Notes on Multi-Linear Algebra and Tensor Calculus

(For the course Geometrical Methods in Mathematical Physics)

by **Valter Moretti**
www.science.unitn.it/∼moretti/homeE.html

Department of Mathematics,
Faculty of Science,
University of Trento
Italy

Academic year 2005-2006

Contents

Acknowledgments.

The author is grateful to R. Aramini who read these notes carefully correcting several misprints and to D. Ansari for helpful suggestions to improve the text.

1 Multi-linear Mappings and Tensors.

Within this section we introduce basic concepts concerning multi-linear algebra and tensors. The theory of vector spaces and linear mappings is assumed to be well known.

1.1 Dual space and conjugate space.

As a first step we introduce the dual space and the conjugate space of a given vector space.

Def.1.1. (Dual Space, Conjugate Dual Space and Conjugate space.) *Let V be a vector space on the field either $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.*
(a) *The* **dual space** *of V, V^*, is the vector space of linear functionals on V, i.e., the linear mappings $f : V \to \mathbb{K}$.*
(b) *If $\mathbb{K} = \mathbb{C}$, the* **conjugate dual space** *of V, $\overline{V^*}$, is the vector space of anti-linear functionals on V, i.e., the antilinear mappings $g : V \to \mathbb{C}$. Finally the* **conjugate space** *of V, $\overline{V}$ is the space $(\overline{V^*})^*$*

Comments.
(1) If V and V' are vector spaces on $\mathbb{C}$, a mapping $f : V \to V'$ is called *anti linear* or *conjugate linear* if it satisfies

$$f(\alpha u + \beta v) = \overline{\alpha} f(u) + \overline{\beta} f(v)$$

for all $u, v \in V$ and $\alpha, \beta \in \mathbb{C}$, $\overline{\lambda}$ denoting the complex conjugate of $\lambda \in \mathbb{C}$. If $V' = \mathbb{C}$ the given definition reduces to the definition of *anti-linear functional*.
(2) V^*, $\overline{V^*}$ and $\overline{V}$ turn out to be vector spaces on the field $\mathbb{K}$ when the composition rule of vectors and the product with elements of the field are defined in the usual way. For instace, if $f, g \in V^*$ or $\overline{V^*}$, and $\alpha \in \mathbb{K}$ then $f + g$ and αf are functions such that:

$$(f + g)(u) := f(u) + g(u)$$

and

$$(\alpha f)(u) := \alpha f(u)$$

for all of $u \in V$.

Def.1.2. (Dual Basis, Conjugate Dual Basis and Conjugate Basis.) *Let V be a finite-dimensional vector space on either $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Let $\{e_i\}_{i\in I}$ be a vector basis of V. The set $\{e^{*j}\}_{j\in I} \subset V^*$ whose elements are defined by*

$$e^{*j}(e_i) := \delta^j_i$$

for all $i, j \in I$ is called the **dual basis** *of $\{e_i\}_{i\in I}$.*
Similarly, if $\mathbb{K} = \mathbb{C}$, the set of elements $\{\overline{e^{*j}}\}_{j\in I} \subset \overline{V^*}$ defined by:

$$\overline{e^{*j}}(e_i) := \delta^j_i$$

for all $i, j \in I$ *is called the* **conjugate dual basis** *of* $\{e_i\}_{i\in I}$.
Finally the set of elements $\{\overline{e}_j\}_{j\in I} \subset \overline{V}$ *defined by:*

$$\overline{e}_p(\overline{e^{*q}}) := \delta_q^q$$

for all $p, q \in I$ *is called the* **conjugate basis** *of* $\{e_i\}_{i\in I}$.

Comments.
(1) We have explicitly assumed that V is finite dimensional. Anyway, each vector space admits a vector basis (i.e. a possibly infinite set of vectors such that each vector of V can be obtained as a finite linear combination of those vectors), therefore one could discuss on the validity of the given definitions in the general case of a non-finite dimensional vector space. We shall not follow that way because we are interested on algebraic features only and the infinite-dimensional case should be approached by convenient topological tools which are quite far form the goals of these introductory notes.
(2) If the vector space V with field $\mathbb{K}$ admits the vector basis $\{e_i\}_{i\in I}$, each linear or anti-linear mapping $f : V \to \mathbb{K}$ is completely defined by giving the values $f(e_i)$ for all of $i \in I$. This is because, if $v \in V$ then $v = \sum_{i\in I_v} c^i e_i$ for some numbers $c^i \in \mathbb{K}$, $I_v \subset I$ being *finite*. Then the linearity of f yelds $f(v) = \sum_{i\in I_v} c^i f(e_i)$. (This fact holds true no matter if $dimV < +\infty$). Using the result above, one realizes that the given definitions of elements e^{*j} and $\overline{e^*}^i$ completely determine these functionals and thus the given definitions are well-posed.
(3) One may wonder whether or not the dual basis and the conjugate basis are *proper* vector bases of respective vector spaces, the following theorem gives a positive answer.
(4) In *spinor* theory, if V is a complex two-dimensional space: The "upper-index spinor space", $\overline{V}$ is the space of "upper-index pointed spinors" or "upper-index primed spinors". $\overline{V^*}$ is the space of "lower-index pointed spinors" or "lower-index primed spinors". Finally V^* is the space of "upper-index spinors".

Theorem 1.1. *Concerning* **Def.1.2** *the dual basis, the conjugate dual basis and the conjugate basis of a base* $\{e_i\}_{i\in I} \subset V$, *with* $dimV < \infty$, *are vector bases for* V^*, $\overline{V^*}$ *and* $\overline{V}$ *respectively. As a consequence* $dimV = dimV^* = dim\overline{V^*} = dim\overline{V}$.

Proof. Consider the dual basis in V^*, $\{e^{*j}\}_{j\in I}$. We have to show that the functionals $e^{*j} : V \to \mathbb{K}$ are generators of V^* and are linearly independent.
(Generators.) Let us show that, if $f : V \mapsto \mathbb{K}$ is linear, then there are numbers $c_j \in \mathbb{K}$ such that $f = \sum_{j\in I} c_j e^{*j}$.
To this end define $f_j := f(e_j)$, $j \in I$, then we argue that $f = f'$ where $f' := \sum_{j\in I} f_j e^{*j}$. Indeed, any $v \in V$ may be decomposed as $v = \sum_{i\in I} v^i e_i$ and, by linearity, we have:

$$f'(v) = \sum_{j\in I} f_j e^{*j}(\sum_{i\in I} v^i e_i) = \sum_{i,j\in I} f_j v^i e^{*j}(e_i) = \sum_{i,j\in I} f_j v^i \delta_i^j = \sum_{j\in I} v^j f_j = \sum_{j\in I} v^j f(e_j) = f(v)\,.$$

(Notice that above we have used the fact that one can extract the summation symbol from the argument of each e^{*j}, this is because the sum on the index i is *finite* by hypotheses it being

$dimV < +\infty$.) Since $f'(v) = f(v)$ holds for all of $v \in V$, we conclude that $f' = f$.
(Linear independence.) We have to show that if $\sum_{k\in I} c_k e^{*k} = 0$ then $c_k = 0$ for all $k \in I$.
To achieve that goal notice that $\sum_{k\in I} c_k e^{*k} = 0$ means $\sum_{k\in I} c_k e^{*k}(v) = 0$ for all $v \in V$. Therefore, putting $v = e_i$ and using the definition of the dual basis, $\sum_{k\in I} c_k e^{*k}(e_i) = 0$ turns out to be equivalent to $c_k \delta_i^k = 0$, namely, $c_i = 0$. This result can be produced for each $i \in I$ and thus $c_i = 0$ for all $i \in I$. The proof for the conjugate dual basis is very similar and is left to the reader. The proof for the conjugate basis uses the fact that $\overline{V}$ is the dual space of $\overline{V^*}$ and thus the first part of the proof applies. The last statement of the thesis holds because, in the four considered cases, the set of the indices I is the same. $\Box$

Notation 1.1. *From now on we take advantage of the following convention. Whenever an index appears twice, once as an upper index and once as a lower index, in whatever expression, the summation over the values of that index is understood in the notation. E.g.,*

$$t_{ijkl} f^{rsil}$$

means

$$\sum_{i,l} t_{ijkl} f^{rsil} .$$

The range of the various indices should be evident from the context or otherwise specified.

We are naturally lead to consider the following issue. It could seem that the definition of dual space, V^* (of a vector space V) may be implemented on V^* itself producing the double dual space $(V^*)^*$ and so on, obtaining for instance $((V^*)^*)^*$ and, by that way, an infinite sequence of dual vector spaces. The theorem below shows that this is not the case because $(V^*)^*$ turns out to be *naturally isomorphic* to the initial space V and thus the apparently infinite sequence of dual spaces ends on the second step. We remind the reader that a vector space isomorphism $F : V \to V'$ is a linear map which is also one-to-one, i.e., injective and surjective. An isomorphism is called *natural* when it is built up using the definition of the involved algebraic structures only and it does not depend on "arbitrary choices". A more precise definition of *natural isomorphism* may be given by introducing the *theory of mathematical categories*.

Theorem 1.2. *If V is a finite-dimensional vector space on the field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, there is a natural isomorphism $F : V \to (V^*)^*$ given by*

$$(F(v))\,(u) := u(v) ,$$

for all $u \in V^*$ and $v \in V$.

Proof. Notice that $F(v) \in (V^*)^*$ because it is a linear functional on V^*:

$$(F(v))(\alpha u + \beta u') := (\alpha u + \beta u')(v) = \alpha u(v) + \beta u'(v) =: \alpha(F(v))(u) + \beta(F(v))(u') .$$

Let us prove that $v \mapsto F(v)$ with $(F(v))(u) := u(v)$ is linear, injective and surjective.
(Linearity.) We have to show that, for all $\alpha, \beta \in \mathbb{K}$, $v, v' \in V$,

$$F(\alpha v + \beta v') = \alpha F(v) + \beta F(v') \, .$$

This is equivalent to, by the definition of F given above,

$$u(\alpha v + \beta v') = \alpha u(v) + \beta u(v') \, ,$$

for all $u \in V^*$. This is obvious because, u is a linear functional on V.
(Injectivity.) We have to show that $F(v) = F(v')$ implies $v = v'$. $(F(v))(u) = (F(v'))(u)$ can be re-written as $u(v) = u(v')$ or $u(v - v') = 0$. In our hypotheses, this holds true for all $u \in V^*$. Then, define $e_1 := v - v'$ and notice that, if $v - v' \neq 0$, one can complete e_1 with other vectors to get a vector basis of V, $\{e_i\}_{i \in I}$. Since u is arbitrary, choosing $u = e^{*1}$ we should have $u(e_1) = 1$ by definition of dual basis but this contradicts $u(e_1) = 0$, i.e., $u(v - v') = 0$. By consequence $v - v' = 0$ and $v = v'$.
(Surjectivity.) We have to show that if $f \in (V^*)^*$, there is $v_f \in V$ such that $F(v_f) = f$.
Fix a basis $\{e_i\}_{i \in I}$ in V and the dual one in V^*. $v_f := f(e^{*i})e_i$ fulfills the requirement $F(v_f) = f$.
□

Def.1.3. (Pairing.) *Let V be a finite-dimensional vector space on $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$ with dual space V^*. The bi-linear map $\langle \, ., \, . \rangle : V \times V^* \to \mathbb{K}$ such that*

$$\langle u, v \rangle := v(u)$$

for all $u \in V$, $v \in V^*$, is called **pairing**.

Comment. Because of the theorem proven above, we may indifferently think $\langle u, v \rangle$ as representing either the action of $u \in (V^*)^*$ on $v \in V^*$ or the action of $v \in V^*$ on $u \in V$.

Notation 1.2. From now on

$$V \simeq W$$

indicates that the vector spaces V and W are isomorphic under some natural isomorphism. If the field of V and W is $\mathbb{C}$,

$$V \overline{\simeq} W$$

indicates that there is a natural anti-isomorphism, i.e. there is an injective, surjective, anti-linear mapping $G : V \to W$ built up using the definition of V and $\overline{V}$ only.

Exercises 1.1.
1.1.1. Show that if $v \in V$ then $v = \langle v, e^{*j} \rangle e_j$, where $\{e_j\}_{j \in I}$ is any basis of the finite dimensional vector space V.
(Hint. Decompose $v = c^i e_i$, compute $\langle v, e^{*k} \rangle$ taking the linearity of the the left entrance into

account. Alternatively, show that $v - \langle v, e^{*j}\rangle e_j = 0$ proving that $f(v - \langle v, e^{*j}\rangle e_j) = 0$ for every $f \in V^*$.)
1.1.2. Show that $V^* \simeq \overline{V^*}$ if $dim V < \infty$ (and the field of V is $\mathbb{C}$). Similarly, show that $V \simeq \overline{V}$ under the same hypotheses.
(Hint. The anti-isomorphism $G : V^* \to \overline{V^*}$, is defined by $(G(v))(u) := \overline{\langle v, u\rangle}$ and the anti-isomorphism $F : V \to \overline{V}$, is defined by $(F(v))(u) := \langle v, G^{-1}(u)\rangle$.)
1.1.3. Show that if the finite-dimensional vector spaces V and V' are isomorphic or anti-isomorphic, then V^* and V'^* are so.
(Hint. If the initial (anti-) isomorphism is $F : V \to V'$ consider $G : V'^* \to V^*$ defined by $\langle F(u), v'\rangle = \langle u, G(v')\rangle$.)
1.1.4. Show that $\overline{V}^* \simeq \overline{V^*}$ if V is a finite-dimensional vector space on $\mathbb{C}$.
(Hint. Use Theorem 1.2.)
1.1.5. Write explicitly the (anti-)isomorphisms in **1.1.2**, **1.1.3**, **1.1.4**, with respect to bases $\{e_i\}_{i\in I} \subset V$, $\{\overline{e}_i\}_{i\in I} \subset \overline{V}$ and the corresponding ones in V^*, $\overline{V^*}$, $\overline{V}$, $\overline{V^*}^*$.

1.2 Multi linearity: tensor product, tensors, universality theorem.

Let us consider $n \geq 1$ vector spaces $V_1, \ldots, V_n$ on the common field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ and another vector space W on $\mathbb{K}$. In the following $\mathcal{L}(V_1, \ldots, V_n|W)$ denotes the vector space (on $\mathbb{K}$) of *multi-linear maps* from $V_1 \times \ldots \times V_n$ to W. We recall the reader that a mapping $f : V_1 \times \ldots \times V_n \to W$ is said to be *multi linear* if, arbitrarily fixing $n-1$ vectors, $v_i \in V_i$ for $i \neq k$, each mapping $v_k \mapsto f(v_1, \ldots, v_k, \ldots, v_n)$ is linear for every $k = 1, \ldots, n$. We leave to the reader the trivial proof of the fact that $\mathcal{L}(V_1, \ldots, V_n|W)$ is a *vector space* on $\mathbb{K}$, with respect to the usual sum of pair of functions and product of an element of $\mathbb{K}$ and a function. If $W = \mathbb{K}$ we use the shorter notation $\mathcal{L}(V_1, \ldots, V_n) := \mathcal{L}(V_1, \ldots, V_n|\mathbb{K})$, $\mathbb{K}$ being the common field of $V_1, \ldots, V_n$.

Exercises 1.2.
1.2.1. Suppose that $\{e_{k,i_k}\}_{i_k\in I_k}$ are bases of the vector spaces V_k, $k = 1, \ldots, n$ on the same field $\mathbb{K}$. Suppose also that $\{e_i\}_{i\in I}$ is a basis of the vector space W on $\mathbb{K}$. Show that each $f \in \mathcal{L}(V_1, \ldots, V_n|W)$ satisfies

$$f(v_1, \ldots, v_n) = v_1^{i_1} \cdots v_n^{i_n} \langle f(e_{1,i_1}, \ldots, e_{n,i_n}), e^{*i}\rangle e_i \,,$$

for all $(v_1, \ldots, v_n) \in V_1 \times \cdots \times V_n$.

We have a first fundamental and remarkable theorem. To introduce it we employ three steps.
(1) Take $f \in \mathcal{L}(V_1, ..., V_n|W)$. This means that f associates every string $(v_1, ..., v_n) \in V_1 \times \cdots \times V_n$ with a corresponding element $f(v_1, \ldots, v_n) \in W$, and the map

$$(v_1, \ldots, v_n) \mapsto f(v_1, \ldots, v_n)$$

is multilinear.
(2) Since, for *fixed* $(v_1, \ldots, v_n)$, the vector $f(v_1, \ldots, v_n)$ is an element of W, the action of $w^* \in W^*$

on $f(v_1,\ldots,v_n)$ makes sense, producing $\langle f(v_1,\ldots,v_n), w^*\rangle \in \mathbb{K}$.
(3) Allowing $v_1,\ldots v_n$ and w^* to range freely in the corresponding spaces, the map Ψ_f with

$$\Psi_f : (v_1,\ldots,v_n,w^*) \mapsto \langle f(v_1,\ldots,v_n), w^*\rangle ,$$

turns out to be multilinear by construction. Hence, by definition $\Psi_f \in \mathcal{L}(V_1,\ldots,V_n,W^*)$.
The theorem concerns the map F which associates f with Ψ_f.

Theorem 1.3. *If $V_1,\ldots,V_n$ are finite-dimensional vector space on the common field $\mathbb{K}$ ($=\mathbb{C}$ or $\mathbb{R}$), and W is another finite-dimensional vector spaces on $\mathbb{K}$, the vector spaces $\mathcal{L}(V_1,\ldots,V_n|W)$ and $\mathcal{L}(V_1,\ldots,V_n,W^*)$ are naturally isomorphic by means of the map $F : \mathcal{L}(V_1,\ldots,V_n|W) \to \mathcal{L}(V_1,\ldots,V_n,W^*)$ with $F : f \mapsto \Psi_f$ defined by*

$$\Psi_f(v_1,\ldots,v_n,w^*) := \langle f(v_1,\ldots,v_n), w^*\rangle ,$$

for all $(v_1,\ldots,v_n) \in V_1 \times \ldots \times V_n$ and $w^* \in W^*$.

Proof. Let us consider the mapping F defined above. We have only to establish that F is linear, injective and surjective. This ends the proof.
(Linearity.) We have to prove that $\Psi_{\alpha f+\beta g} = \alpha\Psi_f + \beta\Psi_g$ for all $\alpha,\beta \in \mathbb{K}$ and $f,g \in \mathcal{L}(V_1,\ldots,V_n|W)$.
In fact, making use of the left-hand linearity of the pairing, one has

$$\Psi_{\alpha f+\beta g}(v_1,\ldots,v_n,w^*) = \langle(\alpha f+\beta g)(v_1,\ldots,v_n), w^*\rangle = \alpha\langle f(v_1,\ldots,v_n), w^*\rangle + \beta\langle g(v_1,\ldots,v_n), w^*\rangle$$

and this is nothing but:

$$\Psi_{\alpha f+\beta g}(v_1,\ldots,v_n,w^*) = (\alpha\Psi_f + \beta\Psi_g)(v_1,\ldots,v_n,w^*) .$$

Since $(v_1,\ldots,v_n,w^*)$ is arbitrary, the thesis is proven.
(Injectivity.) We have to show that if $\Psi_f = \Psi_{f'}$ then $f = f'$.
In fact, if $\Psi_f(v_1,\ldots,v_n,w^*) = \Psi_{f'}(v_1,\ldots,v_n,w^*)$ for all $(v_1,\ldots,v_n,w^*)$, using the definition of Ψ_g we have $\langle f(v_1,\ldots,v_n), w^*\rangle = \langle f'(v_1,\ldots,v_n), w^*\rangle$, or, equivalently

$$\langle f(v_1,\ldots,v_n) - f'(v_1,\ldots,v_n), w^*\rangle = 0 ,$$

for all $(v_1,\ldots,v_n)$ and w^*. Then define $e_1 := f(v_1,\ldots,v_n) - f'(v_1,\ldots,v_n)$, if $e_1 \neq 0$ we can complete it to a basis of W. Fixing $w^* = e^{*1}$ we should have

$$\langle f(v_1,\ldots,v_n) - f'(v_1,\ldots,v_n), w^*\rangle = 1$$

which contradicts the statement above. Therefore $f(v_1,\ldots,v_n) - f'(v_1,\ldots,v_n) = 0$ for all $(v_1,\ldots,v_n) \in V_1 \times \ldots \times V_n$, in other words $f = f'$.
(Surjectivity.) We have to show that for each $\Phi \in \mathcal{L}(V_1,\ldots,V_n,W^*)$ there is a $f_\Phi \in \mathcal{L}(V_1,\ldots,V_n|W)$ with $\Psi_{f_\Phi} = \Phi$.

To this end, fix vector bases $\{e_{r,i_r}\}_{i_r\in I_r} \subset V_r$ and a (dual) basis $\{e^{*k}\}_{k\in I} \subset W^*$. Then, take $\Phi \in \mathcal{L}(V_1,\ldots,V_n,W^*)$ and define the mapping $f_\Phi \in \mathcal{L}(V_1,\ldots,V_n|W)$ given by

$$f_\Phi(v_1,\ldots,v_n) := v_1^{i_1}\cdots v_n^{i_n}\Phi(e_{1,i_1},\ldots,e_{n,i_n},e^{*k})e_k \, .$$

By construction that mapping is multilinear and, using multilinearity and Exercise **1.2.1**, we find

$$\Psi_{f_\Phi}(v_1,\ldots,v_n,w^*) = \langle v_1^{i_1}\cdots v_n^{i_n}\Phi(e_{1,i_1},\ldots,e_{n,i_n},e^{*k})e_k, w^*_h e^{*h}\rangle = \Phi(v_1,\ldots,v_n,w^*) \, ,$$

for all $(v_1,\ldots,v_n,w^*)$ and this is equivalent to $\Psi_{f_\Phi} = \Phi$. □

Important Remark.
An overall difficult point in understanding and trusting in the statement of the theorem above relies upon the fact that the function f has n entries, whereas the function Ψ_f has $n+1$ entries but f is identified to Ψ_f by means of F. This fact may seem quite weird at first glance and the statement of the theorem may seem suspicious by consequence. The difficulty can be clarified from a practical point of view as follows.
If bases $\{e_{1,i_1}\}_{i_1\in I_1}$,...,$\{e_{n,i_n}\}_{i_n\in I_n}$, for V_1,...,V_n respectively, and $\{e_k\}_{k\in I}$ for W are fixed and $\Psi \in \mathcal{L}(V_1,\ldots,V_n,W^*)$, one has, for constant coefficients $P^k_{i_1\ldots i_n}$ depending on Ψ,

$$\Psi(v_1,\ldots,v_n,w^*) = P^k_{i_1\ldots i_n} v_1^{i_1}\ldots v_n^{i_n} w^*_k \, ,$$

where w^*_k are the components of w^* in the dual basis $\{e^{*k}\}_{k\in I}$ and $v_p^{i_p}$ the components of v_p in the basis $\{e_{p,i_p}\}_{i_p\in I_p}$. Now consider the other map $f \in \mathcal{L}(V_1,\ldots,V_n|W)$ whose argument is the string $(v_1,\ldots,v_n)$, *no further vectors being necessary*:

$$f : (v_1,\ldots,v_n) \mapsto P^k_{i_1\ldots i_n} v_1^{i_1}\ldots v_n^{i_n} e_k$$

The former mapping Ψ, which deals with $n+1$ arguments, is associated with f which deals with n arguments. The point is that we have taken advantage of the bases, $\{e_k\}_{k\in I_k}$ in particular.
Within this context, Theorem 1.3 just proves that *the correspondence which associates Ψ to f is linear, bijective and independent from the chosen bases.*

Theorem 1.3 allow us to restrict our study to the spaces of multi-linear *functionals* $\mathcal{L}(V_1,\ldots,V_k)$, since the spaces of multi-linear *maps* are completely encompassed. Let us introduce the concept of *tensor product* of vector spaces. The following definitions can be extended to encompass the case of non-finite dimensional vector spaces by introducing suitable topological notions (e.g., Hilbert spaces). To introduce the tensor product we need a trivial but relevant preliminary definition.

Def.1.4. (Linear action of a vector space.) *Let U and V two finite-dimensional vector space on the same field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. We say that U* **linearly acts on** *V* **by means of** *F, if*

$F : U \to V^*$ is a natural isomorphism.

Examples.
(1) If V is a finite-dimensional vector space on $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, V^* linearly acts on V by $F = id_{V^*}$.
(2) If V is that as in **(1)**, V linearly acts on V^* by means of the natural isomorphism between V and $(V^*)^*$.

In the following if U linearly acts on V by means of F, and $u \in U$, $v \in V$ we write $u(v)$ instead of $(F(u))(v)$ whenever F is unambiguously determined. For instance, in the second example above, $v(u)$ means $\langle v, u\rangle$ when $v \in V$ and $u \in V^*$.

Def.1.5. (Tensor product.) *Let $U_1, \ldots, U_n$ be $n \geq 1$ finite-dimensional vector spaces on the common field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, suppose each U_i linearly acts on V_i (by means of some F_i), for $i = 1, \ldots, n$.*
(1) *if $(u_1, \ldots, u_n) \in U_1 \times \ldots \times U_n$, $u_1 \otimes \ldots \otimes u_n$ denotes the multi linear mapping in $\mathcal{L}(V_1, \ldots, V_n)$ defined by*

$$(u_1 \otimes \ldots \otimes u_n)(v_1, \ldots, v_n) := u_1(v_1) \cdots u_n(v_n) ,$$

for all $(v_1, \ldots, v_n) \in V_1 \times \ldots \times V_n$. $u_1 \otimes \ldots \otimes u_n$ is called **tensor product of vectors** *$u_1, \ldots, u_n$.*
(2) *The mapping $\otimes : U_1 \times \ldots \times U_n \to \mathcal{L}(V_1, \ldots, V_n)$ given by: $\otimes : (u_1, \ldots, u_n) \mapsto u_1 \otimes \ldots \otimes u_n$, is called* **tensor product map**.
(3) *The vector subspace of $\mathcal{L}(V_1, \ldots, V_n)$ generated by all of $u_1 \otimes \ldots \otimes u_n$ for all $(u_1, \ldots, u_n) \in U_1 \times \ldots \times U_n$ is called* **tensor product of spaces** *$U_1, \ldots, U_n$ and is indicated by $U_1 \otimes \ldots \otimes U_n$. The vectors in $U_1 \otimes \ldots \otimes U_n$ are called* **tensors**.

Remarks.
(1) $U_1 \otimes \ldots \otimes U_n$ is made of all the linear combinations of the form $\sum_{j=1}^{N} \alpha_j u_{1,j} \otimes \ldots \otimes u_{n,j}$, where $\alpha_j \in \mathbb{K}$, $u_{k,j} \in U_k$ and $N = 1, 2, \ldots$.
(2) It is trivially proven that the tensor product map:

$$(u_1, \ldots, u_n) \mapsto u_1 \otimes \ldots \otimes u_n ,$$

is *multi linear*.
That is, for any $k \in \{1, 2, \ldots, n\}$, the following identity holds for all $u, v \in V_k$ and $\alpha, \beta \in \mathbb{K}$:

$$u_1 \otimes \ldots \otimes u_{k-1} \otimes (\alpha u + \beta v) \otimes u_{k+1} \otimes \ldots \otimes u_k$$

$$= \alpha(u_1 \otimes \ldots \otimes u_{k-1} \otimes u \otimes u_{k+1} \otimes \ldots \otimes u_k)$$

$$+\beta(u_1 \otimes \ldots \otimes u_{k-1} \otimes v \otimes u_{k+1} \otimes \ldots \otimes u_k) .$$

As a consequence, it holds

$$(\alpha u) \otimes v = u \otimes (\alpha v) = \alpha(u \otimes v) ,$$

and similar identites hold considering whatever number of factor spaces in a tensor product of vector spaces.

(3) From the given definition, if $V_1, \ldots, V_n$ are given finite-dimensional vector spaces on the same field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, it is clear what we mean by $V_1 \otimes \ldots \otimes V_n$ or $V_1^* \otimes \ldots \otimes V_n^*$ but also, for instance, $V_1^* \otimes V_2 \otimes V_3$ or $V_1 \otimes V_2^* \otimes V_3^*$. One simply has to take the natural linear action of V on V^* and V^* on V into account.

The natural question which arises is if $U_1 \otimes \ldots \otimes U_n$ is a *proper* subspace of $\mathcal{L}(V_1, \ldots, V_n)$ or, conversely, recovers the whole space $\mathcal{L}(V_1, \ldots, V_n)$. The following theorem gives an answer to that question.

Theorem 1.4. *Referring to* **Def.1.4**, *it holds:*

$$U_1 \otimes \ldots \otimes U_n = \mathcal{L}(V_1, \ldots, V_n)\,.$$

Proof. It is sufficient to show that if $f \in \mathcal{L}(V_1, \ldots, V_n)$ then $f \in U_1 \otimes \ldots \otimes U_n$. To this end fix bases $\{e_{k,i}\}_{i \in I_k} \subset V_k$ for $k = 1, \ldots, n$. f above is completely determined by coefficients $f_{i_1,\ldots,i_n} := f(e_{1,i_1}, \ldots, e_{n,i_n})$ since, every $v_k \in V_k$ can be decomposed as $v_k = v_k^{i_k} e_{k i_k}$ and thus, by multi linearity:

$$f(v_1, \ldots, v_n) = v_1^{i_1} \cdots v_n^{i_n} f(e_{1,i_1}, \ldots, e_{n,i_n})\,.$$

Thenconsider the tensor $t_f \in U_1 \otimes \ldots \otimes U_n$ defined by:

$$t_f := f_{i_1 \ldots i_n} e_1^{*i_1} \otimes \ldots \otimes e_n^{*i_n}\,,$$

where we have identified each U_k with the corresponding space V_k^* which linearly acts on V_k. Then, by **Def.1.5**, one can directly prove that, by multi linearity

$$t_f(v_1, \ldots, v_n) = v_1^{i_1} \ldots v_n^{i_n} f_{i_1 \ldots i_n} = f(v_1, \ldots, v_n)\,,$$

for all of $(v_1, \ldots, v_n) \in V_1 \times \ldots \times V_n$. This is nothing but $t_f = f$. $\Box$

Another relevant result is stated by the theorem below.

Theorem 1.5. *Referring to* **Def.1.4**, *the following statements hold.*
(a) *The dimension of* $U_1 \otimes \ldots \otimes U_n$ *is:*

$$dim(U_1 \otimes \ldots \otimes U_n) = \prod_{k=1}^{n} dim U_k = \prod_{k=1}^{n} dim V_k\,.$$

(b) *If* $\{e_{k,i_k}\}_{i_k \in I_k}$ *is a basis of* U_k, $k = 1, \ldots, n$, *then :* $\{e_{1,i_1} \otimes \ldots \otimes e_{n,i_n}\}_{(i_1,\ldots,i_n) \in I_1 \times \ldots \times I_n}$ *is a vector basis of* $U_1 \otimes \ldots \otimes U_n$.
(c) *If* $t \in U_1 \otimes \ldots \otimes U_n$, *the components of* t *with respect to a basis* $\{e_{1,i_1} \otimes \ldots \otimes e_{n,i_n}\}_{(i_1,\ldots,i_n) \in I_1 \times \ldots \times I_n}$ *are given by (identifying each* V_i *with the corresponding* U_i^**):*

$$t^{i_1 \ldots i_n} = t(e_1^{*i_1}, \ldots, e_n^{*i_n})$$

and thus it holds

$$t = t^{i_1 \ldots i_n} e_{1,i_1} \otimes \ldots \otimes e_{n,i_n} \,.$$

Proof. Notice that **(b)** trivially implies **(a)** because the elements $e_{1,i_1} \otimes \ldots \otimes e_{n,i_n}$ are exactly $\prod_{k=1}^{n} dimU_k$. (Also $\prod_{k=1}^{n} dimU_k = \prod_{k=1}^{n} dimV_k$ because each U_k is isomorphic to the corresponding V_k^* and $dimV_k = dimV_k^*$ by definition of dual basis.) So it is sufficient to show that the second statement holds true. To this end, since elements $e_{1,i_1} \otimes \ldots \otimes e_{n,i_n}$ are generators of $U_1 \otimes \ldots \otimes U_n$, it is sufficient to show that they are linearly independent. Consider the generic vanishing linear combination

$$C^{i_1 \ldots i_n} e_{1,i_1} \otimes \ldots \otimes e_{n,i_n} = 0 \,,$$

We want to show that all of the coefficients $C^{i_1 \ldots i_n}$ vanish. We may identify each V_k with the corresponding U_k^* by means of the natural isomorphism of U and $(U^*)^*$ and thus we may consider each dual-basis vector $e_r^{*j_r} \in U_j^*$ as an element of V_j. Then the action of that linear combination of multi-linear functionals on the generic element $(e_1^{*j_1}, \ldots, e_n^{*j_n}) \in V_1 \times \cdots \times V_n$ produces the result

$$C^{j_1 \ldots j_n} = 0 \,.$$

Since we can arbitrarily fix the indices $j_1, \ldots j_n$ this proves the thesis.
Concerning **(c)**, by the uniqueness of the components of a vector with respect to a basis, it is sufficient to show that, defining

$$t' := t^{i_1 \ldots i_n} e_{1,i_1} \otimes \ldots \otimes e_{n,i_n} \,,$$

where

$$t^{i_1 \ldots i_n} := t(e_1^{*i_1}, \ldots, e_n^{*i_n}) \,,$$

it holds

$$t'(v_1, \ldots, v_n) = t(v_1, \ldots, v_n) \,,$$

for all $(v_1, \ldots, v_n) \in V_1^* \times \ldots \times V_n^*$. By multi linearity,

$$t(v_1, \ldots, v_n) = v_{1\,i_1} \cdots v_{n\,i_n} t(e_1^{*i_1}, \ldots, e_n^{*i_n}) \,,$$

By multi linearity, the left-hand side is nothing but $t'(v_1, \ldots, v_n)$. This concludes the proof. □

There are two important theorems which, *together with the identification of* V *and* $(V^*)^*$, imply that all of the spaces which one may build up coherently using the symbols $\otimes$, V_k, * and () are naturally isomorphic to spaces which are of the form $V_{i_1}^{(*)} \otimes \ldots \otimes V_{i_n}^{(*)}$. The rule to produce spaces naturally isomorphic to a given initial space is that one has to (1) ignore parentheses, (2) assume that * is *distributive with respect to* $\otimes$ and (3) assume that * is *involutive* (i.e. $(X^*)^* \simeq X$). For instance, one has:

$$((V_1^* \otimes V_2) \otimes (V_3 \otimes V_4^*))^* \simeq V_1 \otimes V_2^* \otimes V_3^* \otimes V_4 \,.$$

Let us state the theorems corresponding to the rules (2) and (1) respectively (the rule (3) being nothing but **Theorem 1.2**).

Theorem 1.6. * *is* **distributive** *with respect to* $\otimes$ *by means of* **natural isomorphisms** $F : V_1^* \otimes \ldots \otimes V_n^* \to (V_1 \otimes \ldots \otimes V_n)^*$. *In other words, if* $V_1 \ldots V_n$ *are finite dimensional vector spaces on the same field* $\mathbb{K} = \mathbb{R}$ *or* $\mathbb{C}$, *it holds:*

$$(V_1 \otimes \ldots \otimes V_n)^* \simeq V_1^* \otimes \ldots \otimes V_n^* \, .$$

under F *and furthermore*

$$\langle u_1 \otimes \ldots \otimes u_n, F(v_1^* \otimes \ldots \otimes v_n^*) \rangle = \langle u_1, v_1^* \rangle \cdots \langle u_n, v_n^* \rangle \, ,$$

for every choice of $v_i^* \in V_i^*$, $u_i \in V_i$ *and* $i = 1, 2, \ldots, n$.

Theorem 1.7. *The tensor product is* **associative** *by means of* **natural isomorphisms**. *In other words considering a space which is made of tensor products of tensor products of finite-dimensional vector spaces on the same field* $\mathbb{R}$ *or* $\mathbb{C}$, *one may omit parenthesis everywhere obtaining a space which is naturally isomorphic to the initial one. So, for instance*

$$V_1 \otimes V_2 \otimes V_3 \simeq V_1 \otimes (V_2 \otimes V_3) \, .$$

where the natural isomprphism $F_1 : V_1 \otimes V_2 \otimes V_3 \to V_1 \otimes (V_2 \otimes V_3)$ *satisfies*

$$F_1 : v_1 \otimes v_2 \otimes v_3 \to v_1 \otimes (v_2 \otimes v_3)$$

for every choice of $v_i \in V_i$, $i = 1, 2, 3$.

Sketch of proof. The natural isomorphisms required in the two theorems above can be built up as follows. In the former case, consider a linear mapping $F : V_n^* \otimes \ldots \otimes V_n^* \to (V_1 \otimes \ldots \otimes V_n)^*$ which satisfies

$$F(v_1^* \otimes \ldots \otimes v_n^*) \in (V_1 \otimes \ldots \otimes V_n)^* \, ,$$

where, for all $u_1 \otimes \ldots \otimes u_n \in V_1 \otimes \ldots \otimes V_n$,

$$F(v_1^* \otimes \ldots \otimes v_n^*)(u_1 \otimes \ldots \otimes u_n) := \langle u_1, v_1^* \rangle \ldots \langle u_n, v_n^* \rangle \, . \tag{1}$$

In the latter case, consider the first proposed example. The required natural isomorphism can be build up as a linear mapping $F_1 : V_1 \otimes V_2 \otimes V_3 \to (V_1 \otimes V_2) \otimes V_3$ such that

$$F_1 : v_1 \otimes v_2 \otimes v_3 \mapsto (v_1 \otimes v_2) \otimes v_3 \, ,$$

for all $v_1 \otimes v_2 \otimes v_3 \in V_1 \otimes V_2 \otimes V_3$.
Using the linearity and the involved definitions the reader can simply prove that the mappings F, F_1 are surjective. Moreover, by the previously proven theorems

$$dim(V_1 \otimes \cdots \otimes V_n)^* = dim(V_1 \otimes \cdots \otimes V_n) = dim(V_1^* \otimes \cdots \otimes V_n^*)$$

and

$$dim((V_1 \otimes V_2) \otimes V_3) = dim(V_1 \otimes V_2) \cdot dim(V_3) = dim(V_1 \otimes V_2 \otimes V_3)\ .$$

As a consequence, F and F_1 must be also injective and thus they are isomorphisms. $\Box$

Very important remark. The central point is that the mappings F and F_1 above have been given by specifying their action on tensor products of elements (e.g, $F(v_1 \otimes \ldots \otimes v_n)$) and not on *linear combinations* of these tensor products of elements. Remind that, for instance $V_1^* \otimes \ldots \otimes V_n^*$ is not the set of products $u_1^* \otimes \ldots \otimes u_n^*$ but it is the set of *linear combinations* of those products. Hence, in order to completely define F and F_1, one must require that F and F_1 admit *uniquely determined linear extensions* on their initial domains in order to encompass the whole tensor spaces generated by linear combinations of simple tensor products. In other words one has to complete the given definition, in the former case, by adding the further requirement

$$F(\alpha\, u_1^* \otimes \ldots \otimes u_n^* + \beta\, v_1^* \otimes \ldots \otimes v_n^*) = \alpha F(u_1^* \otimes \ldots \otimes u_n^*) + \beta F(v_1^* \otimes \ldots \otimes v_n^*)\ ,$$

and similarly in the latter case. Despite it could seem a trivial task it is not the case. Indeed it is worthwhile noticing that each product $v_1^* \otimes \ldots \otimes v_n^*$ can be re-written using linear combinations, e.g., concerning F:

$$v_1^* \otimes \ldots \otimes v_n^* = [(v_1^* + u_1^*) \otimes \ldots \otimes v_n^*]\ -\ [u_1^* \otimes \ldots \otimes v_n^*]\ .$$

Now consider the identities, which has to hold as a consequence of the assumed linearity of F, such as:

$$F(v_1^* \otimes \ldots \otimes v_n^*) = F((v_1^* + u_1^*) \otimes \ldots \otimes v_n^*) - F(u_1^* \otimes \ldots \otimes v_n^*)$$

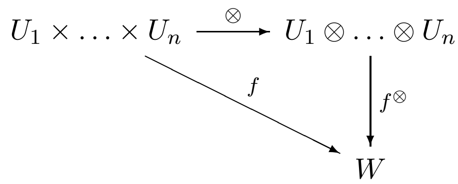
Above $F(v_1^* \otimes \ldots \otimes v_n^*)$, $F(u_1^* \otimes \ldots \otimes v_n^*)$, $F((v_1^* + u_1^*) \otimes \ldots \otimes v_n^*)$ are *independently* defined as we said at the beginning and there is no reason, in principle, for the validity of the constraint:

$$F(v_1^* \otimes \ldots \otimes v_n^*) = F((v_1^* + u_1^*) \otimes \ldots \otimes v_n^*) - F(u_1^* \otimes \ldots \otimes v_n^*)\ .$$

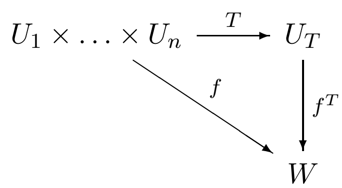
Similar problems arise concerning F_1.
The general problem which arises by the two considered cases can be stated as follows. Suppose we are given a tensor product of vector spaces $V_1 \otimes \ldots \otimes V_n$ and we are interested in the possible *linear extensions* of a mapping f on $V_1 \otimes \ldots \otimes V_n$, with values in some vector space W, when f is initially defined on simple products $v_1 \otimes \ldots \otimes v_n \in V_1 \otimes \ldots \otimes V_n$ only.
Is there any general prescription on the specification of values $f(v_1 \otimes \ldots \otimes v_n)$ which assures that f can be extended, uniquely, *to a linear mapping from $V_1 \otimes \ldots \otimes V_n$ to W?*
An answer is given by the following very important **universality theorem.**

Theorem 1.8. (Universality Theorem.) *Given $n \geq 1$ finite-dimensional vector spaces $U_1, \ldots, U_n$ on the same field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, the following statements hold.*
(a) *For each finite-dimensional vector space W and each multi-linear mapping $f : U_1 \times \ldots \times U_n \rightarrow$*

W, there is a unique linear mapping $f^\otimes : U_1 \otimes \ldots \otimes U_n \to W$ such that the diagram below **commutes** *(in other words, $f^\otimes \circ \otimes = f$).*



(b) *For $U_1, \ldots, U_n$ fixed as above, suppose there is another pair (T, U_T) such that, for each W and each f with $f : U_1 \times \ldots \times U_n \to W$ multi linear, there is a unique linear map $f^T : U_T \to W$ such that the diagram below commute,*



then there is a natural vector space isomorphism $\phi : U_1 \otimes \ldots \otimes U_n \to U_T$ with $f^T \circ \phi = f^\otimes$. In other words, given $U_1, \ldots, U_n$ as in the hypotheses, the pair $(\otimes, U_1 \otimes \ldots \otimes U_n)$ is determined up to vector spaces isomorphisms by the diagram property above.

Remark. The property of the pair $(\otimes, U_1 \otimes \ldots \otimes U_n)$ stated in the second part of the theorem is called **universality property**.

Before proving the theorem, let us explain how it gives a precise answer to the stated question. The theorem says that a *linear extension* of any function f with values in W, initially defined on simple tensor products only, $f(v_1 \otimes \ldots \otimes v_n)$ with $v_1 \otimes \ldots \otimes v_n \in U_1 \otimes \ldots \otimes U_n$, does *exist* on the whole domain space $U_1 \otimes \ldots \otimes U_n$ and it is *uniquely* determined provided $f(v_1 \otimes \ldots \otimes v_n) = g(v_1, \ldots, v_n)$ where $g : U_1 \times \ldots \times U_n \to W$ is some *multi-linear function.*
Concerning the mappings F and F_1 introduced above, we may profitably use the universality theorem to show that they are well-defined on the whole domain made of linear combinations of simple tensor products. In fact, consider F for example. We can define the *multi-linear* mapping $G : V_1^* \times \ldots \times V_n^* \to (V_1 \otimes \ldots \otimes V_n)^*$ such that $G(v_1^*, \ldots, v_n^*)$ it the unique multi-linear function with:

$$(G(v_1^*, \ldots, v_n^*))\,(u_1, \ldots, u_n) := \langle u_1, v_1^* \rangle \cdots \langle u_n, v_n^* \rangle\,, \quad \text{for all } (u_1, \ldots, u_n) \in U_1 \times \cdots \times U_n \qquad (2)$$

(The reader should check on the well-posedness and cd multi linearity of G.) Then the universality theorem with $U_k = V_k^*$ and $W = (V_1 \otimes \ldots \otimes V_n)^*$ assures the existence of a linear mapping $F : V_1^* \otimes \ldots \otimes V_n^* \to (V_1 \otimes \ldots \otimes V_n)^*$ with the required properties because $F := G^\otimes$ is such that

$$F(v_1^* \otimes \ldots \otimes v_n^*) = (G^\otimes \circ \otimes)(v_1^*, \ldots, v_n^*) = G(v_1^*, \ldots, v_n^*)\,,$$

si that (1) is fulfilled due to (2). A similar multi-linear mapping G_1 can be found for F_1:

$$G_1 : (v_1, v_2, v_3) \mapsto (v_1 \otimes v_2) \otimes v_3 \,.$$

$F_1 := G_1^{\otimes}$ can be used in the proof of Theorem 1.7.

Proof of **Theorem 1.8**. **(a)** Fix bases $\{e_{k,i_k}\}_{i_k \in I_k} \subset U_k$, $k = 1, \ldots, n$ and $\{e_j\}_{j \in I} \subset W$. Because of the multi-linearity property, the specification of coefficients $f^j_{i_1 \ldots i_n} \in \mathbb{K}$, $i_k \in I_k, j \in I$, $k = 1, \ldots, n$ uniquely determines a multi-linear mapping $f : U_1 \times \ldots \times U_n \to W$ by defining $f^j_{i_1 \ldots i_n} := \langle f(e_{1,i_1}, \ldots, e_{n,i_n}), e^{*j} \rangle$. Conversely, a multi-linear mapping $f : U_1 \times \ldots \times U_n \to W$ uniquely determines coefficients $f^j_{i_1 \ldots i_n} \in \mathbb{K}$ by the same rule. Similarly, because of the linearity property, a linear mapping $g : U_1 \otimes \ldots \otimes U_n \to W$ is completely and uniquely determined by coefficients $g^j_{i_1 \ldots i_n} := \langle g(e_{i_1} \otimes \ldots \otimes e_{i_n}), e^j \rangle$. Therefore, if $f : U_1 \times \ldots \times U_n \to W$ is given and it is determined by coefficients $f^j_{i_1 \ldots i_n}$ as above, defining $f^{\otimes} : U_1 \otimes \ldots \otimes U_n \to W$ by the coefficients $(f^{\otimes})^j_{i_1 \ldots i_n} := f^j_{i_1 \ldots i_n}$, we have

$$\langle f^{\otimes}(v_1 \otimes \ldots \otimes v_n), e^{*j} \rangle = v^{i_1} \ldots v^{i_n} f^j_{i_1 \ldots i_n} = \langle f(v_1, \ldots, v_n), e^{*j} \rangle,$$

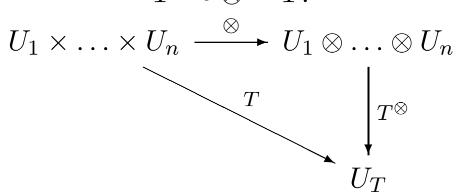
for all $(v_1, \ldots, v_n) \in U_1 \times \ldots \times U_n$ and $j \in I$. This means that $f^{\otimes}(v_1 \otimes \ldots \otimes v_n) = f(v_1, \ldots, v_n)$ for all $(v_1, \ldots, v_n) \in U_1 \times \ldots \times U_n$. In other words $f^{\otimes} \circ \otimes = f$. The uniqueness of the mapping $f^{\otimes}$ is obvious: suppose there is another mapping $g^{\otimes}$ with $g^{\otimes} \circ \otimes = f$ then $(f^{\otimes} - g^{\otimes}) \circ \otimes = 0$. This means in particular that

$$\langle (f^{\otimes} - g^{\otimes})(e_{i_1} \otimes \ldots \otimes e_{i_n}), e^{*j} \rangle = 0 \,,$$

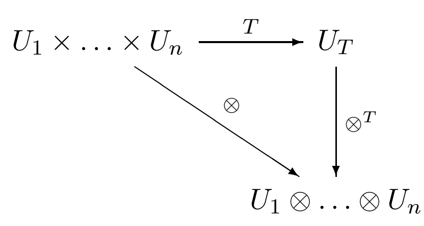
for all $i_k \in I_k$, $k = 1, \ldots, n$ and $j \in I$. Since the coefficients above completely determine a map, the considered mapping must be the null mapping and thus: $f^{\otimes} = g^{\otimes}$.

(b) Suppose that there is a pair (T, U_T) as specified in the hypotheses. Then, using the part **(a)** with $f = T$ and $W = U_T$ we have the diagram below ,where we have a former diagram commutation relation:

$$T^{\otimes} \circ \otimes = T.$$



On the other hand, using the analogous property of the pair (T, U_T) with $f = \otimes$ and $W = U_1 \otimes \ldots \otimes U_n$ we also have the commutative diagram



which involves a second diagram commutation relation:

$$\otimes^T \circ T = \otimes \,.$$

The two obtained relations imply

$$(T^{\otimes} \circ \otimes^T) \circ T = T$$

and

$$(\otimes^T \circ T^{\otimes}) \circ \otimes = \otimes \,,$$

In other words, if $RanT \subset U_T$ is the range of the mapping T and $Ran\otimes \subset U_1 \otimes \ldots \otimes U_n$ is the analogue for $\otimes$:

$$(T^{\otimes} \circ \otimes^T)|_{RanT} = Id_{RanT}$$

and

$$(\otimes^T \circ T^{\otimes})|_{Ran\otimes} = Id_{Ran\otimes}$$

Then consider the latter. Notice that $Ran\otimes$ is not a subspace, but the subspace spanned by $Ran\otimes$ coincides with $U_1 \otimes \ldots \otimes U_n$ by definition of tensorial product. Since $\otimes^T \circ T^{\otimes}$ is linear (because composition of linear maps) and defined on the whole space $U_1 \otimes \ldots \otimes U_n$, we conclude that

$$\otimes^T \circ T^{\otimes} = Id_{U_1 \otimes \ldots \otimes U_n} \,.$$

Concerning the former identity we could conclude analogously, i.e.,

$$T^{\otimes} \circ \otimes^T = Id_{U_T} \,,$$

provided the subspace spanned by $RanT$, $Span(RanT)$, coincides with the whole space U_T. Anyway this holds true by the uniqueness property of f^T for a fixed f (see hypotheses about the pair (T, U_T)). (If $Span(RanT) \neq U_T$, then $U_T = Span(RanT) \oplus S$, $S \neq \{0\}$ being some proper subspace of U_T with $S \cap Span(RanT) = \{0\}$. Then, a mapping f^T could not be uniquely determined by the requirement $f^T \circ T = f$ because such a relation would be preserved under modifications of f^T on the subspace S).
We conclude that the linear mapping $\phi := T^{\otimes} : U_1 \otimes \ldots \otimes U_n \to U_T$ is a natural isomorphism with inverse $\otimes^T$.
As far as the property $f^T \circ \phi = f^{\otimes}$ is concerned we may deal with as follows. Since $\phi : U_1 \otimes \ldots \otimes U_n \to U_T$ is an isomorphism, $f^T \circ T = f$ implies $(f^T \circ \phi) \circ (\phi^{-1} \circ T) = f$, but $\phi^{-1} \circ T = \otimes^T \circ T = \otimes$, therefore $(f^T \circ \phi) \circ \otimes = f$. By the uniqueness of the mapping $f^{\otimes}$ which satisfies the same identity, it has to hold $f^T \circ \phi = f^{\otimes}$. $\Box$

Exercises 1.3.
1.3.1. Consider a finite-dimensional vector space V and its dual V^*. Show by the universality theorem that there is a natural isomorphism such that

$$V \otimes V^* \simeq V^* \otimes V \,.$$

(Hint. Consider the bilinear mapping $f : V \times V^* \to V^* \otimes V$ with $f : (v_1, v_2^*) \mapsto v_2^* \otimes v_1$. Show that $f^{\otimes}$ is injective and thus surjective because $dim(V^* \otimes V) = dim(V \otimes V^*)$.)

2 Tensor algebra. Abstract Index Notation.

2.1 Tensor algebra generated by a vector space.

Let V be a finite-dimensional vector space on the field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Equipped with this algebraic structure, in principle, we may build up several different tensor spaces. Notice that we have to consider also $\mathbb{K}$, $\mathbb{K}^*$ and V^* as admissible tensor factors. (In the following we shall not interested in conjugate spaces.) Obviously, we are interested in tensor products which are not identifiable by some natural isomorphism.
First consider the dual $\mathbb{K}^*$ when we consider $\mathbb{K}$ as a vector space on the field $\mathbb{K}$ itself. $\mathbb{K}^*$ is made of linear functionals from $\mathbb{K}$ to $\mathbb{K}$. Each $c^* \in \mathbb{K}^*$ has the form $c^*(k) := c \cdot k$, for all $k \in \mathbb{K}$, where $c \in \mathbb{K}$ is a fixed field element which completely determines c^*. The mapping $c \mapsto c^*$ is a (natural) vector space isomorphism. Therefore

$$\mathbb{K} \simeq \mathbb{K}^* \,.$$

Then we pass to consider products $\mathbb{K} \otimes \ldots \otimes \mathbb{K} \simeq \mathbb{K}^* \otimes \ldots \otimes \mathbb{K}^* = \mathfrak{L}(\mathbb{K}, \ldots, \mathbb{K})$. Each multi-linear mapping $f \in \mathfrak{L}(\mathbb{K}, \ldots, \mathbb{K})$ is completely determined by the number $f(1, \ldots, 1)$, since $f(k_1, \ldots, k_n) = k_1 \cdots k_n f(1, \ldots, 1)$. One can trivially show that the mapping $f \mapsto f(1, \ldots, 1)$ is a (natural) vector space isomorphism between $\mathbb{K} \otimes \ldots \otimes \mathbb{K}$ and $\mathbb{K}$ itself. Therefore

$$\mathbb{K} \otimes \ldots \otimes \mathbb{K} \simeq \mathbb{K} \,.$$

Also notice that the found isomorphism trivially satisfies $c_1 \otimes \ldots \otimes c_n \mapsto c_1 \cdots c_n$, and thus the tensor product mapping reduces to the ordinary product of the field.
We pass to consider the product $\mathbb{K} \otimes V = \mathfrak{L}(\mathbb{K}^*, V^*) \simeq \mathfrak{L}(\mathbb{K}, V^*)$. Each multi-linear functional f in $\mathfrak{L}(\mathbb{K}, V^*)$ is completely determined by the element of V, $f(1, \cdot) : V^* \to \mathbb{K}$, which maps each $v^* \in V^*$ in $f(1, v^*)$. Once again, it is a trivial task to show that $f \mapsto f(1, \cdot)$ is a (natural) vector space isomorphism between $\mathbb{K} \otimes V$ and V itself.

$$\mathbb{K} \otimes V \simeq V \,.$$

Notice that the found isomorphism satisfies $k \otimes v \mapsto kv$ and thus the tensor product mapping reduces to the ordinary product of a field element and a vector.
Concluding, only the spaces $\mathbb{K}$, V, V^* and whatever products of V and V^* may be significantly different. This result leads us to the following definition.

Def.2.1. (Tensor Algebra generated by V.) *Let V be a finite-dimensional vector space with field $\mathbb{K}$.*
(1) *The* **tensor algebra** *$\mathcal{A}_{\mathbb{K}}(V)$ generated by V with field $\mathbb{K}$ is the class whose elements are the vector spaces: $\mathbb{K}$, V, V^* and all of tensor products of factors V and V^* in whatever order and number.*
(2) *The tensors of $\mathbb{K}$ are called* **scalars***, the tensors of V are called* **contravariant vectors***, the tensors of V^* are called* **covariant vectors***, the tensors of spaces $V^{n\otimes} := V \otimes \ldots \otimes V$,*

where V appears $n \geq 1$ times, are called **contravariant tensors of order** n *or* **tensors of order** $(n, 0)$*, the tensors of spaces* $V^{*n\otimes} := V^* \otimes \ldots \otimes V^*$*, where* V^* *appears* $n \geq 1$ *times, are called* **covariant tensors of order** n *or* **tensors of order** $(0, n)$*, the remaining tensors which belong to spaces containing* n *factors* V *and* m *factors* V^* *are called* tensors of order (n, m) *not depending on the order of factors.*

Remark. Obviously, pairs of tensors spaces made of the same number of factors V and V^* in different order, are naturally isomorphic (see **exercise 1.3.1**). However, for practical reasons it is convenient to consider these spaces as different spaces and use the identifications when and if necessary.

2.2 The abstract index notation and rules to handle tensors.

Let us introduce the **abstract index notation**. Consider a finite-dimensional vector space V with field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. After specification of a basis $\{e_i\}_{i\in I} \subset V$ and the corresponding basis in V^*, each tensor is completely determined by giving its components with respect to the induced basis in the corresponding tensor space in $\mathcal{A}_{\mathbb{K}}(V)$. We are interested in the transformation rule of these components under change of the base in V. Suppose to fix another basis $\{e'_j\}_{j\in I} \subset V$ with $e_i = A^j{}_i e'_j$. The coefficients $A^j{}_i$ determine a matrix $A := [A^j{}_i]$ in the matrix group $GL(dimV, \mathbb{K})$, i.e., the group (see the next section) of $dimV \times dimV$ matrices with coefficients in $\mathbb{K}$ and non-vanishing determinant. First consider a contravariant vector $t = t^i e_i$, passing to the other basis, we have $t = t^i e_i = t'^j e'_j$ and thus $t'^j e'_j = t^i A^j{}_i e'_j$. This is equivalent to $(t'^j - A^j{}_i t^i)e'_j = 0$ which implies

$$t'^j = A^j{}_i t^i \ ,$$

because of the linear independence of vectors e'_j. Similarly, if we specify a set of components in $\mathbb{K}$, $\{t^i\}_{i\in I}$ for each basis $\{e_i\}_{i\in I} \subset V$ and these components, changing the basis to $\{e'_j\}_{j\in I}$, transform as

$$t'^j = A^j{}_i t^i \ ,$$

where the coefficients $A^j{}_i$ are defined by

$$e_i = A^j{}_i e'_j \ ,$$

then a contravariant tensor t is defined. It is determined by $t := t^i e_i$ in each basis. The proof is self evident.
Concerning covariant vectors, a similar result holds. Indeed a covariant vector $u \in V^*$ is completely determined by the specification of a set of components $\{u_i\}_{i\in I}$ for each basis $\{e^{*i}\}_{i\in I} \subset V^*$ (dual basis of $\{e_i\}_{i\in I}$ above) when these components, changing the basis to $\{e'^{*j}\}_{j\in I}$ (dual base of $\{e'_j\}_{j\in I}$ above), transform as

$$u'_j = B_j{}^i u_i \ ,$$

where the coefficients ${B_l}^r$ are defined by

$$e^{*i} = {B_j}^i e'^{*j} .$$

What is the relation between the matrix $A = [{A^j}_i]$ and the matrix $B := [{B_k}^h]$? The answer is clear: it must be

$$\delta^i_j = \langle e_j, e^{*i}\rangle = {A^l}_j {B_k}^i \langle e'_l, e'^{*k}\rangle = {A^l}_j {B_k}^i \delta^k_l = {A^k}_j {B_k}^i .$$

In other words it has to hold $I = A^t B$, which is equivalent to

$$B = A^{-1t} .$$

(Notice that ^{t} and $^{-1}$ commute.)
Proper tensors have components which transform similarly to the vector case. For instance, consider $t \in V \otimes V^*$, fix a basis $\{e_i\}_{i\in I} \subset V$, the dual one $\{e^{*i}\}_{i\in I} \subset V^*$ and consider that induced in $V \otimes V^*$, $\{e_j \otimes e^{*j}\}_{(i,j)\in I\times I}$. Then $t = {t^i}_j e_i \otimes e^{*j}$. By bi linearity of the tensor product map, if we pass to consider another basis $\{e'_i\}_{i\in I} \subset V$ and those associated in the relevant spaces as above, concerning the components ${t'^k}_l$ of t in the new tensor space basis, one trivially gets

$${t'^k}_l = {A^k}_i {B_l}^j {t^i}_j ,$$

where the matrices $A = [{A^j}_i]$ and $B := [{B_k}^h]$ are those considered above. It is clear that the specification of a tensor of $V \otimes V^*$ is completely equivalent to the specification of a set of components for each basis of $V \otimes V^*$, $\{e_j \otimes e^{*j}\}_{(i,j)\in I\times I}$, provided these components transform as specified above under change of basis.
We can generalize the obtained results after a definition.

Def. 2.2 (Canonical bases.) *Let $\mathcal{A}_{\mathbb{K}}(V)$ be the tensor algebra generated by the finite-dimensional vector space V on the field $\mathbb{K}$ ($= \mathbb{C}$ or $\mathbb{R}$). If $B = \{e_i\}_{i\in I}$ is a basis in V with dual basis $B^* = \{e^{*i}\}_{i\in I} \subset V^*$, the* **canonical bases** *associated to the former are the bases in the tensor spaces of $\mathcal{A}_{\mathbb{K}}(V)$ obtained by tensor products of elements of B and B^*.*

Remark. Notice that also $\{e_i \otimes e'_j\}_{i,j\in I}$ is a basis of $V \otimes V$ if $\{e_i\}_{i\in I}$ and $\{e'_j\}_{j\in I}$ are bases of V. However, $\{e_i \otimes e'_j\}_{i,j\in I}$ is *not* canonical unless $e_i = e'_i$ for all $i \in I$.

Theorem 2.1. *Consider the tensor algebra $\mathcal{A}_{\mathbb{K}}(V)$ generated by a finite-dimension vector space V with field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ and take a tensor space $V^{n\otimes} \otimes V^{*m\otimes} \in \mathcal{A}_{\mathbb{K}}(V)$. The specification of a tensor $t \in V^{n\otimes} \otimes V^{*m\otimes}$ is completely equivalent to the specification of a set of components*

$$\{{t^{i_1\ldots i_n}}_{j_1\ldots j_m}\}_{i_1,\ldots,i_n,j_1,\ldots,j_m\in I}$$

with respect to each canonical basis of $V^{n\otimes} \otimes V^{*m\otimes}$,

$$\{e_{i_1} \otimes \ldots \otimes e_{i_n} \otimes e^{*j_1} \otimes \ldots \otimes e^{*j_n}\}_{i_1,\ldots i_n,j_1,\ldots j_m\in I}$$

which, under change of basis:

$$\{e'_{i_1} \otimes \ldots \otimes e'_{i_n} \otimes e'^{*j_1} \otimes \ldots \otimes e'^{*j_n}\}_{i_1,\ldots i_n, j_1,\ldots j_m \in I}$$

transform as:

$$t'^{i_1\ldots i_n}{}_{j_1\ldots j_m} = A^{i_1}{}_{k_1} \ldots A^{i_n}{}_{k_n} B_{j_1}{}^{l_1} \ldots B_{j_m}{}^{l_m} t^{k_1\ldots k_n}{}_{l_1\ldots l_m} \,,$$

where

$$e_i = A^j{}_i e'_j \,,$$

and the coefficients $B_j{}^l$ are those of the matrix:

$$B = A^{-1t} \,,$$

with $A := [A^j{}_i]$. The associated tensor t is represented by

$$t = t^{i_1\ldots i_n}{}_{j_1\ldots j_m} e_{i_1} \otimes \ldots \otimes e_{i_n} \otimes e^{*j_1} \otimes \ldots \otimes e^{*j_n}$$

for each considered canonical basis. Analogous results hold for tensor spaces whose factors V and V^* take different positions.

Notation 2.1. *In the* **abstract index notation** *a tensor is indicated by writing its generic component in a non-specified basis. E.g. $t \in V^* \otimes V$ is indicated by $t_i{}^j$.*

Somewhere in the following we adopt a cumulative index notation, i.e., letters $A, B, C, \ldots$ denote set of covariant, contravariant or mixed indices. For instance $t^{ijk}{}_{lm}$ can be written as t^A with ${}^A = {}^{ijk}{}_{lm}$. Similarly e_A denotes the element of a canonical basis $e_i \otimes e_j \otimes e_k \otimes e^{*l} \otimes e^{*m}$. Moreover, if ^{A} and ^{B} are cumulative indices, the indices of the cumulative index ^{AB} are those of ^{A} immediately followed by those of ^{B}, e.g, if ${}^A = {}^{ijk}{}_{lm}$, and ${}^B = {}^{pq}{}_u{}^n$, ${}^{AB} = {}^{ijk}{}_{lm}{}^{pq}{}_u{}^n$.

Let us specify the allowed mathematical rules to produces tensors from given tensors. To this end we shall use both the synthetic and the index notation.

Linear combinations of tensors of a fixed tensor space. Take a tensor space $S \in \mathcal{A}_{\mathbb{K}}(V)$. This is a vector space by definition, and thus picking out $s, t \in S$, and $\alpha, \beta \in \mathbb{K}$, linear combinations can be formed which still belong to S. in other words we may define the tensor of S

$$u := \alpha s + \beta t \quad \text{or, in the abstract index notation} \quad u^A = \alpha s^A + \beta t^A \,.$$

The definition of u above given by the abstract index notation means that the components of u are related to the components of s and t by a linear combination which has *the same form in whatever canonical basis of the space S* and the coefficients α,β *do not depend on the basis.*

Products of tensors of generally different tensor spaces. Take two tensor spaces $S, S' \in \mathcal{A}_{\mathbb{K}}(V)$ and pick out $t \in S$, $t' \in S'$. Then the tensor $t \otimes t' \in S \otimes S'$ is well defined. Using the associativity of the tensor product by natural isomorphisms, we found a unique tensor space $S'' \in \mathcal{A}_{\mathbb{K}}(V)$ which is isomorphic to $S \otimes S'$ and thus we can identify $t \otimes t'$ with a tensor in S'' which we shall indicate by $t \otimes t'$ once again with a little misuse of notation. $t \otimes t'$ is called the **product** of tensors t and t'. Therefore, the product of tensors coincides with the usual tensor product of tensors up to a natural isomorphism. What about the abstract index notation? We leave to the reader the trivial proof of the following fact based on **Theorem 1.7**,

$$(t \otimes t')^{AB} = t^A t'^B \,.$$

Thus, for instance, if $S = V \otimes V^*$ and $S' = V^*$ the tensors t and t' are respectively indicated by $t^i{}_j$ and s_k and thus $(t \otimes s)^i{}_{jk} = t^i{}_j s_k$.

Contractions. Consider a tensor space of $\mathcal{A}_{\mathbb{K}}(V)$ of the form

$$U_1 \otimes \ldots \otimes U_k \otimes V \otimes U_{k+1} \otimes \ldots \otimes U_l \otimes V^* \otimes U_{l+1} \otimes \ldots \otimes U_n$$

where U_i denotes either V or V^*. Everything we are going to say can be re-stated for the analogous space

$$U_1 \otimes \ldots \otimes U_k \otimes V^* \otimes U_{k+1} \otimes \ldots \otimes U_l \otimes V \otimes U_{l+1} \otimes \ldots \otimes U_n \,.$$

Then consider the *multi-linear* mapping C with domain

$$U_1 \times \ldots \times U_k \times V \times U_{k+1} \times \ldots \times U_l \times V^* \times U_{l+1} \times \ldots \times U_n \,,$$

and values in

$$U_1 \otimes \ldots \otimes U_k \otimes U_{k+1} \otimes \ldots \otimes U_l \otimes U_{l+1} \otimes \ldots \otimes U_n$$

defined by:

$$(u_1, \ldots, u_k, v, u_{k+1}, \ldots, u_l, v^*, v_{l+1}, \ldots, v_n) \mapsto \langle v, v^* \rangle u_1 \otimes \ldots \otimes u_k \otimes u_{k+1} \otimes \ldots \otimes u_l \otimes u_{l+1} \otimes \ldots \otimes u_n \,.$$

By the universality theorem there is a *linear* mapping $C^{\otimes}$, called **contraction** of V and V^*, defined on the whole tensor space

$$U_1 \otimes \ldots \otimes U_k \otimes V \otimes U_{k+1} \otimes \ldots \otimes U_l \otimes V^* \otimes U_{l+1} \otimes \ldots \otimes U_n$$

taking values in

$$U_1 \otimes \ldots \otimes U_k \otimes U_{k+1} \otimes \ldots \otimes U_l \otimes U_{l+1} \otimes \ldots \otimes U_n$$

such that, on simple products of vectors reduces to

$$u_1 \otimes \ldots \otimes u_k \otimes v \otimes u_{k+1} \otimes \ldots \otimes u_l \otimes v^* \otimes u_{l+1} \otimes \ldots \otimes u_n \mapsto \langle v, v^* \rangle u_1 \otimes \ldots \otimes u_k \otimes u_{k+1} \otimes \ldots \otimes u_l \otimes u_{l+1} \otimes \ldots \otimes u_n .$$

This linear mapping takes tensors in a tensor product space with $n+2$ factors and produces tensors in a space with n factors. The simplest case arises for $n = 0$. In that case $C : V \times V^* \to \mathbb{K}$

is nothing but the bilinear pairing $C : (v, v^*) \mapsto \langle v, v^* \rangle$ and $C^\otimes$ is the linear associated mapping by the universality theorem.
Finally, let us represent the contraction mapping within the abstract index picture. It is quite simple to show that, $C^\otimes$ takes a tensor $t^{AiB}{}_j{}^C$ where ^{A}, ^{B} and ^{C} are arbitrary cumulative indices, and produces the tensor $(C^\otimes t)^{ABC} := t^{AkB}{}_k{}^C$ where we remark the convention of *summation of the double repeated index* k. To show that the abstract-index representation of contractions is that above notice that the contractions are linear and thus

$$C^\otimes(t^{AiB}{}_j{}^C e_A \otimes e_i \otimes e_B \otimes e^{*j} \otimes e_C) = t^{AiB}{}_j{}^C C^\otimes(e_A \otimes e_i \otimes e_B \otimes e^{*j} \otimes e_C) = t^{AiB}{}_j{}^C \delta^j_i e_A \otimes e_B \otimes e_C \,,$$

and thus

$$C^\otimes(t^{AiB}{}_j{}^C e_A \otimes e_i \otimes e_B \otimes e^{*j} \otimes e_C) = t^{AkB}{}_k{}^C e_A \otimes e_B \otimes e_C \,.$$

This is nothing but:

$$(C^\otimes t)^{ABC} := t^{AkB}{}_k{}^C \,.$$

To conclude we pass to consider a final theorem which shows that there is a one-to-one correspondence between linear mappings on tensors and tensors them-selves.

Theorem 2.2 (Linear mappings and tensors.) *Let S, S' be a pair of tensor spaces of a tensor algebra $\mathcal{A}_{\mathbb{K}}(V)$. The vector space of linear mappings from S to S' is naturally isomorphic to $S^* \otimes S'$ (which it is naturally isomorphic to the corresponding tensor space of $\mathcal{A}_{\mathbb{K}}(V)$). The isomorphism $F : \mathcal{L}(S|S') \to S^* \otimes S'$ is given by $f \mapsto t(f)$, where the tensor $t_f \in S^* \otimes S'$ is that (uniquely) determined by the requirement*

$$C^\otimes(s \otimes t_f) = f(s) \,, \quad \textit{for all } s \in S$$

Moreover, fixing a basis $\{e_i\}_{i\in I}$ in V, let $\{e_A\}$ denote the canonical basis induced in S, $\{e'_B\}$ that induced in S' and $\{e'^{*C}\}$ that induced in S'^*. With those definitions

$$t(f)_A{}^C = \langle f(e_A), e'^{*C} \rangle$$

and, if $s = s^A E_A \in S$,

$$f(s)^C = s^A t(f)_A{}^C \,.$$

The isomorphism F is the composition of the isomorphisms in **Theorem 1.3** *and* **Theorem 1.2**: $\mathcal{L}(S|S') \simeq \mathcal{L}(S, S'^*) = S^* \otimes (S'^*)^* \simeq S^* \otimes S'$.

Skech of proof. Defining t_f in components as said above one has a linear map $F : f \mapsto t_f$ satisfying $C^\otimes(s \otimes t_f) = f(s)$, for all $s \in S$. Since $S^* \otimes S'$ and $\mathcal{L}(S|S') \simeq \mathcal{L}(S, S'^*) \simeq S^* \otimes S'$ have the same dimension, injectivity of F would imply surjectivity. Injectivity can be straightforwardly proved using the definition of t_f itself. Similarly, using the rules to change bases as in **Theorem 2.1**, one finds the independence on the used canonical bases. The last statement can be checked by direct inspection. □

Let us illustrate how one may use that theorem. For instance consider a linear mapping $f : V \to V$ where V is finite-dimensional with field $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$. Then f defines a tensor of $V^* \otimes V$. In fact, fixing a basis $\{e_i\}_{i\in I}$in V and considering those canonically associated, by the linearity of f and the pairing:

$$f(v) = \langle f(v), e^{*k}\rangle e_k = v^i \langle f(e_i), e^{*k}\rangle e_k \, .$$

We may define the tensor $t(f) \in V^* \otimes V$ such that

$$t(f) := \langle f(e_i), e^{*k}\rangle e^{*i} \otimes e_k \, .$$

Employing **Theorem 2.1** one can trivially show that the given definition is well posed. We leave to the reader the proof that the obtained maps which associates linear functions with tensors is the isomorphism of the theorem above.
The action of f on v can be represented in terms of $t(f)$ and v using the rules presented above. Indeed, by the abstract index notation, one has

$$(f(v))^k = v^i t(f)_i{}^k \, .$$

In other words the action of f on v reduces to (1) a product of the involved tensors:

$$v^i t(f)_j{}^k \, ,$$

(2) followed by a convenient contraction:

$$(f(v))^k = v^i t(f)_i{}^k \, .$$

More complicate cases can be treated similarly. For example, linear mappings f from $V^* \otimes V$ to $V^* \otimes V \otimes V$ are determined by tensors $t(f)^i{}_{jk}{}^{lm}$ of $V \otimes V^* \otimes V^* \otimes V \otimes V$ and their action on tensors $u_p{}^q$ of $V^* \otimes V$ is

$$f(u)_k{}^{lm} = u_i{}^j t(f)^i{}_{jk}{}^{lm}$$

i.e., a product of tensors and two contractions. Obviously

$$t(f)^i{}_{jk}{}^{lm} = \left(f(e^{*i} \otimes e_j)\right)(e_k, e^{*l}, e^{*m})$$

where we have used the multi-linear action of elements of $V^* \otimes V \otimes V$ on elements of $V \times V^* \times V^*$.

2.3 Physical invariance of the form of laws and tensors.

A physically relevant result is that the rules given above to produce a new tensor from given tensors have the same form whatever basis one use to handle tensors. For that reason the abstract index notation makes sense. In physics the choice of a basis is associated with the choice of a reference frame. As is well known, various relativity principles (*Galileian Principle*, *Special*

Relativity Principle and *General Relativity Principle*) assume that
"*the law of Physics can be written in such a way that they preserve their form whatever reference frame is used*".
The allowed reference frames range in a class singled out by the considered relativity principle, for instance in Special Relativity the relevant class is that of inertial reference frames.
It is clear that the use of tensors and rules to compose and decompose tensors to represent physical laws is very helpful in implementing relativity principles. In fact the theories of Relativity can be profitably formulated in terms of tensors and operations of tensors just to assure the invariance of the physical laws under change of the reference frame. When physical laws are given in terms of tensorial relations one says that those laws are *covariant*. It is worthwhile stressing that covariance is *not* the only way to state physical laws which preserve their form under changes of reference frames. For instance the Hamiltonian formulation of mechanics, in relativistic contexts, is invariant under change of reference frame but it is not formuled in terms of tensor relations (in spacetime).

3 Some Applications.

In this section we present a few applications of the theory previously developed.

3.1 Tensor products of group representations.

As is known a *group* is an algebraic structure, $(G, \circ)$, where G is a set and $\circ : G \times G \to G$ is a mapping called the *composition rule* of the group. Moreover the following three conditions have to be satisfied.
(1) $\circ$ is *associative*, i.e.,

$$g_1 \circ (g_2 \circ g_3) = (g_1 \circ g_2) \circ g_3 \;, \quad \text{for all} \quad g_1, g_2, g_3 \in G \;.$$

(2) There is a *group unit*, i.e., there is $e \in G$ such that

$$e \circ g = g \circ e = g \;, \quad \text{for all} \quad g \in G \;.$$

(3) Each element $g \in G$ admits an *inverse element*, i.e.,

$$\text{for each} \quad g \in G \quad \text{there is} \quad g^{-1} \in G \quad \text{with} \quad g \circ g^{-1} = g^{-1} \circ g = e \;.$$

We remind the reader that the unit element turns out to be unique and so does the inverse element for each element of the group (the reader might show those uniqueness properties as an exercise). A group $(G, \circ)$ is said to be *commutative* or *Abelian* if $g \circ g' = g' \circ g$ for each pair of elements, $g, g' \in G$; otherwise it is said to be *non-commutative* or *non-Abelian*. A subset $G' \subset G$ of a group is called *subgroup* if it is a group with respect to the restriction to $G' \times G'$ of the composition rule of G.

If $(G_1, \circ_1)$ and $(G_2, \circ_2)$ are groups, a (group) *homomorphism* from G_1 to G_2 is a mapping $h : G_1 \to G_2$ which *preserves the group structure*, i.e., the following requirement has to be fulfilled:

$$h(g \circ_1 g') = h(g) \circ_2 h(g') \quad \text{for all} \quad g, g' \in G_1 \;,$$

As a consequence, they also hold with obvious notations:

$$h(e_1) = e_2 \;,$$

and

$$h(g^{-1_1}) = (h(g))^{-1_2} \quad \text{for each} \quad g \in G_1 \;.$$

Indeed, if $g \in G_1$ one has $h(g) \circ e_2 = h(g) = h(g \circ e_1) = h(g) \circ h(e_1)$. Applying $h(g)^{-1}$ on the left, one finds $e_2 = h(e_1)$. On the other hand $h(g)^{-1} \circ h(g) = e_2 = h(e_1) = h(g^{-1} \circ g) = h(g^{-1}) \circ h(g)$ implies $h(g)^{-1} = h(g^{-1})$.
A group *isomorphism* is a *bijective* group homomorphism.
A simple example of group is $GL(n, \mathbb{K})$ which is the set of the $n \times n$ matrices A with components in the field $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$ and $detA \neq 0$. The group composition rule is the usual product

of matrices. If $n > 0$ the group is non-commutative. An important subgroup of $GL(n, \mathbb{K})$ is $SL(n, \mathbb{K})$, i.e., the set of matrices B in $GL(n, \mathbb{K})$ with $detB = 1$. (The reader might show that $SL(n, \mathbb{K})$ is a subgroup of $GL(n, \mathbb{K})$.) However there are groups which are not defined as group of matrices, e.g., $(\mathbb{Z}, +)$. An non completely trivial example is given by the *group of permutations of n elements* which we shall consider in the next subsection.

Exercises 3.1.
3.1.1. Prove the uniqueness of the unit element and the inverse element in any group.
3.1.2. Show that in any group G the unique element e such that $e^2 = e$ is the unit element.
3.1.3. Show that if G' is a subgroup of G the unit element of G' must coincide with the unit element of G, and, if $g \in G'$, the inverse element g^{-1} in G' coincides with the inverse element in G.
3.1.4. Show that if $h : G_1 \to G_2$ is a group homomorphism, then $h(e_1) = e_2$ and $h(g^{-1_1}) = (h(g))^{-1_2}$ for each $g \in G_1$.
(Hint. Use **3.1.2.** to achieve the former statement.)
3.1.5. Show that if $h : G_1 \to G_2$ is a group homomorphism, then $h(G_1)$ is a subgroup of G_2.

We are interested in the concept of (linear) *representation of a group on a vector space.* In order to state the corresponding definition, notice that, if V is a (not necessarily finite-dimensional) vector space, $\mathcal{L}(V|V)$ contains an important group. This is $GL(V)$ which is the set of both injective and surjective elements of $\mathcal{L}(V|V)$ equipped with the usual composition rule of maps. We can give the following definition.

Def.3.1. (Linear group on a vector space.) *If V is a vector space, $GL(V)$ denotes the group of linear mappings $f : V \to V$ such that f is injective and surjective, with group composition rule given by the usual mappings composition. $GL(V)$ is called the* **linear group on** V.

Remarks.
(a) If $V := \mathbb{K}^n$ then $GL(V) = GL(n, \mathbb{K})$.
(b) If $V \neq V'$ it is not possible to define the analogue of $GL(V)$ considering some subset of $\mathcal{L}(V|V')$. (The reader should explain the reason.)

Def.3.2. (Linear group representation on a vector space.) *Let $(G, \circ)$ be a group and V a vector space. A (linear group)* **representation of** G **on** V *is a homomorphism $\rho : G \to GL(V)$. Moreover a representation $\rho : G \to GL(V)$ is called:*
(1) faithful *if is injective,*
(2) free *if, for any $v \in V$, the subgroup of G made of the elements h_v such that $\rho(h_v)v = v$ contains only the unit element of G,*
(3) transitive *if for each pair $v, v' \in V$ there is $g \in G$ with $v' = \rho(g)v$.*
(4) irreducible *if there is no* proper *vector subspace $S \subset V$ which is* **invariant** *under the action of $\rho(G)$, i.e., which satisfies $\rho(g)S \subset S$ for all $g \in G$.*

Equipped with the above-given definitions we are able to study the simplest interplay of tensors and group representations. We want to show that the notion of tensor product allows the definitions of *tensor products of representations.* That mathematical object is of fundamental importance in applications to Quantum Mechanics, in particular as far as systems with many components are concerned.
Consider a group G (from now on we omit to specify the symbol of the composition rule whenever it does not produces misunderstandings) and several representations of the group $\rho_i : G \to GL(V_i)$, where $V_1, \ldots, V_n$ are finite-dimensional vector spaces on the same field $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$. For each $g \in G$, we may define a multi-linear mapping $[\rho_i(g), \ldots, \rho_n(g)] \in \mathcal{L}(V_1, \ldots, V_n | V_1 \otimes \ldots \otimes V_n)$ given by, for all $(\rho_i(g), \ldots, \rho_n(g)) \in V_1 \times \ldots \times V_n$,

$$[\rho_i(g), \ldots, \rho_n(g)] : (v_1, \ldots, v_n) \mapsto (\rho_1(g)v_1) \otimes \ldots \otimes (\rho_n(g)v_n) .$$

That mapping is multi linear because of the multi linearity of the tensor-product mapping and the linearity of the operators $\rho_k(g)$. Using the universality theorem, we *uniquely* find a linear mapping which we indicate by $\rho_1(g) \otimes \ldots \otimes \rho_n(g) : V_1 \otimes \ldots \otimes V_n \to V_1 \otimes \ldots \otimes V_n$ such that:

$$\rho_1(g) \otimes \ldots \otimes \rho_n(g)(v_1 \otimes \ldots \otimes v_n) = (\rho_1(g)v_1) \otimes \ldots \otimes (\rho_n(g)v_n) .$$

Def.3.3. (Tensor product of representations.) *Let $V_1, \ldots V_n$ be finite-dimensional vector spaces on the same field $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$ and suppose there are n representations $\rho_i : G \to GL(V_k)$ of the same group G on the given vector spaces. The set of linear maps*

$$\{\rho_1(g) \otimes \ldots \otimes \rho_n(g) : V_1 \otimes \ldots \otimes V_n \to V_1 \otimes \ldots \otimes V_n \mid g \in G\} ,$$

defined above is called **tensor product of representations** $\rho_1, \ldots, \rho_n$.

The relevance of the definition above is evident because of the following theorem,

Theorem 3.1. *Referring to the definition above, the elements of a tensor product of representations $\rho_1, \ldots \rho_n$ of the group G on spaces $V_1, \ldots, V_n$ define a* **linear group representation of G on the tensor-product space** $V_1 \otimes \ldots \otimes V_n$.

Proof. We have to show that the mapping

$$g \mapsto \rho_1(g) \otimes \ldots \otimes \rho_n(g) ,$$

is a group homomorphism from G to $GL(V_1 \otimes \ldots \otimes V_n)$. Taking account of the fact that each ρ_i is a group homomorphism, if $g, g' \in G$, one has

$$\rho_1(g')\otimes\ldots\otimes\rho_n(g')(\rho_1(g)\otimes\ldots\otimes\rho_n(g)(v_1\otimes\ldots\otimes v_n)) = \rho_1(g')\otimes\ldots\otimes\rho_n(g')((\rho_1(g)v_1)\otimes\ldots\otimes(\rho_n(g)v_n))$$

and this is

$$(\rho_1(g' \circ g)v_1) \otimes \ldots \otimes (\rho_n(g' \circ g)v_n) .$$

The obtained result holds true also using a canonical basis for $V_1 \otimes \ldots \otimes V_n$ made of usual elements $e_{1,i_1} \otimes \ldots e_{n,i_n}$ in place of $v_1 \otimes \ldots \otimes v_n$. By linearity, it means that

$$(\rho_1(g') \otimes \ldots \otimes \rho_n(g'))(\rho_1(g) \otimes \ldots \otimes \rho_n(g)) = \rho_1(g' \circ g) \otimes \ldots \otimes \rho_n(g' \circ g) .$$

To conclude notice that $\rho_1(g) \otimes \ldots \otimes \rho_n(g) \in GL(V_1 \otimes \ldots \otimes V_n)$ because $\rho_1(g) \otimes \ldots \otimes \rho_n(g)$ is (1) linear and furthermore it is (2) bijective. The latter can be proved as follows: $\rho_1(g^{-1}) \otimes \ldots \otimes \rho_n(g^{-1}) \circ (\rho_1(g) \otimes \ldots \otimes \rho_n(g)) = (\rho_1(g) \otimes \ldots \otimes \rho_n(g)) \circ \rho_1(g^{-1}) \otimes \ldots \otimes \rho_n(g^{-1}) = \rho_1(e) \otimes \ldots \otimes \rho_n(e) = I$. The last identity follows by linearity form $(\rho_1(e) \otimes \ldots \otimes \rho_n(e))(v_1 \otimes \cdots \otimes v_n) = v_1 \otimes \cdots \otimes v_n$. $\Box$

More generally, if $A_k : V_k \to U_k$ are n linear mappings (operators), and all involved vector spaces are finite dimensional and with the same field $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$, it is defined the *tensor product of operators*.

Def.3.4 (Tensor Product of Operators.) *If $A_k : V_k \to U_k$, $k = 1, \ldots, n$ are n linear mappings (operators), and all the vector spaces U_i, V_j are finite dimensional with the same field $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$, the* **tensor product of** *$A_1, \ldots, A_n$ is the linear mapping*

$$A_1 \otimes \ldots \otimes A_n : V_1 \otimes \ldots \otimes V_n \to U_1 \otimes \ldots \otimes U_n$$

uniquely determined by the universality theorem and the requirement:

$$(A_1 \otimes \ldots \otimes A_n) \circ \otimes = A_1 \times \ldots \times A_n ,$$

where

$$A_1 \times \ldots \times A_n : V_1 \times \ldots \times V_n \to U_1 \otimes \ldots \otimes U_n ,$$

is the multi-linear mapping such that, for all $(v_1, \ldots, v_n) \in V_1 \times \ldots \times V_n$:

$$A_1 \times \ldots \times A_n : (v_1, \ldots, v_n) \mapsto (A_1 v_1) \otimes \ldots \otimes (A_n v_n) .$$

Remark. Employing the given definition and the same proof used for the relevant part of the proof of **Theorem 3.1**, it is simply shown that if $A_i : V_i \to U_i$ and $B_i : U_i \to W_i$, $i = 1, \ldots, n$ are $2n$ linear maps and all involved spaces V_i, U_j, W_k are finite dimensional with the same field $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$, then

$$B_1 \otimes \ldots \otimes B_n \;\; A_1 \otimes \ldots \otimes A_n = (B_1 A_1) \otimes \ldots \otimes (B_n A_n) .$$

3.2 A quantum physical example.

Physicsts are involved with Hilbert spaces whenever they handle quantum mechanics. A Hilbert space is noting but a complex vector space equipped with a Hermitean scalar product (see the next chapter) such that it is complete with respect to the norm topology induced by that scalar product. As far as we are concerned we need only the structure of vector space. Physically

speaking, the vectors of the Hilbert space represent the states of the considered physical system (actually things are more complicated but we do not matter). To consider the simplest case we assume that the vector space which describes the states of the system is finite-dimensional (that is the case for the spin part of a quantum particle). Moreover, physics implies that the space $\mathcal{H}$ of the states of a composit system S made of two systems S_1 and S_2 associated with Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, is the tensor product $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. There should be several remarks in the infinite dimensional case since our definition of tensor product works for finite-dimensional spaces only, however suitable and well-behaved generalizations are, in fact, possible. Let the system S_1 be described by a state $\psi \in \mathcal{H}_1$ and suppose to transform the system by the action of an element R of some physical group of transformations $\mathcal{G}$ (e.g. $SO(3)$). The transformed state ψ' is given by $U_R^{(1)}\psi$ where $\mathcal{G} \ni R \mapsto U_R^{(1)}$ is a representation of $\mathcal{G}$ in terms of linear transformations $U_R^{(1)} : \mathcal{H}_1 \to \mathcal{H}_1$. Actually, physics and the celebrated *Wigner's theorem* in particular, requires that every $U_R^{(1)}$ be a *unitary* (or *anti unitary*) transformation but this is not relevant for our case. The natural question concerning the representation of the action of $\mathcal{G}$ on the composit system S is:
"If we know the representations $\mathcal{G} \ni R \mapsto U_R^{(1)}$ and $\mathcal{G} \ni R \mapsto U_R^{(2)}$, what about the representation of the action of $\mathcal{G}$ on S in terms of linear transformations in the Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$?"
The answer given by physics, at least when the systems S_1 and S_2 do not interact, is that $U_g := U_g^{(2)} \otimes U_g^{(1)}$.

3.3 Permutation group and symmetry of tensors.

We remind the definition of the *group of permutations of n objects* and give some know results of basic group theory whose proofs may be found in any group-theory textbook.

Def.3.5. (Group of permutations.) *Consider the set $I_n := \{1, \ldots, n\}$, the* **group of permutations of** n **objects**, *$\mathcal{P}_n$ is the set of the bijective mappings $\sigma : I_n \to I_n$ equipped with the composition rule given by the usual composition rule of functions. Moreover,*
(a) *the elements of $\mathcal{P}_n$ are called* **permutations** *(of n objects);*
(b) *a permutation of $\mathcal{P}_n$ with $n \geq 2$ is said to be a* **transposition** *if differs from the identity mapping and reduces to the identity mapping when restricted to some subset of I_n containing $n-2$ elements.*

Comments.
(1) $\mathcal{P}_n$ contains $n!$ elements.
(2) Each permutation $\sigma \in \mathcal{P}_n$ can be represented by a corresponding string $(\sigma(1), \ldots, \sigma(n))$.
(3) If, for instance $n = 5$, with the notation above $(1, 2, 3, 5, 4)$, $(5, 2, 3, 4, 1)$, $(1, 2, 4, 3, 5)$ are transpositions, $(2, 3, 4, 5, 1)$, $(5, 4, 3, 2, 1)$ are not.
(3) It is possible to show that each permutation $\sigma \in \mathcal{P}_n$ can be decomposed as a product of transpositions $\sigma = \tau_1 \circ \ldots \circ \tau_k$. In general there are several different transposition-product decompositions for each permutation, however it is possible to show that if $\sigma = \tau_1 \circ \ldots \circ \tau_k =$

$\tau'_1 \circ \ldots \circ \tau'_r$, where τ_i and τ'_j are transpositions, then $r+k$ is even. Equivalently, r is even or odd if and only if k is so. This defines the **parity**, $\epsilon_\sigma \in \{-1,+1\}$, of a permutation σ, where, $\epsilon_\sigma = +1$ if σ can be decomposed as a product of an *even* number of transpositions and $\epsilon_\sigma = -1$ if σ can be decomposed as a product of an *odd* number of transpositions.
(4) If $A = [A_{ij}]$ is a real or complex $n \times n$ matrix, it is possible to show (by induction) that:

$$detA = \sum_{\sigma \in \mathcal{P}_n} \epsilon_\sigma A_{1\sigma(1)} \cdots A_{n\sigma(n)} .$$

Alternatively, the identity above may be used to define the determinant of a matrix.

We pass to consider the action of $\mathcal{P}_n$ on tensors. Fix $n > 1$, consider the tensor algebra $\mathcal{A}_{\mathbb{K}}(K)$ and single out the tensor space $V^{n\otimes} := V \otimes \ldots \otimes V$ where the factor V appears n times Then consider the following action of $\mathcal{P}_n$ on $V^{n\times} := V \times \ldots \times V$ where the factor V appears n times. For each $\sigma \in \mathcal{P}_n$ consider the mapping

$$\hat{\sigma} : V^{n\times} \to V^{n\otimes} \quad \text{such that} \quad \hat{\sigma}(v_1, \ldots, v_n) \mapsto v_{\sigma^{-1}(1)} \otimes \ldots \otimes v_{\sigma^{-1}(n)} .$$

It is quite straightforward to show that $\hat{\sigma}$ is multi linear. Therefore, let

$$\sigma^\otimes : V^{n\otimes} \to V^{n\otimes} ,$$

be the linear mapping uniquely determined by $\hat{\sigma}$ by means of the universality theorem. By definition, it is completely determined by linearity and the requirement

$$\sigma^\otimes : v_1 \otimes \ldots \otimes v_n \mapsto v_{\sigma^{-1}(1)} \otimes \ldots \otimes v_{\sigma^{-1}(n)} .$$

Theorem 3.2. *The above-defined linear mapping*

$$\sigma \mapsto \sigma^\otimes$$

with $\sigma \in \mathcal{P}_n$*, is a group representation of* $\mathcal{P}_n$ *on* $V^{n\otimes}$*.*

Proof. First we show that if $\sigma, \sigma' \in \mathcal{P}_n$ then

$$\sigma^\otimes(\sigma'^\otimes(v_1 \otimes \ldots \otimes v_n)) = (\sigma \circ \sigma')^\otimes (v_1 \otimes \ldots \otimes v_n) ,$$

This follows from the definition: $\sigma^\otimes(\sigma'^\otimes(v_1 \otimes \ldots \otimes v_n)) = \sigma^\otimes(v_{\sigma'^{-1}(1)} \otimes \ldots \otimes v_{\sigma'^{-1}(n)})$. Re-defining $u_i := v_{\sigma'^{-1}(i)}$ so that $u_{\sigma^{-1}(j)} := v_{\sigma'^{-1}(\sigma^{-1}(j))}$, one finds the identity $\sigma^\otimes(\sigma'^\otimes(v_1 \otimes \ldots \otimes v_n)) = u_{\sigma^{-1}(1)} \otimes \ldots \otimes u_{\sigma^{-1}(n)} = v_{\sigma'^{-1}\circ\sigma^{-1}(1)} \otimes \ldots \otimes v_{\sigma'^{-1}\circ\sigma^{-1}(n)} = v_{(\sigma\circ\sigma')^{-1}(1)} \otimes \ldots \otimes v_{(\sigma\circ\sigma')^{-1}(n)} = (\sigma \circ \sigma')^\otimes (v_1 \otimes \ldots \otimes v_n)$. In other words

$$\sigma^\otimes(\sigma'^\otimes(v_1 \otimes \ldots \otimes v_n)) = (\sigma \circ \sigma')^\otimes (v_1 \otimes \ldots \otimes v_n) .$$

In particular, that identity holds alsofor a canonical basis of elements $e_{i_1} \otimes \ldots \otimes e_{i_n}$

$$\sigma^\otimes(\sigma'^\otimes(e_{i_1} \otimes \ldots \otimes e_{i_n})) = (\sigma \circ \sigma')^\otimes (e_{i_1} \otimes \ldots \otimes e_{i_n}) .$$

By linearity such an identity will hold true for all arguments in $V^{n\otimes}$ and thus we have

$$\sigma^{\otimes}\sigma'^{\otimes} = (\sigma \circ \sigma')^{\otimes} .$$

The mappings $\sigma^{\otimes}$ are linear by constructions and are bijective because

$$\sigma^{\otimes}\sigma'^{\otimes} = (\sigma \circ \sigma')^{\otimes}$$

implies

$$\sigma^{\otimes}\sigma^{-1\otimes} = \sigma^{-1\otimes}\sigma^{\otimes} = e^{\otimes} = I .$$

The identity $e^{\otimes} = I$ can be proven by noticing that $e^{\otimes} - I$ is linear and vanishes when evaluated on any canonical base of $V^{n\otimes}$. We have shown that $\sigma^{\otimes} \in GL(V^{n\otimes})$ for all $\sigma \in \mathcal{P}_n$ and the mapping $\sigma \mapsto \sigma^{\otimes}$ is a homomorphism. This concludes the proof. □

Let us pass to consider the abstract index notation and give a representation of the action of $\sigma^{\otimes}$ within that picture.

Theorem 3.3. *If t is a tensor in $V^{n\otimes} \in \mathcal{A}_{\mathbb{K}}(V)$ with $n \geq 2$ and $\sigma \in \mathcal{P}_n$, then the components of t with respect to any canonical basis of $V^{n\otimes}$ satisfy*

$$(\sigma^{\otimes}t)^{i_1\cdots i_n} = t^{i_{\sigma(1)}\cdots i_{\sigma(n)}} .$$

Proof.

$$\sigma^{\otimes}t = \sigma^{\otimes}(t^{j_1\cdots j_n}e_{j_1} \otimes \ldots \otimes e_{j_n}) = t^{j_1\cdots j_n}e_{j_{\sigma^{-1}(1)}} \otimes \cdots \otimes e_{j_{\sigma^{-1}(n)}} .$$

Since $\sigma : I_n \to I_n$ is bijective, if we define $i_k := j_{\sigma^{-1}(k)}$, it holds $j_k = i_{\sigma(k)}$. Using this identity above we find

$$\sigma^{\otimes}t = t^{i_{\sigma(1)}\cdots i_{\sigma(n)}}e_{i_1} \otimes \ldots \otimes e_{i_n} .$$

That is nothing but the thesis. □

To conclude we introduce the concept of symmetric or anti-symmetric tensor.

Def. 3.6 (Symmetric and anti-symmetric tensors.) *Let V be a finite-dimensional vector space with field $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$. Consider the space $V^{n\otimes} \in \mathcal{A}_{\mathbb{K}}(V)$ of tensors of order $(n, 0)$ with $n \geq 2$.*
(a) *$t \in V^{n\otimes}$ is said to be* **symmetric** *if*

$$\sigma^{\otimes}t = t ,$$

for all of $\sigma \in \mathcal{P}_n$, or equivalently, using the abstract index notation,

$$t^{j_1\cdots j_n} = t^{j_{\sigma(1)}\cdots j_{\sigma(n)}} ,$$

for all of $\sigma \in \mathcal{P}_n$.
(b) $t \in V^{n\otimes}$ *is said to be* **anti symmetric** *if*

$$\sigma^{\otimes} t = \epsilon_\sigma \, t \, ,$$

for all of $\sigma \in \mathcal{P}_n$*, or equivalently using the abstract index notation:*

$$t^{j_1 \cdots j_n} = \epsilon_\sigma \, t^{j_{\sigma(1)} \cdots j_{\sigma(n)}} \, ,$$

for all of $\sigma \in \mathcal{P}_n$.

Remark. Concerning the definitition in **(b)** notice that $\epsilon_\sigma = \epsilon_{\sigma^{-1}}$.

Examples 3.1.
3.1.1. Suppose $n = 2$, then a symmetric tensor $s \in V \otimes V$ satisfies $s^{ij} = s^{ji}$ and an antisymmetric tensor $a \in V \otimes V$ satisfies $a^{ij} = -a^{ji}$.
3.1.2. Suppose $n = 3$, then it is trivially shown that $\sigma \in \mathcal{P}_3$ has parity 1 if and only if σ is a **cyclic permutation**, i.e., $(\sigma(1), \sigma(2), \sigma(3)) = (1, 2, 3)$ or $(\sigma(1), \sigma(2), \sigma(3)) = (2, 3, 1)$ or $(\sigma(1), \sigma(2), \sigma(3)) = (3, 1, 2)$.
Now consider the vector space V with $dimV = 3$. It turns out that a tensor $e \in V \otimes V \otimes V$ is anti symmetric if and only if

$$e^{ijk} = 0 \, ,$$

if (i, j, k) is *not* a permutation of $(1, 2, 3)$ and, otherwise,

$$e^{ijk} = \pm e^{123} \, ,$$

where the sign $+$ takes place if the permutation $(\sigma(1), \sigma(2), \sigma(3)) = (i, j, k)$ is cyclic and the sign $-$ takes place otherwise. That relation between parity of a permutation and cyclicity does *not* hold true for $n > 3$.

Remarks.
(1) Consider a generic tensor space $S \in \mathcal{A}_{\mathbb{K}}(V)$ which contains $n \geq 2$ spaces V as factors. We may suppose for sake of simplicity $S = S_1 \otimes V^{n\otimes} \otimes S_2$ where $S_1 = U_1 \otimes \ldots \otimes U_{n-1}$, $S_2 = U_{n+1} \otimes \ldots \otimes U_m$ and $U_i = V$ or $U_i = V^*$. Anyway all what we are going to say holds true also if the considered n spaces V do not define a unique block $V^{n\otimes}$. We may define the action of $\sigma \in \mathcal{P}_n$ on the whole space S starting by a multi linear mapping

$$\hat{\sigma} : U_1 \times \ldots \times U_k \times V^{n\times} \times U_{k+1} \times \ldots \times U_m \to U_1 \otimes \ldots \otimes U_k \otimes V^{n\otimes} \otimes U_{k+1} \otimes \ldots \otimes U_m \, ,$$

such that reduces to the tensor-product mapping on $U_1 \times \ldots \times U_k$ and $U_{k+1} \times \ldots \times U_m$:

$$\sigma : (u_1, \ldots, u_k, v_1, \ldots v_n, u_k, \ldots, u_m) \mapsto u_1 \otimes \ldots \otimes u_k \otimes v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(n)} \otimes u_{k+1} \otimes \ldots \otimes u_m \, .$$

Using the universality theorem as above, we build up a representation of $\mathcal{P}_n$ on S, $\sigma \mapsto \sigma^{\otimes}$ which "acts on $V^{n\otimes}$ only". Using the abstract index notation the action of $\sigma^{\otimes}$ is well represented:

$$\sigma^{\otimes} : t^{A\ i_1 \ldots i_n\ B} \mapsto t^{A\ i_{\sigma(1)} \ldots i_{\sigma(n)}\ B} .$$

This allows one to define and study the symmetry of a tensor referring to a few indices singled out among the complete set of indices of the tensors. E.g., a tensor $t^{ij}{}_k{}^r$ may be symmetric or anti symmetric, for instance, with respect to the indices i and j or j and r or ijr.

(2) Everything we have obtained and defined may be similarly re-obtained and re-defined considering spaces of tensors of order $(0, n)$ with $n \geq 2$, i.e. covariant tensors. In that case the *antisymmetric tensors* of order $(0, n)$ are called n**-forms**.

(3) Notice that no discussion on the symmetry of indices of different kind (one covariant and the other contravariant) is possible.

Exercises 3.2.

3.2.1. Let t be a tensor in $V^{n\otimes}$ (or $V^{*n\otimes}$). Show that t is symmetric or anti symmetric if there is a canonical basis where the components have symmetric or anti symmetric indices, i.e., $t^{i_1 \ldots i_n} = t^{i_{\sigma(1)} \cdots i_{\sigma(n)}}$ or respectively $t^{i_1 \cdots i_n} = \epsilon_\sigma\, t^{i_{\sigma(1)} \cdots i_{\sigma(n)}}$ for all $\sigma \in \mathcal{P}_n$.

Note. The result implies that, to show that a tensor is symmetric or anti symmetric, it is sufficient to verify the simmetry or anti symmetry of its components within a *single* canonical basis.

3.2.2. Show that the sets of symmetric tensors of order $(n, 0)$ and $(0, n)$ are vector subspaces of $V^{n\otimes}$ and $V^{*n\otimes}$ respectively.

3.2.3. Show that the subspace of anti-symmetric tensors of order $(0, n)$ (the space of n-forms) in $V^{*n\otimes}$ has dimension $\binom{dimV}{n}$ if $n \leq dimV$. What about $n > dimV$?

3.2.4. Consider a tensor $t^{i_1 \cdots i_n}$, show that the tensor is symmetric if and only if it is symmetric with respect to each arbitrary chosen pair of indices, i.e.

$$t^{\cdots i_k \cdots i_p \cdots} = t^{\cdots i_p \cdots i_k \cdots} ,$$

for all $p, k \in \{1, \ldots n\}$, $p \neq k$.

3.2.5. Consider a tensor $t^{i_1 \cdots i_n}$, show that the tensor is anti symmetric if and only if it is anti symmetric with respect to each arbitrarily chosen pair of indices, i.e.

$$t^{\cdots i_k \cdots i_p \cdots} = -t^{\cdots i_p \cdots i_k \cdots} ,$$

for all $p, k \in \{1, \ldots n\}$, $p \neq k$.

3.2.6. Show that $V \otimes V = A \oplus S$ where $\oplus$ denotes the direct sum and A and S are respectively the space of anti-symmetric and symmetric tensors in $V \otimes V$. Does such a direct decomposition hold if considering $V^{n\otimes}$ with $n > 2$?

4 Scalar Products and Metric Tools.

This section concerns the introduction of the notion of scalar product and several applications on tensors.

4.1 Scalar products.

First of all we give the definition of a *pseudo scalar product* and *semi scalar products* which differ from the notion of *scalar product* for the positivity and the non-degenerateness requirement respectively. In fact, a pseudo scalar product is a generalization of the usual definition of scalar product which has many applications in mathematical physics, relativistic theories in particular. Semi scalar products are used in several applications of quantum field theory (for instance in the celebrated *GNS theorem*).

Def.4.1. (Pseudo Scalar Product.) *Let V be a vector space on the field either $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.*
(a) *A* **pseudo scalar product** *is a mapping $(\,|\,) : V \times V \to \mathbb{K}$ which is:*
(i) **bi linear**, *i.e., for all $u \in V$ both $(u|\cdot) : v \mapsto (u|v)$ and $(\cdot|u) : v \mapsto (v|u)$ are linear functionals on V;*
(ii) **symmetric**, *i.e., $(u|v) = (v|u)$ for all $u, v \in V$;*
(iii) **non-degenerate**, *i.e., $(u|v) = 0$ for all $v \in V$ implies $u = 0$.*
(b) *If $\mathbb{K} = \mathbb{C}$, a* **Hermitian pseudo scalar product** *is a mapping $(\,|\,) : V \times V \to \mathbb{K}$ which is:*
(i) **sesquilinear**, *i.e., for all $u \in V$, $(u|\cdot)$ and $(\cdot|u)$ are a linear functional and an anti-linear functional on V respectively;*
(ii) **Hermitian**, *i.e., $(u|v) = \overline{(v|u)}$ for all $u, v \in V$;*
(iii) **non-degenerate**.

Def.4.2 (Semi Scalar Product.) *Let V be a vector space on the field either $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.*
(a) *If $\mathbb{K} = \mathbb{R}$, a* **semi scalar product** *is a mapping $(\,|\,) : V \times V \to \mathbb{R}$ which satisfies* (ai),(aii) *above and is*
(iv) **semi-defined positive**, *i.e., $(u|u) \geq 0$ for all $u \in V$.*
(b) *If $\mathbb{K} = \mathbb{C}$, a* **Hermitian semi scalar product** *is a mapping $(\,|\,) : V \times V \to \mathbb{K}$ which satisfies* (bi),(bii) *above and is*
(iv) **semi-defined positive.**

Finally we give the definition of scalar product.

Def.4.3. (Scalar Product.) *Let V be a vector space on the field $\mathbb{K} = \mathbb{R}$ (resp. $\mathbb{C}$) endowed with a pseudo scalar product (resp. Hermitian pseudo scalar product) $(\,|\,)$. $(\,|\,)$ is called* **scalar product** *(***resp. Hermitian scalar product***) if $(\,|\,)$ is also a semi scalar product, i.e., if it is* **semi-defined positive**.

Comments.
(1) Notice that all given definitions do not require that V is finite dimensional.
(2) If $\mathbb{K} = \mathbb{C}$ and $(|)$ is not Hermitian, in general, any requirement on positivity of $(u|u)$ does not make sense because $(u|u)$ may *not* be real. If instead Hermiticity holds, we have $(u|u) = \overline{(u|u)}$ which assures that $(u|u) \in \mathbb{R}$ and thus positivity may be investigated.
(3) Actually, a (Hermitian) scalar product is *positive defined*, i.e,

$$(u|u) > 0 \quad \text{if} \quad u \in V \setminus \{0\} \,,$$

because of Cauchy-Schwarz' inequality

$$|(u|v)|^2 \leq (u|u)(v|v) \,,$$

which we shall prove below for semi scalar products.
(4) A **semi norm** on a vector space V with field $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$, is a mapping $|| \ || : V \to \mathbb{K}$ such that:
(i) $||v|| \in \mathbb{R}$ and in particular $||v|| \geq 0$ for all $v \in V$;
(ii) $||\alpha v|| = |\alpha| ||v||$ for all $\alpha \in \mathbb{K}$ and $v \in V$;
(iii) $||u + v|| \leq ||u|| + ||v||$ for all $u, v \in V$.
A semi norm $|| \ || : V \to \mathbb{K}$ is a **norm** if
(iv) $||v|| = 0$ implies $v = 0$.
Notice that for semi norms it holds: $||0|| = 0$ because of (ii) above. With the given definitions, it quite simple to show (the reader might try to give a proof) that if V with field $\mathbb{K} = \mathbb{R}$ $(\mathbb{C})$ is equipped by a (Hermitian) semi scalar product then $||v|| := \sqrt{(v|v)}$ for all $v \in V$ defines a semi norm. Furthermore if $(\ |\ )$ is a scalar product, then the associated semi norm is a norm.
(5) If a vector space V is equipped with a norm $|| \ ||$ it becomes a *metric space* by defining the *distance* $d(u, v) := ||u - v||$ for all $u, v \in V$. A **Banach space** $(V, || \ ||)$ is a vector space equipped with a norm such that the associates metric space is *complete*, i.e, all Cauchy's sequences converge. A **Hilbert space** is a Banach space with norm given by a (Hermitean if the field is $\mathbb{C}$) scalar product as said above. Hilbert spaces are the central mathematical objects used in quantum mechanics.

Exercices 4.1.
4.1.1. Show that if $(\ |\ )$ is a (Hermitian) semi scalar product on V with field $\mathbb{R}$ $(\mathbb{C})$ then the mapping on V, $v \mapsto ||v|| := \sqrt{(v|v)}$, satisfies $||u + v|| \leq ||u|| + ||v||$ as a consequence of Cauchy-Schwarz' inequality

$$|(u|v)|^2 \leq (u|u)(v|v) \,,$$

which holds true by all (Hermitian) semi scalar product.
(Hint. Compute $||u + v||^2 = (u + v|u + v)$ using bi linearity or sesquilinearity property of $(\ |\ )$, then use Cauchy-Schwarz' inequality.)

Theorem 4.1. (Cauchy-Schwarz' inequality.) *Let V be a vector space with field $\mathbb{R}$ ($\mathbb{C}$) equipped with a (Hermitian) semi scalar product $(\,|\,)$. Then, for all $u, v \in V$, Cauchy-Schwarz' inequality holds:*

$$|(u|v)|^2 \leq (u|u)(v|v)\,.$$

Proof. Consider the complex case with a Hermitian semi scalar product. Take $u, v \in V$. For all $z \in \mathbb{C}$ it must hold $(zu + v|zu + v) \geq 0$ by definition of Hermitean semi scalar product. Using sesquilinearity and Hermiticity :

$$0 \leq \overline{z}z(u|u) + (v|v) + \overline{z}(u|v) + z(v|u) = |z|^2(u|u) + (v|v) + \overline{z}(u|v) + z\overline{(u|v)}\,,$$

which can be re-written as

$$|z|^2(u|u) + (v|v) + 2Re\{\overline{z}(u|v)\} \geq 0\,.$$

Then we pass to the polar representation of z, $z = re^{i\,\alpha}$ with $r, \alpha \in \mathbb{R}$ arbitrarily and independently fixed. Decompose also $(u|v)$, $(u|v) = |(u|v)|e^{i\ arg(u|v)}$. Inserting above we get:

$$F(r,\alpha) := r^2(u|u) + 2r|(u|v)|Re[e^{i(arg(u|v)\ -\ \alpha)}] + (v|v) \geq 0\,,$$

for all $r \in \mathbb{R}$ when $\alpha \in R$ is fixed arbitrarily. Since the right-hand side above is a second-order polynomial in r, the inequlity implies that, for all $\alpha \in \mathbb{R}$,

$$\left\{2|(u|v)|Re[e^{i(arg(u|v)\ -\ \alpha)}]\right\}^2 - 4(v|v)(u|u) \leq 0\,,$$

which is equivalent to

$$|(u|v)|^2\cos(arg(u|v)\ -\ \alpha) - (u|u)(v|v) \leq 0\,,$$

for all $\alpha \in \mathbb{R}$. Choosing $\alpha = arg(u|v)$, we get Cauchy-Schwarz' inequality:

$$|(u|v)|^2 \leq (u|u)(v|v)\,.$$

The real case can be treated similarly, replacing $z \in \mathbb{C}$ with $x \in \mathbb{R}$ and the proof is essentially the same. $\Box$.

Corollary. *A bilinear symmetric (resp. sesquilinear Hermitian) mapping $(\,|\,) : V \times V \to \mathbb{R}$ (resp. $\mathbb{C}$) is a scalar product (resp. Hermitian scalar product) if and only if it is* **positive defined**, *that is $(u|u) > 0$ for all $u \in V \setminus \{\mathbf{0}\}$.*

Proof. Assume that $(\,|\,)$ is a (Hermitian) scalar product. Hence $(u|u) \geq 0$ by definition and $(\,|\,)$ is non-degenerate. Moreover it holds $|(u|v)|^2 \leq (u|u)(v|v)$. As a consequence, if $(u|u) = 0$ then $(u|v) = 0$ for all $v \in V$ and thus $u = 0$ because $(\,|\,)$ is non-degenerate. We have proven that $(u|u) > 0$ if $u \neq 0$. That is, a scalar product is a positive-defined bilinear symmetric (resp. sesquilinear Hermitian) mapping $(\,|\,) : V \times V \to \mathbb{R}$ (resp. $\mathbb{C}$). Now assume that $(\,|\,)$ is positive-defined bilinear symmetric (resp. sesquilinear Hermitian). By definition it is a semi scalar product since positive definiteness implies positive semi-definiteness. Let us prove that $(|)$ is non-degenerate and this concludes the proof. If $(u|v) = 0$ for all $v \in V$ then, choosing $v = u$, the positive definiteness implies $u = \mathbf{0}$. $\Box$

4.2 Natural isomorphism between V and V^* and metric tensor.

Let us show that if V is a finite-dimensional vector space endowed with a pseudo scalar product, V is isomorphic to V^*. That isomorphism is *natural* because it is built up using the structure of vector space with scalar product only, specifying nothing further.

Theorem 4.2. (Natural (anti)isomorphism between V and V^*.) *Let V be a finite-dimensional vector space with field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.*
(a) *If V is endowed with a pseudo scalar product $(\,|\,)$ (also if $\mathbb{K} = \mathbb{C}$),*
(i) *the mapping defined on V, $h : u \mapsto (u|\cdot)$, where $(u|\cdot)$ is the linear functional, $(u|\cdot) : v \mapsto (u|v)$, is an isomorphism;*
(ii) *$(h(u)|h(v))^* := (u|v)$ defines a pseudo scalar product on V^*.*
(b) *If $\mathbb{K} = \mathbb{C}$ and V is endowed with a Hermitean pseudo scalar product $(\,|\,)$,*
(i) *the mapping defined on V, $h : u \mapsto (u|\cdot)$, where $(u|\cdot)$ is the linear functional, $(u|\cdot) : v \mapsto (u|v)$, is an anti isomorphism;*
(ii) *$(h(u)|h(v))^* := \overline{(u|v)}$ $(= (v|u))$ defines a Hermitian pseudo scalar product on V^*.*

Proof. First consider (i) in the cases (a) and (b). It is obvious that $(u|\cdot) \in V^*$ in both cases. Moreover the linearity or antilinearity of the mapping $u \mapsto (u|\cdot)$ is a trivial consequence of the definition of pseudo scalar product and Hermitian pseudo scalar product respectively.
Then remind the well-known theorem, $dim(Kerf) + dimf(V) = dimV$, which holds true for linear and anti-linear mappings from some finite-dimensional vector space V to some vector space V'. Since $dimV = dimV^*$, it is sufficient to show that $h : V \to V^*$ defined by $u \mapsto (u|\cdot)$ has trivial kernel, i.e., is injective: this also assures the surjectivity of the map. Therefore, we have to show that $(u|\cdot) = (u'|\cdot)$ implies $u = u'$. This is equivalent to show that $(u - u'|v) = 0$ for all $v \in V$ implies $u - u' = 0$. This is nothing but the non-degenerateness property, which holds by definition of (Hermitean) scalar product.
Statements (ii) cases are obvious in both by definition of (Hermitian) pseudo scalar products using the fact that h is a (anti) isomorphism. $\Box$

Remarks.
(1) Notice that, with the definitions given above it holds also $(u|v)^* = (h^{-1}u|h^{-1}v)$ and, for the Hermitian case, $(u|v)^* = \overline{(h^{-1}u|h^{-1}v)}$ for all $u, v \in V^*$. This means that h and h^{-1} *(anti)preserve* the scalar products.
(2) The theorem above holds also considering a Hilbert space and its topological dual space

(i.e., the subspace of the dual space consisting of continuous linear functionals on the Hilbert space). That is the mathematical content of celebrated Riesz' representation theorem.

Exercise 4.2.
4.2.1. Show that, for all $u, v \in V$:

$$(u|v) = \langle u, h(v)\rangle = (h(v)|h(u))^* ,$$

no matter if $(\,|\,)$ is Hermitian or not.

From now on we specialize to the pseudo-scalar-product case dropping the Hermitian case. Suppose $(|)$ is a pseudo scalar product on a finite-dimensional vector space V with field $\mathbb{K}$ either $\mathbb{R}$ or $\mathbb{C}$. The mapping $(u, v) \mapsto (u|v)$ is bi linear on $V \times V$ and thus is a tensor $g \in V^* \otimes V^*$. Fixing a canonical basis in $V^* \otimes V^*$ induced by a basis $\{e_i\}_{i\in I} \subset V$, we can write:

$$\mathbf{g} = g_{ij} e^{*i} \otimes e^{*j} ,$$

where, by **Theorem 1.5**,

$$g_{ij} = (e_i|e_j) .$$

Def.4.4. (Metric Tensor.) *A pseudo scalar product* $(|) = \mathbf{g} \in V^* \otimes V^*$ *on a finite-dimensional vector space* V *with field* $\mathbb{K}$ *either* $\mathbb{R}$ *or* $\mathbb{C}$ *is called* **pseudo-metric tensor**. *If* $\mathbb{K} = \mathbb{R}$*, a pseudo-metric tensor is called* **metric tensor** *if it defines a scalar product.*

Remarks.
(1) By **Theorem 2.2**, the isomorphism $h : V \to V^*$ is represented by a tensor of $V^* \otimes V^*$ which acts on elements of V by means of a product of tensors and a contraction. The introduction of the pseudo-metric tensor allows us to represent the isomorphism $h : V \to V^*$ by means of the abstract index notation determining the tensor representing h. Indeed, since $h : u \mapsto (u|\ )$ and $(u|v) = (e_i|e_j)u^i v^j = g_{ij} v^i u^j$ we trivially have:

$$(hu)_i = g_{ij} u^j .$$

Hence h is represented by g itself.
(2) Notice that pseudo-metric tensors are **symmetric** because of the symmetry of pseudo scalar products:

$$\mathbf{g}(u, v) = (u|v) = (v|u) = \mathbf{g}(v, u) .$$

Components of pseudo-metric tensors with respect to canonical basis enjoy some simple but important properties which are listed below.

Theorem 4.3. (Properties of the metric tensor.) *Referring to* **Def.4.4**, *the components of any pseudo-metric tensor* $\mathbf{g}$, $g_{ij} := \mathbf{g}(e_i, e_j)$ *with respect to the canonical basis induced in*

$V^* \otimes V^*$ by any basis $\{e_i\}_{i=1,\ldots,n} \subset V$, enjoy the following properties:
(1) *define a symmetric matrix $[g_{ij}]$, i.e,*

$$g_{ij} = g_{ji}\,;$$

(2) *$[g_{ij}]$ is non singular, i.e., it satisfies:*

$$det[g_{ij}] \neq 0\,;$$

(3) *if $\mathbb{K} = \mathbb{R}$ and g is a scalar product, the matrix $[g_{ij}]$ is positive defined.*

Proof. **(1)** It is obvious: $g_{ij} = (e_i|e_j) = (e_j|e_i) = g_{ji}$.
(2) Suppose $det[g_{ij}] = 0$ and define $n = dimV$. The linear mapping $\mathbb{K}^n \to \mathbb{K}^n$ determined by the matrix $g := [g_{ij}]$ has a non-trivial kernel. In other words, there are n reals u^j, $j = 1, \ldots, n$ defining a $\mathbb{K}^n$ vector $[u] := (u^1, \ldots, u^n)^t$ with $g[u] = 0$ and $[u] \neq 0$. In particular $[v]^t g[u] = 0$ for whatever choice of $[v] \in \mathbb{K}^n$. Defining $u := u^j e_j$, the obtained result implies that there is $u \in V \setminus \{0\}$ with $(u|v) = (v|u) = 0$ for all $v \in V$. This is impossible because $(|)$ is non degenerate by hypothesis.
(3) The statement, $(u|u) > 0$ if $u \in V \setminus \{\mathbf{0}\}$, reads, in the considered canonical basis $[u]^t g[u] > 0$ for $[u] \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. That is one of the equivalent definitions of a positive defined matrix g. $\Box$

The following theorem shows that a (pseudo) scalar product can be given by the assignment of a convenient tensor which satisfies some properties when represented in some canonical bases. The important point is that there is no need to check on these properties for *all* canonical bases, verification for a single canonical basis is sufficient.

Theorem 4.4 (Assignment of a (pseudo) scalar product.) *Let V be a finite-dimensional vector space with field $\mathbb{K}$ either $\mathbb{R}$ or $\mathbb{C}$. Suppose $\mathbf{g} \in V^* \otimes V^*$ is a tensor such that there is a canonical basis of $V^* \otimes V^*$ where the components g_{ij} of $\mathbf{g}$ define a symmetric matrix $g := [g_{ij}]$ with non-vanishing determinant. Then $\mathbf{g}$ is a pseudo-metric tensor, i.e. a pseudo scalar product. Furthermore, if $\mathbb{K} = \mathbb{R}$ and $[g_{ij}]$ is positive defined, the pseudo scalar product is a scalar product.*

Proof. If $\mathbf{g}$ is represented by a symmetric matrix of components in a canonical basis then it holds in all remaining bases and the tensor is symmetric (see **Exercise 3.2.1**). This implies that $(u|v) := \mathbf{g}(u, v)$ is a bi-linear symmetric functional. Suppose $(\,|\,)$ is degenerate, then there is $u \in V$ such that $u \neq \mathbf{0}$ and $(u|v) = 0$ for all $v \in V$. Using notations of the proof of the item (2) of **Theorem 4.3**, we have in components of the considered canonical bases, $[u]^t g[v] = 0$ for all $[v] = (v^1, \ldots, v^n)^t \in \mathbb{K}^n$ where $n = dimV$. Choosing $[v] = g[u]$, it also holds $[u]^t gg[u] = 0$. Since $g = g^t$, this is equivalent to $(g[u])^t g[u] = 0$ which implies $g[u] = 0$. Since $[u] \neq \mathbf{0}$, g cannot be injective and $detg = 0$. This is not possible by hypotheses, thus $(\,|\,)$ is non-degenerate. We conclude that $(u|v) := g(u, v)$ define a pseudo scalar product.
Finally, if $\mathbb{K} = \mathbb{R}$ and g is also positive defined, $(\,|\,)$ itself turns out to be positive defined, i.e., it is a scalar product since $(u|u) = [u]^t g[u] > 0$ if $[u] \neq \mathbf{0}$ (which is equivalent to $u \neq \mathbf{0}$). $\Box$

Let us introduce the concept of *signature* of a pseudo-metric tensor in a vector space with field $\mathbb{R}$ by reminding Sylvester's theorem whose proof can be found in any linear algebra textbook. The definition is interlaced with the definition of an *orthonormal basis.*

Theorem 4.5 (Sylvester's theorem.) *Let A be a real symmetric $n \times n$ matrix.*
(a) *There is a non-singular (i.e., with non vanishing determinant) real $n \times n$ matrix D such that:*

$$DAD^t = diag(0, \ldots, 0, -1, \ldots, -1, +1, \ldots, +1) ,$$

where the reals $0, -1, +1$ appear $v \geq 0$ times, $m \geq 0$ times and $p \geq 0$ times respectively with $v + m + p = n$.
(b) *the triple (v, m, p) does not depend on D. In other words, if, for some non-singular real $n \times n$ matrix matrix $E \neq D$, EAE^t is diagonal and the diagonal contains reals $0, -1, +1$ only (in whatever order), then $0, -1, +1$ respectively appear v times, m times and p times.*

If $\mathbf{g} \in V^* \otimes V^*$ is a pseudo-metric tensor on the finite-dimensional vector space v with field $\mathbb{R}$, the transformation rule of the components of g with respect to canonical bases (see **Theorem 2.1**) induced by bases $\{e_i\}_{i\in I}$, $\{e'_j\}_{j\in I}$ of V are

$$g'_{pq} = B_p{}^i B_q{}^j g_{ij} .$$

Defining $g' := [g_{pq}]$, $g := [g_{ij}]$, $B := [B_h{}^k]$, they can be re-written as

$$g' = BgB^t .$$

We remind (see **Theorem 2.1**) that the non-singular matrices B are defined by $B = A^{-1t}$, where $A = [A^i{}_j]$ and $e_m = A^l{}_m e'_l$. Notice that the specification of B is completely equivalent to the specification of A because $A = B^{-1t}$.
Hence, since g is real and symmetric by **Theorem 4.3**, Sylvester's theorem implies that, starting from any basis $\{e_i\}_{i\in I} \subset V$ one can find another basis $\{e'_j\}_{j\in I}$, which induces a canonical basis in $V^* \otimes V^*$ where the pseudo-metric tensor is represented by a diagonal matrix. It is sufficient to pick out a transformation matrix B as specified in **(a)** of **Theorem 4.5**. In particular, one can find B such that each element on the diagonal of g' turns out to be either -1 or $+1$ only. The value 0 is not allowed because it would imply that the matrix has vanishing determinant and this is not possible because of **Theorem 4.3**. Moreover the pair (m, p), where (m, p) are defined in **Theorem 4.5**, does not depend on the basis $\{e'_j\}_{j\in I}$. In other words, it is an *intrisic* property of the pseudo-metric tensor: that is the *signature* of the pseudo-metric tensor.

Def.4.5 (Pseudo Orthonormal Bases and Signature). *Let $\mathbf{g} \in V^* \otimes V^*$ be a pseudo-metric tensor on the finite-dimensional vector space V with field $\mathbb{R}$.*
(a) *A basis $\{e_i\}_{i\in I} \subset V$ is called* **pseudo orthonormal** *with respect to* **g** *if the components of* **g** *with respect to the canonical basis induced in $V^* \otimes V^*$ form a diagonal matrix with eigenvalues in $\{-1, +1\}$. In other words, $\{e_i\}_{i\in I}$ is pseudo orthonormal if*

$$(e_i, e_j) = \pm\delta_{ij} .$$

If the pseudo-metric tensor is a metric tensor the pseudo-orthonormal bases are called orthonormal bases.
(b) *The pair (m, p), where m is the number of eigenvalues -1 and p is the number of eigenvalues $+1$ of a matrix representing the components of* $\mathbf{g}$ *in an orthonormal basis is called* **signature** *of* $\mathbf{g}$.
(c) $\mathbf{g}$ *and its signature are said* **elliptic** *or* **Euclidean** *or* **Riemannian** *if $m = 0$,* **hyperbolic** *if $m > 0$ and $p \neq 0$,* **Lorentzian** *or* **normally hyperbolic** *if $m = 1$ and $p \neq 0$.*
(d) *If* $\mathbf{g}$ *is hyperbolic, an orthonormal basis $\{e_i\}_{i\in I}$ is said to be* **canonical** *if the matrix of the components of* $\mathbf{g}$ *takes the form:*

$$diag(-1, ..., -1, +1, ..., +1) \ .$$

Exercises. 4.2.
4.2.1. Show that a *pseudo-metric* tensor $\mathbf{g}$ is a *metric* tensor if and only if its signature is elliptic.

Remark. If $\{e_i\}_{i\in I}$ is an orthonormal basis with respect to a hyperbolic pseudo-metric tensor g, one can trivially re-order the vectors of the basis giving rise to a canonical orthonormal basis.

Comment. Let us consider a pseudo-metric tensor $\mathbf{g}$ in V with field $\mathbb{R}$. Let (m, p) be the signature of $\mathbf{g}$ and let $\mathcal{N}_g$ be the class of all of the canonical pseudo-orthonormal bases in V with respect to $\mathbf{g}$. In the following we shall indicate by η the matrix $diag(-1, ..., -1, +1, ..., +1)$ which represents the components of $\mathbf{g}$ with respect to each basis of $\mathcal{N}_{\mathbf{g}}$. If A is a matrix corresponding to a change of basis in $\mathcal{N}_{\mathbf{g}}$, and $B := A^{-1t}$ is the associated matrix concerning change of basis in V^*, it has to hold

$$\eta = B\eta B^t \ .$$

Conversely, each real $n \times n$ matrix B which satisfies the identity above determines $A = B^{-1t}$ which represents a change of basis in $\mathcal{N}_{\mathbf{g}}$. In particular each B which satisfies the identity above must be *non singular* because A is such. Furthermore, taking the determinant of both sides in the identity above and taking $det\eta = (-1)^m$ into account, we have that $(detB)^2 = 1$ and thus $detB = \pm 1$. Where we have also used $detB = detB^t$.
Noticing that $\eta = \eta^{-1}$ and $A = B^{-1t}$, the identity above can be equivalently re-written in terms of the matrices A:

$$\eta = A\eta A^t \ .$$

The equation above completely determines the set $O(m, p) \subset GL(n, \mathbb{R})$ $(n = m + p)$ of all real non-singular $n \times n$ matrices which correspond to changes of bases in $\mathcal{N}_{\mathbf{g}}$.
It is possible to show that $O(m, p)$ is a subgroup of $GL(n, \mathbb{R})$ called the **pseudo orthogonal** group of order (m, p). Notice that, if $m = 0$, $O(0, p) = O(n)$ reduces to the usual orthogonal group of order n. $O(1, 3)$ is the celebrated **Lorentz group** which is the central mathematical object in relativistic theories.

Exercises. 4.3.

4.3.1. Show that if $A \in O(m,p)$ then A^{-1} exists and

$$A^{-1} = \eta A^t \eta \,.$$

4.3.2. Show that $O(m,p)$ is a group with respect to the usual multiplication of matrices.

Note. This implies that $O(m,p)$ is a subgroup of $GL(n,\mathbb{R})$ with $n = p+m$.

(Hint. You have to prove that, (1) the identity matrix I belongs to $O(m,p)$, (2) if A and A' belong to $O(m,p)$, AA' belongs to $O(m,p)$, (3) if A belongs to $O(m,p)$, then A^{-1} exists and belongs to $O(m,p)$.)

4.3.3. Show that $SO(m,p) := \{A \in O(m,p) \mid detA = 1\}$ is not the empty set and is a subgroup of $O(m,p)$.

Note. $SO(m,p)$ is called the **special pseudo orthogonal group** of order (m,p).

4.3.4. Consider the special Lorentz group $SO(1,3)$ and show that the set

$$S\!\!\uparrow\!\! O(m,p) := \{A \in SO(1,3) \ \mid \ {A^1}_1 > 0\}$$

is a not empty subgroup.

Note. $S\!\!\uparrow\!\! O(m,p)$ is called the **special orthocronous Lorentz group**.

4.3 Raising and lowering of indices of tensors.

Consider a finite dimensional vector space V with field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ endowed with a pseudo-metric tensor $\mathbf{g}$. As we said above, there is a natural isomorphism $h : V \to V^*$ defined by $h : u \mapsto (u|\cdot) = \mathbf{g}(u,\cdot)$. This isomorphism may be extended to the whole tensor algebra $\mathcal{A}_{\mathbb{K}}(V)$ using the universality theorem and **Def.3.4**.

Indeed, consider a space $S \in \mathcal{A}_{\mathbb{K}}(V)$ of the form $A \otimes V \otimes B$, where A and B are tensor spaces of the form $U_1 \otimes \ldots \otimes U_k$, and $U_{k+1} \otimes \ldots \otimes U_m$ respectively, U_i being either V or V^*. We may define the operators:

$$h^{\otimes} := I_1 \otimes \ldots \otimes I_k \otimes h \otimes I_{k+1} \otimes \ldots \otimes I_m : A \otimes V \otimes B \to A \otimes V^* \otimes B \,,$$

and

$$(h^{-1})^{\otimes} := I_1 \otimes \ldots \otimes I_k \otimes h^{-1} \otimes I_{k+1} \otimes \ldots \otimes I_m : A \otimes V^* \otimes B \to A \otimes V \otimes B \,,$$

where $I_j : U_j \to U_j$ is the identity operator. Using *Remark* after **Def.3.4**, one finds

$$(h^{-1})^{\otimes} h^{\otimes} = I_1 \otimes \ldots \otimes I_k \otimes (h^{-1}h) \otimes I_{k+1} \otimes \ldots \otimes I_m = Id_{A \otimes V \otimes B} \,,$$

and

$$h^{\otimes} (h^{-1})^{\otimes} = I_1 \otimes \ldots \otimes I_k \otimes (hh^{-1}) \otimes I_{k+1} \otimes \ldots \otimes I_m = Id_{A \otimes V^* \otimes B} \,.$$

Therefore $h^{\otimes}$ is an isomorphism with inverse $(h^{-1})^{\otimes}$.

The action of $h^{\otimes}$ and $(h^{-1})^{\otimes}$ is that of **lowering** and **raising indices** respectively. In fact, in abstract index notation, one has:

$$h^{\otimes} : t^{AiB} \mapsto t^{A}{}_{j}{}^{B} := t^{AiB} g_{ij} \,,$$

and

$$(h^{-1})^{\otimes} : u^{A}{}_{i}{}^{B} \mapsto u^{AjB} := t^{A}{}_{i}{}^{B} \tilde{g}^{ij} \,.$$

Above g_{ij} represents the pseudo-metric tensor as specified in *Remark* **1** after **Def.4.4**. What about the tensor $\tilde{g} \in V \otimes V$ representing h^{-1} via **Theorem 2.2**?

Theorem 4.6. *Let $h : V \to V^*$ be the isomorphism determined by a pseudo scalar product, i.e. a pseudo-metric tensor* $\mathbf{g}$ *on the finite-dimensional vector space V with field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.*
(a) *The inverse mapping $h^{-1} : V^* \to V$ is represented via* **Theorem 2.2** *by a* **symmetric** *tensor $\tilde{\mathbf{g}} \in V \otimes V$ such that, if $\{e_i\}_{i\in I}$ is a basis of V, $\tilde{g}^{rs} := \tilde{g}(e^{*r}, e^{*s})$ and $g_{ij} := \mathbf{g}(e_i, e_j)$, then the matrix $[\tilde{g}^{ij}]$ is the* inverse matrix *of $[g_{ij}]$.*
(b) *The tensor $\tilde{\mathbf{g}}$ coincides with the pseudo-metric tensor* **with both indices raised**.

Proof. **(a)** By **Theorem 2.2**, h^{-1} determines a tensor $\tilde{\mathbf{g}} \in V \otimes V$ with $h^{-1}(u^*) = \tilde{\mathbf{g}}(u^*, \cdot)$. In components $(h^{-1}u^*)^i = u^*_k \tilde{g}^{ki}$. On the other hand it must be

$$h(h^{-1}u^*) = u^*$$

or,

$$u^*_k \tilde{g}^{ki} g_{ir} = u^*_r \,,$$

for all $u^* \in V^*$. This is can be re-written

$$[u^*]^t (\tilde{g}g - I) = 0 \,,$$

for all $\mathbb{K}^n$ vectors $[u^*] = (u^*_1, \ldots, u^*_n)$. Then the matrix $(\tilde{g}g - I)^t$ is the null matrix. This implies that

$$\tilde{g}g = I \,,$$

which is the thesis. $\tilde{g}$ is symmetric because is the inverse of a symmetric matrix and thus also the tensor $\tilde{\mathbf{g}}$ is symmetric.
(b) Let g^{ij} be the pseudo-metric tensor with both indices raised, i.e.,

$$g^{ij} := g_{rk} \; \tilde{g}^{kj} \tilde{g}^{ri} \,.$$

By **(a)**, the right-hand side is equal to:

$$\delta^j_r \tilde{g}^{ri} = \tilde{g}^{ji} = \tilde{g}^{ij} \,.$$

That is the thesis. □

Remark. Another result which arises from the proof of the second part of the theorem is that

$$g_i^j = \delta_i^j \,.$$

Comments.
(1) When a vector space is endowed with a pseudo scalar product, tensors can be viewed as abstract objects which may be represented either as covariant or contravariant concrete tensors using the procedure of raising and lowering indices. For instance, a tensor t^{ij} of $V \otimes V$ may be viewed as a covariant tensor when "represented" in its covariant form $t_{pq} := g_{pi}\, g_{qj}\, t^{ij}$. Also, it can be viewed as a mixed tensor $t_p{}^j := g_{pi}\, t^{ij}$ or $t^i{}_q := g_{qj}\, t^{ij}$.
(2) Now consider a finite-dimensional vector space on $\mathbb{R}$, V, endowed with a metric tensor $\mathbf{g}$, i.e., with *elliptic signature*. In orthonormal bases the contravariant and covariant components numerically coincides because $g_{ij} = \delta_{ij} = g^{ij}$. This is the reason because, using the usual scalar product of vector spaces isomorphic to $\mathbb{R}^n$ and working in orthonormal bases, the difference between covariant and contravariant vectors does not arise.
Conversely, in relativistic theories where a Lorentzian scalar product is necessary, the difference between covariant and contravariant vectors turns out to be evident also in orthonormal bases, since the diagonal matrix $[g_{ij}]$ takes an eigenvalue -1.

5 Pseudo tensors, Ricci's pseudotensor and tensor densities.

This section is devoted to introduce very important tools either in theoretical/mathematical physics and in pure mathematics: pseudo tensors and tensor densities.

5.1 Orientation and pseudo tensors.

The first example of "pseudo" object we go to discuss is a *orientation* of a *real* vector space.

Consider a finite-dimensional vector space V with field $\mathbb{R}$. In the following $\mathcal{B}$ indicates the set of all the vector bases of V. Consider two bases $\{e_i\}_{i\in I}$ and $\{e'_j\}_{j\in I}$ in $\mathcal{B}$. Concerning the determinant of the transformation matrix $A := [A^r{}_s]$, with $e_i = A^j{}_i e'_j$, there are two possibilities only: $detA > 0$ or $detA < 0$. It is a trivial task to show that the relation in $\mathcal{B}$:

$$\{e_i\}_{i\in I} \sim \{e'_j\}_{j\in I} \quad \text{iff} \quad detA > 0$$

where A indicates the transformation matrix as above, is an *equivalence relation*. Since there are the only two possibilities above, the partition of $\mathcal{B}$ induced by $\sim$ is made of two *equivalence classes* $\mathcal{B}_1$ and $\mathcal{B}_2$. Hence if a basis belongs to $\mathcal{B}_1$ or $\mathcal{B}_2$ any other basis belongs to the same set if and only if the transformation matrix has positive determinant.

Def.5.1. (Orientation of a vector space.) *Consider a finite-dimensional vector space V with field $\mathbb{R}$, an* **orientation** *of V is a bijective mapping $\mathcal{O} : \{\mathcal{B}_1, \mathcal{B}_2\} \to \{-1, +1\}$. If V has an orientation $\mathcal{O}$, is said to be* **oriented** *and a basis $\{e_i\}_{i\in I} \in \mathcal{B}_k$ is said to be* **positive oriented** *if $\mathcal{O}(\mathcal{B}_k) = +1$ or* **negative oriented** *if $\mathcal{O}(\mathcal{B}_k) = -1$.*

Comment. The usual physical vector space can be orietend "by hand" using the natural basis given by our own right hand. When we use the right hand to give an orientation we determine $\mathcal{O}^{-1}(+1)$ by the exhibition of a basis contained therein.

The given definition can be, in some sense, generalized with the introduction of the concept of *pseudo tensor*.

Def.5.2. (Pseudotensors.) *Let V be a finite-dimensional vector space with field $\mathbb{R}$. Let S be a tensor space of $\mathcal{A}_{\mathbb{R}}(V)$. A* **pseudo tensor of** *S is a bijective mapping $t_s : \{\mathcal{B}_1, \mathcal{B}_2\} \to \{s, -s\}$, where $s \in S$. Moreover:*
(a) *the various tensorial properties enjoyed by both s and $-s$ are attributed to t_s. (So, for instance, if s, and thus $-s$, is symmetric, t_s is said to be symmetric);*
(b) *if $\{e_i\}_{i\in I} \in \mathcal{B}_i$, the* **components of** *t_s with respect to the canonical bases induced by that basis are the components of $t_s(\mathcal{B}_i)$.*

Remarks.
(1) The given definition encompasses the definition of *pseudo scalar*.
(2) It is obvious that the assignment of a pseudo tensor t_s of, for instance, $V^{n\otimes} \otimes V^{*m\otimes}$, is equivalent to the assignment of components

$$t^{i_1 \dots i_n}{}_{j_1 \dots j_m}$$

for each canonical basis

$$\{e_{i_1} \otimes \dots e_{i_n} \otimes e^{*j_1} \otimes \dots \otimes e^{*j_m}\}_{i_1,\dots,i_n,j_1,\dots,j_m \in I}$$

such that the transformation rules passing to the basis

$$\{e'_{r_1} \otimes \dots e'_{r_n} \otimes e'^{*l_1} \otimes \dots \otimes e'^{*l_m}\}_{r_1,\dots,r_n,l_1,\dots,l_m \in I} \ ,$$

are given by:

$$t'^{k_1 \dots k_n}{}_{h_1 \dots h_m} = \frac{detA}{|detA|} \ A^{k_1}_{\ i_1} \cdots A^{k_n}_{\ i_n} \ B_{h_1}^{\ j_1} \cdots B_{h_m}^{\ j_m} \ t^{i_1 \dots i_n}{}_{j_1 \dots j_n} \ ,$$

where $e_l = A^m_{\ l} e'_m$ and $B = A^{-1t}$ with $B := [B_k^{\ j}]$ and $A := [A^p{}_q]$.
In fact, $t^{i_1 \dots i_n}{}_{j_1 \dots j_m} = s^{i_1 \dots i_n}{}_{j_1 \dots j_m}$ if the considered base is in $t_s^{-1}(+1)$ or $t^{i_1 \dots i_n}{}_{j_1 \dots j_m} = (-s)^{i_1 \dots i_n}{}_{j_1 \dots j_m}$ if the considered base is in $t_s^{-1}(-1)$.

Example 5.1. Consider the **magnetic field** $B = B^i e_i$ where e_1, e_2, e_3 is a right-hand orthonormal basis of the space V_3 of the vectors with origin in a point of the physical space E_3. Actually, as every physicist knows, changing basis, the components of B changes as usual only if the new basis is a right-hand basis, otherwise a sign $-$ appears in front of each component. That is a physical requirement due to the Lorentz law. This means that the magnetic field has to be represented in terms of **pseudo vectors**.

5.2 Ricci's pseudo tensor.

A particular pseudo tensor is Ricci's one which is very important in physical applications. The definition of this pseudo tensor requires a preliminary discussion.
Consider a finite-dimensional vector space V with a pseudo-metric tensor $\mathbf{g}$. We know that, changing basis $\{e_i\}_{i\in I} \to \{e'_j\}_{j \in I}$, the components of the pseudo-metric tensor referred to the corresponding canonical bases, transform as:

$$g' = BgB^t \ ,$$

where $g = [g_{ij}]$, $g' = [g'_{pq}]$ and $B = A^{-1t}$, $A := [A^p{}_q]$, $e_l = A^m{}_l e'_m$. This implies that

$$detg' = (detB)^2 \ detg \ , \tag{3}$$

which is equivalent to

$$\sqrt{|detg'|} = |detA|^{-1} \ \sqrt{|detg|} \tag{4}$$

Now fix a basis $\{e_i\}_{i\in I}$ in V and consider the canonical basis induced in $V^{*n\otimes}$ where $n = dimV$, $\{e^{*i_1} \otimes \ldots \otimes e^{*i_n}\}_{i_1,\ldots,i_n \in I}$. Then consider components $\eta_{i_1\ldots i_n}$ referred to the considered basis, given by:

$$\eta_{i_1\ldots i_n} = \epsilon_{\sigma_{i_1\ldots i_n}} \ ,$$

if $(i_1, \ldots, i_n)$ is a permutated string of $(1, 2, \ldots, n)$, otherwise

$$\eta_{i_1\ldots i_n} = 0 \ ,$$

where $\epsilon_{\sigma_{i_1\ldots i_n}}$ is the *parity* of the *permutation* $\sigma_{i_1\ldots i_n} \in \mathcal{P}_n$ defined by:

$$(\sigma(1), \ldots, \sigma(n)) = (i_1, \ldots, i_n) \ .$$

Finally define the components:

$$\varepsilon_{i_1\ldots i_n} := \sqrt{|detg|} \ \eta_{i_1\ldots i_n} \ .$$

We want to show that, if we define analogous components in each canonical basis of $V^{*n\otimes}$, an anti-symmetric $(0, n)$-order pseudo tensor turns out to be defined by the whole assignment of components.
Taking *Remark* **(2)** above into account, it is sufficient to show that, under a change of basis,

$$\sqrt{|detg'|} \ \eta_{i_1\ldots i_n} = \frac{detA}{|detA|} \ B_{i_1}^{\ j_1} \cdots B_{i_n}^{\ j_n} \ \sqrt{|detg|} \ \eta_{j_1\ldots j_n} \ . \tag{5}$$

We start by noticing that:

$$B_{i_1}^{\ j_1} \cdots B_{i_n}^{\ j_n} \ \eta_{j_1\ldots j_n} = \sum_{(j_1,\ldots,j_n)} B_{i_1}^{\ j_1} \cdots B_{i_n}^{\ j_n} \epsilon_{\sigma_{j_1\ldots j_n}} \ ,$$

where $(j_1, \ldots, j_n)$ ranges over the set of all the permutated strings of $(1, 2, \ldots, n)$. We consider the various cases separately.
(1) Suppose that $i_p = i_q$ with $p \neq q$. To fix the extent assume $i_1 = i_2$. Therefore

$$\sum_{(j_1,\ldots,j_n)} B_{i_1}^{\ j_1} B_{i_2}^{\ j_2} \cdots B_{i_n}^{\ j_n} \epsilon_{\sigma_{j_1 j_2\cdots j_n}} = \sum_{(j_1,\ldots,j_n)} B_{i_2}^{\ j_1} B_{i_1}^{\ j_2} \cdots B_{i_n}^{\ j_n} \epsilon_{\sigma_{j_1 j_2\cdots j_n}} \ .$$

The right-hand side can be re-written

$$\sum_{(j_1,\ldots,j_n)} B_{i_1}^{\ j_2} B_{i_2}^{\ j_1} \cdots B_{i_n}^{\ j_n} \epsilon_{\sigma_{j_1 j_2\cdots j_n}} \ ,$$

that is, interchanging the names of j_1 and j_2,

$$\sum_{(j_1,\ldots,j_n)} B_{i_1}^{\ j_2} B_{i_2}^{\ j_1} \cdots B_{i_n}^{\ j_n} \epsilon_{\sigma_{j_2 j_1\cdots j_n}} \ ,$$

But

$$\epsilon_{\sigma_{j_2 j_1 \cdots j_n}} = -\epsilon_{\sigma_{j_1 j_2 \cdots j_n}} ,$$

so that we finally get

$$\sum_{(j_1,\ldots,j_n)} B_{i_1}^{\ j_1} B_{i_2}^{\ j_2} \cdots B_{i_n}^{\ j_n} \epsilon_{\sigma_{j_1 j_2 \cdots j_n}} = - \sum_{(j_1,\ldots,j_n)} B_{i_1}^{\ j_1} B_{i_2}^{\ j_2} \cdots B_{i_n}^{\ j_n} \epsilon_{\sigma_{j_1 j_2 \cdots j_n}} .$$

In other words, *if* $i_1 = i_2$,

$$B_{i_1}^{\ j_1} \cdots B_{i_n}^{\ j_n} \ \sqrt{|detg|} \ \eta_{j_1 \ldots j_n} = 0$$

and consequently (5) holds:

$$\sqrt{|detg'|} \ \eta_{i_1 \ldots i_n} = \frac{detA}{|detA|} \ B_{i_1}^{\ j_1} \cdots B_{i_n}^{\ j_n} \ \sqrt{|detg|} \ \eta_{j_1 \ldots j_n} .$$

because the left-hand side vanishes too. The remaining subcases of $i_p = i_q$ with $p \neq q$, can be treated analogously.

(2) Next we pass to consider the case of $i_k = k$ for $k = 1, 2, \ldots, n$. In that case (5), that we need to prove, reduces to

$$\sqrt{|detg'|} = \sqrt{|detg|} \frac{detA}{|detA|} \ B_1^{\ j_1} \cdots B_n^{\ j_n} \quad \eta_{j_1 \ldots j_n} ,$$

that is

$$\frac{1}{detA} = B_1^{\ j_1} \cdots B_n^{\ j_n} \quad \eta_{j_1 \ldots j_n} , \tag{6}$$

where we have used (4). (6) can equivalently be re-written

$$detB = B_1^{\ j_1} \cdots B_n^{\ j_n} \quad \eta_{j_1 \ldots j_n} ,$$

which is trivially true by the properties of the determinant.

(3) To conclude, it remains to consider the case of a permutation $(i_1, i_2, \ldots, i_n)$ of $(1, 2, \ldots, n)$. In that case, taking (4) into account, the identity (5) which has to be proven, reduces to

$$detB \epsilon_{\sigma_{i_1 \ldots i_n}} = B_{i_1}^{\ j_1} \cdots B_{i_n}^{\ j_n} \quad \eta_{j_1 \ldots j_n} . \tag{7}$$

Let us prove (7). Suppose for instance that $i_1 = 2$ and $i_2 = 1$ while $i_k = k$ if $k > 2$. In that case (7) holds true because the left-hand side is noting but $-detB$ and the right-hand side can be re-written

$$\sum_{(j_1,\ldots,j_n)} B_2^{\ j_1} B_1^{\ j_2} \cdots B_n^{\ j_n} \epsilon_{\sigma_{j_1 j_2 \cdots j_n}} = \sum_{(j_1,\ldots,j_n)} B_1^{\ j_2} B_2^{\ j_1} \cdots B_n^{\ j_n} \epsilon_{\sigma_{j_1 j_2 \cdots j_n}} ,$$

that is, interchanging the names of j_1 and j_2,

$$\sum_{(j_1,\ldots,j_n)} B_1^{\ j_1} B_2^{\ j_2} \cdots B_n^{\ j_n} \epsilon_{\sigma_{j_2 j_1 \cdots j_n}}$$

which, in turn, equals just

$$-\sum_{(j_1,\ldots,j_n)} B_1^{\ j_1} B_2^{\ j_2} \cdots B_n^{\ j_n} \epsilon_{\sigma_{j_1 j_2 \cdots j_n}} = -detB \ .$$

If $(i_1, \ldots, i_n)$ is different from $(1, \ldots, 2)$ just for a transposition only, the same procedure can be adapted trivially. If $(i_1, \ldots, i_n)$ is a proper permutation of $(1, \ldots, 2)$, it can be decomposed as a product of N transpositions. Concerning the right hand side of (7), using the procedure above for each transposition, we get in the end that it can be re-written :

$$B_{i_1}^{\ j_1} \cdots B_{i_n}^{\ j_n} \quad \eta_{j_1 \ldots j_n} = detB(-1)^N \ .$$

On the other hand, in the considered case, the left-hand side of (7) equals

$$detB(-1)^N \ ,$$

so that (7) holds true once again. This concludes the proof.

Def.5.3. (Ricci's Pseudo tensor). *Let V be a vector space with field $\mathbb{R}$ and dimension $n < +\infty$, endowed with a pseudo-metric tensor* **g**. **Ricci's pseudo tensor** *is the anti-symmetric $(0, n)$ pseudo tensor ε represented in each canonic basis by components*

$$\varepsilon_{i_1 \ldots i_n} := \sqrt{|detg|} \ \eta_{i_1 \ldots i_n} \ ,$$

where $g = [g_{ij}]$, g_{ij} being the components of g in the considered basis and

$$\eta_{i_1 \ldots i_n} = \epsilon_{\sigma_{i_1 \ldots i_n}} \ ,$$

if $(i_1, \ldots, i_n)$ is a permutated string of $(1, 2, \ldots, n)$, otherwise

$$\eta_{i_1 \ldots i_n} = 0 \ ,$$

and $\epsilon_{\sigma_{i_1 \ldots i_n}}$ is the parity of the permutation $\sigma_{i_1 \ldots i_n} \in \mathcal{P}_n$ defined by:

$$(\sigma(1), \ldots, \sigma(n)) = (i_1, \ldots, i_n) \ .$$

Ricci's pseudo tensor has various applications in mathematical physics in particular when it is used as a linear operator which produces pseudo tensors when acts on tensors. In fact, consider $t \in V^{n\otimes}$ and take an integer $m \leq n$. Fix a basis in V and, in the canonical bases induced by that basis, consider the action of ε on t:

$$t^{i_1 \ldots i_n} \mapsto \tilde{t}_{j_1 \ldots j_{n-m}} := \varepsilon_{j_1 \ldots j_{n-m} i_1 \ldots i_n} t^{i_1 \ldots i_n} \ .$$

We leave to the reader the proof of the fact that the components $\tilde{t}_{j_1,\ldots,j_{n-m}}$ define a anti-symmetric pseudo tensor of order $(0, n-m)$ which is called the **conjugate** pseudo tensor of t.

Example 5.1. As a trivial but very important example consider the **vector product** in a three-dimensional vector space V on the field $\mathbb{R}$ endowed with a metric tensor g. If $u, v \in V$ we may define the pseudo vector of order $(1,0)$:

$$(u \wedge v)^r := g^{ri} \varepsilon_{ijk} u^j v^k \,.$$

If $\{e_i\}_{i=1,2,3}$ is an orthonormal basis in V, everything strongly semplificates. In fact, the Ricci tensor is represented by components

$$\varepsilon_{ijk} := 0$$

if (i,j,k) is not a permutation of $(1,2,3)$ and, otherwise,

$$\varepsilon_{ijk} := \pm 1 \,,$$

where $+1$ corresponds to *cyclic* permutations of $1,2,3$ and -1 to *non-cyclic* permutations (see **Examples 3.1.2**). In such a basis:

$$(u \wedge v)_i = \varepsilon_{ijk}\, u^j v^k \,,$$

because $g^{ij} = \delta^{ij}$ in each orthonormal bases.

If V is oriented, it is possible to define ε as a *proper tensor* instead a pseudo tensor. In this case one defines the components in a canonical basis associated with a positive-oriented basis ov V as

$$\varepsilon_{i_1\ldots i_n} := \sqrt{|detg|}\ \eta_{i_1\ldots i_n} \,,$$

and

$$\varepsilon_{i_1\ldots i_n} := -\sqrt{|detg|}\ \eta_{i_1\ldots i_n} \,,$$

if the used basis is associated with a basis of V which is negative-orientd. One can prove staight-forwardly that the defined components give rise to a tensor of $V^{*n\otimes}$ called **Ricci tensor**.
This alternative point of view is equivalent, in the practice, to the other point of view corresponding to Definition 5.3.
An important feature of Ricci's pseudotensor is the formula, with obvious notations:

$$t_{i_1\ldots i_p} = \frac{|detg|}{detg}\ \frac{(-1)^{p(n-p)}}{p!(n-p)!}\ \varepsilon_{i_1\ldots i_p j_1\ldots j_n} \varepsilon^{j_1\ldots j_n r_1\ldots r_p}\ t_{r_1\ldots t_p} \,,$$

which holds for *anti-symmetric* tensors $t \in V^{*p\otimes}$, $0 \leq p \leq n = dimV$.
This implies that, if the tensor t is anti symmetric, then its conjugate pseudo tensor $\tilde{t}$ takes the same information than t itself.

Exercises 5.1.
5.1.1 Often, the definition of vector product in $\mathbb{R}^3$ is given, in orthonormal basis, as

$$(u \wedge v)_i = \varepsilon_{ijk}\, u^i v^k \,,$$

where it is assumed that the basis is *right oriented*. Show that it defines a proper vector (and not a pseudo vector) if a convenient definition of $\wedge$ is given in *left oriented* basis.
5.1.2. Is it possible to define a sort of vector product (which maps pair of vectors in vectors) in $\mathbb{R}^4$ generaluzing the vector product in $\mathbb{R}^3$?
5.1.3. In physics literature one may find the statement "Differently from the impulse $\vec{p}$ which is a **polar vector**, the angular momentum $\vec{l}$ is an **axial vector**". What does it mean?
(Solution. Polar vector = vector, Axial vector = pseudo vector.)
5.1.4. Consider the **parity inversion**, $P \in O(3)$, as the active transformation of vectors of physical space defined by $P := -I$ when acting in components of vectors in any orthonormal basis. What do physicists mean when saying "Axial vectors transform differently from polar vectors under parity inversion"?
(Hint. interpret P as a passive transformation, i.e. a changing of basis and extend the result to the active interpretation.)
5.1.5. Can the formula defining the conjugate pseudo tensor of t:

$$t^{i_1 \ldots i_n} \mapsto \tilde{t}_{j_1 \ldots j_{n-m}} := \varepsilon_{j_1 \ldots j_{n-m} i_1 \ldots i_m} t^{i_1 \ldots i_n} \,,$$

be generalized to the case where t is a pseudo tensor? If yes, what sort of geometric object is $\tilde{t}$?
5.1.6. Consider a vector product $u \wedge v$ in $\mathbb{R}^3$ using an orthonormal basis. In that basis there is an anti-symmetric matrix which takes the same information as $u \wedge v$ and can be written down using the components of the vectors u and v. Determine that matrix and explain the tensorial meaning of the matrix.
5.1.7. Prove the formula introduced above:

$$t_{i_1 \ldots i_p} = \frac{|detg|}{detg} \frac{(-1)^{p(n-p)}}{p!(n-p)!} \; \varepsilon_{i_1 \ldots i_p j_1 \ldots j_n} \varepsilon^{j_1 \ldots j_n r_1 \ldots r_p} \; t_{r_1 \ldots t_p} \,,$$

for anti-symmetric tensors $t \in V^{*p\otimes}$, $0 \leq p \leq n = dimV$.

5.3 Tensor densities.

In section 5.2 we have seen that the determinant of the matrix representing a pseudo-metric tensor g transforms, under change of basis with the rule

$$detg' = |detA|^{-2} detg$$

where the pseudo-metric tensor is $g'_{ij} e'^{*i} \otimes e'^{*i} = g_{pq} e^{*p} \otimes e^{*q}$ and $A = [A^i{}_j]$ is the matrix used in the change of basis for controvariant vectors $t^i e_i = t'^p e'_p$, that is $t^i = A^i{}_p t^p$. Thus the assignment of the numbers $detg$ for each basis in $\mathcal{B}$ does not define a scalar because of the presence of the factor $|detA|^{-1}$. Similar mathematical objects plays a relevant role in mathematical/theoretical physics and thus deserve a precise definition.

Def.5.4. (Tensor densities.) *Let V be a finite-dimensional vector space with field $\mathbb{R}$ and $\mathcal{B}$ the class of all bases of V.*
If $S \in \mathcal{A}_{\mathbb{R}}(V)$, a **tensor densities of** S **with weight** $w \in \mathbb{Z} \setminus \{0\}$ *is a mapping $d : \mathcal{B} \to S$ such that, if $B = \{e_i\}_{i\in I}$ and $B' = \{e'_j\}_{j\in I}$ are two bases in $\mathcal{B}$ with $e_k = A^i{}_k e'_i$ then*

$$d(B') = |detA|^w d(B) \, .$$

where $A = [A^i{}_k]$. Furthermore:
(a) *the various tensorial properties enjoyed by all $d(B)$ are attributed to d. (So, for instance, if a $d(B)$ is symmetric (and thus all $d(B)$ with $B \in \mathcal{B}$ are symmetric), d is said to be symmetric);*
(b) *if $B \in \mathcal{B}$, the* **components of** *d with respect to the canonical bases induced by B are the components of $d(B)$ in those bases.*

If $\mathbf{g}$ is a pseudo-metric tensor on V a trivial example of a density with weight w in, for instance $S = V \otimes V^* \otimes V$, can be built up as follows. Take $t \in V \otimes V^* \otimes V$ and define

$$d_t(\{e_i\}_{i\in I}) := (\sqrt{|detg|})^{-w} t \, ,$$

where g is the matrix of the coefficients of $\mathbf{g}$ in the canonical basis associated with $\{e_i\}_{i\in I}$. In components, in the sense of (b) of the definition above:

$$(d_t)^i{}_j{}^k = (\sqrt{|detg|})^{-w} t^i{}_j{}^k \, .$$

To conclude we give the definition of pseudo tensor density which is the straightforward extension of the definition given above.

Def.5.5. (Pseudo-tensor densities.) *Let V be a finite-dimensional vector space with field $\mathbb{R}$ and $\mathcal{B}$ the class of all bases of V.*
If $S \in \mathcal{A}_{\mathbb{R}}(V)$, a **pseudo-tensor densities of** S **with weight** $w \in \mathbb{Z} \setminus \{0\}$ *is a mapping $d : \mathcal{B} \to S$ such that, if $B = \{e_i\}_{i\in I}$ and $B' = \{e'_j\}_{j\in I}$ are two bases in $\mathcal{B}$ with $e_k = A^i{}_k e'_i$ then*

$$d(B') = \frac{detA}{|detA|} |detA|^w d(B) \, .$$

where $A = [A^i{}_k]$. Furthermore:
(a) *the various tensorial properties enjoyed by all $d(B)$ are attributed to d. (So, for instance, if a $d(B)$ is symmetric (and thus all $d(B)$ with $B \in \mathcal{B}$ are symmetric), d is said to be symmetric);*
(b) *if $B \in \mathcal{B}$, the* **components of** *d with respect to the canonical bases induced by B are the components of $d(B)$ in those bases.*

Remark. *It is clear that the sets of tensor densities and pseudo-tensor densities of a fixed space S and with fixed weight form linear spaces with composition rule which reduces to usual linear composition rule of components.*

There is an important property of densities with weight -1 which is very useful in integration theory on manifolds. The property is stated in the following theorem.

Theorem 5.1. *Let V be a vector space on $\mathbb{R}$ with dimension $n < +\infty$. There is a natural isomorphism G from the space of scalar densities of weight -1 and the space of antisymmetric covariant tensors of order n. In components, using notation as in Definition 5.3,*

$$G : \alpha \mapsto \alpha \eta_{i_1 \ldots i_n} .$$

Proof. Fix a canonical basis of $V^{*n\otimes}$ associated with a basis of V. Any nonvanishing tensor t in space, $\Lambda_n(V)$, of antisymmetric covariant tensors of order $n = dimV$ must have the form $t_{i_1 \ldots i_n} = \alpha \eta_{i_1 \ldots i_n}$ in components because different nonvanishing components can be differ only for the sign due to antisymmetry properties of t. Therefore the dimension of $\Lambda_n(V)$ is 1 which is also the dimension of the space of scalar densities of weight -1. The application G defined above in components from $\mathbb{R}$ to $\Lambda_n(V)$ is linear and surjective and thus is injective. Finally, re-adapting straightforwardly the relevant part of the discussion used to define ϵ, one finds that the coefficient α in $t_{i_1 \ldots i_n} = \alpha \eta_{i_1 \ldots i_n}$ transforms as a scalar densities of weight -1 under change of basis. $\Box$

6 Appendix: Square Roots of operators and Polar Decomposition Theorem.

We racall some basic definitions and results which should be known by the reader from elemetary courses of linear algebra.

If $A : V \to V$ is a (linear) operator on any finite-dimensional vector space V with field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, an **eigenvalue** $\lambda \in \mathbb{K}$ of A is a scalar such that

$$(A - \lambda I)u = 0$$

for some $u \in V \setminus \{0\}$. In that case u is called **eigenvector** associated with λ. The set $\sigma(A)$ containing all of the eigenvalues of A is called the **spectrum** of A. The **eigenspace** E_λ associated with $\lambda \in \sigma(A)$ is the subspace of V spanned by the eigenvectors associated with λ.

Proposition A.1. *Let V be a real (complex) finite-dimensional equipped with a (resp. Hermitean) scalar product $(|)$. For every operator $A : V \to V$ there exists exactly one of operator $A^\dagger : V \to V$, called the* **adjoint operator** *of A, such that*

$$(A^\dagger u|v) = (u|Av) ,$$

for all $u, v \in V$.

Proof. Fix $u \in V$, the mapping $v \mapsto (u|Av)$ is a linear functional and thus an element of V^*. By Theorem 4.2 there is a unique element $w_{u,A} \in V$ such that $(u|Av) = (w_{u,A}|v)$ for all $v \in V$. Consider the map $u \mapsto w_{u,A}$. It holds, if a, b are scalars in the field of V and $u, u' \in V$

$$(w_{au+bu',A}|v) = (au+bu'|Av) = \overline{a}(u|Av)+\overline{b}(u'|Av) = \overline{a}(w_{u,A}|v)+\overline{b}(w_{u',A}|v) = (aw_{u,A}+bw_{u',A}|v).$$

Hence, for all $v \in V$:

$$(w_{au+bu',A} - aw_{u,A} - bw_{u',A}|v) = 0 ,$$

The scalar product is nondegenerate by definition and this implies

$$w_{au+bu',A} = aw_{u,A} + bw_{u',A} .$$

We have obtained that the mapping $A^\dagger : u \mapsto w_{u,A}$ is linear, in other words it is an operator. The uniqueness is trivially proven: if the operator B satisfies $(Bu|v) = (u|Av)$ for all $u, v \in V$, it must holds $((B - A^\dagger)u|v) = 0$ for all $u, v \in V$ which, exactly as we obtained above, entails $(B - A^\dagger)u = 0$ for all $u \in V$. In other words $B = A^\dagger$. $\Box$

There are a few simple properties of the adjoint operator whose proofs are straightforward. Below A, B are operators in a real (complex) finite-dimensional vector space V equipped with a (resp. Hermitean) scalar product $(|)$ and a, b belong to the field of V.
(1) $(A^\dagger)^\dagger = A$,

(2) $(aA + bB)^\dagger = \overline{a}A^\dagger + \overline{b}B^\dagger$,
(3) $(AB)^\dagger = B^\dagger A^\dagger$,
(4) $(A^\dagger)^{-1} = (A^{-1})^\dagger$ (if A^{-1} exists).

In a real finite-dimensional vector space V equipped with a scalar product $(|)$, a linear operator $A : V \to V$ is said to be **symmetric** if $(Au|v) = (u|Av)$ for all $u, v \in V$. It is simply proven that A is symmetric if and only if $A = A^\dagger$.

In a complex finite-dimensional vector space V equipped with a Hermitean scalar product $(|)$, a linear operator $A : V \to V$ is said to be **Hermitean** if $(Au|v) = (u|Av)$ for all $u, v \in V$. It is simply proven that $\sigma(A) \subset \mathbb{R}$ if A is Hermitean. It is simply proven that A is Hermitean if and only if $A = A^\dagger$.

In a complex finite-dimensional vector space V equipped with a Hermitean scalar product $(|)$, a linear operator $A : V \to V$ is said to be **unitary** if $(Au|Av) = (u|v)$ for all $u, v \in V$. It is simply shown that every unitary operator is bijective (the space has finite dimension and the operator is injective, $(|)$ being positive). It is simply proven that $\lambda \in \sigma(A)$ entails $|\lambda| = 1$ if A is unitary. It is simply proven that A is unitary if and only if $AA^\dagger = I$ or equivalently $A^\dagger A = I$.

In a real (complex) finite-dimensional vector space V equipped with a (resp. Hermitean) scalar product $(|)$, a linear operator $A : V \to V$ is said to be **normal** if $AA^\dagger = A^\dagger A$. It is clear that symmetric, Hermitean and unitary opertors are normal.

In a complex finite-dimensional vector space V equipped with a Hermitean scalar product $(|)$, If V is that above, and $U \subset V$ is a subspace, the **orthogonal** of U, $U^\perp$, is the subspace of V made of all the vectors which are orthogonal to U, i.e., $v \in U^\perp$ if and only if $(u|v) = 0$ for all $u \in U$. If $w \in V$, the decomposition $w = u + v$ with $u \in U$ and $v \in U^\perp$ is uniquely determined, and the map $P_U : w \mapsto u$ is linear and it is called **orthogonal projector** onto U.

It is possible to show that an operator $P : V \to V$ (V being a complex finite-dimensional vector space V equipped with a Hermitean scalar product $(|)$) is an orthogonal projector onto some subspace $U \subset V$ if and only if both the conditions below hold
(1) $PP = P$,
(2) $P = P^\dagger$.
In that case P is the orthogonal projector onto $U = \{Pv \mid v \in V\}$.
Another pair of useful results concerning orthogonal projectors is the following. Let V be a space as said above, let U, U' be subspaces of V, with P, P' are the corresponding orthogonal projectors P, P'.
(a) U and U' are **orthogonal** to each other, i.e., $U' \subset U^\perp$ (which is equivalent to $U \subset U'^\perp$) if and only if $PP' = P'P = 0$.
(b) $U \subset U'$ if and only if $PP' = P'P = P$.

If V is as above and it has finite dimension n and $A : V \to V$ is normal, there exist a well-known spectral decomposition theorem (the finite-dimensional version of the "spectral theorem").

Proposition A.2 (**Spectral decomposition for normal operators in complex spaces**.) *Let V be a complex finite-dimensional vector space equipped with a Hermitean scalar product $(|)$.*

If the operator $A : V \to V$ is normal (i.e., $A^\dagger A = AA^\dagger$), the following expansion holds:

$$A = \sum_{\lambda \in \sigma(A)} \lambda P_\lambda \,,$$

where P_λ is the orthogonal projector onto the eigenspace associated with λ. Moreover the mapping $\sigma(A) \ni \lambda \mapsto P_\lambda$ satisfies the following two properties:
(1) $I = \sum_\lambda P_\lambda$,
(2) $P_\lambda P_\mu = P_\mu P_\lambda = 0$ *for* $\mu \neq \lambda$.
A **spectral measure**, *i.e. a mapping $B \ni \mu \mapsto P'_\mu$ with $B \subset \mathbb{C}$ finite, P'_μ othogonal projectors and:*
(1)' $I = \sum_{\mu \in B} P'_\mu$,
(2)' $P'_\lambda P'_\mu = P'_\mu P'_\lambda = 0$ *for* $\mu \neq \lambda$,
coincides with $\sigma(A) \ni \lambda \mapsto P_\lambda$ if
(3)' $A = \sum_{\mu \in B} \mu P'_\mu$.

The polar decomposition theorems has many applications in mathematics and physics. In the following we make use of some definitions introduced above. We need two relevant definitions.

Def.A.1. *If V is a real (complex) vector space equipped with a (resp. Hermitean) scalar product $(|)$, an operator $A : V \to V$ is said to be* **positive** *(or* positive semidefined*) if*

$$(u|Au) \geq 0 \quad \textit{for all } u \in V \,.$$

A positive operator A is said to be **strictly positive** *(or* positive defined*) if*

$$(u|Au) = 0 \quad \textit{entails } u = 0 \,.$$

A straightforward consequence of the given definition is the following lemma.

Lemma A.1. *Let V be a complex vector space equipped with a Hermitean scalar product $(|)$. Any positive operator $A : V \to V$, is Hermitean.*
Moreover, if $dimV < \infty$, a normal operator $A : V \mapsto V$
(a) *is positive if and only if $\sigma(A) \subset [0, +\infty)$;*
(b) *is strictly positive if and only if $\sigma(A) \subset (0, +\infty)$.*

Proof. As $(v|Av) \geq 0$, by complex conjugation $(Av|v) = \overline{(v|Av)} = (v|Av)$ and thus

$$((A^\dagger - A)v|v) = 0$$

for all $v \in V$. In general we have:

$$2(Bu|w) = (B(u+w)|(u+w)) + i(B(w+iu)|(w+iu)) - (1+i)(Bw|w) - (1+i)(Bu|u) \,.$$

So that, taking $B = A^\dagger - A$ we get $(Bu|w) = 0$ for all $u, w \in V$ because $(Bv|v) = 0$ for all $v \in V$. $(Bu|w) = 0$ for all $u, w \in V$ entails $B = 0$ or $A^\dagger = A$.
Let us prove (a). Suppose A is positive. We know that $\sigma(A) \subset \mathbb{R}$. Suppose there is $\lambda < 0$ in $\sigma(A)$. Let u be an eigenvector associated with λ. $(u|Au) = \lambda(u|u) < 0$ because $(u|u) > 0$ since $u \neq 0$. This is impossible.
Now assume that A is normal with $\sigma(A) \subset [0, +\infty)$. By Proposition A.2:

$$(u|Au) = \left(\sum_\mu P_\mu u \middle| \sum_\lambda \lambda P_\lambda \sum_\nu P_\nu u \right) = \sum_{\mu,\lambda,\nu} \lambda \left(P_\nu^\dagger P_\lambda^\dagger P_\mu u \middle| u \right) = \sum_{\mu,\lambda,\nu} \lambda \left(P_\nu P_\lambda P_\mu u \right| u)$$

because, if P is an orthogonal projector, $P = P^\dagger$. Using Proposition A.2 once again, $P_\nu P_\lambda P_\mu = \delta_{\nu\mu}\delta_{\mu\lambda} P_\lambda$ and thus

$$(u|Au) = \sum_\lambda \lambda \left(P_\lambda u \right| u) = \sum_\lambda \lambda \left(P_\lambda u \right| P_\lambda u) \ ,$$

where we have used the property of orthogonal projectors $PP = P$. $\lambda \left(P_\lambda u \right| P_\lambda u) \geq 0$ if $\lambda \geq 0$ and thus $(u|Au) \geq 0$ for all $u \in V$.
Concerning (b), assume that A is strictly positive (so it is positive and Hermitean). If $0 \in \sigma(A)$ there must exist $u \neq 0$ with $Au = 0u = 0$. That entails $(u, Au) = (u, 0) = 0$ which is not allowed. Therefore $\sigma(A) \subset [0, +\infty)$. Conversely if A is normal with $\sigma(A) \subset (0, +\infty)$, A is posive by (a). If A is not strictly positive, there is $u \neq 0$ such that $(u|Au) = 0$ and thus, using the same procedure as above,

$$(u|Au) = \sum_\lambda \lambda \left(P_\lambda u \right| P_\lambda u) = 0 \, .$$

Since $\lambda > 0$ and $\left(P_\lambda u \right| P_\lambda u) \geq 0$, it must be $\left(P_\lambda u \right| P_\lambda u) = 0$ for all $\lambda \in \sigma(A)$. This means $P_\lambda u = 0$ for all $\lambda \in \sigma(A)$. This is not possible because, using (1) in Proposition 4.2, $0 \neq u = Iu = \sum_\lambda P_\lambda u = 0$. □

Def.A.2. *If V is a complex vector space equipped with a Hermitean scalar product $(|)$, Let $A : V \to V$ a positive operator. If $B : V \to V$ is another positive operator such that $B^2 = A$, B is called a* **square root of** *A.*

Notice that square roots, if they exist, are Hermitean by Lemma A.1. The next theorem proves that the square root of a positive operator exist and is uniquely determined.

Theorem A.1. *If V is a finite-dimensional complex vector space equipped with a Hermitean scalar product $(|)$. Every positive operator $A : V \to V$ admits a unique square root indicated by $\sqrt{A}$. $\sqrt{A}$ is Hermitian and*

$$\sigma(\sqrt{A}) = \{\sqrt{\lambda} \mid \lambda \in \sigma(A)\} \, .$$

Proof. A is Hermitean by Lemma A.1. Using Proposition A.2

$$A = \sum_{\lambda \in \sigma(A)} \lambda P_\lambda \,.$$

Since $\lambda \geq 0$ we can define the linear operator

$$\sqrt{A} := \sum_{\lambda \in \sigma(A)} \sqrt{\lambda} P_\lambda \,.$$

By Proposition A.2 we have

$$\sqrt{A}\sqrt{A} = \sum_{\lambda \in \sigma(A)} \sqrt{\lambda} P_\lambda \sum_{\mu \in \sigma(A)} \sqrt{\mu} P_\mu = \sum_{\lambda\mu} \sqrt{\lambda\mu} P_\lambda P_\mu \,.$$

Using property (2)

$$\sqrt{A}\sqrt{A} = \sum_{\lambda\mu} \sqrt{\lambda\mu}\,\delta_{\mu\nu} P_\lambda = \sum_{\lambda} (\sqrt{\lambda})^2 P_\lambda = \sum_{\lambda} \lambda P_\lambda = A \,.$$

Notice that $\sqrt{A}$ is Hermitean by construction:

$$\sqrt{A}^\dagger := \sum_{\lambda \in \sigma(A)} \sqrt{\lambda} P_\lambda^\dagger = \sum_{\lambda \in \sigma(A)} \sqrt{\lambda} P_\lambda = \sqrt{A} \,.$$

Moreover, if $B = \{\mu = \sqrt{\lambda}\ |\lambda \in \sigma(A)\}$ and $P'_\mu := P_\lambda$ with $\mu = \sqrt{\lambda}$, it holds

$$\sqrt{A} := \sum_{\mu \in B} \mu P'_\mu \,,$$

and $B \ni \mu \mapsto P'_\mu$ satisfy the properties (1)',(2)',(3)' in Proposition A.2. As a consequence it coincides with the spectral measure associated with $\sqrt{A}$,

$$\sqrt{A} := \sum_{\lambda \in \sigma(A)} \sqrt{\lambda} P_\lambda$$

is the unique spectral decomposition of A,

$$\sigma(\sqrt{A}) = \{\mu = \sqrt{\lambda}\ |\lambda \in \sigma(A)\} \,,$$

and thus $\sqrt{A}$ is positive by Lemma A.1. $\sqrt{A}$ is a Hermitean square root of A with the requested spectral properties.
Let us pass to prove the uniqueness property. Suppose there is another square root S of A. As S is positive, it is Hermitean with $\sigma(S) \subset [0, +\infty)$ and it admits a (unique) spectral decomposition

$$S = \sum_{\nu \in \sigma(S)} \nu P'_\mu \,.$$

Define $B := \{\nu^2 \mid \nu \in \sigma(S)\}$. It is simply proven that the mapping $B \ni \lambda \mapsto P'_{\sqrt{\lambda}}$ satisfy:
(1)' $I = \sum_{\lambda \in B} P'_{\sqrt{\lambda}}$,
(2)' $P'_{\sqrt{\lambda}} P'_{\sqrt{\mu}} = P'_{\sqrt{\mu}} P'_{\sqrt{\lambda}} = 0$ for $\mu \neq \lambda$,
(3)' $A = S^2 = \sum_{\lambda \in B} \lambda P'_{\sqrt{\lambda}}$.
Proposition 4.2 entails that the spectral measure of A and $B \ni \lambda \mapsto P'_{\sqrt{\lambda}}$ coincides: $P'_{\sqrt{\lambda}} = P_\lambda$ for all $\lambda \in \sigma(A) = B$. In other words

$$S = \sum_{\nu \in \sigma(S)} \nu P'_\nu = \sum_{\lambda \in \sigma(A)} \lambda P_\lambda = \sqrt{A}\,.$$

□

Theorem A.2. (**Polar Decomposition of operators**.) *Let V be a finite-dimensional complex vector space equipped with a Hermitean scalar product $(|)$. Every injective (or surjective) operator $A : V \to V$ can be uniquely decomposed as a product of two operators:*

$$A = UM\,,$$

where M is strictly positive and U is unitary.

Proof. First we prove the uniqueness of the decomposition. Assume that $A = UM = U'M'$ where M, M are strictly positive and U, U' are unitary. $A^\dagger = MU^\dagger = M'U'^\dagger$, where we have used the fact that M and M' are hermitean by Lemma 4.1. Thus

$$A^\dagger A = MU^\dagger UM = M'U'^\dagger U'M'\,.$$

But $U^\dagger U = U'^\dagger U' = I$ because U and U' are unitary. As a consequence

$$M^2 = M'^2 = A^\dagger A\,.$$

The latter operator is positive: $(u|A^\dagger Au) = (Au|Au) \geq 0$ and thus $A^\dagger A = M^2 = M'^2$ admits a unique square root. In particular $\sqrt{M'^2} = \sqrt{M^2}$. As M and M' are posive, by uniqueness of square roots, they have to coincide with the $\sqrt{M'^2}$ and $\sqrt{M^2}$ respectively. Hence $M = M'$. From $UM = U'M'$ it also follows $U = U'$.
To conclude the proof, let us build up operators U and M.
We know that $A^\dagger A$ is positive, we can reinforce this property proving that it also is strictly positive. By Lemma A.1, it is sufficient to show that $0 \notin \sigma(A^\dagger A)$ or, in other words, if $A^\dagger Au = 0$, $u = 0$. let us prove that fact. If $A^\dagger Au = 0$, $(v|A^\dagger Au) = 0$ for all $v \in V$ and thus $(Av|Au) = 0$ for all $v \in V$. Since A is injective and $dimV < \infty$, A is surjective and $(Av|Au) = 0$ for all $v \in V$ can equivalently be re-written $(w|Au) = 0$ for all $w \in V$ which ientails $u = 0$. We have proven that $\sigma(A^\dagger A) \subset (0, +\infty)$, as a consequence $|A| := \sqrt{A^\dagger A}$ is strictly positive too by Theorem A.1 and Lemma A.1. Since $\sigma(|A|) \not\ni 0$, $|A|$ is injective and surjective and $|A|^{-1}$ is well-defined.
To conclude, we define

$$M := |A|$$

and

$$U := A|A|^{-1}\,.$$

By construction $UM = A$. Moreover: M is strictly positive (and thus Hermitean) by construction and U is unitary, indeed

$$UU^\dagger = (A|A|^{-1})(A|A|^{-1})^\dagger = A|A|^{-1}(|A|^{-1})^\dagger A^\dagger = A\,|A|^{-1}|A|^{-1}\,A^\dagger$$

where we have used the fact that $|A|^{-1}$ is Hermitean because $|A|$ is so. But $|A|^{-1}|A|^{-1} = (|A|^2)^{-1} = (A^\dagger A)^{-1} = A^{-1}(A^\dagger)^{-1}$ and thus

$$UU^\dagger = A\,A^{-1}(A^\dagger)^{-1}\,A^\dagger = II = I\,.$$

□