$ for rigid deformations. Note that $2\delta_i^{(2)} = \tilde{\delta}_i$ and $2\delta_i^{(-2)} = \tilde{\Delta}_i$, leading to the Green-St Venant and Almansi-Hamel strain tensors when $\mathbf{U}_{\overline{u}}$ and $\mathbf{V}_{\overline{u}}$ are respectively considered, as we have done above. Table 7.1 lists the most common strain tensors available, called the **Doyle-Ericksen tensors**. For $k \neq 0$, they are based on the generic strain tensor

$$\mathbf{E}_{\mathbf{S}}^{(k)} := \frac{1}{k}(\mathbf{S}^k - \mathbf{I}) \tag{7.23}$$

where tensor $\mathbf{S}$ is a positive-definite symmetric tensor equal to or defined from the stretch tensors. Professor HILL[30] generalized the proposition of SETH[54] even further by defining an extension $f(\lambda_i)$, where mapping $f : \mathbb{R}_*^+ \mapsto \mathbb{R}$ is smooth and monotonically increasing, that is, $f'(x) > 0$ for all $x > 0$. Moreover, function f must obviously observe $f(1) = 0$ and, in order to make all strain measures equivalent in deformations almost rigid or small, he imposed $f'(1) = 1$.

Now, in order to describe mathematically the context of **small or infinitesimal deformations** cited above, we need first to define a bijective Lagrangian mapping $\boldsymbol{u}_t : \overline{\mathfrak{B}} \mapsto U_{\mathbb{R}}$ where

$$\boldsymbol{u}_t(\overline{x}) = \boldsymbol{\varphi}_t(\overline{x}) - \overline{x}. \tag{7.24}$$

The function $\boldsymbol{u}_t$ is called **displacement function** whose value $\boldsymbol{u}_t(\overline{u})$ is said to be the **displacement vector** of an arbitrary material point $\overline{u}$ at instant t. From the previous rule, it is straightforward to obtain that tensor

$$\nabla \boldsymbol{u}_t(\overline{u}) = \mathbf{F}_{\overline{u}}^T - \mathbf{I}, \tag{7.25}$$

k	S	Doyle-Ericksen Tensor	Name
-2	$\mathbf{V}_{\overline{u}}$	$\mathsf{E}^{(-2)}_{\mathbf{V}_{\overline{u}}} = \boldsymbol{e}_{\overline{u}} = (\mathbf{I} - \mathbf{B}^{-1}_{\overline{u}})/2$	Almansi-Hamel
-2	$\mathbf{V}^{-1}_{\overline{u}}$	$\mathsf{E}^{(-2)}_{\mathbf{V}^{-1}_{\overline{u}}} = (\mathbf{I} - \mathbf{B}_{\overline{u}})/2$	Finger
-1	$\mathbf{V}_{\overline{u}}$	$\mathsf{E}^{(-1)}_{\mathbf{V}_{\overline{u}}} = \mathbf{I} - \mathbf{V}^{-1}_{\overline{u}}$	Swainger
0	-	$\mathsf{E}^{(0)} = \sum_{i=1}^{3} \ln \lambda_i \hat{x}_i \otimes \hat{x}_i$	Henky
1	$\mathbf{U}_{\overline{u}}$	$\mathsf{E}^{(1)}_{\mathbf{U}_{\overline{u}}} = \mathbf{U}_{\overline{u}} - \mathbf{I}$	Biot
2	$\mathbf{U}^{-1}_{\overline{u}}$	$\mathsf{E}^{(2)}_{\mathbf{U}^{-1}_{\overline{u}}} = (\mathbf{C}^{-1}_{\overline{u}} - \mathbf{I})/2$	Piola
2	$\mathbf{U}_{\overline{u}}$	$\mathsf{E}^{(2)}_{\mathbf{U}_{\overline{u}}} = \mathsf{E}_{\overline{u}} = (\mathbf{C}_{\overline{u}} - \mathbf{I})/2$	Green-St Venant

Table 7.1 – Examples of Doyle-Ericksen Tensors.

called the **displacement gradient** of point $\overline{u}$ at instant t. Thereby, the rule of the inverse of the displacement function and its corresponding gradient are respectively

$$\boldsymbol{u}_t^{-1}(x) = x - \boldsymbol{\varphi}_t^{-1}(x) \qquad \text{and} \qquad \nabla \boldsymbol{u}_t^{-1}(u) = \mathbf{I} - \mathsf{F}_{\overline{u}}^{-\mathrm{T}}. \tag{7.26}$$

As a consequence, the following equalities can be obtained:

$$\begin{aligned}
\mathbf{C}_{\overline{u}} &= \nabla \boldsymbol{u}_t(\overline{u}) \odot_1 \nabla \boldsymbol{u}_t(\overline{u})^{\mathrm{T}} + \nabla \boldsymbol{u}_t(\overline{u}) + \nabla \boldsymbol{u}_t(\overline{u})^{\mathrm{T}} + \mathbf{I};\\
\mathbf{B}_{\overline{u}} &= \nabla \boldsymbol{u}_t(\overline{u})^{\mathrm{T}} \odot_1 \nabla \boldsymbol{u}_t(\overline{u}) + \nabla \boldsymbol{u}_t(\overline{u}) + \nabla \boldsymbol{u}_t(\overline{u})^{\mathrm{T}} + \mathbf{I};\\
\mathbf{C}^{-1}_{\overline{u}} &= \nabla \boldsymbol{u}_t^{-1}(u)^{\mathrm{T}} \odot_1 \nabla \boldsymbol{u}_t^{-1}(u) - \nabla \boldsymbol{u}_t^{-1}(u) - \nabla \boldsymbol{u}_t^{-1}(u)^{\mathrm{T}} + \mathbf{I};\\
\mathbf{B}^{-1}_{\overline{u}} &= \nabla \boldsymbol{u}_t^{-1}(u) \odot_1 \nabla \boldsymbol{u}_t^{-1}(u)^{\mathrm{T}} - \nabla \boldsymbol{u}_t^{-1}(u) - \nabla \boldsymbol{u}_t^{-1}(u)^{\mathrm{T}} + \mathbf{I}.
\end{aligned} \tag{7.27}$$

Proof. First, we verify the second expression of (7.26) from the definition of affine deformation gradient and equality (7.9), from which we conclude that $\mathsf{F}_{\overline{u}}^{-\mathrm{T}} = \nabla \boldsymbol{\varphi}_t^{-1}(u)$. Now, considering (7.17), proof of the above equalities is trivial by isolating $\mathsf{F}_{\overline{u}}^{\mathrm{T}}$, $\mathsf{F}_{\overline{u}}^{-1}$ and $\mathsf{F}_{\overline{u}}^{-\mathrm{T}}$ in (7.25) and (7.26). □

A deformation is called small when $\|\nabla \boldsymbol{u}_t(\overline{u})\| = \|\mathsf{F}_{\overline{u}}^{\mathrm{T}} - \mathbf{I}\| \ll 1$, which leads to $\mathsf{F}_{\overline{u}} \approx \mathbf{I}$ and then, from (7.26), we can conclude that $\nabla \boldsymbol{u}_t^{-1}(u) \approx \nabla \boldsymbol{u}_t(\overline{u})$. Moreover, concerning the equalities above, the four nonlinear terms involving partial inner products of two small displacement gradients become negligible, from which equality $\mathsf{C}_{\overline{\mathrm{u}}} = \mathsf{B}_{\overline{\mathrm{u}}}$ can be considered valid. But since $\nabla \boldsymbol{u}_{\mathrm{t}}^{-1}(\mathbf{u}) \approx \nabla \boldsymbol{u}_{\mathrm{t}}(\overline{\mathbf{u}})$, we can also conclude that $\mathsf{C}_{\overline{\mathrm{u}}} \approx \mathsf{C}_{\overline{\mathrm{u}}}^{-1}$ and $\mathsf{B}_{\overline{\mathrm{u}}} \approx \mathsf{B}_{\overline{\mathrm{u}}}^{-1}$. Therefore, concerning the Doyle–Ericksen tensors of table 7.1, we have

$$\mathsf{E}_{\overline{\mathrm{u}}} \approx \boldsymbol{e}_{\overline{\mathrm{u}}} \approx \mathsf{E}^{(2)}_{\mathsf{U}_{\overline{\mathrm{u}}}^{-1}} \approx \mathsf{E}^{(-2)}_{\mathsf{V}_{\overline{\mathrm{u}}}}$$

in the context of small deformations. If translations are disregarded or simply not present in the affine deformation rule, another consequence of $\mathsf{F}_{\overline{\mathrm{u}}} \approx \mathbf{I}$ is that the shape $\overline{\mathfrak{B}} \approx \mathfrak{B}_{\mathrm{t}}$. Therefore, we can assume that $\overline{\mathfrak{B}}$ is the locus where small deformations take place and tensor

$$\boldsymbol{\epsilon}_{\overline{\mathrm{u}}} := (\nabla \boldsymbol{u}_{\mathrm{t}}(\overline{\mathbf{u}}) + \nabla \boldsymbol{u}_{\mathrm{t}}(\overline{\mathbf{u}})^{\mathrm{T}})/2 \tag{7.28}$$

is their strain measure, called **infinitesimal strain tensor**, which is clearly taken from the Green–St Venant tensor written in terms of displacement gradients after neglecting the very small nonlinear term. In those cases where these nonlinear terms cannot be neglected, that is, cases where small deformations are not valid, the context under study is said to be of **large or finite deformations**. From the above definition, it is evident that the matrix representation of $\boldsymbol{\epsilon}_{\overline{\mathrm{u}}}$, which measures length changes in infinitesimal deformations, is the symmetric part of matrix $[\nabla \boldsymbol{u}_{\mathrm{t}}(\overline{\mathbf{u}})]$. But if the symmetric part of $[\nabla \boldsymbol{u}_{\mathrm{t}}(\overline{\mathbf{u}})]$ refers to the length changing portion of infinitesimal deformation, we can state that its antisymmetric part $[\boldsymbol{W}_{\overline{\mathrm{u}}}]$, where

$$\boldsymbol{W}_{\overline{\mathrm{u}}} := (\nabla \boldsymbol{u}_{\mathrm{t}}(\overline{\mathbf{u}}) - \nabla \boldsymbol{u}_{\mathrm{t}}(\overline{\mathbf{u}})^{\mathrm{T}})/2, \tag{7.29}$$

refers to the rotation portion[5]. Speaking of antisymmetric tensors of second order, we recall that theorem 19 establishes a biunivocal relationship between them and three dimensional Euclidean vectors. Therefore, considering the conditions of this theorem, the axial vector defined by $\boldsymbol{W}_{\overline{\mathrm{u}}}$ is

$$\boldsymbol{w}_{\overline{\mathrm{u}}} := \mathsf{A}_{\mathrm{B}} \hat{\odot}_2 \boldsymbol{W}_{\overline{\mathrm{u}}}. \tag{7.30}$$

7.3 Motion

In this section, time is also considered variable in the mathematical study of motion, when new fundamental concepts arise. But firstly, as we did in the previous section for a certain time t, it is convenient to fix an arbitrary particle $\overline{\mathbf{u}} \in \overline{\mathfrak{B}}$ and then define a surjective mapping $\boldsymbol{\chi}_{\overline{\mathrm{u}}} : \mathbb{T} \mapsto \mathfrak{T}_{\overline{\mathrm{u}}}$ where

$$\boldsymbol{\chi}_{\overline{\mathrm{u}}}(\mathrm{t}) = \boldsymbol{\chi}(\overline{\mathbf{u}}, \mathrm{t}), \forall \mathrm{t} \in \mathbb{T}. \tag{7.31}$$

[5]In Continuum Mechanics literature, $\boldsymbol{W}_{\overline{\mathrm{u}}}$ is commonly called the infinitesimal rotation tensor, but we shall not use this nomenclature here in order to avoid confusion.

Moreover, the univariate motion $\boldsymbol{\chi}_{\overline{\mathbf{u}}}$, which results a $\mathcal{C}^3$ surjection, must also observe restriction $\boldsymbol{\chi}_{\overline{\mathbf{u}}}(\overline{t}) = \overline{\mathbf{u}}$. Therefore, $\boldsymbol{\chi}_{\overline{\mathbf{u}}}$ describes the temporal deformation of particle $\overline{\mathbf{u}}$ and, given a period $\mathbb{P} \subset \mathbb{T}$, the collection of places

$$T_{\overline{\mathbf{u}}} := \{\boldsymbol{\chi}_{\overline{\mathbf{u}}}(t) : t \in \mathbb{P}\} \tag{7.32}$$

is called the **trajectory** or **pathline** of $\overline{\mathbf{u}}$ during $\mathbb{P}$. Additionally to this collection of places, it is also important to define the collection of particles that "occupy" a fixed place during a period of time. In order to do this, from the codomain of mapping (7.4), we define the surjective mapping $\mathbf{p} : \mathcal{S} \times \mathbb{T} \mapsto \overline{\mathcal{B}}$, where $\mathbf{p}$ is called **streak function** if it is a $\mathcal{C}^3$ surjection whith rule

$$\mathbf{p}(\mathbf{x}, t) = \boldsymbol{\chi}(\boldsymbol{\chi}_t^{-1}(\mathbf{x}), t) . \tag{7.33}$$

Following the same procedure adopted for deformations, by fixing a position $\mathbf{u}$ we define univariate surjection $\mathbf{p}_{\mathbf{u}}$, where $\mathbf{p}_{\mathbf{u}}(t) = \mathbf{p}(\mathbf{u}, t)$. In this context, the collection of particles

$$S_{\mathbf{u}} := \{\mathbf{p}_{\mathbf{u}}(t) : t \in \mathbb{P}\} \tag{7.34}$$

is said to be the **streakline** of place $\mathbf{u}$ during $\mathbb{P}$.

The functions in mappings $\boldsymbol{\omega} : \overline{\mathcal{B}} \times \mathbb{T} \mapsto W$ and $\boldsymbol{\Omega} : \mathcal{B}_t \times \mathbb{T} \mapsto W$, where codomain $W \in \{\mathcal{L}_{\mathbb{R}}(U^m), U_{\mathbb{R}}, \mathbb{R}\}$ are specified to be continuously differentiable. In our study, they are called tensor, vector or scalar **distributions**[6] according to W. In order to simplify notation, it is sometimes very useful to also define

$$\boldsymbol{\Omega}_L(\overline{\mathbf{x}}, t) = \boldsymbol{\Omega}(\boldsymbol{\chi}_t(\overline{\mathbf{x}}), t) \qquad \text{and} \qquad \boldsymbol{\omega}_E(\mathbf{x}, t) = \boldsymbol{\omega}(\boldsymbol{\chi}_t^{-1}(\mathbf{x}), t) , \tag{7.35}$$

commonly called the Lagrangian and Eulerian descriptions of $\boldsymbol{\Omega}$ and $\boldsymbol{\omega}$. In genereal terms, notations $\bullet_L$ and $\bullet_E$ can be used on any Eulerian and Lagrangian function. It is then straightforward to conclude that that streak function is the Eulerian description of the deformation, that is, $\mathbf{p} = \boldsymbol{\chi}_E$. Generically, since each of the distributions $\boldsymbol{\Omega}$ and $\boldsymbol{\omega}$ can be described by two domains, the differentiation chain rules force us to clearly specify the domain of derivation, namely, whether these functions will be derived on the material or spatial domain. Therefore, derivatives here are called Lagrangian or Eulerian when performed on particles or places respectively. In this context, *it is specified that the time derivative of* $\bullet$ *is said to be Lagrangian, represented by* $\bar{\partial}_t \bullet$*, or Eulerian, represented by* $\tilde{\partial}_t \bullet$*, when respectively a particle or a place are fixed*. Whenever possible, notations $\bar{\partial}_t^k \bullet$ and $\tilde{\partial}_t^k \bullet$ define Lagrangian and Eulerian time derivatives of order $k \geqslant 2$, while time derivatives $d\bullet/dt$ and $d^2\bullet/dt^2$ of univariate functions of time are simplified by overdotted symbols $\dot{\bullet}$ and $\ddot{\bullet}$ respectively. Concerning distributions $\boldsymbol{\omega}$ and $\boldsymbol{\Omega}$, we have the Lagrangian time derivatives at particle $\overline{\mathbf{u}}$

$$\bar{\partial}_t \boldsymbol{\omega}(\overline{\mathbf{u}}, t) := \frac{d\boldsymbol{\omega}_{\overline{\mathbf{u}}}}{dt}(t) = \dot{\boldsymbol{\omega}}_{\overline{\mathbf{u}}}(t) \qquad \text{and} \qquad \bar{\partial}_t \boldsymbol{\Omega}(\mathbf{u}, t) := \boldsymbol{\Omega}'(\boldsymbol{\chi}_{\overline{\mathbf{u}}}(t), t) \tag{7.36}$$

[6]In Continuum Mechanics classical literature, functions like $\boldsymbol{\omega}$ and $\boldsymbol{\Omega}$ are usually called fields, but we shall not use this nomenclature here in order to avoid confusion.

as well as the Eulerian time derivatives at place $\mathbf{u}$

$$\tilde{\partial}_t \boldsymbol{\Omega}(\mathbf{u},t) := \frac{\mathrm{d}\boldsymbol{\Omega}_{\mathbf{u}}}{\mathrm{d}t}(t) = \dot{\boldsymbol{\Omega}}_{\mathbf{u}}(t) \qquad \text{and} \qquad \tilde{\partial}_t \boldsymbol{\omega}(\overline{\mathbf{u}},t) := \boldsymbol{\omega}'(\mathbf{p}_{\mathbf{u}}(t),t)\,, \tag{7.37}$$

Considering these previous definitions and the Chain Rule for Partial Derivatives (5.16), it is possible to obtain the Lagrangian time derivative

$$\bar{\partial}_t \boldsymbol{\Omega}(\mathbf{x},t) = \dot{\boldsymbol{\Omega}}_{\mathbf{x}}(t) + \dot{\boldsymbol{\Omega}}_t(\boldsymbol{\chi}_{\overline{\mathbf{x}}}(t)) = \tilde{\partial}_t \boldsymbol{\Omega}(\mathbf{x},t) + \boldsymbol{\Omega}_t'(\mathbf{x}) \circ \dot{\boldsymbol{\chi}}_{\overline{\mathbf{x}}}(t) \tag{7.38}$$

and Eulerian time derivative

$$\tilde{\partial}_t \boldsymbol{\omega}(\overline{\mathbf{x}},t) = \dot{\boldsymbol{\omega}}_{\overline{\mathbf{x}}}(t) + \dot{\boldsymbol{\omega}}_t(\mathbf{p}_{\mathbf{x}}(t)) = \bar{\partial}_t \boldsymbol{\omega}(\overline{\mathbf{x}},t) + \boldsymbol{\omega}_t'(\overline{\mathbf{x}}) \circ \dot{\mathbf{p}}_{\mathbf{x}}(t)\,. \tag{7.39}$$

Moreover, in the particular cases of $\boldsymbol{\omega} = \boldsymbol{\Omega}_{\mathrm{L}}$ and $\boldsymbol{\Omega} = \boldsymbol{\omega}_{\mathrm{E}}$,

$$[\dot{\boldsymbol{\Omega}}_t]_{\mathrm{L}} + [\dot{\boldsymbol{\Omega}_{\mathrm{L}}}]_t = 0 \qquad \text{and} \qquad [\dot{\boldsymbol{\omega}}_t]_{\mathrm{E}} + [\dot{\boldsymbol{\omega}_{\mathrm{E}}}]_t = 0\,. \tag{7.40}$$

The specification of the codomains of distributions $\boldsymbol{\omega}$ and $\boldsymbol{\Omega}$ leads to the following important conclusions, where Lagrangian and Eulerian gradients, related to material and spatial derivatives, are represented by symbols $\bar{\bullet}$ and $\tilde{\bullet}$ respectively.

i. If $W = \mathcal{L}_{\mathbb{R}}(\mathrm{U}^m)$, distributions are tensor valued and then, from (5.24),

$$\begin{aligned} \dot{\boldsymbol{\Omega}}_t(\boldsymbol{\chi}_{\overline{\mathbf{x}}}(t)) &= \boldsymbol{\Omega}_t'(\mathbf{x}) \circ \dot{\boldsymbol{\chi}}_{\overline{\mathbf{x}}}(t) = [\dot{\boldsymbol{\chi}}_{\overline{\mathbf{x}}}(t)]^* \odot_1 \tilde{\nabla}\boldsymbol{\Omega}_t(\mathbf{x})\,, \\ \dot{\boldsymbol{\omega}}_t(\mathbf{p}_{\mathbf{x}}(t)) &= \boldsymbol{\omega}_t'(\overline{\mathbf{x}}) \circ \dot{\mathbf{p}}_{\mathbf{x}}(t) = [\dot{\mathbf{p}}_{\mathbf{x}}(t)]^* \odot_1 \bar{\nabla}\boldsymbol{\omega}_t(\overline{\mathbf{x}})\,; \end{aligned} \tag{7.41}$$

ii. If $W = \mathrm{U}_{\mathbb{R}}$, distributions are vector valued and then, from (5.29),

$$\begin{aligned} \dot{\boldsymbol{\Omega}}_t(\boldsymbol{\chi}_{\overline{\mathbf{x}}}(t)) &= \boldsymbol{\Omega}_t'(\mathbf{x}) \circ \dot{\boldsymbol{\chi}}_{\overline{\mathbf{x}}}(t) = \tilde{\nabla}\boldsymbol{\Omega}_t(\mathbf{x})^{\mathrm{T}} \hat{\odot}_1 [\dot{\boldsymbol{\chi}}_{\overline{\mathbf{x}}}(t)]^*\,, \\ \dot{\boldsymbol{\omega}}_t(\mathbf{p}_{\mathbf{x}}(t)) &= \boldsymbol{\omega}_t'(\overline{\mathbf{x}}) \circ \dot{\mathbf{p}}_{\mathbf{x}}(t) = \bar{\nabla}\boldsymbol{\omega}_t(\overline{\mathbf{x}})^{\mathrm{T}} \hat{\odot}_1 [\dot{\mathbf{p}}_{\mathbf{x}}(t)]^*\,; \end{aligned} \tag{7.42}$$

iii. If $W = \mathbb{R}$, distributions are scalar valued and then, from (5.34),

$$\begin{aligned} \dot{\boldsymbol{\Omega}}_t(\boldsymbol{\chi}_{\overline{\mathbf{x}}}(t)) &= \boldsymbol{\Omega}_t'(\mathbf{x}) \circ \dot{\boldsymbol{\chi}}_{\overline{\mathbf{x}}}(t) = \widetilde{\mathrm{grad}}\,\boldsymbol{\Omega}_t(\mathbf{x}) \cdot \dot{\boldsymbol{\chi}}_{\overline{\mathbf{x}}}(t)\,, \\ \dot{\boldsymbol{\omega}}_t(\mathbf{p}_{\mathbf{x}}(t)) &= \boldsymbol{\omega}_t'(\overline{\mathbf{x}}) \circ \dot{\mathbf{p}}_{\mathbf{x}}(t) = \overline{\mathrm{grad}}\,\boldsymbol{\omega}_t(\overline{\mathbf{x}}) \cdot \dot{\mathbf{p}}_{\mathbf{x}}(t)\,. \end{aligned} \tag{7.43}$$

For future purposes, it is now important to present properties involving Lagrangian and Eulerian gradients of respectively Eulerian and Lagrangian functions. The following equalities are straightforward consequences of property (5.41), valid for tensor valued composite functions, which can be easily extended for vector and scalar valued composite functions:

$$\begin{aligned} \bar{\nabla}[\boldsymbol{\Omega}_t]_{\mathrm{L}}(\overline{\mathbf{x}}) &= \bar{\nabla}\boldsymbol{\Omega}_t(\boldsymbol{\chi}_t(\overline{\mathbf{x}})) = \bar{\nabla}\boldsymbol{\chi}_t(\overline{\mathbf{x}}) \odot_1 \tilde{\nabla}[\boldsymbol{\Omega}_t]_{\mathrm{L}}(\overline{\mathbf{x}})\,; \\ \tilde{\nabla}[\boldsymbol{\omega}_t]_{\mathrm{E}}(\mathbf{x}) &= \tilde{\nabla}\boldsymbol{\omega}_t(\mathbf{p}_t(\mathbf{x})) = \tilde{\nabla}\mathbf{p}_t(\mathbf{x}) \odot_1 \bar{\nabla}[\boldsymbol{\omega}_t]_{\mathrm{E}}(\mathbf{x})\,. \end{aligned} \tag{7.44}$$

Proof. From the chain rules (5.7) and (5.16), the following development proves (7.38); equality (7.39) is similarly verified.

$$\begin{aligned}\bar{\partial}_t\boldsymbol{\Omega}(x,t) &= \boldsymbol{\Omega}'(\boldsymbol{\chi}_{\overline{x}}(t),t)\\ &= \dot{\boldsymbol{\Omega}}_x(t)+\dot{\boldsymbol{\Omega}}_t(\boldsymbol{\chi}_{\overline{x}}(t))\\ &= \tilde{\partial}_t\boldsymbol{\Omega}(x,t)+[\boldsymbol{\Omega}_t'\circ\boldsymbol{\chi}_{\overline{x}}(t)]\circ\dot{\boldsymbol{\chi}}_{\overline{x}}(t)\\ &= \tilde{\partial}_t\boldsymbol{\Omega}(x,t)+\boldsymbol{\Omega}_t'(x)\circ\dot{\boldsymbol{\chi}}_{\overline{x}}(t).\end{aligned}$$

Now, we prove the first of properties (7.40); the other is similarly verified. From (7.39), the sum of the last equality of development

$$\begin{gathered}\tilde{\partial}_t\boldsymbol{\Omega}_L(\overline{x},t) = [\dot{\boldsymbol{\Omega}}_L]_{\overline{x}}(t)+[\dot{\boldsymbol{\Omega}}_L]_t(\boldsymbol{\chi}_x(t))\\ \dot{\boldsymbol{\Omega}}_x(t) = \bar{\partial}_t\boldsymbol{\Omega}(x,t)+[\dot{\boldsymbol{\Omega}}_L]_t(\overline{x})\end{gathered}$$

with (7.38) proves the property. □

The vector $\dot{\boldsymbol{\chi}}_{\overline{u}}(t)$, which measures the temporal sensitivity of the deformation at a fixed particle $\overline{u}$, is called the **velocity** of $\overline{u}$ at t. If the particle is variable, the Lagrangian distribution in mapping $\boldsymbol{v}:\overline{\mathcal{B}}\times\mathbb{T}\mapsto U_{\mathbb{R}}$ is called velocity if it is $\mathcal{C}^2$ and vector

$$\boldsymbol{v}(\overline{x},t) = \bar{\partial}_t\boldsymbol{\chi}(\overline{x},t). \tag{7.45}$$

In the context of affine deformations, if instant t is variable for displacements described by (7.24), that is, $\boldsymbol{u}(\overline{x},t)=\boldsymbol{\varphi}(\overline{x},t)-\overline{x}$, then it is straightforward to obtain

$$\bar{\partial}_t\boldsymbol{u}(\overline{x},t)=\boldsymbol{v}(\overline{x},t), \tag{7.46}$$

which is obviously a consequence, not a definition, of velocity. Now, the temporal sensitivity of the velocity of particle $\overline{u}$ at instant t is the vector $\dot{\boldsymbol{v}}_{\overline{u}}(t)=\ddot{\boldsymbol{\chi}}_{\overline{u}}(t)$, called the **acceleration** of $\overline{u}$ at t. The distribution in mapping $\boldsymbol{a}:\overline{\mathcal{B}}\times\mathbb{T}\mapsto U_{\mathbb{R}}$ is called acceleration if it is continuously differentiable and

$$\boldsymbol{a}(\overline{x},t) = \bar{\partial}_t\boldsymbol{v}(\overline{x},t), \tag{7.47}$$

from which the following equality is evident:

$$\bar{\partial}_t^2\boldsymbol{u}(\overline{x},t)=\boldsymbol{a}(\overline{x},t). \tag{7.48}$$

Considering property (7.39), we can conclude that vector $\dot{\mathbf{p}}_u(t)$ measures the temporal sensitivity of the Eulerian description $\boldsymbol{\chi}_E$ of place u at instant t. In this context, equalities

$$\dot{\mathbf{p}}_x(t)=\tilde{\partial}_t\mathbf{p}(x,t)=\tilde{\partial}_t\mathbf{p}_L(\overline{x},t) \qquad \text{and} \qquad \dot{\boldsymbol{\chi}}_{\overline{x}}(t)=\boldsymbol{v}(\overline{x},t)=\boldsymbol{v}_E(x,t)$$

enable us to write from (7.38) and (7.39) the following expressions, solely described by places and particles respectively:

$$\bar{\partial}_t\boldsymbol{\Omega}(x,t)=\tilde{\partial}_t\boldsymbol{\Omega}(x,t)+\boldsymbol{\Omega}_t'(x)\circ\boldsymbol{v}_E(x,t) \tag{7.49}$$

and

$$\tilde{\partial}_t \boldsymbol{\omega}(\overline{x},t) = \bar{\partial}_t \boldsymbol{\omega}(\overline{x},t) + \boldsymbol{\omega}'_t(\overline{x}) \circ \tilde{\partial}_t \mathbf{p}_L(\overline{x},t)\,. \tag{7.50}$$

By definition, acceleration is the Lagrangian time derivative of the velocity $\mathbf{v}$, but what about its Eulerian time derivative? From the second of the previous qualities and expressions (7.42), it is clear that the vector

$$\tilde{\partial}_t \mathbf{v}(\overline{x},t) = \mathbf{a}(\overline{x},t) + \bar{\nabla}\mathbf{v}_t(\overline{x})^T \hat{\odot}_1 [\tilde{\partial}_t \mathbf{p}_L(\overline{x},t)]^*\,. \tag{7.51}$$

Similarly, the Lagrangian time derivative of the Eulerian description $\mathbf{v}_E$ of the velocity results the Eulerian description

$$\mathbf{a}_E(x,t) = \tilde{\partial}_t \mathbf{v}_E(x,t) + \tilde{\nabla}[\mathbf{v}_t]_E(x)^T \hat{\odot}_1 [\mathbf{v}_E(x,t)]^*\,, \tag{7.52}$$

from which second order tensor

$$\mathbf{L}_{\mathbf{u}} := \tilde{\nabla}[\mathbf{v}_t]_E(\mathbf{u})^T \tag{7.53}$$

is said to be the **Eulerian velocity gradient** of $\mathbf{v}_E$ at place $\mathbf{u}$, while

$$\mathbf{M}_{\overline{\mathbf{u}}} := \bar{\nabla}\mathbf{v}_t(\overline{\mathbf{u}})^T \tag{7.54}$$

is the **Lagrangian velocity gradient** of $\mathbf{v}$ at particle $\overline{\mathbf{u}}$. From the Eulerian description of the velocity, a motion of the body $\overline{\mathcal{B}}$ is said to be **steady** in a period $[t_0, t_f]$ if both $\mathbf{v}_E$ and $\mathcal{B}_t$ are time independent, that is, if

$$\tilde{\partial}_t \mathbf{v}_E(\mathbf{u},t) = \mathbf{0} \qquad \text{and} \qquad \mathcal{B}_t = \mathcal{B}_{t_0}\,, \tag{7.55}$$

for all $\mathbf{u} \in \mathcal{B}_{t_0}$ and $t \in [t_0, t_f]$. In this context, we can conclude that every streakline $S_{\mathbf{u}}$ is constituted by particles that have the same velocity $\mathbf{v}$ and the distinct places of every trajectory $T_{\overline{w}}$ constitute a subset of shape $\mathcal{B}_{t_0}$ as shown schematically in figure 7.5.

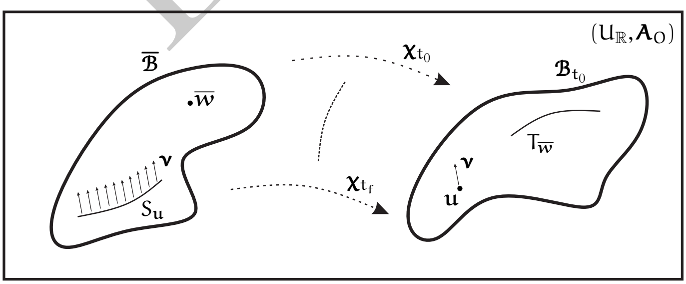

Figure 7.5 – Steady motion.

For the next type of motion, we define the mapping $j : \overline{\mathcal{B}} \times \mathbb{T} \mapsto \mathbb{R}$ whose function is described by the rule $j(\overline{x},t) = \mathrm{Det}[\nabla \boldsymbol{\chi}_t(\overline{x})]$. From this definition, a motion is said to be **isochoric** in a period of time $[t_0, t_f]$ if

$$\bar{\partial}_t j(\overline{\mathbf{u}},t) = 0\,, \tag{7.56}$$

for all $\overline{\mathrm{u}} \in \overline{\mathcal{B}}$ and $\mathrm{t} \in [\mathrm{t}_0, \mathrm{t}_\mathrm{f}]$. From the Chain Rule (5.7), derivative (5.11) and property (5.48), the Lagrangian time derivative on the left hand side can be developed as follows:

$$\begin{aligned}
\bar{\partial}_\mathrm{t} j(\overline{\mathrm{u}}, \mathrm{t}) &= \bar{\partial}_\mathrm{t}\, \mathrm{Det} \circ \nabla \boldsymbol{\chi}_\mathrm{t}(\overline{\mathrm{u}}) \\
&= \mathrm{Det}[\nabla \boldsymbol{\chi}_\mathrm{t}(\overline{\mathrm{u}})] \operatorname{tr}(\nabla \boldsymbol{\chi}_\mathrm{t}(\overline{\mathrm{u}})^{-1} \odot_1 \bar{\partial}_\mathrm{t} \nabla \boldsymbol{\chi}_\mathrm{t}(\overline{\mathrm{u}})) \\
&= \mathrm{Det}[\nabla \boldsymbol{\chi}_\mathrm{t}(\overline{\mathrm{u}})] \operatorname{tr}(\nabla \boldsymbol{\chi}_\mathrm{t}(\overline{\mathrm{u}})^{-1} \odot_1 \nabla_{\overline{\mathrm{u}}} \bar{\partial}_\mathrm{t} \boldsymbol{\chi}(\overline{\mathrm{u}}, \mathrm{t})) \\
&= \mathrm{Det}[\nabla \boldsymbol{\chi}_\mathrm{t}(\overline{\mathrm{u}})] \operatorname{tr}(\nabla \boldsymbol{\chi}_\mathrm{t}(\overline{\mathrm{u}})^{-1} \odot_1 \mathbf{M}_{\overline{\mathrm{u}}}^\mathrm{T}),
\end{aligned} \tag{7.57}$$

from which we conclude that in isochoric motions, $\operatorname{tr}(\nabla \boldsymbol{\chi}_\mathrm{t}(\overline{\mathrm{u}})^{-1} \odot_1 \mathbf{M}_{\overline{\mathrm{u}}}^\mathrm{T}) = 0$ for all $\overline{\mathrm{u}} \in \overline{\mathcal{B}}$ and $\mathrm{t} \in [\mathrm{t}_0, \mathrm{t}_\mathrm{f}]$, due to inequality (7.10).

Considering definitions (7.53) and (7.54) in the context of affine motions, the following important properties can be obtained:

$$\begin{aligned}
\mathbf{M}_{\overline{\mathrm{x}}} &= \mathbf{L}_{\boldsymbol{\chi}_\mathrm{t}(\overline{\mathrm{x}})} \odot_1 \mathbf{F}_{\overline{\mathrm{x}}}; \\
\bar{\nabla} \boldsymbol{a}_\mathrm{t}(\overline{\mathrm{x}}) &= \tilde{\nabla} \boldsymbol{a}_\mathrm{t}(\overline{\mathrm{x}}) \odot_1 \mathbf{F}_{\overline{\mathrm{x}}}.
\end{aligned} \tag{7.58}$$

Based on the tensors $\boldsymbol{\epsilon}_{\overline{\mathrm{u}}}$ and $\boldsymbol{W}_{\overline{\mathrm{u}}}$ defined by (7.28) and (7.29) in the context of infinitesimal deformations, whose locus was chosen to be $\overline{\mathcal{B}}$, since time derivative $\bar{\partial}_\mathrm{t} \boldsymbol{u} = \mathbf{v}$, symmetric tensor

$$\dot{\boldsymbol{\epsilon}}_{\overline{\mathrm{u}}} := (\mathbf{M}_{\overline{\mathrm{u}}}^\mathrm{T} + \mathbf{M}_{\overline{\mathrm{u}}})/2 \tag{7.59}$$

is called the **strain-rate tensor** at $\overline{\mathrm{u}}$, which measures the rate of local changes in length, and antisymmetric tensor

$$\dot{\boldsymbol{W}}_{\overline{\mathrm{u}}} := (\mathbf{M}_{\overline{\mathrm{u}}}^\mathrm{T} - \mathbf{M}_{\overline{\mathrm{u}}})/2 \tag{7.60}$$

is the **spin tensor** at $\overline{\mathrm{u}}$, which measures the rate of local rotation. Moreover, from (7.58), the context of infinitesimal deformations, where $\mathbf{F}_{\overline{\mathrm{x}}} \approx \mathbf{I}$, leads to the conclusion that $\mathbf{M}_{\overline{\mathrm{x}}} \approx \mathbf{L}_\mathrm{x}$. The neglecting conditions for the above definitions, which are the same for (7.28) and (7.29), allow us to assume that $\mathbf{M}_{\overline{\mathrm{x}}} = \mathbf{L}_\mathrm{x}$, from which the following rate tensors can be defined:

$$\dot{\boldsymbol{\epsilon}}_\mathrm{u} := (\mathbf{L}_\mathrm{u}^\mathrm{T} + \mathbf{L}_\mathrm{u})/2 = \dot{\boldsymbol{\epsilon}}_{\overline{\mathrm{u}}} \qquad \text{and} \qquad \dot{\boldsymbol{W}}_\mathrm{u} := (\mathbf{L}_\mathrm{u}^\mathrm{T} - \mathbf{L}_\mathrm{u})/2 = \dot{\boldsymbol{W}}_{\overline{\mathrm{u}}}, \tag{7.61}$$

from which matrices $[\dot{\boldsymbol{\epsilon}}_\mathrm{u}]$ and $[\dot{\boldsymbol{W}}_\mathrm{u}]$ are respectively the symmetric and antisymmetric parts of $[\mathbf{L}_\mathrm{u}]$. In the classical Continuum Mechanics literature, both of these Eulerian measures are more commonly used than their Lagrangian counterparts. In this sense, according to theorem 19, since $\dot{\boldsymbol{W}}_\mathrm{u}$ is antisymmetric, its correspondent axial vector

$$\dot{\boldsymbol{w}}_\mathrm{u} := \mathbf{A}_\mathrm{B} \hat{\odot}_2 \dot{\boldsymbol{W}}_\mathrm{u} \tag{7.62}$$

is called **vorticity vector**, from which it is possible to obtain that

$$\dot{\boldsymbol{w}}_\mathrm{u} = \tilde{\nabla} \times [\mathbf{v}_\mathrm{t}]_\mathrm{E}(\mathbf{u}) = \bar{\nabla} \times \mathbf{v}_\mathrm{t}(\overline{\mathrm{u}}). \tag{7.63}$$

If the spin tensor is zero at $\mathbf{u}$, the vorticity vector is consequently zero at this place, when the motion is said to be **irrotational** or **non-vortical** at $\mathbf{u}$. In this context, from the definition of $\dot{\mathbf{W}}_{\mathbf{u}}$, it is obvious that the Eulerian velocity gradient $\mathbf{L}_{\mathbf{u}}$ results symmetric and therefore no rotation is present since the antisymmetric part of matrix $[\mathbf{L}_{\mathbf{u}}]$ is zero. In the case of zero spin tensor at every place $\mathbf{u}$, then the motion as a whole is called irrotational or non-vortical.

Proof. The following development proves only the first of equalities (7.58) because the second is similarly verified. Making $\boldsymbol{\Omega} = \boldsymbol{\nu}_{\mathrm{E}}$ in property (7.44),

$$\begin{aligned}
\bar{\nabla}\{[\boldsymbol{\nu}_{\mathrm{E}}]_{\mathrm{t}}\}_{\mathrm{L}}(\overline{\mathbf{x}}) &= \bar{\nabla}\boldsymbol{\varphi}_{\mathrm{t}}(\overline{\mathbf{x}}) \odot_1 \tilde{\nabla}\{[\boldsymbol{\nu}_{\mathrm{E}}]_{\mathrm{t}}\}_{\mathrm{L}}(\overline{\mathbf{x}}) \\
\bar{\nabla}\boldsymbol{\nu}_{\mathrm{t}}(\overline{\mathbf{x}})^{\mathrm{T}} &= \tilde{\nabla}[\boldsymbol{\nu}_{\mathrm{E}}]_{\mathrm{t}}(\mathbf{x})^{\mathrm{T}} \odot_1 \bar{\nabla}\boldsymbol{\varphi}_{\mathrm{t}}(\overline{\mathbf{x}})^{\mathrm{T}} \\
\mathbf{M}_{\overline{\mathbf{x}}} &= \mathbf{L}_{\mathbf{x}} \odot_1 \mathbf{F}_{\overline{\mathbf{x}}}.
\end{aligned}$$

Now, considering an arbitrary non zero vector $\mathbf{x} \in \mathrm{U}_{\mathbb{R}}$ and tensor $\mathbf{L}_{\mathbf{u}}^{\mathrm{T}} = \mathbf{g}_1^* \otimes \mathbf{g}_2^*$, as well as definitions (5.54) and (7.61), we prove the first of equalities (7.63) from (7.62) through the following development:

$$\begin{aligned}
2\dot{\boldsymbol{w}}_{\mathbf{u}} &= \mathbf{A}_{\mathrm{B}}\,\hat{\odot}_2 \mathbf{L}_{\mathbf{u}}^{\mathrm{T}} - \mathbf{A}_{\mathrm{B}}\,\hat{\odot}_2 \mathbf{L}_{\mathbf{u}} \\
2\mathbf{x}\cdot\dot{\boldsymbol{w}}_{\mathbf{u}} &= \mathbf{x}\cdot\mathbf{A}_{\mathrm{B}}\,\hat{\odot}_2(\mathbf{g}_1^*\otimes\mathbf{g}_2^*) - \mathbf{x}\cdot\mathbf{A}_{\mathrm{B}}\,\hat{\odot}_2(\mathbf{g}_2^*\otimes\mathbf{g}_1^*) \\
&= \mathbf{A}_{\mathrm{B}}(\mathbf{x},\mathbf{g}_1,\mathbf{g}_2) + \mathbf{A}_{\mathrm{B}}(\mathbf{x},\mathbf{g}_1,\mathbf{g}_2) \\
&= 2\mathbf{x}\cdot\mathbf{A}_{\mathrm{B}}\,\hat{\odot}_2\mathbf{L}_{\mathbf{u}}^{\mathrm{T}} \\
&= 2\mathbf{x}\cdot\tilde{\nabla}\times[\boldsymbol{\nu}_{\mathrm{t}}]_{\mathrm{E}}(\mathbf{u}).
\end{aligned}$$

The second equality of (7.63) is verified similarly. □

For the case of affine isochoric motions, it is valid to write that $\mathrm{tr}(\mathbf{F}_{\overline{\mathbf{u}}}^{-\mathrm{T}} \odot_1 \mathbf{M}_{\overline{\mathbf{u}}}^{\mathrm{T}}) = 0$ for all $\overline{\mathbf{u}} \in \overline{\mathfrak{B}}$ and $\mathrm{t} \in [\mathrm{t}_0, \mathrm{t}_f]$. From the first equality of (7.58), we obtain the Lagrangian description $\mathbf{L}_{\boldsymbol{\chi}_{\mathrm{t}}(\overline{\mathbf{u}})}^{\mathrm{T}} = \mathbf{F}_{\overline{\mathbf{u}}}^{-\mathrm{T}} \odot_1 \mathbf{M}_{\overline{\mathbf{u}}}^{\mathrm{T}}$ and then conclude that, for affine isochoric motions, $\mathrm{tr}(\mathbf{L}_{\boldsymbol{\chi}_{\mathrm{t}}(\overline{\mathbf{u}})}^{\mathrm{T}}) = \mathrm{tr}(\mathbf{L}_{\boldsymbol{\chi}_{\mathrm{t}}(\overline{\mathbf{u}})}) = 0$, since the trace is immune to transpositions, according to equality (3.29).

Chapter

8 Continuum Dynamics

bla bla bla
bla bla bla
bla bla bla
bla bla bla
bla bla bla

8.1 Inertia and Force

In simple terms, Dynamics is here understood as the descriptive study of the causes of motion. In the very first chapter of his *Principia*, Newton inadvertently presents, in the third definition, the point of departure of Dynamics: the concept of inertia. Obscurely defined as an "inherent force" of matter, he states that this "force" is *"the power of resisting by which every body, so far as it is able, perseveres in its state either of resting or of moving uniformly straight forward.[1]"* By this inherent resistance feature of matter, called inertia, presented by every body – which is abstracted devoid of volume[2] – Newton meant resistance to a change of its velocity in time, that is, resistance to acceleration. Therefore, the Newtonian concept of inertia makes the old Aristotelian idea of *"a thing will either be at rest or must be moved* ad infinitum*, unless something more powerful gets in its way[3]"* one of its corollaries: if a volumeless body is subjected to a condition where its inertia does not manifest itself, then this body will perform a non-accelerated motion, that is, it will be either at rest or move uniformly straight forward.

[1] Newton[46], p.50.
[2] Usually called a point mass.
[3] Aristotle[1], p.366.

At this point, it is important to recall from the previous chapter that inertia is only observed from the point of view of an inertial observer, that is, an observer at rest or moving with constant velocity relative to an absolute observer. In this context, one of the measures of the inertia of a volumeless body is a proportional scalar quantity called **mass**, or more precisely, inertial mass: the greater or lesser this scalar, the greater or lesser the inertia. For the case of bodies with volume, which is our concern here, we impose that every body must have an overall regularly distributed mass, which is also deformation independent, that is, the mass calculated from a volume-density mass distribution on any arbitrary shape is always the mass of the body[4]. This last condition, valid only in the context of Newtonian Mechanics, is a manifestation of the **Principle of Mass Conservation**, which states that when there are no mass exchanges with the surroundings of a certain physical system, the total mass of this system remains constant in time. In mathematical terms, by distribution, we mean a tensor, vector or scalar valued vector function in the context of Euclidean spaces: the volume-density mass distribution cited above, for example, is obviously a scalar distribution.

Following the first definition of the *Principia*, "quantity of matter", Newton presented a measure called "quantity of motion", which is one of the many concepts he borrowed from Galileo. Dealing with collisions of solid volumeless bodies, a subject of great interest in his time, Galileo wrote in the sixth day of *Dialogs Concerning Two New Sciences* the following reasoning exposed by the character Salviati[5]: *"It is evident that the property of force in the mover and of resistance [inertia] in the moved is not single and simple, but is compounded from two actions, by which their energy must be measured. One of these is the weight* [mass]*, of the mover as well as of the resistent; the other is the speed* [velocity] *with which the one must move and the other be moved... if with a lesser weight we wish to raise a greater, it will be necessary to arrange the machine in such a way that the smaller moving weight goes in the same time through a greater space* [distance] *than does the other weight; that is to say, the former is moved more swiftly than the latter... Let us say in general, then, that the momentum of the less heavy body balances the momentum of the more heavy when the speed of the lesser has the same ratio to the speed of the greater as the heaviness of the greater has to that of the lesser...*[6]" In modern terms, Galileo's **linear momentum**, or Newton's "quantity of motion", is defined to be a vector multiple of the velocity vector through inertial mass, that is, linear momentum is mass multiplied by velocity. Therefore, if a volumeless body is subjected to a condition where its inertia does manifest itself, the momentum of this body will change in time, a quantity usually called **linear impulse**. Since inertial mass is constant in time, linear impulse is a vector multiple of the acceleration vector through inertial mass, or linear impulse is

[4]In motions where a portion of the mass of a body is burned in order to propel another portion, the above restriction is still valid because any shape of this body as a whole must include not only the propelled portion of its mass, but also that which is burned.

[5]See section 6.1 for details.

[6]Galileo[24], pp.289-290.

mass times acceleration. Similarly to the case of the mass of a body with volume, from a volume-density mass distribution, the total linear momentum and impulse of a body with volume in a certain instant of time can also be calculated, as we shall see later.

The notorious Newton's Second Axiom, as it is currently stated, specifies that the linear impulse of a volumeless body corresponds to the action on this body of an external physical entity called force. In other words, whenever a force is applied to a volumeless body, this body will present a linear impulse equal to the force applied. If force and linear impulse are mathematically identical but physically different entities, what is the definition of force anyway? Unfortunately, there is not an adequate rigorous answer to this question; nevertheless, we are not wrong to consider that force is an abbreviation for "any physical influence, expressed by a vector magnitude, that produces kinematical outcome". For Newton and most of his contemporaries, this influence always resulted from interaction of bodies, just like the attractive influence of one body on another in classical gravitation, as Newton himself mathematically described it also in *Principia*. In the context of Continuum Mechanics, it is specified that forces in bodies with volume, in a given instant of time, are calculated on areas and on volumes, in such a way that two types of forces arise: a) **body force**, calculated from a volume-density force distribution; b) **contact force**, calculated from an area-density force distribution. An example of a body force is the total influence of gravity or electromagnetic field on a body and of a contact force, the resultant force on body surfaces in contact. In this work, force per volume and force per area are called respectively body force density and contact force density, or **traction**. Continuum Dynamics is mainly concerned with traction distributions and their kinematical consequences. The monumental contribution of Cauchy to the study of deformable bodies relies precisely on the proposition of a domain to traction distributions: in simple terms, the so called **Cauchy's Hypothesis** considers that a traction at a given instant of time and at a point of the surface under study depends linearly on the "direction" of the surface at this point and at this instant of time[7]. Moreover, from this linear relationship between traction and surface direction the fundamental concept of stress emerges, expressed mathematically as a second order tensor.

[7] See theoretical development on section 6.6 and appendix A.

Appendix

A

English Translation of Cauchy's *De La Pression ou Tension Dans Les Corps Solides*

The text in the following pages is my English translation of the seminal paper entitled *De La Pression ou Tension Dans Les Corps Solides*, by Cauchy[14], which effectively marks the beginning of Continuum Mechanics[1]. With the exception of minor corrections and adaptations I made, the author's style and mathematical nomenclature were strictly reproduced.

[1] See section 6.6 for historical details.

ON THE PRESSURE OR TENSION
IN A SOLID BODY.

BY Mr. AUGUSTIN-LOUIS CAUCHY

CHIEF ENGINEER AT PONTS ET CHAUSSÉES, PROFESSOR AT ÉCOLE POLYTECHNIQUE, ADJOINT PROFESSOR AT FACULTÉ DES SCIENCES, MEMBER OF THE ACADEMY OF SCIENCES, KNIGHT OF THE LEGION OF HONOUR

Translated by Mr. R.D. ALGARTE.

Geometers who have researched the equations of equilibrium or motion of plates or of elastic or non-elastic surfaces have distinguished two types of forces produced, some by expansion or contraction, others by the flexion of these same surfaces. Moreover, they generally assumed, in their calculations, that the forces of the first type, called tensions, remain perpendicular to the lines against which they are exerted. It seemed to me that these two types of forces could be reduced by a single one, which must constantly be called tension or pressure, acting on each element of a section chosen at will, not only in a flexible surface, but also in an elastic or non-elastic solid, and which is of the same nature of the hydrostatic pressure exerted by a fluid at rest against the exterior surface of a body. However, the new pressure does not always remain perpendicular to the faces which are subjected to it, nor the same in all directions at a given point. In developing this idea, I came to recognize that the pressure or tension exerted against any plane at a given point of a solid body is very easily deduced, both in magnitude and in direction, from the pressures or tensions exerted against three rectangular planes defined through this same point. This proposition, which I have already addressed in the January 1823 edition of the *Bulletin Des Sciences de La Société Philomatique de Paris*, can be established using the following considerations.

If in an elastic or non-elastic solid body, a small element of volume defined by arbitrary faces is made rigid and invariable, this small element will experience on its different faces and at each point of each of them a certain pressure or tension. This pressure or tension will be similar to the pressure that a fluid exerts against an element of the surface of a solid body, with the only difference that the pressure exerted by a fluid at rest, against the surface of a solid body, is directed perpendicular to this surface from outside to inside, and independent at each point of the inclination of the surface with respect to the coordinate planes, while the pressure or tension exerted at a given point of a solid body, through which a very small surface element passes can be directed perpendicular or obliquely to this surface, sometimes from outside to inside, if there is contraction, sometimes from inside to outside, if there is expansion, and may depend on the inclination of the surface in relation to the planes in question. That being said, let υ be the volume of a portion of the body that has become rigid, s, s', s'', ... the areas of the plane or curved surfaces which cover the volume υ; x, y, z the rectangular coordinates of a point taken at random in the surface s; p the pressure or tension exerted at this point against the surface; α, β, γ the angles that the perpendicular to the surface forms with the semi-axes of the positive coordinates; finally λ, μ, ν the angles formed with the same semi-axes by the direction of the force p. If we project onto the axes x, y and z the various pressures or tensions to which the surface will be subjected, the sums of their algebraic projections on these three axes will be represented by the integrals[1]

$$(1) \qquad \begin{cases} \iint p\cos\lambda\sec\gamma\,dydx, \\ \iint p\cos\mu\sec\gamma\,dydx, \\ \iint p\cos\nu\sec\gamma\,dydx, \end{cases}$$

while the sums of the algebraic projections of their linear moments will be respectively

$$(2) \qquad \begin{cases} \iint p(y\cos\nu - z\cos\mu)\sec\gamma\,dydx, \\ \iint p(z\cos\lambda - x\cos\nu)\sec\gamma\,dydx, \\ \iint p(x\cos\mu - y\cos\lambda)\sec\gamma\,dydx, \end{cases}$$

if we take the origin of the coordinates as the center of moments, or, if we transport it to a point with coordinates x_0, y_0, z_0,

$$(3) \qquad \begin{cases} \iint p[(y-y_0)\cos\nu - (z-z_0)\cos\mu]\sec\gamma\,dydx, \\ \iint p[(z-z_0)\cos\lambda - (x-x_0)\cos\nu]\sec\gamma\,dydx, \\ \iint p[(x-x_0)\cos\mu - (y-y_0)\cos\lambda]\sec\gamma\,dydx. \end{cases}$$

In all these integrals, the limits of the integrations relating to variables x, y must be determined from the shape of the contour of the surface s, so that inside these limits

$$(4) \qquad \iint \sec\gamma\,dydx = s\,.$$

If the surface s becomes plane and the volume υ very small, so that each of its dimensions can be considered an infinitely small quantity of the first order, then the variations that the three products

$$(5) \qquad p\cos\lambda, \quad p\cos\mu, \quad p\cos\nu$$

[1] All the subsequent terms $\sec\gamma\,dydx$ fix the term $\cos\gamma\,dydx$ of the original text. (TN)

will experience, in the passage from one point to another in the surface s, will still be infinitely small of the first order; and, by neglecting the infinitely small third order values in the integrals (1), we shall reduce these integrals to the quantities

$$ps\cos\lambda, \quad ps\cos\mu, \quad ps\cos\nu\,. \tag{6}$$

Moreover, if we make the center of moments coincide with a point in volume v, the integrals (3) will be infinitely small quantities of the third order, and it will suffice to neglect, in these integrals, the infinitely small term of the fourth order, so that they are reduced to products

$$\begin{cases} ps[(\eta - y_0)\cos\nu - (\zeta - z_0)\cos\mu], \\ ps[(\zeta - z_0)\cos\lambda - (\xi - x_0)\cos\nu], \\ ps[(\xi - x_0)\cos\mu - (\eta - y_0)\cos\lambda], \end{cases} \tag{7}$$

ξ, η, ζ designating the fractions

$$\frac{\iint x\sec\gamma\,dydx}{s}, \quad \frac{\iint y\sec\gamma\,dydx}{s}, \quad \frac{\iint z\sec\gamma\,dydx}{s}, \tag{8}$$

that is, the coordinates of the center of gravity of the surface s.

Now, let m be the infinitely small mass related to the volume v. Moreover, let us consider that the letter φ represents the accelerating force applied to this mass, if the solid body is in equilibrium, and, on the contrary case, the excess of the accelerating force applied on m which would be able to produce the observed motion of mass m. Finally, let us call **X**, **Y**, **Z** the algebraic projections of the force φ, and ξ_0, η_0, ζ_0 the coordinates of the center of gravity of the mass m. If we suppose that the accelerating force φ remains the same in magnitude and direction in all points of the mass m, there must be equilibrium between the driving force $m\varphi$ applied at the point (ξ_0, η_0, ζ_0) and the forces to which the pressures or tensions exerted on the surfaces s, s', ... are reduced. So the sums of the algebraic projections of all these forces and their linear moments on the axes x, y, z will have to be reduced to zero. So, if we want to place one or more accents after the letters p, λ, μ, ν, ξ, η, ζ, presented in expressions (6) and (7), to indicate the new values that these expressions take when one passes from the surface s to the surface s', or s'', or s''', ... one will find, neglecting, in the sums of the projected forces, the infinitely small of the third order, and in the sums of the projected linear moments, the infinitely small of the fourth order,

$$\begin{cases} ps\cos\lambda + p's'\cos\lambda' + \cdots + m\mathbf{X} = 0, \\ ps\cos\mu + p's'\cos\mu' + \cdots + m\mathbf{Y} = 0, \\ ps\cos\nu + p's'\cos\nu' + \cdots + m\mathbf{Z} = 0; \end{cases} \tag{9}$$

$$
(10)\quad \begin{cases}
ps[(\eta - y_0)\cos\nu - (\zeta - z_0)\cos\mu] + p's'[(\eta' - y_0)\cos\nu' - (\zeta' - z_0)\cos\mu'] + \\
\qquad\qquad + \cdots + m[(\eta_0 - y_0)\mathrm{Z} - (\zeta_0 - z_0)\mathrm{Y}] = 0, \\
ps[(\zeta - z_0)\cos\lambda - (\xi - x_0)\cos\nu] + p's'[(\zeta' - z_0)\cos\lambda' - (\xi' - x_0)\cos\nu'] + \\
\qquad\qquad + \cdots + m[(\zeta_0 - z_0)\mathrm{X} - (\xi_0 - x_0)\mathrm{Z}] = 0 \\
ps[(\xi - x_0)\cos\mu - (\eta - y_0)\cos\lambda] + p's'[(\xi' - x_0)\cos\mu' - (\eta' - y_0)\cos\lambda'] + \\
\qquad\qquad + \cdots + m[(\xi_0 - x_0)\mathrm{Y} - (\eta_0 - y_0)\mathrm{X}] = 0.
\end{cases}
$$

Now, the mass m being itself infinitely small of the third order, the terms which contain it will be of the third order in the formulas (1) and of the fourth order in the formulas (3). We can therefore neglect these terms, and replace the formulas in question by the following

$$
(11)\quad \begin{cases}
ps\cos\lambda + p's'\cos\lambda' + p''s''\cos\lambda'' + p'''s'''\cos\lambda''' + \cdots = 0, \\
ps\cos\mu + p's'\cos\mu' + p''s''\cos\mu'' + p'''s'''\cos\mu''' + \cdots = 0, \\
ps\cos\nu + p's'\cos\nu' + p''s''\cos\nu'' + p'''s'''\cos\nu''' + \cdots = 0;
\end{cases}
$$

$$
(12)\quad \begin{cases}
ps[(\eta - y_0)\cos\nu - (\zeta - z_0)\cos\mu] + p's'[(\eta' - y_0)\cos\nu' - (\zeta' - z_0)\cos\mu'] + \\
\qquad\qquad + \cdots = 0, \\
ps[(\zeta - z_0)\cos\lambda - (\xi - x_0)\cos\nu] + p's'[(\zeta' - z_0)\cos\lambda' - (\xi' - x_0)\cos\nu'] + \\
\qquad\qquad + \cdots = 0 \\
ps[(\xi - x_0)\cos\mu - (\eta - y_0)\cos\lambda] + p's'[(\xi' - x_0)\cos\mu' - (\eta' - y_0)\cos\lambda'] + \\
\qquad\qquad + \cdots = 0.
\end{cases}
$$

If we wanted to take into account the variations that the accelerating force φ and its projections X, Y, Z can experience, when we go from one point to another in the mass m, we would have to replace, in equations (1) and (3), the six quantities

$$
m\mathrm{X},\quad m\mathrm{Y},\quad m\mathrm{Z};
$$

$$
\begin{gathered}
m[(\eta_0 - y_0)\mathrm{Z} - (\zeta_0 - z_0)\mathrm{Y}], \\
m[(\zeta_0 - z_0)\mathrm{X} - (\xi_0 - x_0)\mathrm{Z}], \\
m[(\xi_0 - x_0)\mathrm{Y} - (\eta_0 - y_0)\mathrm{X}]
\end{gathered}
$$

by six integrals of the form

$$
\iiint \rho\mathrm{X}\,dzdydx,\quad \iiint \rho\mathrm{Y}\,dzdydx,\quad \iiint \rho\mathrm{Z}\,dzdydx;
$$

$$
\begin{gathered}
\iiint \rho[(y - y_0)\mathrm{Z} - (z - z_0)\mathrm{Y}]\,dzdydx, \\
\iiint \rho[(z - z_0)\mathrm{X} - (x - x_0)\mathrm{Z}]\,dzdydx, \\
\iiint \rho[(x - x_0)\mathrm{Y} - (y - y_0)\mathrm{X}]\,dzdydx;
\end{gathered}
$$

ρ denoting the density of the solid body at the point (x, y, z), and the integration limits being relative to the limits of the volume v. But, since the first three integrals would be infinitely small of the third order, and the last three infinitely small of the fourth order, we would still find ourselves brought back to formulas (11) and (12). It remains to show how, with the help of these formulas, one can discover the relations that exist between the pressures or tensions exerted in a given point of a solid body against various planes carried out successively by the same point.

Let us first consider that the volume v takes the form of a right prism, the two bases of which are represented by s and by s'. We shall have $s' = s$; and if, the dimensions of each base being considered as infinitely small of the first order, the height of the prism becomes an infinitely small quantity of an order greater than the first, then, neglecting, in formulas (11), the infinitely small order greater than the second, we shall find

$$(p\cos\lambda + p'\cos\lambda')s = 0, \quad (p\cos\mu + p'\cos\mu')s = 0, \quad (p\cos\nu + p'\cos\nu')s = 0,$$

or equally,

$$p'\cos\lambda' = -p\cos\lambda, \quad p'\cos\mu' = -p\cos\mu, \quad p'\cos\nu' = -p\cos\nu,$$

and we conclude that

$$p = p',$$
$$\cos\lambda' = -\cos\lambda, \quad \cos\mu' = -\cos\mu, \quad \cos\nu' = -\cos\nu.$$

These last equations, which take place only in the case where the height of the prism vanishes, comprise a theorem easy to predict, the statement of which is the following:

Theorem I *The pressures or tensions, at a given point of a solid body, exerted against the two faces of an arbitrary plane through this point, are equal and directly opposite forces.*

Now, let

$$(13) \qquad p', \quad p'', \quad p'''$$

be the pressures or tensions exerted at the point (x, y, z) and on the side of the positive coordinates against three planes through this point parallel to the coordinate planes of y, z, of z, x and of x, y. Moreover, λ', μ', ν'; λ'', μ'', ν''; λ''', μ''', ν''' are the angles formed by the directions of the forces p', p'', p''' with the semi-axes of the positive coordinates. Finally, let us consider that the volume v, taking the form of a rectangular parallelepiped, is enclosed between the three planes through the point (x, y, z), and three parallel planes trough a very close point $(x + \Delta x, y + \Delta y, z + \Delta z)$. The pressures or tensions, supported by the faces of the parallelepiped which will end at this last point, will be approximately

$$(14) \qquad p'\Delta y\Delta z, \quad p''\Delta z\Delta x, \quad p''''\Delta x\Delta y,$$

while their algebraic projections on the axes (x, y, z) will clearly reduce to the quantities

$$(15)\qquad \begin{cases} p's' \cos\lambda' \Delta y \Delta z, & p''s'' \cos\lambda'' \Delta z \Delta x, \quad p'''s''' \cos\lambda''' \Delta x \Delta y, \\ p's' \cos\mu' \Delta y \Delta z, & p''s'' \cos\mu'' \Delta z \Delta x, \quad p'''s''' \cos\mu''' \Delta x \Delta y, \\ p's' \cos\nu' \Delta y \Delta z, & p''s'' \cos\nu'' \Delta z \Delta x, \quad p'''s''' \cos\nu''' \Delta x \Delta y. \end{cases}$$

Concerning the pressures or tensions supported by the faces which end at the point (x, y, z), they will be, by virtue of Theorem I, respectively equal, but directly opposite, to those which act on the parallel faces through the point $(x + \Delta x, y + \Delta y, z + \Delta z)$. So the algebraic projections of these new tensions will be numerically equal to the algebraic projections of the other three, but affected by opposite signs, so that each of the formulas (11) will become an identity. Let us add that the centers of gravity of the six faces of the parallelepiped will merge with their geometric centers, and will be located on three lines carried out parallel to the axes (x, y, z) through the center of the parallelepiped, that is to say, through the point which has coordinates

$$x + \frac{1}{2}\Delta x, \quad y + \frac{1}{2}\Delta y, \quad z + \frac{1}{2}\Delta z.$$

That said, it is clear that, if we take this last point as the center of the moments, the first of the formulas (12) will give

$$p'' \cos\nu'' \Delta z \Delta x \frac{\Delta y}{2} - p''' \cos\mu''' \Delta x \Delta y \frac{\Delta z}{2} - (-p'' \cos\nu'')\Delta z \Delta x \frac{\Delta y}{2} + \\ + (-p''' \cos\mu''')\Delta x \Delta y \frac{\Delta z}{2} = 0$$

and consequently

$$(16)\qquad \begin{cases} p'' \cos\nu'' = p''' \cos\mu'''. \\ \text{Similarly, one will find that} \\ p''' \cos\lambda''' = p' \cos\nu', \\ p' \cos\mu' = p'' \cos\lambda''. \end{cases}$$

Since the axes x, y, z are entirely arbitrary, the equations (16) obviously comprise the theorem we are going to state:

Theorem II *If through any point of a solid body we define two axes which intersect at right angles, and if we project onto one of these axes the pressure or tension supported by a plane perpendicular to the other at the point where it acts, the projection thus obtained will not vary when these same axes are exchanged between them.*

Let us now consider that the volume takes the form of a tetrahedron of which three edges coincide with three infinitely small lengths carried from the point

(x, y, z) on lines parallel to the coordinate axes. Consider the point (x, y, z) to be the vertex of this tetrahedron; denote its base by s, and let α, β, γ be the angles formed, with its semi-axes of positive coordinates, by a perpendicular raised through a point of this base, but extended outside the tetrahedron. The three faces which end at the vertex of the tetrahedron will be measured by the numerical values of the products

$$s \cos \alpha, \quad s \cos \beta, \quad s \cos \gamma. \tag{17}$$

Thereby, if one calls p the pressure or tension supported by the base of the tetrahedron, and if one continues to attribute to the quantities p', p'', p''' the values which they received in equations (16), the first of the formulas (11) will obviously give

$$ps \cos \lambda - p' \cos \lambda' s \cos \alpha - p'' \cos \lambda'' s \cos \beta - p''' \cos \lambda''' s \cos \gamma = 0$$

and consequently

$$\begin{cases} p \cos \lambda = p' \cos \lambda' \cos \alpha + p'' \cos \lambda'' \cos \beta + p''' \cos \lambda''' \cos \gamma. \\ \text{Similarly, one will find that} \\ p \cos \mu = p' \cos \mu' \cos \alpha + p'' \cos \mu'' \cos \beta + p''' \cos \mu''' \cos \gamma, \\ p \cos \nu = p' \cos \nu' \cos \alpha + p'' \cos \nu'' \cos \beta + p''' \cos \nu''' \cos \gamma. \end{cases} \tag{18}$$

So, in order to shorten expressions, if we make

$$\begin{cases} \mathrm{A} = p' \cos \lambda', \\ \mathrm{B} = p'' \cos \mu'', \\ \mathrm{C} = p''' \cos \nu''', \\ \mathrm{D} = p'' \cos \nu'' = p''' \cos \mu''', \\ \mathrm{E} = p''' \cos \lambda''' = p' \cos \nu', \\ \mathrm{F} = p' \cos \mu' = p'' \cos \lambda'', \end{cases} \tag{19}$$

we shall simply have

$$\begin{cases} p \cos \lambda = \mathrm{A} \cos \alpha + \mathrm{F} \cos \beta + \mathrm{E} \cos \gamma, \\ p \cos \mu = \mathrm{F} \cos \alpha + \mathrm{B} \cos \beta + \mathrm{D} \cos \gamma, \\ p \cos \nu = \mathrm{E} \cos \alpha + \mathrm{D} \cos \beta + \mathrm{C} \cos \gamma. \end{cases} \tag{20}$$

These last equations show the relations which subsist, for the point (x, y, z), between the algebraic projections

$$\left\{ \begin{array}{ccc} \mathrm{A}, & \mathrm{F}, & \mathrm{E}; \\ \mathrm{F}, & \mathrm{B}, & \mathrm{D}; \\ \mathrm{E}, & \mathrm{D}, & \mathrm{C} \end{array} \right. \tag{21}$$

of pressures p', p'', p''' exerted at this point, on the side of the positive coordinates, against three planes parallel to the coordinate planes, and the algebraic projections

$$p\cos\lambda, \quad p\cos\mu, \quad p\cos\nu$$

of the pressure or tension p exerted at this same point against any plane perpendicular to a straight line which, extended from the side where the force p occurs, forms, with the semi-axes of positive coordinates, the angles α, β, γ.

From equations (20), it is easy to recognize that, if the volume v, instead of having the shape of a tetrahedron, is defined by any number of plane faces, the formulas (11) and (12) will always be verified. Indeed, these different faces being represented by s, s', ..., let us call α, β, γ; α', β', γ'; ... the angles that straight lines perpendicular to the planes of these same faces, and extended outside the volume v, form with the semi-axes of positive coordinates. In order to obtain the first of formulas (11), it will suffice to add the equations (1) on page 55[2], after having respectively multiplied them by A, F, E, taking into account the first of formulas (20), as well as similar formulas. We establish, in the same way, the second and the third of the formulas (11), by adding the equations (1) (p.55), after having respectively multiplied them by the coefficients F, B, D, or by the coefficients E, D, C. Finally, if we combine the formulas (20) and others from the same type, not only with the equations (1) on page 55, but also with the equations (5) on page 56[3], we shall easily arrive at formulas (12).

One can easily deduce from formulas (20): 1° the intensity of the force p; 2° the angle between the direction of this force and the perpendicular to the plane against which it is exerted. Thereby, if we add these formulas, after having squared each of their members, we shall find

$$(22)\qquad \begin{cases} p^2 = & (\mathrm{A}\cos\alpha + \mathrm{F}\cos\beta + \mathrm{E}\cos\gamma)^2 + \\ & (\mathrm{F}\cos\alpha + \mathrm{B}\cos\beta + \mathrm{D}\cos\gamma)^2 + \\ & (\mathrm{E}\cos\alpha + \mathrm{D}\cos\beta + \mathrm{C}\cos\gamma)^2. \end{cases}$$

[2] This reference concerns the article *Sur Quelques Propriétés des Polyèdres*, in the same volume of the current article, from which we extract the following excerpt. (TN).
Theorem I - *The sum of the algebraic projections of the faces of any polyhedron on the coordinate planes results zero. ...*

$$(1)\qquad \begin{cases} s\cos\alpha + s'\cos\alpha' + s''\cos\alpha'' + \cdots = 0, \\ s\cos\beta + s'\cos\beta' + s''\cos\beta'' + \cdots = 0, \\ s\cos\gamma + s'\cos\gamma' + s''\cos\gamma'' + \cdots = 0. \end{cases}$$

[3] Again referring to the same article of the previous footnote, this equation (5) can be expressed by the equality

$$(5)\qquad \begin{bmatrix} \xi & \xi' & \cdots \\ \eta & \eta' & \cdots \\ \zeta & \zeta' & \cdots \end{bmatrix} \begin{bmatrix} s\cos\alpha & s\cos\beta & s\cos\gamma \\ s'\cos\alpha' & s'\cos\beta' & s'\cos\gamma' \\ \vdots & \vdots & \vdots \end{bmatrix} = \begin{bmatrix} v & 0 & 0 \\ 0 & v & 0 \\ 0 & 0 & v \end{bmatrix},$$

according to what is developed in the article. (TN)

Moreover, if we call δ the angle that we have just spoken, we will obviously have

$$\cos\delta = \cos\alpha\cos\lambda + \cos\beta\cos\mu + \cos\gamma\cos\nu, \tag{23}$$

and consequently,

$$\cos\delta = \frac{1}{p}(\mathrm{A}\cos^2\alpha + \mathrm{B}\cos^2\beta + \mathrm{C}\cos^2\gamma + 2\mathrm{D}\cos\beta\cos\gamma + 2\mathrm{E}\cos\gamma\cos\alpha + \\ + 2\mathrm{F}\cos\alpha\cos\beta). \tag{24}$$

Let us add that, if we describe the force p by two components, one of which is included in the plane under consideration, and the other perpendicular to this plane, the second component will be represented, except for the sign, by the product

$$\begin{cases} p\cos\delta = \mathrm{A}\cos^2\alpha + \mathrm{B}\cos^2\beta + \mathrm{C}\cos^2\gamma + \\ \qquad + 2\mathrm{D}\cos\beta\cos\gamma + 2\mathrm{E}\cos\gamma\cos\alpha + 2\mathrm{F}\cos\alpha\cos\beta. \end{cases} \tag{25}$$

Let us observe finally that this second component will be a tension or a pressure according to whether the formula (25) will present a positive or negative value in the second member.

Suppose now that from the point (x, y, z) we define, on the perpendicular to the plane against which the force p acts a length r whose square represents the numerical value of the ratio

$$\frac{1}{p\cos\delta} \tag{26}$$

and denote by $x + \mathrm{x}$, $y + \mathrm{y}$, $z + \mathrm{z}$ the coordinates of the extremity of that same length. We shall have

$$\frac{\mathrm{x}}{\cos\alpha} = \frac{\mathrm{y}}{\cos\beta} = \frac{\mathrm{z}}{\cos\gamma} = \pm\sqrt{\mathrm{x}^2 + \mathrm{y}^2 + \mathrm{z}^2} = \pm r\,, \tag{27}$$

$$\frac{1}{p\cos\delta} = \pm\, r^2; \tag{28}$$

and, therefore, formula (25) will give

$$\mathrm{A}\mathrm{x}^2 + \mathrm{B}\mathrm{y}^2 + \mathrm{C}\mathrm{z}^2 + 2\mathrm{D}\mathrm{y}\mathrm{z} + 2\mathrm{E}\mathrm{z}\mathrm{x} + 2\mathrm{F}\mathrm{x}\mathrm{y} = \pm 1\,. \tag{29}$$

The variables x, y, z, included in equation (29), are the coordinates of the end of the length r, computed from the point (x, y, z) on three rectangular axes; and this equation itself defines a quadratic surface whose center is the point (x, y, z). When the polynomial

$$\begin{cases} \mathrm{A}\cos^2\alpha + \mathrm{B}\cos^2\beta + \mathrm{C}\cos^2\gamma + \\ \qquad + 2\mathrm{D}\cos\beta\cos\gamma + 2\mathrm{E}\cos\gamma\cos\alpha + 2\mathrm{F}\cos\alpha\cos\beta. \end{cases} \tag{30}$$

preserves the same sign, for whatever values assigned to the angles α, β, γ, then equation (25), reduced to one of the following

$$\mathrm{Ax}^2 + \mathrm{By}^2 + \mathrm{Cz}^2 + 2\mathrm{Dyz} + 2\mathrm{Ezx} + 2\mathrm{Fxy} = 1\,, \tag{31}$$

$$\mathrm{Ax}^2 + \mathrm{By}^2 + \mathrm{Cz}^2 + 2\mathrm{Dyz} + 2\mathrm{Ezx} + 2\mathrm{Fxy} = -\,1\,, \tag{32}$$

represents an ellipsoid. But, if the polynomial (30) changes in sign while the angles α, β, γ vary, the ellipsoid in question will give way to the system of two hyperboloids, one of which will be represented by the equation (31), the other by equation (32); and these two hyperboloids, one of which will have a single sheet, the other two distinct sheets, will be conjugated[4] between them, so that they will have the same center with the same axes, and will be touched at infinity by the same conical surface of the second degree. Let us add that, in the first case, the force

$$\pm p \cos\delta = \frac{1}{r^2} \tag{33}$$

will always be a tension, if the polynomial (30) is positive, a pressure if it is negative. In the second case, on the contrary, the force in question will be sometimes a pressure, sometimes a tension, according to whether the end of the vector ray r is located on the surface of one or the other hyperboloid; and the same force will vanish whenever this vector ray is directed along a generatrix of the above mentioned conical surface.

It can be easily demonstrated that the normal, defined at the end of the radius vector r on the surface (31) or (32), forms, with the semi-axes of positive coordinates, angles whose cosines are proportional to the three polynomials

$$\begin{gathered}
\mathrm{A}\cos\alpha + \mathrm{F}\cos\beta + \mathrm{E}\cos\gamma\,,\\
\mathrm{F}\cos\alpha + \mathrm{B}\cos\beta + \mathrm{D}\cos\gamma\,,\\
\mathrm{E}\cos\alpha + \mathrm{D}\cos\beta + \mathrm{C}\cos\gamma\,.
\end{gathered}$$

So this normal will be directed along the same line as the radius vector if we have

$$\begin{cases}
\dfrac{\mathrm{A}\cos\alpha + \mathrm{F}\cos\beta + \mathrm{E}\cos\gamma}{\cos\alpha}\\
=\dfrac{\mathrm{F}\cos\alpha + \mathrm{B}\cos\beta + \mathrm{D}\cos\gamma}{\cos\beta}\\
=\dfrac{\mathrm{E}\cos\alpha + \mathrm{D}\cos\beta + \mathrm{C}\cos\gamma}{\cos\gamma}\,.
\end{cases} \tag{34}$$

Formula (34) is verified, in fact, when the radius vector coincides with one of the axes of the surface (31) or (32). Then, we write equations (20), combined with

[4]Regarding the properties of conjugated hyperboloids, see *Lecons sur les applications du Calcul infinitésimal à la Géométrie*, p. 275, (*OEuvres de Cauchy*, S. II, T.V)

formula (34), as

$$\frac{\cos\lambda}{\cos\alpha} = \frac{\cos\mu}{\cos\beta} = \frac{\cos\nu}{\cos\gamma} = \pm\frac{\sqrt{\cos^2\lambda + \cos^2\mu + \cos^2\nu}}{\sqrt{\cos^2\alpha + \cos^2\beta + \cos^2\gamma}} = \pm 1\,, \tag{35}$$

and it follows that the force p is itself directed along the radius vector r, or along its extension. Consequently, to the three axes of the surface (31) or (32) correspond three pressures or tensions, each of which is perpendicular to the plane against which it is exerted. We shall call them *principal pressures* or *tensions.* It is then easy to ensure that one finds among them the *maximum* pressure or tension and the *minimum* pressure or tension; because, if we equal to zero the value of p taken from formula (22), and if we consider the equation

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1\,, \tag{36}$$

from which one of the three variables α, β, γ becomes a function of the two others considered as independent, we shall immediately be brought back to formula (34).

If, starting from the point (x, y, z), we lead, on the perpendicular to the plane against which the force p acts, a length equivalent, no longer to the square root of the ratio $\pm\frac{1}{p\cos\delta}$, but to the fraction $\frac{1}{p}$, by designating $x + \mathrm{x}$, $y + \mathrm{y}$, $z + \mathrm{z}$ the coordinates of the end of this length, we would find

$$\left.\begin{aligned} &\frac{\mathrm{x}}{\cos\alpha} = \frac{\mathrm{y}}{\cos\beta} = \frac{\mathrm{z}}{\cos\gamma} = \pm\sqrt{\mathrm{x}^2 + \mathrm{y}^2 + \mathrm{z}^2} = \pm r\,,\\ &\qquad\qquad\qquad\quad \frac{1}{p} = r; \end{aligned}\right. \tag{37}$$

and, consequently, the formula (22) would give

$$(\mathrm{Ax} + \mathrm{Fy} + \mathrm{Ez})^2 + (\mathrm{Fx} + \mathrm{By} + \mathrm{Dz})^2 + (\mathrm{Ex} + \mathrm{Dy} + \mathrm{Cz})^2 = 1\,. \tag{38}$$

Equation (38) describes an ellipsoid whose axes correspond to the values of α, β, γ determined by the formula

$$\begin{cases} \dfrac{\mathrm{A}(\mathrm{A}\cos\alpha + \mathrm{F}\cos\beta + \mathrm{E}\cos\gamma) + \mathrm{F}(\mathrm{F}\cos\alpha + \mathrm{B}\cos\beta + \mathrm{D}\cos\gamma) + \mathrm{E}(\mathrm{E}\cos\alpha + \mathrm{D}\cos\beta + \mathrm{C}\cos\gamma)}{\cos\alpha} \\ = \dfrac{\mathrm{F}(\mathrm{A}\cos\alpha + \mathrm{F}\cos\beta + \mathrm{E}\cos\gamma) + \mathrm{B}(\mathrm{F}\cos\alpha + \mathrm{B}\cos\beta + \mathrm{D}\cos\gamma) + \mathrm{D}(\mathrm{E}\cos\alpha + \mathrm{D}\cos\beta + \mathrm{C}\cos\gamma)}{\cos\beta} \\ = \dfrac{\mathrm{E}(\mathrm{A}\cos\alpha + \mathrm{F}\cos\beta + \mathrm{E}\cos\gamma) + \mathrm{D}(\mathrm{F}\cos\alpha + \mathrm{B}\cos\beta + \mathrm{D}\cos\gamma) + \mathrm{C}(\mathrm{E}\cos\alpha + \mathrm{D}\cos\beta + \mathrm{C}\cos\gamma)}{\cos\gamma}\,. \end{cases} \tag{39}$$

However, since this formula is obviously verified by the values of α, β, γ which satisfy the formula (34), one can affirm that the axes of the new ellipsoid are directed along the same lines of the principal pressures or tensions. We would come to the same conclusion by observing that the *maximum* and *minimum* values of the vector radius, that is to say the major axis and the minor axis of the ellipsoid, necessarily correspond, by virtue of equation (37), the first, to the *maximum* pressure or tension, the second, to the *minimum* pressure or tension.

By summarizing the various propositions that we have just established, we shall obtain the following theorem:

Theorem III *If, after having made an arbitrary plane pass through a given point of a solid body, we define, from this point and on each of the half-axes perpendicular to the plane, two equivalent lengths, the first, of the unit divided by the pressure or tension exerted against the plane, the second, of the unit divided by the square root of this force projected on one of the semi-axes which one considers, these two lengths will be the vector rays of two ellipsoids, whose axes will be directed along the same lines. To these axes there will correspond the principal pressures or tensions, each of which will be normal to the plane which will support it, and on which we will always find the* maximum *pressure or tension, as well as the* minimum *pressure or tension. Concerning the other pressures or tensions, they will be distributed symmetrically around the axes of the two ellipsoids. Let us add that, in certain cases, the second ellipsoid will be replaced by two conjugated hyperboloids. These cases are those in which the system of principal pressures or tensions consists of one tension and two pressures or one pressure and two tensions. Then, if we substitute the force which acts against each plane by two rectangular components, one of which is normal to the plane, this component will be a tension or a pressure, depending on whether the vector ray perpendicular to the plane belongs to one or the other of the two hyperboloids and it will vanish when the vector ray is directed along one of the generatrices of the conical surface of the second degree which touches the two hyperboloids at infinity.*

Let us now consider that from the center of the first ellipsoid one defines arbitrarily three vector rays which intersect at right angles. We can easily prove that, if we divide the unit by each of these vector rays, the sum of the squares of the quotients will be a constant value, equal to the sum that would be obtained by making the three vector rays coincide with the three semi-axes of the ellipsoid (see *Lecons sur les applications du Calcul infinitésimal à la Géométrie*, pp. 274-275)[5]. From this remark, together with the third theorem, we immediately deduce the following proposition:

Theorem IV *If through a given point of a solid body we pass three rectangular planes between them, the sum of the squares of the pressures or tensions supported by these same planes will be a constant value, equal to the sum of the squares of the principal pressures or tensions.*

It may happen that the three principal pressures or tensions, or at least two of them, become equivalent. When these forces reduce to three equal pressures, or three equal tensions, the two ellipsoids we have spoken of reduce to two spheres.

[5] *OEuvres de Cauchy*, S. II, T.V.

Then there is equality of pressure or tension in all directions, and each pressure or tension is perpendicular to the plane which supports it. Moreover, it is important to observe that, from these last two conditions, the second can only be attained as far as the first is equally attained. Indeed, if one supposes the force p constantly directed along the line which forms, with the half-axes of the positive coordinates, the angles α, β, γ, the formula (34) or (35) will remain for any position in this line, and, consequently, for all the values of α, β, γ suitable for verifying equation (36). Now we derive from formula (34): 1° assuming two of the values

$$\cos\alpha, \quad \cos\beta, \quad \cos\gamma$$

to be zero, and the third, to be unity,

$$\text{(40)} \qquad \mathrm{D} = 0, \quad \mathrm{E} = 0, \quad \mathrm{F} = 0;$$

2° considering equations (40),

$$\text{(41)} \qquad \mathrm{A} = \mathrm{B} = \mathrm{C}\,.$$

Consequently, under the accepted hypothesis, the formulas (20) become

$$\text{(42)} \qquad p\cos\alpha = \mathrm{A}\cos\alpha, \quad p\cos\mu = \mathrm{A}\cos\beta, \quad p\cos\nu = \mathrm{A}\cos\gamma\,;$$

and, since we conclude from these last formulas that

$$\text{(43)} \qquad \frac{\cos\alpha}{\cos\lambda} = \frac{\cos\beta}{\cos\mu} = \frac{\cos\gamma}{\cos\nu} = \pm\frac{\sqrt{\cos^2\alpha + \cos^2\beta + \cos^2\gamma}}{\sqrt{\cos^2\lambda + \cos^2\mu + \cos^2\nu}} = \pm 1\,,$$

$$\text{(44)} \qquad p = \pm\,\mathrm{A}\,,$$

it is clear that the pressure or tension, denoted by p, will remain the same in all directions. This is precisely what takes place when we consider a fluid mass in equilibrium. If two principal pressures or tensions become equal, the two ellipsoids mentioned in Theorem II. are reduced to two ellipsoids of revolution, the second of which is replaced, in certain cases, by a system of two hyperboloids of revolution conjugated to one another. Then all the planes including the axis of revolution of these ellipsoids or hyperboloids support equivalent pressures or tensions, each of which, being perpendicular to the plane which is be subjected to it, can be considered as a principal pressure or tension.

The supposition we have just made includes the case where the three forces composing the system of the principal pressures or tensions are equivalent but reduced to a pressure and to two tensions, or to two pressures and a tension. It is important only to observe that, in this case, the first ellipsoid would be replaced by a sphere, and consequently all the planes carried out by the point (x, y, z) would support equivalent pressures or tensions, but directed, some by perpendicular lines, others by straight lines oblique to these same planes.

Generally, whenever a principal tension becomes equivalent to a principal pressure, the planes driven by the axis perpendicular to the directions of these two forces will bear equivalent pressures or tensions, but which will remain oblique to the planes in question, as long as they are distinct from these same forces.

It can still be assumed that one or two of the principal tensions or pressures reduce to zero, or that they all disappear. In the first case, the ellipsoids or hyperboloids, mentioned in the third theorem, will turn into right cylinders which will have conjugate ellipses or hyperbolas as their bases. In the second case, each of these cylinders will be replaced by two parallel planes. In the third case, the pressure or tension, exerted against any plane led by the point (x, y, z), will always reduce to zero.

The formulas previously obtained are simplified when we take for coordinate axes lines parallel to the directions of the principal pressures or tensions corresponding to the point (x, y, z). Then, in fact, the surface, represented by equation (29), must be a quadratic surface correspondent to this coordinate axes, not only through its center, but also through its axes; and one must therefore have

$$\mathrm{D} = 0, \quad \mathrm{E} = 0, \quad \mathrm{F} = 0.$$

That said, the numerical values of A, B, C will obviously represent the principal pressures or tensions and the formulas (20), (22), (25) will be reduced to

$$p \cos\lambda = \mathrm{A}\cos\alpha, \quad p\cos\mu = \mathrm{B}\cos\beta, \quad p\cos\nu = \mathrm{C}\cos\gamma, \tag{45}$$

$$p^2 = \mathrm{A}^2\cos^2\alpha + \mathrm{B}^2\cos^2\beta + \mathrm{C}^2\cos^2\gamma, \tag{46}$$

$$p\cos\delta = \mathrm{A}\cos^2\alpha + \mathrm{B}\cos^2\beta + \mathrm{C}\cos^2\gamma, \tag{47}$$

while equations (29) and (38) will become

$$\mathrm{Ax}^2 + \mathrm{By}^2 + \mathrm{Cz}^2 = \pm 1, \tag{48}$$

$$\mathrm{A}^2\mathrm{x}^2 + \mathrm{B}^2\mathrm{y}^2 + \mathrm{C}^2\mathrm{z}^2 = 1, \tag{49}$$

Equations (46) and (47) show the relations which exist: 1° between the main pressures or tensions, and the pressure or tension p supported by an arbitrary plane; 2° between the first three forces and projections of the latter on a line perpendicular to the plane in question. The angles α, β, γ included in these same equations are precisely the angles formed by the perpendicular to the plane with the axes along which the principal pressures or tensions are directed.

In the particular case where we consider only points located in the x, y plane, and where we disregard one of the dimensions of the solid body, the formulas (45), (46), (47), (48), (49) can be replaced by the following

$$p\cos\lambda = \mathrm{A}\cos\alpha, \quad p\cos\mu = \mathrm{B}\cos\beta, \tag{50}$$

$$p^2 = \mathrm{A}^2\cos^2\alpha + \mathrm{B}^2\cos^2\beta, \tag{51}$$

$$p\cos\delta = \mathrm{A}\cos^2\alpha + \mathrm{B}\cos^2\beta \tag{52}$$

$$\mathrm{Ax}^2 + \mathrm{By}^2 = \pm 1, \tag{53}$$

$$\mathrm{A}^2\mathrm{x}^2 + \mathrm{B}^2\mathrm{y}^2 = 1, \tag{54}$$

Then the ellipsoids or hyperboloids, mentioned in Theorems II and III, are also reduced to ellipses or conjugate hyperbolas, represented by equations (53) and (54).

In other articles, I will show how one can deduce from the principles established above the equations which express the state of equilibrium or the internal motion of an elastic or non-elastic solid body.

ADDITION TO THE PREVIOUS ARTICLE

The values of $p \cos \lambda$, $p \cos \mu$, $p \cos \nu$ given by the formulas (20) of the previous article, are entirely similar to the values of the rectangular components of the force which would solicit a material point placed in the presence of several fixed centers of attraction or repulsion, and very little displaced from a position in which it remained in equilibrium in the middle of the centers in question. Indeed, let us consider that the material point, after having coincided, in the equilibrium position, with the origin of the coordinates, has been transported to a very small distance and designated by ϱ. Moreover, let

r, r', ... be the vector rays defined from the origin to the various centers fixed;
R, R', ... be the forces of attraction or repulsion which, emanating from the same centers, solicit the material point in the position of equilibrium;
P be the resultant force of those to which this point is subjected after its displacement.

Finally, let

$$\alpha, \beta, \gamma\,; \quad a, b, c\,; \quad a', b', c'\,; \quad \cdots\,; \quad \lambda, \mu, \nu$$

be the angles the semi-axes of the positive coordinates form with: 1° the radius vector ϱ; 2° the vector rays r, r', ...; the direction of the force P. Assuming the material point brought back to the equilibrium position, we will easily establish the equations

$$(1) \qquad \sum(\pm \mathrm{R} \cos a) = 0, \quad \sum(\pm \mathrm{R} \cos b) = 0, \quad \sum(\pm \mathrm{R} \cos c) = 0,$$

where the symbol $\sum$ indicates a sum of similar terms, but relative to the various fixed centers and the sign $\pm$ reduces, sometimes to the sign $-$, sometimes to the sign $+$, depending on whether the force R is repulsive or attractive. Let $f(r)$ now be the function of the distance r which measures the force R. While the material point is transported from the origin to the end of the radius ϱ, the values r, R,

a, b, c will suffer correspondent increments that we will designate with the aid of the characteristic Δ, and we will obviously have

$$(2)\qquad \begin{cases} (r+\Delta r)\cos(a+\Delta a) = r\cos a - \varrho\cos\alpha, \\ (r+\Delta r)\cos(b+\Delta b) = r\cos b - \varrho\cos\beta, \\ (r+\Delta r)\cos(a+\Delta a) = r\cos a - \varrho\cos\gamma; \end{cases}$$

$$(3)\qquad \mathrm{R} + \Delta\mathrm{R} = f(r+\Delta r);$$

$$(4)\qquad \begin{cases} \mathrm{P}\cos\lambda = \sum[\pm(\mathrm{R}+\Delta\mathrm{R})\cos(a+\Delta a)]\,, \\ \mathrm{P}\cos\mu = \sum[\pm(\mathrm{R}+\Delta\mathrm{R})\cos(b+\Delta b)]\,, \\ \mathrm{P}\cos\nu = \sum[\pm(\mathrm{R}+\Delta\mathrm{R})\cos(c+\Delta c)]\,. \end{cases}$$

Thereby, we will find that

$$(5)\qquad \begin{cases} (r+\Delta r)^2 = (r\cos a - \varrho\cos\alpha)^2 + (r\cos b - \varrho\cos\beta)^2 + (r\cos c - \varrho\cos\gamma)^2 \\ \qquad = r^2 - 2r\varrho(\cos a\cos\alpha + \cos b\cos\beta + \cos c\cos\gamma) + \varrho^2\,; \end{cases}$$

and then, by considering the quantity p as infinitely small of the first order and neglecting the infinitely small of the second order, we will conclude from formulas (2), (3), (5) that

$$(6)\qquad \Delta r = -\varrho(\cos a\cos\alpha + \cos b\cos\beta + \cos c\cos\gamma)\,;$$

$$(7)\qquad \mathrm{R} + \Delta\mathrm{R} = f(r) + f'(r)\Delta r = \mathrm{R} + f'(r)\Delta r\,;$$

$$(8)\qquad \begin{cases} \cos(a+\Delta a) = \dfrac{r\cos a - \varrho\cos\alpha}{r+\Delta r} = \cos a - \varrho\dfrac{\cos\alpha}{r} - \dfrac{\cos a}{r}\Delta r\,, \\ \cos(b+\Delta b) = \dfrac{r\cos b - \varrho\cos\beta}{r+\Delta r} = \cos b - \varrho\dfrac{\cos\beta}{r} - \dfrac{\cos b}{r}\Delta r\,, \\ \cos(c+\Delta c) = \dfrac{r\cos c - \varrho\cos\beta}{r+\Delta r} = \cos c - \varrho\dfrac{\cos\beta}{r} - \dfrac{\cos c}{r}\Delta r\,. \end{cases}$$

That said, considering equations (1), we will recognize that formulas (4) can be reduced to

$$(9)\qquad \begin{cases} \mathrm{P}\cos\lambda = \sum\left\{\pm\left[f'(r) - \dfrac{f(r)}{r}\right]\cos a\Delta r \mp \dfrac{f(r)}{r}\varrho\cos\alpha\right\}\,, \\ \mathrm{P}\cos\mu = \sum\left\{\pm\left[f'(r) - \dfrac{f(r)}{r}\right]\cos b\Delta r \mp \dfrac{f(r)}{r}\varrho\cos\beta\right\}\,, \\ \mathrm{P}\cos\nu = \sum\left\{\pm\left[f'(r) - \dfrac{f(r)}{r}\right]\cos c\Delta r \mp \dfrac{f(r)}{r}\varrho\cos\gamma\right\}\,; \end{cases}$$

then, recalling the value of Δr in formula (6), and defining

$$(10)\qquad \begin{cases} \mathrm{A} = \varrho \sum \left\{ \mp \left[f'(r) - \dfrac{f(r)}{r} \right] \cos^2 a \mp \dfrac{f(r)}{r} \right\} , \\ \mathrm{B} = \varrho \sum \left\{ \mp \left[f'(r) - \dfrac{f(r)}{r} \right] \cos^2 b \mp \dfrac{f(r)}{r} \right\} , \\ \mathrm{C} = \varrho \sum \left\{ \mp \left[f'(r) - \dfrac{f(r)}{r} \right] \cos^2 c \mp \dfrac{f(r)}{r} \right\} ; \end{cases}$$

$$(11)\qquad \begin{cases} \mathrm{D} = \varrho \sum \left\{ \mp \left[f'(r) + \dfrac{f(r)}{r} \right] \cos b \cos c \right\} , \\ \mathrm{E} = \varrho \sum \left\{ \mp \left[f'(r) + \dfrac{f(r)}{r} \right] \cos c \cos a \right\} , \\ \mathrm{F} = \varrho \sum \left\{ \mp \left[f'(r) + \dfrac{f(r)}{r} \right] \cos a \cos b \right\} , \end{cases}$$

in order to shorten expressions, we will definitely find

$$(12)\qquad \begin{cases} \mathrm{P} \cos \lambda = \mathrm{A} \cos \alpha + \mathrm{F} \cos \beta + \mathrm{E} \cos \gamma , \\ \mathrm{P} \cos \mu = \mathrm{F} \cos \alpha + \mathrm{B} \cos \beta + \mathrm{D} \cos \gamma , \\ \mathrm{P} \cos \nu = \mathrm{E} \cos \alpha + \mathrm{D} \cos \beta + \mathrm{C} \cos \gamma . \end{cases}$$

The formulas (12), as well as several propositions which are deduced from them and which are analogous to theorems I, II, III of the previous article, are due to Mr. Fresnel, who presented them in his *Recherches sur la double réfraction*.

Bibliography

[1] Aristotle. **The Complete Works of Aristotle.** 1. ed. Princeton: Princeton University Press, 1984.

[2] Backus, G. **Continuum Mechanics.** Available at <http://samizdat.mines.edu/backus/>, Accessed on feb 2004.

[3] Belhoste, B. **Augustin-Louis Cauchy: A Biography.** 1. ed. Berlim: Springer-Verlag, 1991.

[4] Benvenuto, E. **An Introduction to the History of Structural Mechanics, Part I: Statics and Resistance of Solids**. 1. ed. New York: Springer-Verlag, 1991.

[5] Bernoulli, Jakob. *Curvatura Laminae Elasticae. Ejus Identitas cum Curvatura Lintei a pondere inclusi fluidi expansi, &c.* in **Acta Eruditorum**, 1694, pp. 262–276.

[6] Bernoulli, Jakob. *Explicationes, annotationes et additiones ad ea, quae in actis sup. anni de curva elastica, isochrona paracentrica, et velaria, &c.* in **Acta Eruditorum**, 1695, pp. 537–553.

[7] Bertram, A. **Elasticity and Plasticity of Large Deformations**. 3. ed. Berlim: Springer-Verlag, 2012.

[8] Bradley, R.; Sandifer, C. **Cauchy's Cours d'analyse: An Annotated Translation.** 1. ed. New York: Springer-Verlag, 2009.

[9] Brandon, R. **Large Deformation Kinematics.** Available at <http://www.me.unm.edu/~rmbrann/gobag.html>, Accessed on jan 2007.

[10] Calinger, R. **Leonhard Euler: Mathematical Genius in the Enlightenment.** 1. ed. Princeton: Princeton University Press, 2016.

[11] Cauchy, A. *Recherches sur L'équilibre et Lemouvement Intérieur des Corps Solides ou Fluides, Élastiques ou Non Élastiques* in **Bulletin Des Sciences de La Société Philomatique de Paris**, January, 1823, pp. 9–13.

[12] Cauchy, A. *Sur La Pression ou Tension Dans Les Corps Solides* in **Exercices de Mathématiques**, March, 1827, pp. 42–57.

[13] Cauchy, A. *Sur Quelques Propriétés des Polyèdres* in **Œuvres Complétes d'Augustin Cauchy**, Série II, Tome VII, 1889, pp. 55–59.

[14] Cauchy, A. *De La Pression ou Tension Dans Un Corps Solide* in **Œuvres Complétes d'Augustin Cauchy**, Série II, Tome VII, 1889, pp. 60–81.

[15] Chadwick, P. **Continuum Mechanics: Concise Theory and Problems.** 1. ed. Mineola: Dover Publications, 1999.

[16] Ciarlet, P. G. **Mathematical Elasticity – Volume 1: Three Dimensional Elasticity.** 1. ed. Amsterdam: Elsevier Science Publishers BV, 1988.

[17] Corben, H. C.; Stehle, P. **Classical Mechanics.** 2. ed. New York: Dover Publications, 1994.

[18] Doyle, T.; Ericksen, J. *Non Linear Elasticity* in **Advances in Applied Mechanics**, Vol IV, 1956, pp. 52–115.

[19] Drozdov, A. D. **Finite Elasticity and Viscoelasticity: a Course in the Nonlinear Mechanics of Solids.** 1. ed. Farrer Road: World Scientific Publishing Co., 1996.

[20] Dugas, R. **A History of Mechanics.** 1. ed. New York: Dover Publications, 1988.

[21] Euler, L. **Scientia Navalis, Volume 1.** 1. ed. St. Petersburg: Academy of Sciences Press, 1749.

[22] Fellmann, E. **Leonhard Euler.** 1. ed. Basel: Birkhäuser Verlag, 2007.

[23] Galileo Galilei. **Dialogs Concerning Two New Sciences.** 1. ed. New York: Dover Publications, 1954.

[24] Galileo Galilei. **Two New Sciences Including Centers of Gravity and Force of Percussion.** 2. ed. Toronto: Wall & Thompson, 1989.

[25] Gurtin, M. E. **An Introduction to Continuum Mechanics.** 1. ed. San Diego: Academic Press, 1981.

[26] Hanyga, A. **Mathematical Theory of Non-Linear Elasticity.** Translated by R. W. Ogden. 1. ed. Chichester/New York/Warsaw: Elis Horwood Limited/John Wiley & Sons/PWN-Polish Scientific Publishers, 1985.

[27] Haupt, P. **Continuum Mechanics and Theory of Materials.** 2. ed. Berlim: Springer-Verlag, 2002.

[28] Heilbron, J.L. **Galileo.** 1. ed. New York: Oxford University Press, 2010.

[29] Higdon, A. et al. **Mecânica dos Materiais.** 3. ed. Rio de Janeiro: Guanabara Dois, 1981.

[30] Hill, R. On Constitutive Inequalitites for Simple Materials - I. **J. Mech. Phys. So.** Oxford: Pergamon Press, vol. 16, pp. 229-242, 1968.

[31] Hill, R. Aspects of Invariance in Solid Mechanics. **Advances in Applied Mechanics.** New York: Academic Press, vol. 18, pp. 1-75, 1978.

[32] Hooke, R. **Lectures *de Potentia Restitutiva*, Or of Spring: Explaining the Power of Springing Bodies.** 1. ed. London: John Martyn, 1678.

[33] Hooke, R. **The Posthumous Works of Robert Hooke.** 1. ed. London: Richard Waller, 1705.

[34] Love, A. E. H. **A Treatise on the Mathematical Theory of Elasticity.** 4. ed. New York: Dover Publications, 1944.

[35] Lubarda, V. A. **Elastoplasticity Theory.** 1. ed. Boca Raton: CRC Press, 2002.

[36] Lund, J.; Byrne, J. *Leonardo Da Vinci's Tensile Strength Tests: Implications For The Discovery Of Engineering Mechanics* in **Civil. Eng. and Env. Syst.** Vol. 00, 2000, pp. 1-8.

[37] Malvern, L. E. **Introduction to the Mechanics of a Continuous Medium.** 1. ed. Englewood Cliffs: Prentice-Hall, 1969.

[38] Mariotte, E. **Traité du mouvement des eaux et des autres corps fluides.** 2. ed. Paris: Jean Néaulme, 1700.

[39] Marsden, J. E.; Hughes, T. J. R. **Mathematical Foundations of Elasticity.** 1. ed. Mineola: Dover Publications, 1994.

[40] Maugin, G. **Material Inhomogeneities in Elasticity.** 1. ed. New York: Chapman & Hall, 1993.

[41] Maugin, G. **Continuum Mechanics Through the Twentieth Century: A Concise Historical Perspective.** 1. ed. Dordrecht: Springer, 2013.

[42] Maugin, G. **Continuum Mechanics Through the Eighteenth and Nineteenth Centuries.** 1. ed. Switzerland: Springer, 2014.

[43] Maugin, G. **Continuum Mechanics Through the Ages - From the Renaissance to the Twentieth Century.** 1. ed. Switzerland: Springer, 2016.

[44] Maxwell, J. C. **Matter and Motion.** 1. ed. New York: Dover Publications, 1991.

[45] Mikhailov, G. *Daniel Bernoulli, Hydrodynamica (1738)* in **Landmark Writings in Western Mathematics 1640–1940.** 1. ed. Amsterdam: Elsevier Science, 2005, pp. 131–142.

[46] Newton, I. **The Principia: Mathematical Principles of Natural Philosophy.**, translated by I. Bernard Cohen and Anne Whitman, 1. ed. Berkeley: University of California Press, 1999.

[47] Noll, W. **The Foundations of Mechanics and Thermodynamics: Selected Papers.** 1. ed. Berlim: Springer-Verlag, 1974.

[48] Ogden, R. W. **Non-Linear Elastic Deformations.** 1. ed. Mineola: Dover Publications, 1997.

[49] Oldfather, W. et al. *Leonhard Euler's Elastic Curves* in **Isis**, Vol. 20, No. 1, Chicago: The University of Chicago Press, 1933, pp. 72–160.

[50] Podio-Guidugli, P. A Primer on Elasticity. **Journal of Elasticity.** Berlim: Springer-Verlag, vol. 40, pp. 1–104, 2000.

[51] Popov, E. **Engineering Mechanics of Solids.** 1. ed. Englewood Cliffs: Prentice Hall, 1990.

[52] Prager, W. **Introduction to Mechanics of Continua.** 1. ed. Mineola: Dover Publications, 2004.

[53] Sedov, L. I. **Mechanics of Continuous Media - Volume 1.** Translated by J. P. Nowacki. 1. ed. Farrer Road: World Scientific Publishing Co., 1997.

[54] Seth, B. *Generalized Strain Measure with Applications to Physical Problems* in **MRC Technical Summary Report #248**, July, 1961.

[55] Šilhavý, M. **The Mechanics and Thermomechanics of Continuous Media.** 1. ed. Berlim: Springer-Verlag, 1997.

[56] Slaugther, W. S. **The Linearized Theory of Elasticity**. 1. ed. Boston: Birkhäuser, 2002.

[57] Sokolnikoff, I. S. **Mathematical Theory of Elasticity.** 2. ed. Malabar: Robert Krieger Publishing Company, 1983.

[58] Spivak, M. **A Comprehensive Introduction to Differential Geometry – Volume 1** 3 ed. Huston: Publish or Perish Inc, 2005.

[59] Talpaert, Y. R. **Tensor Analysis and Continuum Mechanics.** 1. ed. Dordrecht: Kluwer Academic Publishers, 2002.

[60] Temam, R. M.; Miranville, A. M. **Mathematical Modeling in Continuum Mechanics.** 2. ed. Cambridge: Cambridge University Press, 2005.

[61] Timoshenko, S. **History of Strength of Materials.** 1 ed. Mineola: Dover Publications, 1983.

[62] Tolstoy, L. **War & Peace.** 1 ed. London: The Folio Society Ltd, 2014.

[63] Truesdell, C. A. *Rational fluid mechanics, 1687–1765: Editor's introduction to vol. II 12 of Euler's works* in **Euleri Opera Omnia, Series II, vol. 12.** 1. ed. Zurich: Füssli, 1954, pp. 1–125.

[64] Truesdell, C. A. **The Rational Mechanics of Flexible or Elastic Bodies: 1638–1788**. 1. ed. Basel: Birkhäuser, 1960.

[65] Truesdell, C. A. **Essays in the History of Mechanics**. 1. ed. New York: Springer-Verlag, 1968.

[66] Truesdell, C. A. **A First Course in Rational Continuum Mechanics – Volume 1.** 1. ed. New York: Academic Press, 1977.

[67] Truesdell, C. A.; Noll, W. **The Non-Linear Field Theories of Mechanics.** 3. ed. New York: Springer-Verlag, 2003.

[68] Wallace, W. *Galileo's Pisan studies in science and philosophy* in **The Cambridge Companion to Galileo.** 1. ed. Cambridge: Cambridge University Press, 1998, pp. 27–52.

[69] Wang, C.-C; Truesdell, C. A. **Introduction to Rational Elasticity.** 1. ed. Leyden: Noordhoff International Publishing, 1973.

[70] Weyl, H. **Space Time Matter.** 4. ed. Mineola: Dover Publications, 1952.

[71] Zammattio, C. et al. **Leonardo The Scientist.** 1. ed. New York: McGraw-Hill, 1980.

Notation and Symbols

$A^{\times n}$	Cartesian product of n sets A_i, 10
A^n	Cartesian product of n sets A, 10
f	Function f, 11
$f(d)$	Value of the function f on d, 11
D_f	Domain of the function f, 11
$\mapsto$	Mapped to..., 11
$\mathbb{R}$	Real numbers, 11
R_f	Image of function f, 12
R_B^{-1}	Preimage of set B, 12
$\Leftrightarrow$	Bidiretional implication, 12
f^{-1}	Inverse function of f, 12
G	Group, 14
$\exists!$	There is one and only one, 14
$\sum_{i=1}^{n}$	Sum of n terms, 16
$\mathbb{N}^+$	Natural numbers without zero, 16
$\mathbb{F}$	Field defined by F, 16
$\mathbb{C}$	Complex numbers, 16
$\Re$	Real part of..., 16
$\Im$	Imaginary part of..., 16
H	Array, 17
$H_{i_1 \cdots i_q}$	Element of array, 17
$A *_q B$	q-product of arrays A and B, 17
$A : B$	Scalar product of arrays A and B, 17
det	Determinant of..., 20
A^{T}	Transpose matrix of A, 21
$\overline{F_{ji}}$	Complex conjugate of scalar F_{ji}, 22
$H^\dagger$	Conjugate transpose of matrix H, 22
i	Imaginary number, 22
tr	Trace of..., 24
sgn	Sign of real number..., 27
adj	Adjugate matrix of..., 28
$\mathbf{x}$	Vector x, 31
$V_{\mathbb{F}}$	Vector space of V on $\mathbb{F}$, 32
$\overline{V_{\mathbb{F}}}$	Conjugate vector space of V on $\mathbb{F}$, 32
$\mathbf{v}^c$	Element of a conjugate vector space, 32
$\mathrm{span}(U)$	Subset spanned by U, 33
$\dim(V_{\mathbb{F}})$	Dimension of $V_{\mathbb{F}}$, 34
$\mathbb{R}^+$	Nonnegative real numbers, 34
$\lVert\mathbf{x}\rVert$	Norm of vector $\mathbf{x}$, 35
$\mathbf{x} \cdot \mathbf{y}$	Inner product of vectors $\mathbf{x}$ and $\mathbf{y}$, 35
$\perp$	Orthogonality, 36
$\hat{\mathbf{u}}_1$	Unitary vector, 36

$\widehat{A}_3$	Interior of set A_3, 38
$\{\hat{\mathbf{e}}_1, \cdots, \hat{\mathbf{e}}_n\}$	Natural basis, 41
$(x_1, \cdots, x_n)$	Natural coordinates of vector $\mathbf{x}$, 41
$V_{\mathbb{F}}^{U}$	Function space whose functions map U to V, 42
$\mathcal{C}_{\mathbb{F}}(U, V)$	Function space of continuous functions that map $U_{\mathbb{F}}$ to $V_{\mathbb{F}}$, 43
$\mathbb{R}_*^+$	Positive real numbers, 43
$\mathcal{L}_{\mathbb{F}}(U, V)$	Function space of linear functions that map U to V, 43
$\mathcal{CL}_{\mathbb{F}}(Z, Y)$	Function space of continuous linear functions that map Z to Y, 44
$\mathbf{f}_i^B$	i-th coordinate functional of basis B, 44
$\mathbf{u}^*$	Covector of $\mathbf{u}$, 46
$\mathcal{G}_{\mathbb{F}}(V)$	Group of unary invertible operators on domain $V_{\mathbb{F}}$, 51
$\mathcal{N}_{\mathbb{F}}(V)$	Unitary group on domain $V_{\mathbb{F}}$, 51
$\mathcal{I}_{\mathbb{F}}(V)$	Isometry group on domain $V_{\mathbb{F}}$, 51
$[\mathbf{u}]^B$	Matrix of vector $\mathbf{u}$ on basis B, 52
$[\mathbf{g}_B]^C$	Matrix of linear function $\mathbf{g}$ on bases B and C, 53
$\mathcal{U}^+{}_{\mathbb{F}}(Y)$	Proper unitary group on domain $Y_{\mathbb{F}}$, 57
$\mathcal{U}^-{}_{\mathbb{F}}(Y)$	Improper unitary group on domain $Y_{\mathbb{F}}$, 57
$\boldsymbol{\mathcal{L}}_{\mathbb{F}}(U^{\times m})$	Tensor space of order m defined by $(U_i)_{\mathbb{F}}$, 68
$\mathbf{T}$	Tensor, 68
$\boldsymbol{\mathcal{L}}_{\mathbb{F}}^{(n)}(U^{\times m})$	Tensor space of order m where $\dim(U_i)_{\mathbb{F}} = n$, 68
$\overline{\boldsymbol{\mathcal{L}}_{\mathbb{F}}(U^{\times m})}$	Conjugate tensor space of $\boldsymbol{\mathcal{L}}_{\mathbb{F}}(U^{\times m})$, 68
$\mathbf{T}^c$	Element of a conjugate tensor space, 68
$\boldsymbol{\mathcal{L}}_{\mathbb{F}}(V^{(p,q)})$	Type (p, q) tensor space, 68
$\boldsymbol{\mathcal{LS}}_{\mathbb{F}}(V^m)$	Symmetric tensor space, 73
$\boldsymbol{\mathcal{LA}}_{\mathbb{F}}(V^m)$	Antisymmetric tensor space, 73
$\mathbf{I}$	Identity tensor, 77
$\mathbf{c}^{\otimes}$	Tensor function lifted from multiantilinear $\mathbf{c}$, 79
$\mathbf{c}_p^{\otimes}$	Contraction of order p of..., 80
$\mathrm{tr}_{r,s}$	r, s trace of..., 80
$\mathbf{A} \diamond_p \mathbf{B}$	p-th order contractive product of tensors, 81
$\mathbf{A} \diamond \mathbf{B}$	Contractive product of tensors, 82
$\mathbf{A} \odot_p \mathbf{B}$	p-th order partial inner product of tensors, 83
$\mathbf{A} \hat{\odot}_p \mathbf{B}$	vector from a partial inner product of tensors, 83
$\mathbf{T}_p^{\circledast}$	p-cotensor of $\mathbf{T}$, 87
$\mathbf{T}^{\circledast}$	Cotensor of $\mathbf{T}$, 88
$\overline{\mathbf{T}^{\circledast}}$	Representative function of $\mathbf{T}$, 89
$\boldsymbol{g}$	Representative function of second order tensor $\mathbf{G}$, 90
$\boldsymbol{\mathcal{LN}}_{\mathbb{F}}(N^{\times m})$	Isotropic tensor space, 100
$\boldsymbol{\mathcal{LN}}_{\mathbb{F}}^{+}(N^{\times m})$	Anti isotropic tensor space, 100
$\mathbf{u} \oplus a$	Point defined by the action of vector $\mathbf{u}$ on point a, 103

Index