A mini-course on tensors

Ivo Terek*

We start with a question: what do Differential Geometry, Algebraic Topology, General Relativity and Quantum Mechanics have in common? Among other things, all of them employ as tools certain objects called *tensors*.

Be it to define the curvature of pseudo-Riemannian manifolds, to state certain results regarding the homology of CW-complexes (such as the Künneth formula), to understand the celebrated Einstein's field equations, or to describe the state space of a composite quantum system, tensors will be there.

The goal of this text is to present a fast and basic introductions to tensors, minimizing the number of pre-requisites: for a good understanding of what's about to be done here, it is recommended a basic familiarity with the concepts of *dual space* and *dual basis*, perhaps seen in a second Linear Algebra course. Moreover, we expect a certain "mathematical maturity" of the reader (which you probably already have, if decided to study this topic seriously).

We start Section 1 defining tensors in vector spaces as certain multilinear maps. We exhibit bases for tensor spaces by using a basis of the initial domain space, and we also introduce Einstein's summation convention (hopefully at the right moment, to avoid bigger traumas).

In Section 2, we'll see how to use a non-degenerate inner product to identify a vector space with its dual in a natural way, via the so-called *musical isomorphisms*. In particular, we'll see how to use them to identify tensors of different types: such identifications, expressed in terms of a basis, are nothing more than the famous process of raising and lowering indices.

We will conclude the discussion in Section 3, where we briefly present an algebraic approach to tensor products, to be characterized by a so-called *universal property*; we also relate such approach with what was done in the previous two sections.

Along the text you will encounter a few exercises, whose simple purpose is help you lose the fear of tensors, and also to help you clarify some issues which require your active participation (yes, get pen and paper).

Remark.

- The following short summary, as the numbers of theorems, propositions and exercises mentioned all have hyperlinks, so that you can navigate through the .pdf file by clicking on them. $\smile$

*terekcouto.1@osu.edu

- The diagrams in Section 3 were produced with the software xfig.

Contents

1 Multilinear maps **3**

2 And what about when we have a scalar product? **15**

3 The universal property **26**

Are you ready? Let's begin.

1 Multilinear maps

Fix, once and for all, a finite-dimensional real vector space V, its dual space $V^* \doteq \{f\colon V \to \mathbb{R} \mid f \text{ is linear}\}$, and denote by $\mathrm{Lin}(V)$ the space of all linear operators[1] of V. For integers $r, s \geq 1$, denote $(V^*)^r \doteq V^* \times \cdots \times V^*$ (r times) and also $V^s \doteq V \times \cdots \times V$ (s times).

Definition 1.1. A *tensor of type* (r, s) on V is a multilinear map $T\colon (V^*)^r \times V^s \to \mathbb{R}$. The set of tensors of type (r, s) on V will be denoted by $\mathcal{T}^r{}_s(V)$.

Remark.

- We'll also say that $T \in \mathcal{T}^r{}_s(V)$ is r times contravariant and s times covariant. The number $r + s$ is called the *rank* or *order* of T. With operations defined pointwise[2], $\mathcal{T}^r{}_s(V)$ becomes a vector space.

- Note that $\mathcal{T}^1{}_0(V) = V^{**} \cong V$ and that $\mathcal{T}^0{}_1(V) = V^*$. Recall that the identification between V and V^{**} is given by $V \ni \boldsymbol{v} \mapsto \hat{\boldsymbol{v}} \in V^{**}$, where $\hat{\boldsymbol{v}}$ is defined by $\hat{\boldsymbol{v}}(f) \doteq f(\boldsymbol{v})$. The reasoning behind such definition of $\hat{\boldsymbol{v}}$ is simple: given only $\boldsymbol{v} \in V$ and $f \in V^*$, there is only one reasonable way to produce a real number.

  Exercise 1.1. Show that $V \ni \boldsymbol{v} \mapsto \hat{\boldsymbol{v}} \in V^{**}$ is indeed a linear isomorphism, in case this is new for you.

  Suggestion. Don't forget to start verifying that for each $\boldsymbol{v} \in V$, $\hat{\boldsymbol{v}}$ is indeed an element of V^{**} (this ensures that the codomain of our isomorphism will be V^{**}).

  We are thus justified in calling linear functionals *covectors*. Also, we set the convention $\mathcal{T}^0{}_0(V) = \mathbb{R}$.

- Sometimes we may encounter multilinear maps whose domains mix the order between the V^* and V factors. For example

$$T\colon V \times (V^*)^3 \times V^2 \times V^* \to \mathbb{R},$$

  and the set of all multilinear maps with such domain will be denoted by $\mathcal{T}_1{}^3{}_2{}^1(V)$; contravariant indices above, and covariant indices below. Let's call those domains *scrambled*.

Example 1.2.

(1) The *evaluation* map $\delta\colon V^* \times V \to \mathbb{R}$ defined by $\delta(f, \boldsymbol{v}) \doteq f(\boldsymbol{v})$; We then have $\delta \in \mathcal{T}^1{}_1(V)$.

[1] Linear transformations from V to itself.
[2] That is, $(T + S)(\boldsymbol{v}) \doteq T(\boldsymbol{v}) + S(\boldsymbol{v})$ and $(\lambda T)(\boldsymbol{v}) \doteq \lambda T(\boldsymbol{v})$.

(2) Given linear operators $T, S \in \mathrm{Lin}(V)$ and $B \in \mathscr{T}^0_{\ 2}(V)$, the *$(T,S)$-pull-back of B*, $(T,S)^*B\colon V \times V \to \mathbb{R}$ is defined by $(T,S)^*B(\boldsymbol{v},\boldsymbol{w}) \doteq B(T\boldsymbol{v}, S\boldsymbol{w})$; Then $(T,S)^*B \in \mathscr{T}^0_{\ 2}(V)$.

(3) The *determinant* $\det\colon (\mathbb{R}^n)^n \to \mathbb{R}$, which takes n vectors and gives as output the determinant of the matrix obtained by placing all vectors in columns; Then $\det \in \mathscr{T}^0_{\ n}(\mathbb{R}^n)$.

(4) The *trace* $\mathrm{tr} : \mathrm{Mat}(n,\mathbb{R}) \to \mathbb{R}$, which maps a matrix to its trace; We then have $\mathrm{tr} \in \mathscr{T}^0_{\ 1}\big(\mathrm{Mat}(n,\mathbb{R})\big)$.

(5) Given $f, g \in V^*$, the *tensor product* of f and g, denoted $f \otimes g\colon V \times V \to \mathbb{R}$, is defined by $(f \otimes g)(\boldsymbol{v},\boldsymbol{w}) \doteq f(\boldsymbol{v})g(\boldsymbol{w})$; Then $f \otimes g \in \mathscr{T}^0_{\ 2}(V)$.

(6) Given a vector $\boldsymbol{v} \in V$ and a covector $f \in V^*$, the *tensor product* of $\boldsymbol{v}$ and f, denoted $\boldsymbol{v} \otimes f\colon V^* \times V \to \mathbb{R}$, is defined by $(\boldsymbol{v} \otimes f)(g, \boldsymbol{w}) \doteq g(\boldsymbol{v})f(\boldsymbol{w})$; Then $\boldsymbol{v} \otimes f \in \mathscr{T}^1_{\ 1}(V)$.

Exercise 1.2. Make sure you understand that the examples above really are tensors.

In the literature, tensors are also described as "n-dimensional arrays of numbers", that is, a sort of generalization of matrices. Let's see how this idea is born:

Definition 1.3. Let $\mathscr{B} = (\boldsymbol{e}_i)_{i=1}^n$ and $\mathscr{B}^* = (\boldsymbol{e}^i)_{i=1}^n$ be dual bases for V and V^*. If $T \in \mathscr{T}^r_{\ s}(V)$, the *components* of T relative to $\mathscr{B}$ are the numbers defined by

$$T^{i_1\ldots i_r}{}_{j_1\ldots j_s} \doteq T(\boldsymbol{e}^{i_1}, \ldots, \boldsymbol{e}^{i_r}, \boldsymbol{e}_{j_1}, \ldots, \boldsymbol{e}_{j_s}),$$

for all $1 \leq i_1, \ldots, i_r, j_1, \ldots, j_s \leq n$.

Remark.

- We will never consider bases of V and V^* which are not dual, when computing the components of a given tensor. This is why we say that the components of T are expressed relatively to $\mathscr{B}$, without mentioning $\mathscr{B}^*$ (which is in fact determined by $\mathscr{B}$). Thus:

  Throughout the text, every time a basis $\mathscr{B} = (\boldsymbol{e}_i)_{i=1}^n$ of V is declared, assume also given the dual basis $\mathscr{B}^* = (\boldsymbol{e}^i)_{i=1}^n$, with this notation.

- For a tensor with scrambled domain, such as $T \in \mathscr{T}^{2\ \ 1}_{\ \ 3}(V)$, we'll write its components as $T^{i_1 i_2}{}_{j_1 j_2 j_3}{}^{i_3} \doteq T(\boldsymbol{e}^{i_1}, \boldsymbol{e}^{i_2}, \boldsymbol{e}_{j_1}, \boldsymbol{e}_{j_2}, \boldsymbol{e}_{j_3}, \boldsymbol{e}^{i_3})$. Keeping track of the spaces between indices will be important when we start raising and lowering indices in the next section, so we might as well start now.

Example 1.4.

(1) Let δ be like in Example 1.2 (p. 3), and $\mathscr{B} = (\boldsymbol{e}_i)_{i=1}^n$ be a basis for V. Thus, by definition, we have

$$\delta^i{}_j = \delta(\boldsymbol{e}^i, \boldsymbol{e}_j) = \boldsymbol{e}^i(\boldsymbol{e}_j) = \begin{cases} 1, \text{ if } i = j, \\ 0, \text{ if } i \neq j. \end{cases}$$

The tensor δ is then called the *Kronecker delta*. Its components receive the same name, for simplicity.

(2) Consider $\det \in \mathscr{T}^0{}_n(\mathbb{R}^n)$. If $\mathscr{B} = (\boldsymbol{e}_i)_{i=1}^n$ is a positive and orthonormal basis of $\mathbb{R}^n$ (equipped with the usual inner product), we denote the components of det by

$$\begin{aligned} \epsilon_{i_1 \ldots i_n} &\doteq \det(\boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_n}) \\ &= \begin{cases} 1, \text{ if } (i_1, \ldots, i_n) \text{ is an even permutation of } (1, \ldots, n) \\ -1, \text{ if } (i_1, \ldots, i_n) \text{ is an odd permutation of } (1, \ldots, n) \\ 0, \text{ else.} \end{cases} \end{aligned}$$

The components $\epsilon_{i_1 \ldots i_n}$ are called the *Levi-Civita permutation symbol*.

Exercise 1.3. If $\mathscr{B} = (\boldsymbol{e}_i)_{i=1}^n$ is a basis for V, show that $\boldsymbol{v} = \sum_{i=1}^n \boldsymbol{e}^i(\boldsymbol{v})\boldsymbol{e}_i$ and that $f = \sum_{i=1}^n f(\boldsymbol{e}_i)\boldsymbol{e}^i$, for any $\boldsymbol{v} \in V$ and $f \in V^*$. Following the component notation previously introduced in Definition 1.3 (p. 4), we may then write $\boldsymbol{v} = \sum_{i=1}^n v^i \boldsymbol{e}_i$ (where $v^i = \boldsymbol{e}^i(\boldsymbol{v})$) and $f = \sum_{i=1}^n f_i \boldsymbol{e}^i$ (where $f_i = f(\boldsymbol{e}_i)$).

Suggestion. Verify that both sides of the equalities proposed give back the same result when applied to a bases of V and V^*, for example, $\mathscr{B}$ and $\mathscr{B}^*$ themselves.

Exercise 1.4. Given a basis $\mathscr{B} = (\boldsymbol{e}_i)_{i=1}^n$ of V and following the notation in Example 1.2 (p. 3), compute:

(a) $((T, S)^* B)_{ij}$ in terms of the components of B and the matrices $(T^i{}_j)_{i,j=1}^n$ and $(S^i{}_j)_{i,j=1}^n$ of T and S relative to $\mathscr{B}$.

(b) $(\boldsymbol{v} \otimes f)^i{}_j$ in terms of the components of $\boldsymbol{v}$ and f.

Proposition 1.5. *Let $\mathscr{B} = (\boldsymbol{e}_i)_{i=1}^n$ be a basis for V. Then*

$$\mathscr{B} \otimes \mathscr{B}^* \doteq \{\boldsymbol{e}_i \otimes \boldsymbol{e}^j \mid 1 \leq i, j \leq n\}$$

is a basis for $\mathscr{T}^1{}_1(V)$. In particular, $\dim \mathscr{T}^1{}_1(V) = n^2$.

Proof: Let's start by checking that $\mathscr{B} \otimes \mathscr{B}^*$ is linearly independent. Suppose that we have

$$\sum_{i,j=1}^n a^i{}_j \boldsymbol{e}_i \otimes \boldsymbol{e}^j = 0.$$

Our goal is to show that all the coefficients $a^i{}_j$ vanish. Evaluating both sides of this equality in the pair (e^k, e_ℓ), we obtain

$$\begin{aligned} 0 &= \left(\sum_{i,j=1}^n a^i{}_j e_i \otimes e^j\right)(e^k, e_\ell) = \sum_{i,j=1}^n a^i{}_j (e_i \otimes e^j)(e^k, e_\ell) \\ &= \sum_{i,j=1}^n a^i{}_j e^k(e_i) e^j(e_\ell) = \sum_{i,j=1}^n a^i{}_j \delta^k{}_i \delta^j{}_\ell \\ &= a^k{}_\ell, \end{aligned}$$

as wanted.

And to see that $\mathcal{B} \otimes \mathcal{B}^*$ spans $\mathcal{T}^1{}_1(V)$, consider $T \in \mathcal{T}^1{}_1(V)$. By Exercise 1.3, we have:

$$\begin{aligned} T(f, v) &= T\left(\sum_{i=1}^n f(e_i) e^i, \sum_{i=1}^n e^j(v) e_j\right) = \sum_{i,j=1}^n f(e_i) e^j(v) T(e^i, e_j) \\ &= \sum_{i,j=1}^n T^i{}_j f(e_i) e^j(v) = \sum_{i,j=1}^n T^i{}_j (e_i \otimes e^j)(f, v) \\ &= \left(\sum_{i,j=1}^n T^i{}_j e_i \otimes e^j\right)(f, v), \end{aligned}$$

for all $f \in V^*$ and $v \in V$, whence $T = \sum_{i,j=1}^n T^i{}_j e_i \otimes e^j$, as wanted. $\square$

Exercise 1.5. Let (e_1, e_2) be a basis for $\mathbb{R}^2$ (not necessarily the standard one) and $T \doteq e_1 \otimes e_1 + e_2 \otimes e_2$.

(a) Show that T cannot be written in the form $v_1 \otimes v_2$, for any $v_1, v_2 \in \mathbb{R}^2$.

(b) Find $v_1, v_2, w_1, w_2 \in \mathbb{R}^2$ such that $T = v_1 \otimes v_2 + w_1 \otimes w_2$, but $v_1 \otimes v_2$ and $w_1 \otimes w_2$ are not multiples of $e_1 \otimes e_1$ and $e_2 \otimes e_2$.

So, a tensor T of type $(1,1)$ can be represented by a matrix $(T^i{}_j)_{i,j=1}^n$, just like an operator $T \in \operatorname{Lin}(V)$. This should make us suspect of the existence of an isomorphism $\mathcal{T}^1{}_1(V) \cong \operatorname{Lin}(V)$. The great thing about this is that not only such isomorphism exists, but that it is natural (in the sense that it does not depend on a choice of basis for V).

Theorem 1.6. *The map* $\Psi\colon \operatorname{Lin}(V) \to \mathcal{T}^1{}_1(V)$ *given by* $\Psi(T)(f, v) \doteq f(T(v))$ *is a linear isomorphism.*

Proof: The verification that indeed $\Psi(T) \in \mathcal{T}^1{}_1(V)$ is the Exercise 1.6 to come. To see that Ψ is an isomorphism, we'll use a basis $\mathcal{B} = (e_i)_{i=1}^n$ for V.

To do so, it suffices to see that if $T^i{}_j$ are the components of $\Psi(T)$ relative to $\mathscr{B}$, then

$$\Psi(T) = \sum_{i,j=1}^{n} T^i{}_j \boldsymbol{e}_i \otimes \boldsymbol{e}^j \quad \text{e}$$

$$T(\boldsymbol{e}_j) = \sum_{i=1}^{n} \boldsymbol{e}^i(T\boldsymbol{e}_j)\boldsymbol{e}_i = \sum_{i=1}^{n} \Psi(T)(\boldsymbol{e}^i, \boldsymbol{e}_j)\boldsymbol{e}_i = \sum_{i=1}^{n} T^i{}_j \boldsymbol{e}_i.$$

In other words, the matrix of the linear operator relative to $\mathscr{B}$ and the components of the associated tensor are the same thing. □

Exercise 1.6. Verify that in the above proof we have $\Psi(T) \in \mathscr{T}^1{}_1(V)$.

Via this isomorphism, it makes sense to talk about the *trace* and the *determinant* of a tensor of type $(1,1)$. Unlike the trace, the determinant is not easily generalized for higher order tensor. So, we will focus on the trace:

Proposition 1.7. *There is a unique linear map* $\operatorname{tr}^1_1 \colon \mathscr{T}^1{}_1(V) \to \mathbb{R}$ *such that*

$$\operatorname{tr}^1_1(\boldsymbol{v} \otimes f) = f(\boldsymbol{v}),$$

for all $f \in V^*$ *and* $\boldsymbol{v} \in V$*. The operation* tr^1_1 *is usually called a* contraction.

Proof: Let's see how this map should work in terms of a basis $\mathscr{B} = (\boldsymbol{e}_i)_{i=1}^n$ for V. Given $T \in \mathscr{T}^1{}_1(V)$, we have

$$\begin{aligned}\operatorname{tr}^1_1(T) &= \operatorname{tr}^1_1\left(\sum_{i,j=1}^{n} T^i{}_j \boldsymbol{e}_i \otimes \boldsymbol{e}^j\right) = \sum_{i,j=1}^{n} T^i{}_j \operatorname{tr}^1_1(\boldsymbol{e}_i \otimes \boldsymbol{e}^j) \\ &= \sum_{i,j=1}^{n} T^i{}_j \boldsymbol{e}^j(\boldsymbol{e}_i) = \sum_{i,j=1}^{n} T^i{}_j \delta^j{}_i = \sum_{i=1}^{n} T^i{}_i.\end{aligned}$$

Naturally, we would like to define $\operatorname{tr}^1_1(T) \doteq \sum_{i=1}^n T^i{}_i$, where $(T^i{}_j)_{i,j=1}^n$ are the components of T relative to a basis for V. But then we are required to show that this definition does not depend on the choice of basis. That is, if $\widetilde{\mathscr{B}} = (\widetilde{\boldsymbol{e}}_i)_{i=1}^n$ is another basis for V (with the corresponding dual basis with upper indices also understood), we have to show that $\sum_{i=1}^n T^i{}_i = \sum_{i=1}^n \widetilde{T}^i{}_i$ holds.

This is done by applying Exercise 1.3 (p. 5) once more:

$$\begin{aligned}\sum_{i=1}^{n} T^i{}_i &= \sum_{i=1}^{n} T(\boldsymbol{e}^i, \boldsymbol{e}_i) = \sum_{i=1}^{n} T\left(\sum_{j=1}^{n} \boldsymbol{e}^i(\widetilde{\boldsymbol{e}}_j)\widetilde{\boldsymbol{e}}^j, \sum_{k=1}^{n} \widetilde{\boldsymbol{e}}^k(\boldsymbol{e}_i)\widetilde{\boldsymbol{e}}_k\right) \\ &= \sum_{i,j,k=1}^{n} \boldsymbol{e}^i(\widetilde{\boldsymbol{e}}_j)\widetilde{\boldsymbol{e}}^k(\boldsymbol{e}_i)T(\widetilde{\boldsymbol{e}}^j, \widetilde{\boldsymbol{e}}_k) = \sum_{j,k=1}^{n} \widetilde{T}^j{}_k \widetilde{\boldsymbol{e}}^k\left(\sum_{i=1}^{n} \boldsymbol{e}^i(\widetilde{\boldsymbol{e}}_j)\boldsymbol{e}_i\right) \\ &= \sum_{j,k=1}^{n} \widetilde{T}^j{}_k \widetilde{\boldsymbol{e}}^k(\widetilde{\boldsymbol{e}}_j) = \sum_{j,k=1}^{n} \widetilde{T}^j{}_k \delta^k{}_j = \sum_{j=1}^{n} \widetilde{T}^j{}_j,\end{aligned}$$

as desired. □

Remark.

- The notation tr^1_1 will become clearer when we discuss contraction for tensors of general type (r,s), soon.
- Observe that, following the notation used in Theorem 1.6 (p. 6), we indeed have $\operatorname{tr}^1_1(T) = \operatorname{tr}(\Psi(T))$, where the second tr of course denotes the usual trace of linear operators.
- In the same fashion, there is a unique linear maps $\overline{\operatorname{tr}}^1_1 \colon \mathscr{T}_1^{\;1}(V) \to \mathbb{R}$ such that $\overline{\operatorname{tr}}^1_1(f \otimes \boldsymbol{v}) = f(\boldsymbol{v})$. We will see in Section 2 that in the presence of a non-degenerate scalar product $\langle \cdot, \cdot \rangle$ in V, $\overline{\operatorname{tr}}^1_1$ is equivalent to tr^1_1.

For the first time we needed to relate components of a tensor relative to different bases. This ends up being important in Physics, where problems are usually dealt with in terms of different coordinate systems. For tensors of type $(1,1)$, we have the:

Proposition 1.8. *Let $T \in \mathscr{T}^1{}_1(V)$ and consider bases $\mathscr{B} = (\boldsymbol{e}_i)_{i=1}^n$ and $\widetilde{\mathscr{B}} = (\widetilde{\boldsymbol{e}}_i)_{i=1}^n$ for V. If*

$$\widetilde{\boldsymbol{e}}_j = \sum_{i=1}^n a^i{}_j \boldsymbol{e}_i \quad \text{and} \quad \boldsymbol{e}_j = \sum_{i=1}^n b^i{}_j \widetilde{\boldsymbol{e}}_i,$$

then

$$\widetilde{T}^i{}_j = \sum_{k,\ell=1}^n b^i{}_k a^\ell{}_j T^k{}_\ell.$$

Proof: Straightforward computation:

$$\widetilde{T}^i{}_j = T(\widetilde{\boldsymbol{e}}^i, \widetilde{\boldsymbol{e}}_j) = T\left(\sum_{k=1}^n b^i{}_k \boldsymbol{e}^k, \sum_{\ell=1}^n a^\ell{}_j \boldsymbol{e}_\ell\right) = \sum_{k,\ell=1}^n b^i{}_k a^\ell{}_j T^k{}_\ell.$$

$\square$

Exercise 1.7. Assume the notation above.

(a) Show that we indeed have $\widetilde{\boldsymbol{e}}^i = \sum_{k=1}^n b^i{}_k \boldsymbol{e}^k$ and that $\boldsymbol{e}^i = \sum_{j=1}^n a^i{}_j \widetilde{\boldsymbol{e}}^j$.

Suggestion. Use Exercise 1.3 (p. 5).

(b) Show that the matrices $(a^i{}_j)_{i,j=1}^n$ and $(b^i{}_j)_{i,j=1}^n$ are inverses.

Suggestion. Do substitutions and apply the definition of matrix product, verifying that $\sum_{k=1}^n a^i{}_k b^k{}_j = \sum_{k=1}^n b^i{}_k a^k{}_j = \delta^i{}_j$.

(c) Mostre that if $T \in \mathscr{T}^1{}_2(V)$, then $\widetilde{T}^i{}_{jk} = \sum_{\ell,p,q=1}^n b^i{}_\ell a^p{}_j a^q{}_k T^\ell{}_{pq}$. Can you guess what will happen for general tensors $T \in \mathscr{T}^r{}_s(V)$? *Spoiler:* Theorem 1.14 (p. 14, to come).

The above exercise actually tells us several things about how tensors behave under base change. In particular, it says that the dual basis transforms in the "opposite" direction of the initial basis. One can use this to argue and justify the names "contravariant/covariant" mentioned before.

Now that we are used to tensors of type $(1,1)$, we will see that the adaptations for the (r,s) case require almost no effort at all. We start with the general definition of the tensor product $\otimes$ operation, mentioned briefly in items (5) and (6) of Example 1.2, p. 3):

Definition 1.9. The *tensor product* of $T \in \mathscr{T}^r{}_s(V)$ and $S \in \mathscr{T}^{r'}{}_{s'}(V)$ is the tensor $T \otimes S \in \mathscr{T}^{r+r'}{}_{s+s'}(V)$ defined by

$$\begin{aligned}(T \otimes S)(f^1, \ldots, f^{r+r'}, v_1, \ldots, v_{s+s'}) \\ \doteq T(f^1, \ldots, f^r, v_1, \ldots, v_s) S(f^{r+1}, \ldots, f^{r+r'}, v_{s+1}, \ldots, v_{s+s'}).\end{aligned}$$

Exercise 1.8.

(a) Show that $\otimes$ is associative.

(b) Give an example (when $r = r'$ and $s = s'$) showing that $\otimes$ is not, in general, commutative.

(c) Show that if $\mathscr{B} = (e_i)_{i=1}^n$ is a basis for V, then

$$(T \otimes S)^{i_1 \ldots i_{r+r'}}{}_{j_1 \ldots j_{s+s'}} = T^{i_1 \ldots i_r}{}_{j_1 \ldots j_s} S^{i_{r+1} \ldots i_{r+r'}}{}_{j_{s+1} \ldots j_{s+s'}}.$$

(d) Besides the result from the item above, we see that $\otimes$ really does behave like a product: show that if $T_1, T_2, T \in \mathscr{T}^r{}_s(V)$, $S_1, S_2, S \in \mathscr{T}^{r'}{}_{s'}(V)$ and $\lambda \in \mathbb{R}$, then:

- $(T_1 + \lambda T_2) \otimes S = T_1 \otimes S + \lambda(T_2 \otimes S)$;
- $T \otimes (S_1 + \lambda S_2) = T \otimes S_1 + \lambda(T \otimes S_2)$.

Remark. One can also define the tensor product of tensors with scrambled domains, providing arguments and feeding the tensors in order until there are no arguments left, also unscrambling the domains. For example, if $T \in \mathscr{T}^1{}_2{}^1(V)$ and $S \in \mathscr{T}^2{}_1(V)$, we can define $T \otimes S \in \mathscr{T}^4{}_3(V)$ by

$$(T \otimes S)(f^1, f^2, f^3, f^4, v_1, v_2, v_3) = T(f^1, v_1, v_2, f^2) S(f^3, f^4, v_3).$$

Such definition gives us the reasonable relation

$$(T \otimes S)^{i_1 i_2 i_3 i_4}{}_{j_1 j_2 j_3} = T^{i_1}{}_{j_1 j_2}{}^{i_2} S^{i_3 i_4}{}_{j_3}.$$

Exercise 1.9. Define $T \otimes S \in \mathscr{T}^3{}_5(V)$ for $T \in \mathscr{T}^1{}_3(V)$ and $S \in \mathscr{T}^2{}_2(V)$, and write its components in terms of the components of T and S.

With this, we can generalize Proposition 1.5 (p. 5):

Proposition 1.10. *Let $\mathscr{B} = (\boldsymbol{e}_i)_{i=1}^n$ be a basis for V. Then*

$$\mathscr{B}^{\otimes r} \otimes (\mathscr{B}^*)^{\otimes s} \doteq \{\boldsymbol{e}_{i_1} \otimes \cdots \otimes \boldsymbol{e}_{i_r} \otimes \boldsymbol{e}^{j_1} \otimes \cdots \otimes \boldsymbol{e}^{j_s} \mid 1 \leq i_1, \ldots, i_r, j_1, \ldots, j_s \leq n\}$$

is a basis for $\mathscr{T}^r{}_s(V)$. In particular, $\dim \mathscr{T}^r{}_s(V) = n^{r+s}$.

Proof: To see that $\mathscr{B}^{\otimes r} \otimes (\mathscr{B}^*)^{\otimes s}$ is linearly independent, consider the linear combination

$$\sum_{i_1,\ldots,i_r,j_1,\ldots,j_s=1}^n a^{i_1\ldots i_r}{}_{j_1\ldots j_r} \boldsymbol{e}_{i_1} \otimes \cdots \otimes \boldsymbol{e}_{i_r} \otimes \boldsymbol{e}^{j_1} \otimes \cdots \otimes \boldsymbol{e}^{j_s} = 0.$$

Evaluate both sides of the equality in $(\boldsymbol{e}^{k_1}, \cdots, \boldsymbol{e}^{k_r}, \boldsymbol{e}_{\ell_1}, \cdots, \boldsymbol{e}_{\ell_s})$ and use the definition of tensor product to obtain

$$\sum_{i_1,\ldots,i_r,j_1,\ldots,j_s=1}^n a^{i_1\ldots i_r}{}_{j_1\ldots j_r} \boldsymbol{e}^{k_1}(\boldsymbol{e}_{i_1}) \cdots \boldsymbol{e}^{k_r}(\boldsymbol{e}_{i_r}) \boldsymbol{e}^{j_1}(\boldsymbol{e}_{\ell_1}) \cdots \boldsymbol{e}^{j_s}(\boldsymbol{e}_{\ell_s}) = 0.$$

Simplifying, it follows that

$$\sum_{i_1,\ldots,i_r,j_1,\ldots,j_s=1}^n a^{i_1\ldots i_r}{}_{j_1\ldots j_r} \delta^{k_1}_{i_1} \cdots \delta^{k_r}_{i_r} \delta^{j_1}_{\ell_1} \cdots \delta^{j_s}_{\ell_s} = 0,$$

whence we get $a^{k_1\ldots k_r}{}_{\ell_1\ldots\ell_s} = 0$. The indices were arbitrary, so we conclude that $\mathscr{B}^{\otimes r} \otimes (\mathscr{B}^*)^{\otimes s}$ is linearly independent.

Now we'll see that $\mathscr{B}^{\otimes r} \otimes (\mathscr{B}^*)^{\otimes s}$ spans $\mathscr{T}^r{}_s(V)$. Indeed, if $T \in \mathscr{T}^r_s(V)$, we claim that

$$T = \sum_{i_1\ldots i_r,j_1,\ldots,j_s=1}^n T^{i_1\ldots i_r}{}_{j_1\ldots j_s} \boldsymbol{e}_{i_1} \otimes \cdots \otimes \boldsymbol{e}_{i_r} \otimes \boldsymbol{e}^{j_1} \otimes \cdots \otimes \boldsymbol{e}^{j_s}.$$

To wit, it suffices to see that both sides give the same result when evaluated in a basis of the domain space $(V^*)^r \times V^s$. So, consider arbitrary indices and the $(r+s)$-uple $(\boldsymbol{e}^{k_1}, \cdots, \boldsymbol{e}^{k_r}, \boldsymbol{e}_{\ell_1}, \cdots, \boldsymbol{e}_{\ell_s})$. Evaluating in such $(r+s)$-uple, in the left-hand side we obtain $T^{k_1\ldots k_r}{}_{\ell_1\ldots\ell_s}$, by definition, and on the right-hand side we also get $T^{k_1\ldots k_r}{}_{\ell_1\ldots\ell_s}$, by a similar computation done (with Kronecker deltas) to verify the linear independence of $\mathscr{B}^{\otimes r} \otimes (\mathscr{B}^*)^{\otimes s}$.

$\square$

Exercise 1.10. Assume the notation from the above proposition.

(a) Show that

$$T = \sum_{i_1\ldots i_r,j_1,\ldots,j_s=1}^n T^{i_1\ldots i_r}{}_{j_1\ldots j_s} \boldsymbol{e}_{i_1} \otimes \cdots \otimes \boldsymbol{e}_{i_r} \otimes \boldsymbol{e}^{j_1} \otimes \cdots \otimes \boldsymbol{e}^{j_s}$$

directly, computing $T(f^1, \ldots, f^r, v_1, \ldots, v_s)$ for arbitrary entries, mimicking the case $(r,s) = (1,1)$ done in the proof of Proposition 1.5 (p. 5).

(b) Compare carefully the proofs of propositions 1.5 (p. 5) and 1.10 and convince yourself that no original idea was employed here (except maybe from what was suggested in item (a)).

Remark (Einstein's convention)**.** To avoid writing unpleasant things such as, for example, $\sum_{i_1\ldots i_r, j_1,\ldots,j_s=1}^n \cdots$, we will adopt the following convention, due to Einstein:

- Usually an agreement is made about the ranges of the indices, and different ranges will correspond to different alphabets. For example, a, b and c will range from 1 to n, i, j and k from 1 to 3, greek letters as μ, ν and λ from 0 to 4 (common in General Relativity, to express the components of the metric tensor of a spacetime as $g_{\mu\nu}$), etc..

- All the summation symbols are omitted, being understood that if the same index appears in a monomial expression once above and once below, we're actually summing the expression over that index. For example:
  - $\boldsymbol{v} = \sum_{i=1}^n v^i \boldsymbol{e}_i$ becomes $\boldsymbol{v} = v^i \boldsymbol{e}_i$ and $f = \sum_{i=1}^n f_i \boldsymbol{e}^i$ becomes $f = f_i \boldsymbol{e}^i$.
  - $\operatorname{tr}^1_1(T) = \sum_{i=1}^n T^i{}_i$ becomes $\operatorname{tr}^1_1(T) = T^i{}_i$.
  - $\widetilde{T}^i{}_j = \sum_{k,\ell=1}^n b^i{}_k a^\ell{}_j T^k{}_\ell$ becomes $\widetilde{T}^i{}_j = b^i{}_k a^\ell{}_j T^k{}_\ell$.
  - $A^i = \sum_{j=1}^n B^i{}_j C^j$ becomes $A^i = B^i{}_j C^j$.
  - $\sum_{i=1}^n A^i{}_{ji}$ becomes $A^i{}_{ji}$ and $\sum_{k=1}^n a^i{}_k b^k{}_j = \delta^i{}_j$ becomes $a^i{}_k b^k{}_j = \delta^i{}_j$.

- One must pay attention to which indices are "free" and which ones are not. It is the same situation when we have dummy variables in definite integrals or summations themselves. For example, we have that

  $$v^i \boldsymbol{e}_i = v^j \boldsymbol{e}_j = v^k \boldsymbol{e}_k = \cdots,$$

  etc., but note that in those situations, the repeated indices indeed appear once above and once below. That said, expressions such as $v^i v^i$ or $v_i + u^i$ are not compatible with Einstein's convention, and in the rare cases one cannot avoid them, the solution is to use the summation symbol explicitly.

- In this setting, the action of the Kronecker delta consists in switching the index being summed by the remaining index in the delta. For example, $\delta^i{}_j v^j = v^i$, also $\delta^i{}_j \delta^k{}_\ell A^{j\ell} = A^{ik}$, etc..

- When performing substitutions, one must be careful to avoid repeating dummy indices. For example, if $p^i = a^i{}_j v^j$ and $v^j = b^j{}_i w^i$, it is not correct to write $p^i = a^i{}_j b^j{}_i w^i$. Note that the index i appears three times in the left-hand side. The correct would be to identify that in $v^j = b^j{}_i w^i$, the index i is a dummy. So, we may write $v^j = b^j{}_k w^k$ and substitute $p^i = a^i{}_j b^j{}_k w^k$, without conflicts (observe the double summation implied).

Exercise 1.11.

(a) Suppose that $v^i = a^i{}_j w^j$ and that $(b^i{}_j)_{i,j=1}^n$ is the inverse matrix of $(a^i{}_j)_{i,j=1}^n$. Show that $w^i = b^i{}_j v^j$.

Suggestion. Multiply both sides of $v^i = a^i{}_j w^j$ by b^k_i and, after simplifying, rename $k \to i$.

(b) Simplify:

- $\delta^i{}_j \delta^j{}_k \delta^k{}_i$.
- $\epsilon_{1jk\ell} \delta^j{}_2 \delta^k{}_4 \delta^\ell{}_3$.
- $\delta^i{}_j v_i u^j$.
- $\delta^2_j \delta^j{}_k v^k$.
- $\delta^3_j \delta^j{}_1$;
- $\epsilon_{i3k} \delta^i{}_p v^k$.

(c) The Kronecker delta and the Levi-Civita symbol are very particular tensors, which after certain identifications (to be seen in Section 2), may be identified with tensors whose components are, in adequate bases, δ_{ij} and $\epsilon^i{}_{jk}$, assuming the same values as $\delta^i{}_j$ and ϵ_{ijk} (revisit Example 1.4, p. 5). Show that

$$\epsilon^i{}_{jk} \epsilon^k{}_{\ell m} = \delta^i{}_\ell \delta_{jm} - \delta^i{}_m \delta_{j\ell}.$$

Suggestion. Be patient and analyse cases, this is a "combinatorics" problem. Observe that there are no summations on the right-hand side, only on the left-hand side.

(d) Redo Exercise 1.7 (p. 8) using Einstein's convention.

(e) (Challenge) Show that if $\boldsymbol{n} = (n^1, n^2, n^3) \in \mathbb{R}^3$ is a unit vector,

$$N^{ij} \doteq \delta^{ij} - \epsilon^{ij}{}_k n^k + n^i n^j \quad \text{e} \quad M_i^{\ j} \doteq \delta^{ij} + \epsilon_i{}^j{}_k n^k,$$

then $N^{ij} M_j^{\ k} = 2\delta^{ik}$.

Suggestion. Again, the numerical values of δ^{ij} and δ_{ij} are the same of $\delta^i{}_j$, and similarly for $\epsilon^{ij}{}_k$, $\epsilon_i{}^j{}_k$ and ϵ_{ijk}. By item (c), we have that $\epsilon^{ij}{}_r \epsilon_j{}^k{}_s = \delta^i{}_s \delta^k{}_r - \delta^{ik} \delta_{rs}$. Moreover, $\epsilon_j{}^k{}_s n^j n^s = 0$ (why?), and $\boldsymbol{n}$ being a unit vector says that $\delta_{rs} n^r n^s = 1$.

$\boxed{\textit{We will adopt Einstein's convention from here on.}}$

Back to tensors. The same argument given in the proof of Proposition 1.10 (p. 10) allows us to find bases for tensor spaces with scrambled domains. For example:

Exercise 1.12. Let $\mathscr{B} = (e_i)_{i=1}^n$ be a basis for V. Show that

$$\mathscr{B}^* \otimes \mathscr{B} \otimes \mathscr{B}^* \doteq \{e^i \otimes e_j \otimes e^k \mid 1 \leq i,j,k \leq n\}$$

is a basis for the space $\mathscr{T}_1{}^1{}_1(V)$, with operations defined pointwise. In particular, $\dim \mathscr{T}_1{}^1{}_1(V) = n^3$.

In general, all the spaces of tensors for a given order are isomorphic, no matter how the domains might be scrambled. It is possible to exhibit such isomorphisms without choosing bases for V and V^*, once we have a non-degenerate scalar product $\langle \cdot, \cdot \rangle$ on V. We will see how this works in Section 2.

In Proposition 1.7 (p. 7), we introduced the contraction $\operatorname{tr}{}^1_1 \colon \mathscr{T}^1{}_1(V) \to \mathbb{R}$. For tensors of type (r,s), the contraction we will define from $\operatorname{tr}{}^1_1$ produces not a real number, but yet another tensor:

Definition 1.11. Let $r,s \geq 1$. The *contraction* in the a-th contravariant slot and b-th covariant slot is the map $\operatorname{tr}{}^a_b \colon \mathscr{T}^r{}_s(V) \to \mathscr{T}^{r-1}{}_{s-1}(V)$ given by

$$\begin{aligned}\operatorname{tr}{}^a_b(T)(f^1,\ldots,f^{r-1},v_1,\ldots,v_{s-1}) &\doteq \\ &\doteq \operatorname{tr}{}^1_1\big(T(f^1,\ldots,f^{a-1},\bullet,f^a,\ldots,f^{r-1},v_1,\ldots,v_{b-1},\bullet,v_b,\ldots,v_s)\big).\end{aligned}$$

Remark. The $\bullet$ notation indicates exactly which arguments remain free. For example, if $T \in \mathscr{T}^1{}_1(V)$ and $f \in V^*$ is fixed, $T(f,\bullet)$ denotes the element of $\mathscr{T}^0{}_1(V)$ given by $V \ni v \mapsto T(f,v) \in \mathbb{R}$.

Essenially, we freeze all possible arguments so to obtain a tensor of type $(1,1)$, and then we apply the usual $\operatorname{tr}{}^1_1$ contraction.

Example 1.12. Fix a basis $\mathscr{B} = (e_i)_{i=1}^n$ for V.

(1) If $\delta \colon V^* \times V \to \mathbb{R}$ is given by $\delta(f,v) = f(v)$, like in Example 1.2 (p. 3), then $\operatorname{tr}{}^1_1(\delta) = \delta^i{}_i = n$.

(2) If $B \in \mathscr{T}^1{}_2(V)$ is given by $B = B^i{}_{jk} e_i \otimes e^j \otimes e^k$, we have two possible contractions. Namely:
$$\operatorname{tr}{}^1_1(B) = B^i{}_{ik} e^k \quad \text{and} \quad \operatorname{tr}{}^1_2(B) = B^i{}_{ji} e^j.$$

(3) If $W \in \mathscr{T}^2{}_3(V)$ is given by $W = W^{ij}{}_{k\ell m} e_i \otimes e_j \otimes e^k \otimes e^\ell \otimes e^m$, we have $2 \cdot 3 = 6$ possible contractions (all of them order 3 tensors). Some of them are
$$\begin{aligned}\operatorname{tr}{}^1_1(W) &= W^{ij}{}_{i\ell m} e_j \otimes e^\ell \otimes e^m, \\ \operatorname{tr}{}^1_2(W) &= W^{ij}{}_{kim} e_j \otimes e^k \otimes e^m, \quad \text{and} \\ \operatorname{tr}{}^2_3(W) &= W^{ij}{}_{k\ell j} e_i \otimes e^k \otimes e^\ell.\end{aligned}$$

Exercise 1.13. Find the remaining contractions $\operatorname{tr}^1_3(W)$, $\operatorname{tr}^2_1(W)$ and $\operatorname{tr}^2_2(W)$.

Once the above examples are understood, the statement of the general case is not surprising (it is actually an automatic corollary of Proposition 1.7, p. 7):

Proposition 1.13. *Let $\mathscr{B} = (\boldsymbol{e}_i)_{i=1}^n$ be a basis for V and $T \in \mathscr{T}^r{}_s(V)$. So given $1 \leq a \leq r$ and $1 \leq b \leq s$, we have*

$$(\operatorname{tr}^a_b(T))^{i_1\ldots i_{r-1}}{}_{j_1\ldots j_{s-1}} = T^{i_1\ldots i_{a-1}k i_a\ldots i_{r-1}}{}_{j_1\ldots j_{b-1}k j_b\ldots j_{s-1}}.$$

In an analogous way to what we did previously when discussing tensors of type $(1,1)$, let's see how the components of a general tensor behave under base change:

Theorem 1.14. *Let $T \in \mathscr{T}^r{}_s(V)$ and consider bases $\mathscr{B} = (\boldsymbol{e}_i)_{i=1}^n$ and $\widetilde{\mathscr{B}} = (\widetilde{\boldsymbol{e}}_i)_{i=1}^n$ for V. If $\widetilde{\boldsymbol{e}}_j = a^i{}_j\boldsymbol{e}_i$ and $\boldsymbol{e}_j = b^i{}_j\widetilde{\boldsymbol{e}}_i$, then*

$$\widetilde{T}^{i_1\ldots i_r}{}_{j_1\ldots j_s} = b^{i_1}_{k_1}\cdots b^{i_r}_{k_r}a^{\ell_1}_{j_1}\cdots a^{\ell_s}_{j_s}T^{k_1\ldots k_r}{}_{\ell_1\ldots\ell_s}.$$

Remark. One way to think of this transformation law is in terms of the $(a^i{}_j)_{i,j=1}^n$, bearing in mind that $(b^i{}_j)_{i,j=1}^n$ is the inverse matrix (we saw that on Exercise 1.7, p. 8): for each covariant (lower) index, an a term contributes, while for each contravariant (upper) index, a b term contrbutes. That is, "co" terms correspond to the "direct" coefficient matrix, and "contra" terms to its inverse.

Proof: It is almost automatic, from Exercise 1.7 (p. 8) and Einstein's convention:

$$\begin{aligned}
\widetilde{T}^{i_1\ldots i_r}{}_{j_1\ldots j_s} &= T(\widetilde{\boldsymbol{e}}^{i_1},\ldots,\widetilde{\boldsymbol{e}}^{i_r},\widetilde{\boldsymbol{e}}_{j_1},\ldots,\widetilde{\boldsymbol{e}}_{j_s}) \\
&= T(b^{i_1}_{k_1}\boldsymbol{e}^{k_1},\ldots,b^{i_r}_{k_r}\boldsymbol{e}^{k_r},a^{\ell_1}_{j_1}\boldsymbol{e}_{\ell_1},\ldots,a^{\ell_s}_{j_s}\boldsymbol{e}_{\ell_s}) \\
&= b^{i_1}_{k_1}\cdots b^{i_r}_{k_r}a^{\ell_1}_{j_1}\cdots a^{\ell_s}_{j_s}T(\boldsymbol{e}^{k_1},\ldots,\boldsymbol{e}^{k_r},\boldsymbol{e}_{\ell_1},\ldots,\boldsymbol{e}_{\ell_s}) \\
&= b^{i_1}_{k_1}\cdots b^{i_r}_{k_r}a^{\ell_1}_{j_1}\cdots a^{\ell_s}_{j_s}T^{k_1\ldots k_r}{}_{\ell_1\ldots\ell_s}.
\end{aligned}$$

$\square$

2 And what about when we have a scalar product?

We know what a positive-definite inner product in a vector space is. The condition that such product is positive-definite may be weakened without loss of relevant algebraic properties for what we're doing here. Let's register this in the following definition, using the language we have established so far:

Definition 2.1. A (pseudo-Euclidean) *scalar product* on V is a tensor $\langle \cdot, \cdot \rangle \in \mathcal{T}^0_2(V)$ satisfying:

(i) $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$ for all $\boldsymbol{x}, \boldsymbol{y} \in V$. That is, $\langle \cdot, \cdot \rangle$ is *symmetric*.

(ii) if $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$ for all $\boldsymbol{y} \in V$, then necessarily $\boldsymbol{x} = \boldsymbol{0}$. That is, $\langle \cdot, \cdot \rangle$ is *non-degenerate*.

If instead of (ii), the stronger condition

(iii) $\langle \boldsymbol{x}, \boldsymbol{x} \rangle > 0$ for all non-zero $\boldsymbol{x} \in V$,

holds, then $\langle \cdot, \cdot \rangle$ is called an (Euclidean, positive-definite) *inner product* on V.

Remark. Note that (iii) does imply (ii) – it suffices to take $\boldsymbol{y} = \boldsymbol{x}$. In more general setting, such a scalar product $\langle \cdot, \cdot \rangle$ is also called a *metric tensor*.

$$\boxed{\textit{We fix from here on a scalar product } \langle \cdot, \cdot \rangle \textit{ on } V.}$$

Since $\langle \cdot, \cdot \rangle$ is a tensor, in particular we knoow what are its components relative to a basis $\mathscr{B} = (\boldsymbol{e}_i)_{i=1}^n$ of V: $g_{ij} \doteq \langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle$. The conditions imposed in the definition of $\langle \cdot, \cdot \rangle$ will give us good properties of such components. Namely:

- Condition (i) ensures that the matrix $(g_{ij})_{i,j=1}^n$ is symmetric;
- Condition (ii) ensures that the matrix $(g_{ij})_{i,j=1}^n$ is *non-singular* (that is, has an inverse).

The inverse matrix of $(g_{ij})_{i,j=1}^n$ is usually denoted by $(g^{ij})_{i,j=1}^n$. The advantages of having a scalar product start when we use it to obtain new natural identifications (which do not depend on any choice of basis for V).

Exercise 2.1. Simplify:

- $\delta^i{}_j g_{ki} g^{j\ell} \delta^k{}_\ell$.
- $\epsilon_{1\ell m} g_{ij} g^{jk} \delta^1{}_k \delta^\ell{}_2 \delta^m_3$.

Let's begin by properly identifying V with its dual V^*, using $\langle \cdot, \cdot \rangle$:

Proposition 2.2 (Musical isomorphisms).

(i) The flat *map* $\flat\colon V \to V^*$ *defined by* $\boldsymbol{v}_\flat(\boldsymbol{w}) \doteq \langle \boldsymbol{v}, \boldsymbol{w} \rangle$ *is an isomorphism.*

(ii) Given $f \in V^*$, there is a unique $f^\sharp \in V$ such that $f(\boldsymbol{v}) = \langle f^\sharp, \boldsymbol{v}\rangle$, for all $\boldsymbol{v} \in V$. Thus, it is well defined the sharp *map $\sharp\colon V^* \to V$, which is the inverse isomorphism of flat.*

Proof: Clearly $\flat$ is linear. Since V has finite dimension, it suffices to show that $\flat$ is injective. But if $\boldsymbol{v} \in \ker \flat$, we have $\langle \boldsymbol{v}, \boldsymbol{w}\rangle = 0$ for all $\boldsymbol{w} \in V$. Since $\langle \cdot, \cdot \rangle$ is non-degenerate, it follows that $\boldsymbol{v} = \boldsymbol{0}$. So $\flat$ is an isomorphism, and we denote its inverse by $\sharp$. ☐

The next natural step is to analyze how those musical isomorphisms are expressed in terms of a basis:

Proposition 2.3. *Let $\mathscr{B} = (\boldsymbol{e}_i)_{i=1}^n$ be a basis for V. Suppose that $\boldsymbol{v} = v^i \boldsymbol{e}_i \in V$ and $f = f_i \boldsymbol{e}^i \in V^*$. So $\boldsymbol{v}_\flat = v_i \boldsymbol{e}^i$ and $f^\sharp = f^i \boldsymbol{e}_i$, where*

$$v_i = g_{ij} v^j \quad \text{and} \quad f^i = g^{ij} f_j.$$

Remark.

- Note the abuses of notation $(\boldsymbol{v}_\flat)_i = v_i$ and $(f^\sharp)^i = f^i$.
- The product $\langle \cdot, \cdot \rangle$ is used to raise and lower the indices of the components of $\boldsymbol{v}$ and f. Such operation is common in Physics and Geometry. Note also the similarity of the action of the coefficients g_{ij} and g^{ij} with the Kronecker delta.
- This justify the name "musical isomorphisms": the $\flat$ lowers the pitch of a scale (lowers the indices of the components of $\boldsymbol{v}$, $v^i \to v_i$), while the $\sharp$ raises the pitch of a scale (raises the indices of the components of f, $f_i \to f^i$).
- Mnemonic: vectors have "sharp tips" (sharp $\nearrow$), so $f^\sharp$ is always a vector.

Proof: On one hand, the equality $\boldsymbol{v}_\flat(\boldsymbol{e}_j) = \langle \boldsymbol{v}, \boldsymbol{e}_j \rangle$ reads as $v_j \boldsymbol{e}^j(\boldsymbol{e}_i) = \langle v^j \boldsymbol{e}_j, \boldsymbol{e}_i \rangle$, that is, $v_i = g_{ji} v^j = g_{ij} v^j$, as desired. On the other hand, $f(\boldsymbol{e}_j) = \langle f^\sharp, \boldsymbol{e}_j \rangle$ becomes $f_i \boldsymbol{e}^i(\boldsymbol{e}_j) = \langle f^i \boldsymbol{e}_i, \boldsymbol{e}_j \rangle$, whence $f_j = g_{ij} f^i$. Multiplying everything by g^{kj} (and summing over j, of course[3]), we have

$$g^{kj} f_j = g^{kj} g_{ij} f^i = \delta^k_{\ i} f^i = f^k,$$

and renaming $k \to i$ gives us that $f^i = g^{ij} f_j$, exactly as stated. ☐

Proposition 2.4. *Let $\mathscr{B} = (\boldsymbol{e}_i)_{i=1}^n$ be a basis for V. So*

$$(\boldsymbol{e}_i)_\flat = g_{ij} \boldsymbol{e}^j \quad \text{and} \quad (\boldsymbol{e}^i)^\sharp = g^{ij} \boldsymbol{e}_j.$$

Proof: Let's check the first one. As $\boldsymbol{e}_i = \delta^j_{\ i} \boldsymbol{e}_j$, the previous proposition gives us that

$$(\boldsymbol{e}_i)_\flat = g_{jk} \delta^k_{\ i} \boldsymbol{e}^j = g_{ji} \boldsymbol{e}^j = g_{ij} \boldsymbol{e}^j,$$

as desired. ☐

[3] Some texts call that "contracting against g^{kj}".

Exercise 2.2.

(a) Make sure you understood the computations done using Einstein's convention in the last two proofs.

(b) Verify that $(\boldsymbol{e}^i)^\sharp = g^{ij}\boldsymbol{e}_j$, completing the previous proof.

Changing gears a bit, what we have just now amounts to using $\langle \cdot, \cdot \rangle$ to identify $\mathcal{T}^0{}_1(V)$ and $\mathcal{T}^1{}_0(V)$. This can be done to identify $\mathcal{T}^r{}_s(V)$ with $\mathcal{T}^{r'}{}_{s'}(V)$, provided $r + s = r' + s'$. Let's see some low order cases, to begin with.

Proposition 2.5. *The map* $\sharp_1 \colon \mathcal{T}^0{}_2(V) \to \mathcal{T}^1{}_1(V)$ *given by* $T^{\sharp_1}(f, \boldsymbol{v}) \doteq T(f^\sharp, \boldsymbol{v})$ *is an isomorphism.*

Proof: Clearly $\sharp_1$ is linear, and since $\dim \mathcal{T}^0_2(V) = \dim \mathcal{T}^1{}_1(V)$, it suffices to see that $\sharp_1$ is injective. Let $\mathcal{B} = (\boldsymbol{e}_i)_{i=1}^n$ be a basis for V. Note that

$$(T^{\sharp_1})^i{}_j = T^{\sharp_1}(\boldsymbol{e}^i, \boldsymbol{e}_j) = T((\boldsymbol{e}^i)^\sharp, \boldsymbol{e}_j) = T(g^{ik}\boldsymbol{e}_k, \boldsymbol{e}_j) = g^{ik}T(\boldsymbol{e}_k, \boldsymbol{e}_j) = g^{ik}T_{kj}.$$

If $T^{\sharp_1} = 0$, then $g^{ik}T_{kj} = 0$. Multiplying everything by $g_{\ell i}$, we have

$$0 = g_{\ell i}g^{ik}T_{kj} = \delta^k{}_\ell T_{kj} = T_{\ell j},$$

and from the arbitrariety of indices it follows that $T = 0$. $\square$

Remark. Instead of writing $(T^{\sharp_1})^i{}_j = g^{ik}T_{kj}$ every time, we usually write just $T^i{}_j = g^{ik}T_{kj}$, just like we have been doing with vectors and covectors.

Exercise 2.3.

(a) Show that the map $\flat_1 \colon \mathcal{T}^2{}_0(V) \to \mathcal{T}^1{}_1(V)$ defined by $T_{\flat_1}(f, \boldsymbol{v}) \doteq T(f, \boldsymbol{v}_\flat)$ is an isomorphism.

(b) In case you have already not checked this while doing item (a), verify also that $T^i{}_j = g_{jk}T^{ik}$.

Exercise 2.4.

(a) Show that $\flat_{1,2} \colon \mathcal{T}^2{}_0(V) \to \mathcal{T}^0{}_2(V)$ given by $T_{\flat_{1,2}}(\boldsymbol{v}, \boldsymbol{w}) = T(\boldsymbol{v}_\flat, \boldsymbol{w}_\flat)$ is an isomorphism.

(b) In case you have already not checked this while doing item (a), verify also that $T_{ij} = g_{ik}g_{j\ell}T^{k\ell}$.

With that in hands, we could think that we have exhausted all the possible identifications between order 2 tensors. But let's not forget about tensors with scrambled domains:

Proposition 2.6. *The map* $\sharp_1{}_{\flat_1} : \mathcal{T}^1_1(V) \to \mathcal{T}_1{}^1(V)$ *given by*

$$T^{\sharp_1}{}_{\flat_1}(f, \boldsymbol{v}) = T(f^\sharp, \boldsymbol{v}_\flat)$$

is an isomorphism.

Proof: Like in Proposition 2.5 (p. 17), it suffices to show that $\sharp_1{}_{\flat_1}$ is injective. Let's do this again using a basis $\mathcal{B} = (\boldsymbol{e}_i)_{i=1}^n$ for V. We have:

$$T^i{}_j = T^{\sharp_1}{}_{\flat_1}(\boldsymbol{e}^i, \boldsymbol{e}_j) = T((\boldsymbol{e}^i)^\sharp, (\boldsymbol{e}_j)_\flat) = T(g^{ik}\boldsymbol{e}_k, g_{j\ell}\boldsymbol{e}^\ell) = g^{ik}g_{j\ell}T_k{}^\ell.$$

If $T^{\sharp_1}{}_{\flat_1} = 0$, then $g^{ik}g_{j\ell}T_k{}^\ell = 0$. Multiplying all by $g_{pi}g^{qj}$, we obtain

$$0 = g_{pi}g^{qj}g^{ik}g_{j\ell}T_k{}^\ell = \delta^k{}_p\delta^q{}_\ell T_k{}^\ell = T_p{}^q,$$

and the arbitrariety of indices gives us that $T = 0$. ☐

Remark.

- In the previous proof, multiplying everything by $g_{pi}g^{qj}$ was not just some *deus ex machina* – this is precisely "what was missing" to make the necessary Kronecker deltas appear, to conclude the argument, also being careful to not repeat dummy indices (i.e., that is why we needed new indices p and q). This is hopefully clear, in case you did exercises 2.3 and 2.4 (p. 17).

- All the facets of a given tensor end up being represented by the same kernel letter, in our case, T. In which space T lives depends on the position of the indices (alternatively, on its arguments), and on the context.

- Raising and lowering indices of the Kronecker delta and the Levi-Civita symbol by using a positive-definite inner product, we conclude the numerical equalities $\delta^i{}_j = \delta_{ij} = \delta^{ij} = \delta_j{}^i$, as well as $\epsilon^i{}_{jk} = \epsilon_{ijk} = \epsilon^{ijk}$ for the Levi-Civita symbol, justifying the notation adopted in Exercise 1.11 (p. 12). For example, $\epsilon^i{}_{jk} = \delta^{i\ell}\epsilon_{\ell jk}$.

Let's then register the general case:

Theorem 2.7. *If* $r + s = r' + s'$*, then we have a natural identification* $\mathcal{T}^r{}_s(V) \cong \mathcal{T}^{r'}{}_{s'}(V)$.

Remark. There are plenty of isomorphisms between those spaces. For example, one could lower all the contravariant indices, obtaining isomorphisms of both spaces with $\mathcal{T}^0{}_{r+s}(V)$.

In the previous section, we have seen a generalization of the concept of trace: the contraction between contravariant and covariant indices. Combining that with musical isomorphisms will allow us to define contractions between indices of the same type, such as $\operatorname{tr}^{a,b}$ e $\operatorname{tr}_{a,b}$. Let's see next how to do this, again starting with low order cases to gain intuition:

Definition 2.8.

(i) The *covariant contraction* is the map $\operatorname{tr}_{1,2}\colon \mathcal{T}^0_2(V) \to \mathbb{R}$ given by

$$\operatorname{tr}_{1,2}(T) \doteq \operatorname{tr}^1_1(T^{\sharp_1}),$$

where $\sharp_1\colon \mathcal{T}^0_2(V) \to \mathcal{T}^1_1(V)$ is the isomorphism given in Proposition 2.5 (p. 17).

(ii) The *contravariant contraction* is the map $\operatorname{tr}^{1,2}\colon \mathcal{T}^2_0(V) \to \mathbb{R}$ given by

$$\operatorname{tr}^{1,2}(T) \doteq \operatorname{tr}^1_1(T_{\flat_1}),$$

where $\flat_1\colon \mathcal{T}^2_0(V) \to \mathcal{T}^1_1(V)$ is the isomorphism given in Exercise 2.3 (p. 17).

Exercise 2.5. Show that if T is a order 2 tensor and $\mathcal{B} = (e_i)_{i=1}^n$ is a basis for V, then $T^i_{\ i} = T_i^{\ i}$.

Remark. In general $T^i_{\ j} \neq T_i^{\ j}$. This exercise shows that we could also have "transferred" all the situation to $\mathcal{T}_1^{\ 1}(V)$ instead and used the map $\overline{\operatorname{tr}}^1_1$, briefly mentioned in Section 1.

Proposition 2.9. *Let $\mathcal{B} = (e_i)_{i=1}^n$ be a basis for V.*

(a) If $T \in \mathcal{T}^0_2(V)$, then $\operatorname{tr}_{1,2}(T) = g^{ij}T_{ij}$.

(b) If $T \in \mathcal{T}^2_0(V)$, then $\operatorname{tr}_{1,2}(T) = g_{ij}T^{ij}$.

Proof: Note that $T^i_{\ i} = g^{ij}T_{ij} = g_{ij}T^{ij}$. □

Corollary 2.10. *Let $T \in \operatorname{Lin}(V)$ be a linear operator. Defining a tensor $\widetilde{T} \in \mathcal{T}^0_2(V)$ by $\widetilde{T}(x, y) = \langle T(x), y\rangle$, we have that $\operatorname{tr}(T) = \operatorname{tr}_{1,2}(\widetilde{T})$.*

Remark. The content of this corollary is the fact that $\operatorname{tr}(T)$ does not depend on the choice of scalar product. In other words, another process to compute the trace of T is to *choose* some scalar product $\langle\cdot,\cdot\rangle$ on V, define the associated $\widetilde{T}$, and then apply the contraction $\operatorname{tr}_{1,2}$.

Exercise 2.6. Complete the details of the proof of Proposition 2.9 (p. 19), in case you're not convinced. Also show the previous corollary.

Example 2.11. Suppose that $\mathcal{B} = (e_i)_{i=1}^n$ is an *orthonormal* basis for V. When $\langle\cdot,\cdot\rangle$ is not necessarily positive-definite, "orthonormal" means that $\langle e_i, e_j\rangle = 0$ if $i \neq j$, and for every $1 \leq i \leq n$ we have $\epsilon_i \doteq \langle e_i, e_i\rangle \in \{-1, 1\}$. That is, in matrix form[4] we have $(g_{ij})_{i,j=1}^n = \operatorname{diag}(\epsilon_1, \ldots, \epsilon_n)$ and, in particular, $(g^{ij})_{i,j=1}^n = (g_{ij})_{i,j=1}^n$. It follows that:

[4] The quantity of negative ϵ_i's is the same for every orthonormal basis of V. This (perhaps non-trivial) result is known as *Sylvester's Law of Inertia*).

1. If $T \in \mathscr{T}^0_2(V)$, then $\operatorname{tr}_{1,2}(T) = \sum_{i=1}^n \epsilon_i T(\boldsymbol{e}_i, \boldsymbol{e}_i)$.

2. If $T \in \mathscr{T}^2_0(V)$, then $\operatorname{tr}^{1,2}(T) = \sum_{i=1}^n \epsilon_i T(\boldsymbol{e}^i, \boldsymbol{e}^i)$.

Recall that the generalized trace $\operatorname{tr}^a_b \colon \mathscr{T}^r{}_s(V) \to \mathscr{T}^{r-1}_{s-1}(V)$ was a map reducing the order of a given tensor by 2. This will remain true here:

Definition 2.12. Let $a \leq b$ be non-negative integers.

(i) Let $s \geq 2$. The covariant contraction on the indices a and b is the map $\operatorname{tr}_{a,b} \colon \mathscr{T}^r{}_s(V) \to \mathscr{T}^r{}_{s-2}(V)$ defined by:

$$\begin{aligned}\operatorname{tr}_{a,b}(T)(f^1, \ldots, f^r, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_{s-2}) &\doteq \\ &\doteq \operatorname{tr}_{1,2}\big(T(f^1, \ldots, f^r, \boldsymbol{v}_1, \ldots, \bullet, \ldots, \bullet, \ldots, \boldsymbol{v}_{s-2})\big),\end{aligned}$$

where the $\bullet$s are placed in the a-th and b-th covariant slots.

(ii) Let $r \geq 2$. The contravariant contraction on the indices a and b is the map $\operatorname{tr}^{a,b} \colon \mathscr{T}^r{}_s(V) \to \mathscr{T}^{r-2}_s(V)$ defined by:

$$\begin{aligned}\operatorname{tr}^{a,b}(T)(f^1, \ldots, f^{r-2}, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_s) &\doteq \\ &\doteq \operatorname{tr}^{1,2}\big(T(f^1, \ldots, \bullet, \ldots, \bullet, \ldots, f^{r-2}, \boldsymbol{v}_1, \ldots, , \boldsymbol{v}_s)\big),\end{aligned}$$

where the $\bullet$s are places in the a-th and b-th contravariant slots.

Example 2.13. Fix a basis $\mathscr{B} = (\boldsymbol{e}_i)_{i=1}^n$ for V.

(1) Since $\langle \cdot, \cdot \rangle \in \mathscr{T}^0_2(V)$, it makes sense to compute $\operatorname{tr}_{1,2}(\langle \cdot, \cdot \rangle)$. Proposition 2.9 (p. 19) gives us that $\operatorname{tr}_{1,2}(\langle \cdot, \cdot \rangle) = g^{ij} g_{ij} = g^{ji} g_{ij} = \delta^j{}_j = n$.

(2) If $\det \in \mathscr{T}^0_n(\mathbb{R}^n)$ and $\langle \cdot, \cdot \rangle$ is the standard inner product in $\mathbb{R}^n$, then the trace $\operatorname{tr}_{a,b}(\det)$ is the zero tensor, for all choices of a and b, as det is totally skew-symmetric.

(3) If $\boldsymbol{v}, \boldsymbol{w} \in V$, then $\operatorname{tr}^{1,2}(\boldsymbol{v} \otimes \boldsymbol{w}) = \langle \boldsymbol{v}, \boldsymbol{w} \rangle$. Indeed, by Proposition 2.9 we have

$$\operatorname{tr}^{1,2}(\boldsymbol{v} \otimes \boldsymbol{w}) = g_{ij}(\boldsymbol{v} \otimes \boldsymbol{w})^{ij} = g_{ij} v^i w^j = \langle \boldsymbol{v}, \boldsymbol{w} \rangle.$$

(4) Similarly to the previous item, if $f, g \in V^*$ are covectors, we have the relation $\operatorname{tr}_{1,2}(f \otimes g) = \langle f^\sharp, g^\sharp \rangle$. To wit, in coordinates we have:

$$\begin{aligned}\operatorname{tr}_{1,2}(f \otimes g) &= g^{ij}(f \otimes g)_{ij} = g^{ij} f_i g_j \\ &= g^{ij} g_{ik} f^k g_{j\ell} g^\ell = \delta^i{}_\ell g_{ik} f^k g^\ell \\ &= g_{k\ell} f^k g^\ell = \langle f^\sharp, g^\sharp \rangle.\end{aligned}$$

Exercise 2.7. Since we mentioned the quantity $\langle f^\sharp, g^\sharp\rangle$, let's define a new product $\langle\cdot,\cdot\rangle^*\colon V^*\times V^*\to\mathbb{R}$ by $\langle f,g\rangle^* \doteq \langle f^\sharp, g^\sharp\rangle$. Show that $\langle\cdot,\cdot\rangle^*$ is a non-degenerate scalar product in V^*, which is positive-definite if $\langle\cdot,\cdot\rangle$ is. What are its components, in terms of the $(g_{ij})_{i,j=1}^n$?

Note that $\mathscr{T}^0_2(V^*) = \mathscr{T}^2_0(V)$, so who we call "vectors" or "covectors" actually depends on which space we started with.

(5) If $W\in\mathscr{T}^2_3(V)$ is given (in terms of a basis) by $W = W^{ij}{}_{k\ell m}\boldsymbol{e}_i\otimes\boldsymbol{e}_j\otimes\boldsymbol{e}^k\otimes\boldsymbol{e}^\ell\otimes\boldsymbol{e}^m$, like in Example 1.12 (p. 13), we have that $\operatorname{tr}^{1,2}(W)\in\mathscr{T}^0_3(V)$ is given by

$$\operatorname{tr}^{1,2}(W) = g_{ij}W^{ij}{}_{k\ell m}\boldsymbol{e}^k\otimes\boldsymbol{e}^\ell\otimes\boldsymbol{e}^m.$$

We can also compute $\operatorname{tr}_{1,3}(W)\in\mathscr{T}^2_1(V)$:

$$\operatorname{tr}_{1,3}(W) = g^{km}W^{ij}{}_{k\ell m}\boldsymbol{e}_i\otimes\boldsymbol{e}_j\otimes\boldsymbol{e}^\ell.$$

Exercise 2.8. What is $\operatorname{tr}_{1,2}(W)\in\mathscr{T}^2_1(V)$?

In a similar way done in Proposition 1.13 (p. 14), we register the coordinate expressions for the general contractions:

Proposition 2.14. *Let $\mathscr{B} = (\boldsymbol{e}_i)_{i=1}^n$ be a basis for V and $T\in\mathscr{T}^r_s(V)$.*

(i) If $s\geq 2$, then

$$(\operatorname{tr}_{a,b}(T))^{i_1\ldots i_r}{}_{j_1\ldots j_{s-2}} = g^{k\ell}T^{i_1\ldots i_r}{}_{j_1\ldots k\ldots\ell\ldots j_{s-2}}.$$

(ii) If $r\geq 2$, then

$$(\operatorname{tr}^{a,b}(T))^{i_1\ldots i_{r-2}}{}_{j_1\ldots j_s} = g_{k\ell}T^{i_1\ldots k\ldots\ell\ldots i_r}{}_{j_1\ldots j_s}.$$

The indices indicated in red on T are in the a-th and b-th slots.

The next two exercises have the goal to make sure you got used enough to the sort of manipulations we have been doing so far:

Exercise 2.9 (A taste of Geometry). A tensor $R\in\mathscr{T}^0_4(V)$ is called *curvature-like* if it satisfies

(i) $R(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{w}) = -R(\boldsymbol{y},\boldsymbol{x},\boldsymbol{z},\boldsymbol{w}) = -R(\boldsymbol{x},\boldsymbol{y},\boldsymbol{w},\boldsymbol{z})$;

(ii) $R(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{w}) = R(\boldsymbol{z},\boldsymbol{w},\boldsymbol{x},\boldsymbol{y})$.

(iii) $R(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\cdot) + R(\boldsymbol{y},\boldsymbol{z},\boldsymbol{x},\cdot) + R(\boldsymbol{z},\boldsymbol{x},\boldsymbol{y},\cdot) = 0$;

Suppose that V has a scalar product $\langle\cdot,\cdot\rangle$.

(a) Show that:

$$\begin{aligned}&\operatorname{tr}_{1,2}(R) = \operatorname{tr}_{3,4}(R) = 0,\\ &\operatorname{tr}_{2,4}(R) = \operatorname{tr}_{1,3}(R) \quad \text{and}\\ &\operatorname{tr}_{1,4}(R) = \operatorname{tr}_{2,3}(R) = -\operatorname{tr}_{1,3}(R).\end{aligned}$$

That is, it suffices to know $\operatorname{tr}_{1,3}(R)$ to actually know all possible contractions of R.

Suggestion. How are symmetries (i) and (ii) described in coordinates?

(b) Show that $R^0\colon V^4 \to \mathbb{R}$ defined by

$$R^0(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{w}) \doteq \langle \boldsymbol{y},\boldsymbol{z}\rangle\langle \boldsymbol{x},\boldsymbol{w}\rangle - \langle \boldsymbol{x},\boldsymbol{z}\rangle\langle \boldsymbol{y},\boldsymbol{w}\rangle$$

is a curvature-like tensor. It is called the *fundamental curvature of* $\langle\cdot,\cdot\rangle$.

(c) If R is a curvature-like tensor, since $\langle\cdot,\cdot\rangle$ is non-degenerate, it is well defined a map $R\colon V^3 \to V$ such that $R(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{w}) = \langle R(\boldsymbol{x},\boldsymbol{y})\boldsymbol{z},\boldsymbol{w}\rangle$. The *Ricci tensor* associated to R is $\operatorname{Ric} \in \mathcal{T}^0_2(V)$, given by

$$\operatorname{Ric}(\boldsymbol{x},\boldsymbol{y}) \doteq \operatorname{tr}\,(R(\cdot,\boldsymbol{x})\boldsymbol{y}).$$

Show that for the fundamental curvature R^0 of $\langle\cdot,\cdot\rangle$, the relation

$$\operatorname{Ric}^0(\boldsymbol{x},\boldsymbol{y}) = (n-1)\langle \boldsymbol{x},\boldsymbol{y}\rangle$$

holds.

Remark. It is usual in Geometry to write simply $R(\boldsymbol{x},\boldsymbol{y})\boldsymbol{z}$ instead of $R(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})$. This is not a mistake.

(d) The *scalar curvature* associated to R is defined by $S \doteq \operatorname{tr}_{1,2}(\operatorname{Ric})$. For the fundamental curvature R^0 of $\langle\cdot,\cdot\rangle$, use the previous item and conclude that $S^0 = n(n-1)$.

Remark. Some curiosities:

- In the definition of a curvature-like tensor, actually symmetries (i) and (iii) together imply symmetry (ii)! That is, condition (ii) is superfluous. The standard argument for this is known as "Milnor's octahedron".
- Moreover, it is possible to show that the subspace $\mathcal{R}(V) \subseteq \mathcal{T}^0_4(V)$ formed by the curvature-like tensors on V has dimension $\dim \mathcal{R}(V) = n^2(n^2-1)/12$.

Exercise 2.10. Let $T, S \in \mathcal{T}^0_2(V)$. The *Kulkarni-Nomizu product* of T and S is defined as

$$\begin{aligned}&(T \owedge S)(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{w}) \doteq\\ &\qquad \doteq T(\boldsymbol{x},\boldsymbol{z})S(\boldsymbol{y},\boldsymbol{w}) + T(\boldsymbol{y},\boldsymbol{w})S(\boldsymbol{x},\boldsymbol{z}) - T(\boldsymbol{x},\boldsymbol{w})S(\boldsymbol{y},\boldsymbol{z}) - T(\boldsymbol{y},\boldsymbol{z})S(\boldsymbol{x},\boldsymbol{w}).\end{aligned}$$

(a) Verify that $T \owedge S \in \mathscr{T}^0_4(V)$ and that it satisfies symmetries (i) and (ii) from the definition of a curvature-like tensor.

(b) Show that if T and S are symmetric, then $T \owedge S$ also satisfies symmetry (iii), and thus is a curvature-like tensor.

(c) Suppose that V has a scalar product $\langle \cdot, \cdot \rangle$. Show that

$$\operatorname{tr}_{1,3}(T \owedge S) = \operatorname{tr}_{1,2}(T)S + \operatorname{tr}_{1,2}(S)T - \operatorname{tr}_{2,3}(S \otimes T) - \operatorname{tr}_{2,3}(T \otimes S).$$

(d) Suppose that V has a scalar product $\langle \cdot, \cdot \rangle$. Show that

$$\langle \cdot, \cdot \rangle \owedge \langle \cdot, \cdot \rangle = -2R^0,$$

where R^0 is the fundamental curvature of $\langle \cdot, \cdot \rangle$.

Before we proceed to Section 3 to discuss some Algebra, let's see how to use the language of tensors to establish in a simple way some formulas with the cross product in $\mathbb{R}^3$. Suppose until the end of this section that $V = \mathbb{R}^3$ and that $\langle \cdot, \cdot \rangle$ is the standard inner product. We begin with a coordinate-free definition of cross product:

Definition 2.15. Let $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^3$. The *cross product* of $\boldsymbol{v}$ and $\boldsymbol{w}$ is the unique vector $\boldsymbol{v} \times \boldsymbol{w} \in \mathbb{R}^3$ such that $\langle \boldsymbol{v} \times \boldsymbol{w}, \boldsymbol{x} \rangle = \det(\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{x})$, for all $\boldsymbol{x} \in \mathbb{R}^3$.

Lemma 2.16. *Let $\mathscr{B} = (\boldsymbol{e}_i)_{i=1}^3$ be a positive and orthonormal basis for $\mathbb{R}^3$. So we have $(\boldsymbol{v} \times \boldsymbol{w})^i = \epsilon^i{}_{jk} v^j w^k$, where $\epsilon^i{}_{jk}$ is the Levi-Civita symbol (seen in Example 1.4, p. 5 and in Exercise 1.11, p. 12) .*

Proof: Making $\boldsymbol{x} = \boldsymbol{e}_i$ in the definition of $\boldsymbol{v} \times \boldsymbol{w}$, on one side we have

$$\langle \boldsymbol{v} \times \boldsymbol{w}, \boldsymbol{e}_i \rangle = \langle (\boldsymbol{v} \times \boldsymbol{w})^j \boldsymbol{e}_j, \boldsymbol{e}_i \rangle = (\boldsymbol{v} \times \boldsymbol{w})^j \delta_{ji} = (\boldsymbol{v} \times \boldsymbol{w})_i.$$

On the other side:

$$\det(\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{e}_i) = \det(v^j \boldsymbol{e}_j, w^k \boldsymbol{e}_k, \boldsymbol{e}_i) = v^j w^k \epsilon_{jki} = \epsilon_{ijk} v^j w^k.$$

Thus $(\boldsymbol{v} \times \boldsymbol{w})_i = \epsilon_{ijk} v^j w^k$. Raising the index i on both sides, we conclude the proof. □

Remark. With this you can write the expression for $\boldsymbol{v} \times \boldsymbol{w}$ explicitly and convince yourself that the definition given here coincides with the definition you learned as a child.

An example of application for this expression is the following:

Proposition 2.17. *Let $\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{z} \in \mathbb{R}^3$. Then*

$$(\boldsymbol{v} \times \boldsymbol{w}) \times \boldsymbol{z} = \langle \boldsymbol{z}, \boldsymbol{v} \rangle \boldsymbol{w} - \langle \boldsymbol{z}, \boldsymbol{w} \rangle \boldsymbol{v}.$$

Proof: We can proceed at the coordinate level, using a positive and orthonormal basis for $\mathbb{R}^3$ and Exercise 1.11 (p. 12). We have:

$$\begin{aligned}((\boldsymbol{v}\times\boldsymbol{w})\times\boldsymbol{z})^i &= \epsilon^i{}_{jk}(\boldsymbol{v}\times\boldsymbol{w})^j z^k \\ &= \epsilon^i{}_{jk}\epsilon^j{}_{\ell m}v^\ell w^m z^k \\ &= \epsilon^i{}_{kj}\epsilon^j{}_{m\ell}v^\ell w^m z^k \\ &= (\delta^i{}_m\delta_{k\ell} - \delta^i{}_\ell\delta_{km})v^\ell w^m z^k \\ &= \delta^i{}_m\delta_{k\ell}v^\ell w^m z^k - \delta^i{}_\ell\delta_{km}v^\ell w^m z^k \\ &= \delta_{k\ell}z^k v^\ell w^i - \delta_{km}z^k w^m v^i \\ &= \langle \boldsymbol{z},\boldsymbol{v}\rangle w^i - \langle \boldsymbol{z},\boldsymbol{w}\rangle v^i,\end{aligned}$$

as wanted. □

Corollary 2.18 (Jacobi identity)**.** *Let $\boldsymbol{v},\boldsymbol{w},\boldsymbol{z}\in\mathbb{R}^3$. Then*

$$(\boldsymbol{v}\times\boldsymbol{w})\times\boldsymbol{z} + (\boldsymbol{w}\times\boldsymbol{z})\times\boldsymbol{v} + (\boldsymbol{z}\times\boldsymbol{v})\times\boldsymbol{w} = \boldsymbol{0}.$$

Exercise 2.11. Show the Jacobi identity.

Such manipulations also allow us to establish some identities regarding differential operators, such as gradient, curl and divergence. If can $= (\boldsymbol{e}_i)_{i=1}^3$ is the standard basis and $\nabla = (\partial_1,\partial_2,\partial_3)$ is the usual derivative vector, recall that if $\varphi\colon\mathbb{R}^3\to\mathbb{R}$ is smooth and $\boldsymbol{F}\colon\mathbb{R}^3\to\mathbb{R}^3$ is a smooth vector field in $\mathbb{R}^3$, then

$$\begin{aligned}\operatorname{grad}\varphi &= \nabla\varphi = (\partial_1\varphi,\partial_2\varphi,\partial_3\varphi), \\ \operatorname{div}\boldsymbol{F} &= \langle\nabla,\boldsymbol{F}\rangle = \partial_1F^1+\partial_2F^2+\partial_3F^3 \quad \text{e} \\ \operatorname{curl}\boldsymbol{F} &= \nabla\times\boldsymbol{F} = \epsilon^{ij}{}_k\partial_jF^k\boldsymbol{e}_i.\end{aligned}$$

Proposition 2.19. *Let $\varphi\colon\mathbb{R}^3\to\mathbb{R}$ be smooth. Then* $\operatorname{curl}\operatorname{grad}\varphi = \boldsymbol{0}$.

Proof: It suffices to note that $(\operatorname{curl}\operatorname{grad}\varphi)^i = \epsilon^{ij}{}_k\partial_j\partial_k\varphi$. Since the indices j and k are dummy, we have that

$$\epsilon^{ij}{}_k\partial_j\partial_k\varphi = \epsilon^{ik}{}_j\partial_k\partial_j\varphi = \epsilon^{ik}{}_j\partial_j\partial_k\varphi,$$

using in the last step that second order partial derivatives commute. On the other hand, since the Levi-Civita symbol is skew-symmetric, we get

$$\epsilon^{ij}{}_k\partial_j\partial_k\varphi = -\epsilon^{ik}{}_j\partial_j\partial_k\varphi.$$

Thus $\epsilon^{ij}{}_k\partial_j\partial_k\varphi = 0$. □

Exercise 2.12. Make a similar argument (paying attention to dummy indices) and show that if $\boldsymbol{F}\colon\mathbb{R}^3\to\mathbb{R}^3$ is a smooth vector field in $\mathbb{R}^3$, then $\operatorname{div}\operatorname{curl}\boldsymbol{F} = 0$.

Proposition 2.20. *Let $\boldsymbol{F}\colon \mathbb{R}^3 \to \mathbb{R}^3$ be a smooth vector field in $\mathbb{R}^3$. Then*

$$\operatorname{curl}\operatorname{curl}\boldsymbol{F} = \operatorname{grad}(\operatorname{div}\boldsymbol{F}) - \nabla^2\boldsymbol{F},$$

where $\nabla^2\boldsymbol{F}$ denotes the (vector) Laplacian of $\boldsymbol{F}$.

Proof: Let's proceed just like in Proposition 2.17 (p. 23), using the identity seen in Exercise 1.11 (p. 12) with the correct index balance. We have that:

$$\begin{aligned}
(\operatorname{curl}\operatorname{curl}\boldsymbol{F})^i &= \epsilon^{ij}{}_k\partial_j(\operatorname{rot}\boldsymbol{F})^k \\
&= \epsilon^{ij}{}_k\partial_j\epsilon^{k\ell}{}_m\partial_\ell F^m \\
&= \epsilon^{ij}{}_k\epsilon^{k\ell}{}_m\partial_j\partial_\ell F^m \\
&= (\delta^{i\ell}\delta^j{}_m - \delta^i{}_m\delta^{j\ell})\partial_j\partial_\ell F^m \\
&= \delta^{i\ell}\delta^j{}_m\partial_j\partial_\ell F^m - \delta^i{}_m\delta^{j\ell}\partial_j\partial_\ell F^m \\
&= \delta^{i\ell}\partial_j\partial_\ell F^j - \delta^{j\ell}\partial_j\partial_\ell F^i \\
&= \delta^{i\ell}\partial_\ell(\operatorname{div}\boldsymbol{F}) - \delta^{j\ell}\partial_j\partial_\ell F^i \\
&= (\operatorname{grad}(\operatorname{div}\boldsymbol{F}))^i - (\nabla^2\boldsymbol{F})^i \\
&= (\operatorname{grad}(\operatorname{div}\boldsymbol{F}) - \nabla^2\boldsymbol{F})^i,
\end{aligned}$$

as wanted. $\square$

Exercise 2.13. Let $\boldsymbol{F}, \boldsymbol{G}\colon \mathbb{R}^3 \to \mathbb{R}^3$ be smooth vector fields in $\mathbb{R}^3$.

(a) Show that $\operatorname{div}(\boldsymbol{F}\times\boldsymbol{G}) = \langle\operatorname{curl}\boldsymbol{F}, \boldsymbol{G}\rangle - \langle\operatorname{curl}\boldsymbol{G}, \boldsymbol{F}\rangle$.

(b) Show that

$$\operatorname{curl}(\boldsymbol{F}\times\boldsymbol{G}) = \langle\boldsymbol{G},\nabla\rangle\boldsymbol{F} + (\operatorname{div}\boldsymbol{G})\boldsymbol{F} - \langle\boldsymbol{F},\nabla\rangle\boldsymbol{G} - (\operatorname{div}\boldsymbol{F})\boldsymbol{G},$$

where $\langle\boldsymbol{F},\nabla\rangle\boldsymbol{G}$ denotes the differential operator $F^1\partial_1 + F^2\partial_2 + F^3\partial_3$ acting componentwise in $\boldsymbol{G}$, etc..

(c) Show that

$$\operatorname{grad}\langle\boldsymbol{F},\boldsymbol{G}\rangle = \langle\boldsymbol{G},\nabla\rangle\boldsymbol{F} + \langle\boldsymbol{F},\nabla\rangle\boldsymbol{G} - (\operatorname{curl}\boldsymbol{F})\times\boldsymbol{G} - (\operatorname{curl}\boldsymbol{G})\times\boldsymbol{F}.$$

3 The universal property

There is another approach to tensors, usually prefered by algebraist, in Mathematics. Let's briefly discuss it and relate it with everything we have done so far. Let V and W be two finite-dimensional vector spaces, until the end of the text.

Definition 3.1. A *tensor product* of V and W is a pair $(\mathscr{T}, \overline{\otimes})$, where $\mathscr{T}$ is a vector space and $\overline{\otimes}\colon V \times W \to \mathscr{T}$ is a bilinear map satisfying the following *universal property*: given any real vector space Z and any bilinear map $B\colon V \times W \to Z$, there is a unique linear map $\hat{B}\colon \mathscr{T} \to Z$ such that $\hat{B} \circ \overline{\otimes} = B$. In other words, the following diagram can always be uniquely completed:

$$\begin{tikzcd}
\mathscr{T} \arrow[dr, dashed, "\hat{B}"] & \\
V \times W \arrow[u, "\overline{\otimes}"] \arrow[r, "B"'] \arrow[ur, phantom, "\circlearrowleft"] & Z
\end{tikzcd}$$

Figure 1: The universal property of $(\mathscr{T}, \overline{\otimes})$.

Remark.

- Usually one writes $\otimes$ instead of $\overline{\otimes}$. We use the notation $\overline{\otimes}$ for pedagogical reasons, until we establish the necessay isomorphisms to identify $\overline{\otimes}$ with the actual operation $\otimes$ studied in the previous sections.
- In this setting, we'll say that $\hat{B}$ is a *linearization* of B (via $\overline{\otimes}$).

That is, a tensor product of V and W is a vector space equipped with a map $\overline{\otimes}$ which universally linearizes all bilinear maps defined in $V \times W$. Such tensor product actually acts as a "translator", converting bilinear maps B into linear maps $\hat{B}$ which are, in a certain sense, equivalents to B, and $\overline{\otimes}$ is the dictionary.

The issue with this definition is that it is not clear, a priori, whether a tensor product between two given spaces even exists to begin with. Or if it is unique. As it is frequent in Mathematics, verifying uniqueness is easier than verifying the existence:

Proposition 3.2. *Let $(\mathscr{T}_1, \overline{\otimes}_1)$ and $(\mathscr{T}_2, \overline{\otimes}_2)$ be two tensor products of V and W. Then there exists a linear isomorphism $\Phi\colon \mathscr{T}_1 \to \mathscr{T}_2$ such that $\Phi \circ \overline{\otimes}_1 = \overline{\otimes}_2$.*

$$\begin{tikzcd} \mathscr{T}_1 \arrow[rr, "\Phi"] & & \mathscr{T}_2 \\ & V \times W \arrow[ul, "\overline{\otimes}_1"] \arrow[ur, "\overline{\otimes}_2"'] \arrow[u, phantom, "\circlearrowleft"] & \end{tikzcd}$$

Figure 2: Uniqueness of the tensor product up to isomorphism.

Proof: Since $\overline{\otimes}_2 \colon V \times W \to \mathscr{T}_2$ is bilinear, the universal property of $(\mathscr{T}_1, \overline{\otimes}_1)$ will give us a unique linear map $\Phi \colon \mathscr{T}_1 \to \mathscr{T}_2$ such that $\Phi \circ \overline{\otimes}_1 = \overline{\otimes}_2$.

$$\begin{tikzcd} \mathscr{T}_1 \arrow[dr, dashed, "\Phi"] & \\ V \times W \arrow[u, "\overline{\otimes}_1"] \arrow[r, "\overline{\otimes}_2"'] \arrow[ur, phantom, "\circlearrowleft"] & \mathscr{T}_2 \end{tikzcd}$$

Figure 3: The existence of Φ.

To see that this Φ is already the desired isomorphism, we need to exhibit its inverse, running the same proof "backwards": since $\overline{\otimes}_1 \colon V \times W \to \mathscr{T}_1$ is bilinear, the universal property of $(\mathscr{T}_2, \overline{\otimes}_2)$ gives us a unique linear map $\Psi \colon \mathscr{T}_2 \to \mathscr{T}_1$ such that $\Psi \circ \overline{\otimes}_2 = \overline{\otimes}_1$.

$$\begin{tikzcd} \mathscr{T}_2 \arrow[dr, dashed, "\Psi"] & \\ V \times W \arrow[u, "\overline{\otimes}_2"] \arrow[r, "\overline{\otimes}_1"'] \arrow[ur, phantom, "\circlearrowleft"] & \mathscr{T}_1 \end{tikzcd}$$

Figure 4: The existence of Ψ.

Now we have to check that this Ψ is indeed the inverse of Φ. For this end, we'll use the uniqueness part of those universal properties. Let's explore the universal property of $(\mathscr{T}_1, \overline{\otimes}_1)$ for the bilinear map $\overline{\otimes}_1 \colon V \times W \to \mathscr{T}_1$. Clearly $\mathrm{Id}_{\mathscr{T}_1} \colon \mathscr{T}_1 \to \mathscr{T}_1$ satisfies $\mathrm{Id}_{\mathscr{T}_1} \circ \overline{\otimes}_1 = \overline{\otimes}_1$, but on the other hand

$$(\Psi \circ \Phi) \circ \overline{\otimes}_1 = \Psi \circ (\Phi \circ \overline{\otimes}_1) = \Psi \circ \overline{\otimes}_2 = \overline{\otimes}_1,$$

so that $\Psi \circ \Phi = \mathrm{Id}_{\mathscr{T}_1}$.

$$\begin{tikzcd}
\mathscr{T}_1 \arrow[dr, dashed, "\mathrm{Id}_{\mathscr{T}_1}"] & \\
V \times W \arrow[u, "\overline{\otimes}_1"] \arrow[r, "\overline{\otimes}_2"'] & \mathscr{T}_1
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathscr{T}_1 \arrow[dr, dashed, "\Psi \circ \Phi"] & \\
V \times W \arrow[u, "\overline{\otimes}_1"] \arrow[r, "\overline{\otimes}_2"'] & \mathscr{T}_1
\end{tikzcd}$$

Figure 5: Proving that $\Psi \circ \Phi = \mathrm{Id}_{\mathscr{T}_1}$.

Similarly, one shows that $\Phi \circ \Psi = \mathrm{Id}_{\mathscr{T}_2}$. $\square$

And for our psychological comfort, we should show the existence of a tensor product of V and W. We'll employ a slightly more general construction:

Example 3.3 (Free vector space)**.** Let S be a non-empty set and consider the collection $\mathscr{F}(S)$ of the functions $f\colon S \to \mathbb{R}$ such that $f(s) \neq 0$ only for finitely many elements $s \in S$. The operations of addition and scalar multiplication in $\mathscr{F}(S)$, defined pointwise, make it a vector space. For each $a \in S$, define $\delta_a\colon S \to \mathbb{R}$ by $\delta_a(x) = 1$ if $x = a$ and 0 otherwise.

Exercise 3.1. Show that $\{\delta_a \mid a \in S\}$ is a basis for $\mathscr{F}(S)$.

Moreover, the map $S \ni s \mapsto \delta_s \in \{\delta_a \mid a \in S\}$ is a bijection, and so we identify a with δ_a, for each $a \in S$. So we have proven that given any non-empty set S, there is a vector space with S as a basis.

To exhibit a tensor product of V and W, we consider the quotient of $\mathscr{F}(V \times W)$ by the subspace $\mathscr{F}_0$ spanned by the elements of the form

$$\begin{gathered}
(\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{w}) - (\boldsymbol{v}_1, \boldsymbol{w}) - (\boldsymbol{v}_2, \boldsymbol{w}), \\
(\boldsymbol{v}, \boldsymbol{w}_1 + \boldsymbol{w}_2) - (\boldsymbol{v}, \boldsymbol{w}_1) - (\boldsymbol{v}, \boldsymbol{w}_2), \\
(\lambda \boldsymbol{v}, \boldsymbol{w}) - \lambda(\boldsymbol{v}, \boldsymbol{w}) \quad \text{e} \\
(\boldsymbol{v}, \lambda \boldsymbol{w}) - \lambda(\boldsymbol{v}, \boldsymbol{w}).
\end{gathered}$$

Suppose that $\overline{\otimes}\colon V \times W \to \mathscr{F}(V \times W)/\mathscr{F}_0$ is the function assigning to each $(\boldsymbol{v}, \boldsymbol{w})$ its class $\boldsymbol{v} \mathbin{\overline{\otimes}} \boldsymbol{w}$ in $\mathscr{F}(V \times W)/\mathscr{F}_0$. Such quotient gives us the relations

$$\begin{aligned}
(\boldsymbol{v}_1 + \boldsymbol{v}_2) \mathbin{\overline{\otimes}} \boldsymbol{w} &= \boldsymbol{v}_1 \mathbin{\overline{\otimes}} \boldsymbol{w} + \boldsymbol{v}_2 \mathbin{\overline{\otimes}} \boldsymbol{w} \\
\boldsymbol{v} \mathbin{\overline{\otimes}} (\boldsymbol{w}_1 + \boldsymbol{w}_2) &= \boldsymbol{v} \mathbin{\overline{\otimes}} \boldsymbol{w}_1 + \boldsymbol{v} \mathbin{\overline{\otimes}} \boldsymbol{w}_2 \\
(\lambda \boldsymbol{v}) \mathbin{\overline{\otimes}} \boldsymbol{w} &= \lambda(\boldsymbol{v} \mathbin{\overline{\otimes}} \boldsymbol{w}) = \boldsymbol{v} \mathbin{\overline{\otimes}} (\lambda \boldsymbol{w}).
\end{aligned}$$

Compare them with the Exercise 1.8 (p. 9). Finally:

Proposition 3.4. *The pair $(\mathscr{F}(V \times W)/\mathscr{F}_0, \overline{\otimes})$ is a tensor product of V and W.*

Proof: By construction, $\overline{\otimes}$ is bilinear. Let Z be a real vector space and take a bilinear map $B\colon V \times W \to Z$. Since $V \times W$ is a basis for $\mathscr{F}(V \times W)$, there is a unique linear extension $\widetilde{B}\colon \mathscr{F}(V \times W) \to Z$ of B. But as B is bilinear, $\widetilde{B}$ vanishes on the elements of $\mathscr{F}_0$, and thus passes to the quotient as a linear map $\hat{B}\colon \mathscr{F}(V \times W)/\mathscr{F}_0 \to Z$, satisfying $\hat{B}(\boldsymbol{v} \overline{\otimes} \boldsymbol{w}) = B(\boldsymbol{v}, \boldsymbol{w})$, as wanted. □

This way, we denote $V \overline{\otimes} W = (\mathscr{F}(V \times W)/\mathscr{F}_0, \overline{\otimes})$. But once the existence of $V \overline{\otimes} W$ is establish, we never need to worry about the its construction again. Before proceeding, we note two important details:

- A generic element of $V \overline{\otimes} W$ is not necessarily of the form $\boldsymbol{v} \overline{\otimes} \boldsymbol{w}$ for certain $\boldsymbol{v} \in V$ and $\boldsymbol{w} \in W$, but it is instead a sum of elements of this form (then called *decomposable*, or *pure tensors*).
- To define linear maps in $V \overline{\otimes} W$, we necessarily have to use the universal property. This is made clear by the quotient construction given above, for example, but the issue is that a given element of $V \overline{\otimes} W$ may be represented in more than one way. In other words, it may be expressed as two sum, of different elements. The universal property ensures that all of those possibilities were considered at the same time.

With this in mind, the identification we're looking for is given in the following:

Proposition 3.5. *The map*

$$B\colon V \times W \to \mathrm{Lin}_2(V^* \times W^*, \mathbb{R})$$

given by $B(\boldsymbol{v}, \boldsymbol{w})(f, g) \doteq f(\boldsymbol{v})g(\boldsymbol{w})$ induces an isomorphism

$$V \overline{\otimes} W \cong \mathrm{Lin}_2(V^* \times W^*),$$

where $\mathrm{Lin}_2(V^* \times W^*, \mathbb{R}) = \{T\colon V^* \times W^* \to \mathbb{R} \mid T \text{ is bilinear}\}$.

Definition 3.6. Let $V_1, \ldots, V_p$ be vector space. A *tensor product* of $V_1, \ldots, V_p$ is a pair $(\mathscr{T}, \otimes)$, where $\mathscr{T}$ is a vector space and

$$\otimes\colon V_1 \times \cdots \times V_p \to \mathscr{T}$$

is a multilinear map satisfying the following *universal property*: given any vector space Z and any multilinear map

$$B\colon V_1 \times \cdots \times V_p \to Z,$$

there is a unique linear map $\hat{B}\colon \mathscr{T} \to Z$ such that $\hat{B} \circ \otimes = B$. In other words, the following diagram can always be uniquely completed:

$$\begin{tikzcd}
\mathcal{T} \arrow[dr, dashed, "\hat{B}"] & \\
V_1 \times \cdots \times V_p \arrow[u, "\otimes"] \arrow[r, "B"'] \arrow[ur, phantom, "\circlearrowright"] & Z
\end{tikzcd}$$

Figure 6: The universal property of $(\mathcal{T}, \overline{\otimes})$, again.

It is the same history as before: a tensor product universally linearizes all multilinear maps, using a single map $\otimes$ as a dictionary. The philosophy behind all the ideas given so far will not change.

Exercise 3.2. Let $V_1, \ldots, V_p$ be vector spaces, and $(\mathcal{T}_1, \otimes_1)$ and $(\mathcal{T}_2, \otimes_2)$ be two tensor products of $V_1, \ldots, V_p$. Show that there is a linear isomorphism $\Phi\colon \mathcal{T}_1 \to \mathcal{T}_2$ such that $\Phi \circ \otimes_1 = \otimes_2$.

Regarding the existence of a tensor product in this setting, we again consider a quotient of the form $\mathcal{F}(V_1 \times \cdots \times V_p)/\mathcal{F}_0$, where $\mathcal{F}_0$ is the subspace spanned by certain elements, to make the quotient projection restricted to $V_1 \times \cdots \times V_p$ multilinear.

Exercise 3.3. Try to describe the subspace $\mathcal{F}_0$ when $p = 3$.

And, as before, one can show that the tensor product of $V_1, \ldots, V_p$ so defined, denoted by $\bigotimes_{i=1}^p V_i$ or $V_1 \otimes \cdots \otimes V_p$, is isomorphic to the space of multilinear maps from $V_1^* \times \cdots \times V_p^*$ to $\mathbb{R}$. In the case where $V_1 = \cdots = V_p = V$, we can simply write $V^{\otimes p} = \bigotimes_{i=1}^p V_i = \bigotimes^p V$. With this notation, we then conclude that

$$\mathcal{T}^r{}_s(V) = V^{\otimes r} \otimes (V^*)^{\otimes s}.$$

Exercise 3.4. Write $\mathcal{T}^1{}_2{}^2(V)$ and $\mathcal{T}_1{}^2{}_3(V)$ as tensor products of V and V^*, in the correct order.

To conclude the discussion, let's see a couple of other applications of the universal property:

Proposition 3.7 (Commutativity)**.** *Let V and W be vector spaces. Then we have that $V \otimes W \cong W \otimes V$.*

Proof: By now, it should be reasonably clear that the map we're looking for is $\hat{B}\colon V \otimes W \to W \otimes V$ given by $\hat{B}(\boldsymbol{v} \otimes \boldsymbol{w}) = \boldsymbol{w} \otimes \boldsymbol{v}$. To define it rigorously, consider $B\colon V \times W \to W \otimes V$ given by $B(\boldsymbol{v}, \boldsymbol{w}) = \boldsymbol{w} \otimes \boldsymbol{v}$. Since B is bilinear, the universal property of $V \otimes W$ will give us $\hat{B}$.

$$\begin{tikzcd}
V \otimes W \arrow[dr, dashed, "\hat{B}"] & \\
V \times W \arrow[u, "\otimes"] \arrow[r, "B"'] \arrow[ur, phantom, "\circlearrowright"] & W \otimes V
\end{tikzcd}$$

Figure 7: Formalizing $\hat{B}(\boldsymbol{v} \otimes \boldsymbol{w}) = \boldsymbol{w} \otimes \boldsymbol{v}$.

To show that $\hat{B}$ is an isomorphism, we repeat the same argument "backwards" to define its inverse. $\square$

Exercise 3.5. Define formally the inverse of $\hat{B}$.

Suggestion. Do not forget to use the uniqueness of linearizations given by the universal property to ensure that the inverse you have defined indeed works, just like what we have done in Proposition 3.2 (p. 26).

$$\begin{tikzcd}
W \otimes V \arrow[dr, dashed] & \\
W \times V \arrow[u, "\otimes"] \arrow[r] \arrow[ur, phantom, "\circlearrowright"] & V \otimes W
\end{tikzcd}$$

Figure 8: Hint.

We also have the:

Proposition 3.8 (Associativity). *Let V_1, V_2 and V_3 be vector spaces. Then*

$$(V_1 \otimes V_2) \otimes V_3 \cong V_1 \otimes V_2 \otimes V_3.$$

Proof: We want to construct the map taking $(\boldsymbol{v}_1 \otimes \boldsymbol{v}_2) \otimes \boldsymbol{v}_3$ to $\boldsymbol{v}_1 \otimes \boldsymbol{v}_2 \otimes \boldsymbol{v}_3$, and the idea for that is to use universal properties "backwards". Fixed $\boldsymbol{v}_3 \in V_3$, define $\Phi_{\boldsymbol{v}_3} \colon V_1 \times V_2 \to V_1 \otimes V_2 \otimes V_3$ by $\Phi_{\boldsymbol{v}_3}(\boldsymbol{v}_1, \boldsymbol{v}_2) \doteq \boldsymbol{v}_1 \otimes \boldsymbol{v}_2 \otimes \boldsymbol{v}_3$. Note that $\Phi_{\boldsymbol{v}_3}$ is bilinear, so the universal property of $V_1 \otimes V_2$ yields a linear map

$$\widehat{\Phi_{\boldsymbol{v}_3}} \colon V_1 \otimes V_2 \to V_1 \otimes V_2 \otimes V_3$$

satisfying $\widehat{\Phi_{\boldsymbol{v}_3}}(\boldsymbol{v}_1 \otimes \boldsymbol{v}_2) = \boldsymbol{v}_1 \otimes \boldsymbol{v}_2 \otimes \boldsymbol{v}_3$.

$$\begin{tikzcd}
V_1 \otimes V_2 \arrow[dr, dashed, "\widehat{\Phi_{v_3}}"] & \\
V_1 \times V_2 \arrow[u, "\otimes"] \arrow[r, "\Phi_{v_3}"'] \arrow[ur, phantom, "\circlearrowright"] & V_1 \otimes V_2 \otimes V_3
\end{tikzcd}$$

Figure 9: The first step.

Thus, it is well-defined the map $\Phi\colon (V_1 \otimes V_2) \times V_3 \to V_1 \otimes V_2 \otimes V_3$ given by

$$\Phi(\boldsymbol{v}_1 \otimes \boldsymbol{v}_2, \boldsymbol{v}_3) \doteq \boldsymbol{v}_1 \otimes \boldsymbol{v}_2 \otimes \boldsymbol{v}_3.$$

But since $\widehat{\Phi_{v_3}}$ is linear, we have that Φ is bilinear, and so the universal property of $(V_1 \otimes V_2) \otimes V_3$ gives a linear map $\widehat{\Phi}\colon (V_1 \otimes V_2) \otimes V_3 \to V_1 \otimes V_2 \otimes V_3$ satisfying

$$\widehat{\Phi}((\boldsymbol{v}_1 \otimes \boldsymbol{v}_2) \otimes \boldsymbol{v}_3) = \boldsymbol{v}_1 \otimes \boldsymbol{v}_2 \otimes \boldsymbol{v}_3.$$

$$\begin{tikzcd}
(V_1 \otimes V_2) \otimes V_3 \arrow[dr, dashed, "\widehat{\Phi}"] & \\
(V_1 \times V_2) \times V_3 \arrow[u, "\otimes"] \arrow[r, "\Phi"'] \arrow[ur, phantom, "\circlearrowright"] & V_1 \otimes V_2 \otimes V_3
\end{tikzcd}$$

Figure 10: Concluding the definition of $\widehat{\Phi}$.

The construction of the inverse map is simpler and requires a single step: define $\Psi\colon V_1 \times V_2 \times V_3 \to (V_1 \otimes V_2) \otimes V_3$ by

$$\Psi(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3) \doteq (\boldsymbol{v}_1 \otimes \boldsymbol{v}_2) \otimes \boldsymbol{v}_3.$$

Since Ψ is trilinear, the universal property of $V_1 \otimes V_2 \otimes V_3$ finally gives us a linear map $\widehat{\Psi}\colon V_1 \otimes V_2 \otimes V_3 \to (V_1 \otimes V_2) \otimes V_3$ satisfying

$$\widehat{\Psi}(\boldsymbol{v}_1 \otimes \boldsymbol{v}_2 \otimes \boldsymbol{v}_3) = (\boldsymbol{v}_1 \otimes \boldsymbol{v}_2) \otimes \boldsymbol{v}_3.$$

$$\begin{tikzcd}
V_1 \otimes V_2 \otimes V_3 \arrow[dr, dashed, "\widehat{\Psi}"] & \\
V_1 \times V_2 \times V_3 \arrow[u, "\otimes"] \arrow[r, "\Psi"'] & (V_1 \otimes V_2) \otimes V_3
\end{tikzcd}$$

Figure 11: Defining the inverse $\widehat{\Psi}$.

Clearly $\widehat{\Phi}$ and $\widehat{\Psi}$ are inverses, which establishes the desired isomorphism. $\square$

Practice some writing:

Exercise 3.6. Let V_1, V_2 and V_3 be vector spaces. Show that

$$V_1 \otimes (V_2 \otimes V_3) \cong V_1 \otimes V_2 \otimes V_3.$$

In general, this associativity holds for the tensor product of finitely many vector spaces. Note that there is no ambiguity then in writing $\mathscr{T}^r{}_s(V) = V^{\otimes r} \otimes (V^*)^{\otimes s}$.

Exercise 3.7. Let V_1, W_1, V_2 and W_2 be vector spaces, and take $T\colon V_1 \to W_1$ and $S\colon V_2 \to W_2$ two linear maps. Show that there is a unique linear transformation $T \otimes S\colon V_1 \otimes V_2 \to W_1 \otimes W_2$ such that

$$(T \otimes S)(\boldsymbol{v}_1 \otimes \boldsymbol{v}_2) = T(\boldsymbol{v}_1) \otimes S(\boldsymbol{v}_2),$$

for all $\boldsymbol{v}_1 \in V_1$ and $\boldsymbol{v}_2 \in V_2$.

Remark.

- This indicates how to generalize the tensor product defined in Section 1 to multilinear maps with codomains more complicated than $\mathbb{R}$. Bear in mind that $\mathbb{R}^{\otimes p} \cong \mathbb{R}$, for all p.

- One can also show that if $\mathscr{B}_1, \mathscr{B}_2, \mathscr{C}_1$ and $\mathscr{C}_2$ are bases for V_1, V_2, W_1 and W_2, respectively (with dimensions n_1, n_2, m_1 and m_2), and $[T]_{\mathscr{B}_1,\mathscr{C}_1} = A = (a^i{}_j)$ and $[S]_{\mathscr{B}_2,\mathscr{C}_2} = B = (b^i{}_j)$, then

$$[T \otimes S]_{\mathscr{B}_1 \otimes \mathscr{B}_2, \mathscr{C}_1 \otimes \mathscr{C}_2} \equiv A \otimes B \doteq \begin{pmatrix} a^1{}_1 B & \cdots & a^1{}_{n_1} B \\ \vdots & \ddots & \vdots \\ a^{m_1}{}_1 B & \cdots & a^{m_1}{}_{n_1} B \end{pmatrix}.$$

  The matrix $A \otimes B$ is called the *Kronecker product* of A and B. Such product has some interesting properties. For example, if A and B are square matrices of orders n and m, respectively, then we have the identity[5]

$$\det(A \otimes B) = (\det A)^m (\det B)^n.$$

[5] Yes, the order of A is the exponent of $\det B$ and vice-versa. This is not a mistake.

There are further universal properties than the one seen for the tensor product here. Let's see an example:

Definition 3.9 (Complexification)**.** Let V be a real vector space. A *complexification of* V is a pair $(V_{\mathbb{C}}, \iota)$, where $V_{\mathbb{C}}$ is a complex vector space and $\iota\colon V \to V_{\mathbb{C}}$ is a $\mathbb{R}$-linear map satisfying the following universal property: given any complex vector space Z and a $\mathbb{R}$-linear map $T\colon V \to Z$, there is a unique $\mathbb{C}$-linear map $T_{\mathbb{C}}\colon V_{\mathbb{C}} \to Z$ such that $T_{\mathbb{C}} \circ \iota = T$. In other words, the following diagram can always be uniquely completed:

$$\begin{tikzcd} V_{\mathbb{C}} \arrow[dr, dashed, "T_{\mathbb{C}}"] & \\ V \arrow[u, "\iota"] \arrow[r, "T"'] \arrow[ur, phantom, "\circlearrowleft"] & Z \end{tikzcd}$$

Figure 12: The universal property of $(V_{\mathbb{C}}, \iota)$.

The next result and its proof should not be surprising by now:

Proposition 3.10. *Let V be a real vector space, and $(V_{\mathbb{C}1}, \iota_1)$ and $(V_{\mathbb{C}2}, \iota_2)$ be two complexifications of V. Then there is a $\mathbb{C}$-linear isomorphism $\Phi\colon V_{\mathbb{C}1} \to V_{\mathbb{C}2}$ such that $\Phi \circ \iota_1 = \iota_2$.*

$$\begin{tikzcd} V_{\mathbb{C}1} \arrow[rr, "\Phi"] & \arrow[d, phantom, "\circlearrowleft"] & V_{\mathbb{C}2} \\ & V \arrow[ul, "\iota_1"] \arrow[ur, "\iota_2"'] & \end{tikzcd}$$

Figure 13: The uniqueness of the complexification up to isomorphism.

Proof: The proof is similar to what we did for Proposition 3.2 (p. 26). Since $\iota_2\colon V \to V_{\mathbb{C}2}$ is $\mathbb{R}$-linear, the universal property of $(V_{\mathbb{C}1}, \iota_1)$ will give us a unique $\mathbb{C}$-linear map $\Phi\colon V_{\mathbb{C}1} \to V_{\mathbb{C}2}$ such that $\Phi \circ \iota_1 = \iota_2$.

$$\begin{tikzcd}
V_{\mathbb{C}1} \arrow[dr, dashed, "\Phi"] \arrow[r, phantom, "\circlearrowleft"] & \\
V \arrow[u, "\iota_1"] \arrow[r, "\iota_2"'] & V_{\mathbb{C}2}
\end{tikzcd}$$

Figure 14: Defining Φ.

Similarly, using that $\iota_1 \colon V \to V_{\mathbb{C}1}$ is $\mathbb{R}$-linear, the universal property of $(V_{\mathbb{C}2}, \iota_2)$ will give us a unique $\mathbb{C}$-linear map $\Psi \colon V_{\mathbb{C}2} \to V_{\mathbb{C}1}$ such that $\Psi \circ \iota_2 = \iota_1$.

$$\begin{tikzcd}
V_{\mathbb{C}2} \arrow[dr, dashed, "\Psi"] \arrow[r, phantom, "\circlearrowleft"] & \\
V \arrow[u, "\iota_2"] \arrow[r, "\iota_1"'] & V_{\mathbb{C}1}
\end{tikzcd}$$

Figure 15: Defining Ψ.

Then Φ and Ψ are inverses.

$$\begin{tikzcd}
V_{\mathbb{C}1} \arrow[dr, dashed, "\mathrm{Id}_{V_{\mathbb{C}1}}"] \arrow[r, phantom, "\circlearrowleft"] & \\
V \arrow[u, "\iota_1"] \arrow[r, "\iota_1"'] & V_{\mathbb{C}1}
\end{tikzcd}
\qquad
\begin{tikzcd}
V_{\mathbb{C}1} \arrow[dr, dashed, "\Psi \circ \Phi"] \arrow[r, phantom, "\circlearrowleft"] & \\
V \arrow[u, "\iota_1"] \arrow[r, "\iota_1"'] & V_{\mathbb{C}1}
\end{tikzcd}$$

Figure 16: Proving that $\Psi \circ \Phi = \mathrm{Id}_{V_{\mathbb{C}1}}$.

Similarly, we show that $\Phi \circ \Psi = \mathrm{Id}_{V_{\mathbb{C}2}}$. $\square$

Remark. The sort of argument given above works, in general, for proving the uniqueness (up to isomorphism) of any object characterized by a universal property.

Exercise 3.8. The following objects also have universal properties. Look them up:

(a) The direct sum $\bigoplus_{i\in I} V_i$ of a family $(V_i)_{i\in I}$ of vector spaces.

(b) The direct product $\prod_{i\in I} V_i$ of a family $(V_i)_{i\in I}$ of vector spaces.

(c) The quotient V/W of a vector space V by a subspace W.

(d) The free vector space $\mathscr{F}(S)$ over a non-empty set S (seen in Example 3.3, p. 28).

Universal properties also appear in several other areas of Mathematics. They are studied rigorously in *Category Theory*.

Fortunately, the construction of complexifications is easier. One of the most usual constructions, which you might already know, is outlined in the:

Exercise 3.9. Let V be a real vector space. Define in the cartesian product $V \times V$ the following multiplication by complex scalar:

$$(a+bi)(\boldsymbol{u},\boldsymbol{v}) \doteq (a\boldsymbol{u}-b\boldsymbol{v}, b\boldsymbol{u}+a\boldsymbol{v}).$$

(a) With this multiplication and the usual addition performed componentwise, show that $V\times V$ is a complex vector space. Note that $(\boldsymbol{u},\boldsymbol{v}) = (\boldsymbol{u},\boldsymbol{0}) + i(\boldsymbol{0},\boldsymbol{v})$. This way, we write $V \oplus iV \doteq V\times V$.

   Remark. If you feel comfortable doing it, you might as well identify $\boldsymbol{u} \equiv (\boldsymbol{u},\boldsymbol{0})$ and $i\boldsymbol{v} \equiv (\boldsymbol{0},\boldsymbol{v})$, and start writing $(\boldsymbol{u},\boldsymbol{v}) = \boldsymbol{u} + i\boldsymbol{v}$, just like you were working in $\mathbb{C}$. As a matter of fact, this construction applied to $V=\mathbb{R}$ indeed gives $\mathbb{C}$.

(b) Show that if $(\boldsymbol{v}_i)_{i=1}^n$ is a $\mathbb{R}$-basis for V, then $((\boldsymbol{v}_i,\boldsymbol{0}))_{i=1}^n$ is a $\mathbb{C}$-basis for $V\oplus iV$. Thus $\dim_{\mathbb{C}}(V\oplus iV) = \dim_{\mathbb{R}} V$.

(c) If $\iota\colon V\to V\oplus iV$ is given by $\iota(\boldsymbol{u}) = (\boldsymbol{u},\boldsymbol{0})$, show that $(V\oplus iV,\iota)$ satisfies the universal property of the complexification.

(d) Bonus: suppose that $\langle\cdot,\cdot\rangle$ is a scalar product on V. Show that

$$\langle \boldsymbol{u}_1 + i\boldsymbol{v}_1, \boldsymbol{u}_2 + i\boldsymbol{v}_2\rangle_{\mathbb{C}} \doteq \langle \boldsymbol{u}_1,\boldsymbol{u}_2\rangle + \langle \boldsymbol{v}_1,\boldsymbol{v}_2\rangle + i(\langle \boldsymbol{v}_1,\boldsymbol{u}_2\rangle - \langle \boldsymbol{u}_1,\boldsymbol{v}_2\rangle)$$

   is a *hermitian and sesquilinear*[6] product on $V\oplus iV$, that is, it is linear in the first entry, anti-linear in the second entry, and it satisfies $\langle \boldsymbol{z},\boldsymbol{w}\rangle_{\mathbb{C}} = \overline{\langle \boldsymbol{w},\boldsymbol{z}\rangle_{\mathbb{C}}}$.

Several properties of complexifications can be proven via the uniqueness given by the universal property characterizing them. For instance:

Exercise 3.10. Let $(V_{\mathbb{C}},\iota)$ be a complexification of V. Show that if Z is a complex vector space, $T,S\colon V\to Z$ are $\mathbb{R}$-linear and $\lambda\in\mathbb{R}$, then:

(a) $(T+S)_{\mathbb{C}} = T_{\mathbb{C}} + S_{\mathbb{C}}$;

[6] Better than linear, weaker than bilinear: 1.5-linear.

(b) $(\lambda T)_{\mathbb{C}} = \lambda T_{\mathbb{C}}$.

Another way to explicitly construct complexification is via tensor products:

Proposition 3.11. *The pair* $(\mathbb{C} \otimes V, 1 \otimes \_)$ *is a complexification of* V.

Proof: Before anything else, note that the multiplication by complex scalar defined by $\mu(\lambda \otimes \boldsymbol{v}) \doteq (\mu\lambda) \otimes \boldsymbol{v}$ turns $\mathbb{C} \otimes V$ into a complex vector space. And clearly the map

$$V \ni \boldsymbol{v} \overset{\iota}{\longmapsto} 1 \otimes \boldsymbol{v} \in \mathbb{C} \otimes V$$

is $\mathbb{R}$-linear. So, let Z be a complex vector space, and take a $\mathbb{R}$-linear map $T\colon V \to Z$. Define $\widetilde{T}\colon \mathbb{C} \times V \to Z$ by $\widetilde{T}(\lambda, \boldsymbol{v}) \doteq \lambda T(\boldsymbol{v})$.

$$\begin{tikzcd}
\mathbb{C} \otimes V \arrow[rd, dashed, "T_{\mathbb{C}}"] & \\
\mathbb{C} \times V \arrow[u, "\otimes"] \arrow[r, "\widetilde{T}"'] \arrow[ru, phantom, "\circlearrowright"] & Z
\end{tikzcd}$$

Figure 17: Constructing $T_{\mathbb{C}}$ via the propriedade universal of $\mathbb{C} \otimes V$.

Since $\widetilde{T}$ is $\mathbb{R}$-bilinear, there is a unique $\mathbb{R}$-linear map $T_{\mathbb{C}}\colon \mathbb{C} \otimes V \to Z$ such that

$$T_{\mathbb{C}}(\lambda \otimes \boldsymbol{v}) = \lambda T(\boldsymbol{v}),$$

for all $\lambda \in \mathbb{C}$ and $\boldsymbol{v} \in V$. In particular, $T_{\mathbb{C}}(\iota(\boldsymbol{v})) = T(\boldsymbol{v})$, and thus the only thing left to show is that $T_{\mathbb{C}}$ is actually $\mathbb{C}$-linear. But

$$T_{\mathbb{C}}(\mu(\lambda \otimes \boldsymbol{v})) = T_{\mathbb{C}}((\mu\lambda) \otimes \boldsymbol{v}) = (\mu\lambda)T(\boldsymbol{v}) = \mu(\lambda T(\boldsymbol{v})) = \mu T_{\mathbb{C}}(\lambda \otimes \boldsymbol{v}),$$

as wanted. $\square$