Tensor Products: The Phantom Menace

Nikolay Nowaczyk Jesko Hüttenhain

Once upon a time, in Bonn

"If one is to understand the great mystery, one must study all its aspects, not just the dogmatic narrow view of the Jedi."
— Chancellor Palpatine

Contents

1 Algebra – A new Hope

We will define the tensor product in terms of its universal property in the first subsection, then give some further intuition, related to polynomial functions, in the second part.

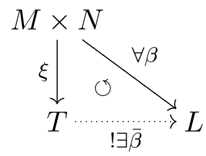
1.1 Definition and Existence

For the entire subsection, let R be a unitary, commutative ring, which might serve your intuition best to think of as a field.

1.1 Definition. Let M and N be R-modules. A **tensor product for M and N over R** is an R-module T together with a bilinear map[1] $\xi : M \times N \to T$ such that the following universal property is satisfied:

[1] i.e. for all $(m, n) \in M \times N$, the maps $\xi(\bullet, m)$ and $\xi(n, \bullet)$ are both R-linear.

If $\beta : M \times N \to L$ is any bilinear map, then there exists a unique R-module homomorphism $\bar{\beta} : T \to L$ with $\bar{\beta} \circ \xi = \beta$. In diagram language,



In other words, the bilinear maps from $M \times N$ to L are in bijection with the homomorphisms from T to L, for all R-modules L.

1.2 Theorem. The tensor product T of any two R-modules exists and is generated, as an R-module, by the image of the associated bilinear map $M \times N \to T$.

Proof. Since we have to prove the existence in the most general case, we have to resort to drastic means and force the issue. We denote by $F := R[M \times N]$ the free R-module generated by all $(m, n) \in M \times N$. Now, let $Q \subseteq F$ be the submodule generated by the elements

- $(m, n + n') - (m, n) - (m, n')$,
- $(m + m', n) - (m, n) - (m', n)$,
- $(m, \alpha \cdot n) - \alpha \cdot (m, n)$ and
- $(\alpha \cdot m, n) - \alpha \cdot (m, n)$

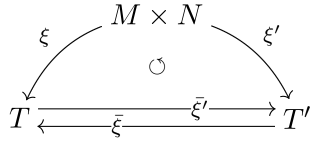
for all $\alpha \in R$, $m, m' \in M$ and $n, n' \in N$. We then claim that $T = F/Q$ defines a tensor product together with the map $\xi : M \times N \to T$, defined by sending (m, n) to its residue class in T. This map is bilinear by definition of Q.

We now verify that it satisfies the universal property. Assume that $\beta : M \times N \to L$ is a bilinear map of R-modules. Then, we get a unique homomorphism $\hat{\beta} : F \to L$ with the property that $\hat{\beta}(m, n) = \beta(m, n)$ for all $(m, n) \in M \times N$. Since β is bilinear, $\ker(\hat{\beta}) \supseteq Q$. Denote by $\pi : F \twoheadrightarrow T$ the canonical projection. By the fundamental theorem of homomorphisms, there exists a unique $\bar{\beta} : T \to L$ with the property that $\bar{\beta} \circ \pi = \hat{\beta}$. Thus, $\bar{\beta}$ is unique with the property that

$$\bar{\beta}(\xi(m, n)) = \bar{\beta}(\pi(m, n)) = \hat{\beta}(m, n) = \beta(m, n). \qquad \square$$

1.3 Proposition. Any two tensor products for M and N over R are uniquely isomorphic. Consequently, we write $M \otimes_R N$ for "the" tensor product.

Proof. Assume that T and T' are both tensor products for M and N over R.



If we denote by ξ and ξ' the corresponding bilinear maps, they each induce unique R-module homomorphisms $\bar{\xi}'$ and $\bar{\xi}$ by the universal property of the other tensor product, such that

$$\bar{\xi} \circ \bar{\xi}' \circ \xi = \bar{\xi} \circ \xi' = \xi.$$

Now, we may apply the universal property of the tensor product T to ξ itself and infer that the R-module homomorphism $\mathrm{id}_T : T \to T$ is unique with the property $\mathrm{id}_T \circ \xi = \xi$. Thus, $\bar{\xi} \circ \bar{\xi}' = \mathrm{id}_T$. We proceed analogously to verify $\bar{\xi}' \circ \bar{\xi} = \mathrm{id}_{T'}$. □

1.4 Remark. Let M and N be R-modules. Let $\xi : M \times N \to M \otimes_R N$ denote the bilinear map associated to the tensor product. We usually avoid working with this map and write $m \otimes n := \xi(m, n)$ instead. These elements are said to be **fully decomposable**. By 1.2, $M \otimes_R N$ is generated by its fully decomposable elements – but certainly not every element of the tensor product is fully decomposable.

In general, an element $x \in M \otimes_R N$ can be written as a finite linear combination of fully decomposable elements, i.e. $x = \sum_{i=1}^{N} m_i \otimes n_i$. Note that this representation is not unique – the fully decomposable elements are subject to the obvious "bilinear" relations.

1.2 Algebraic Applications

Although it can be defined in such great generality, let us look at some special applications of the Tensor product.

1.5 Proposition. Let M, N and L be R-modules. Then, we have the following canonical isomorphisms:

(i). Commutativity:

$$\begin{aligned} M \otimes_R N &\longrightarrow N \otimes_R M \\ m \otimes n &\longmapsto n \otimes m \end{aligned}$$

(ii). Neutral Element:

$$\begin{aligned} R \otimes_R M &\longrightarrow M \\ a \otimes m &\longmapsto a \cdot m \end{aligned}$$

(iii). Associativity:

$$\begin{aligned} (M \otimes_R N) \otimes_R L &\longrightarrow M \otimes_R (N \otimes_R L) \\ (m \otimes n) \otimes l &\longmapsto m \otimes (n \otimes l) \end{aligned}$$

In other words, the set of equivalence classes of R-modules modulo isomorphism form an abelian monoid with neutral element R and binary operation $\otimes_R$.

Proof. For all three statements, one proceeds equivalently. We prove part (iii): The map

$$\begin{aligned}(M \otimes_R N) \times L &\longrightarrow M \otimes_R (N \otimes_R L)\\ (m \otimes n, l) &\longmapsto m \otimes (n \otimes l)\end{aligned}$$

is bilinear, so it induces the unique linear map (iii). Equivalently, we obtain the obvious inverse homomorphism. In the same fashion, one proves parts (i) and (ii). □

1.6 Proposition. Assume that $\{ M_i \mid i \in I \}$ is a family of R-modules and N any R-module. Then, there exists an isomorphism

$$\begin{aligned}\Big(\bigoplus\nolimits_{i\in I} M_i\Big) \otimes_R N &\longrightarrow \bigoplus\nolimits_{i\in I}(M_i \otimes_R N)\\ (x_i)_{i\in I} \otimes n &\longmapsto (x_i \otimes n)_{i\in I}\end{aligned}$$

Proof. The existence of the above map follows from the universal property of the tensor product. Conversely, the maps

$$\begin{aligned}M_j \times N &\longrightarrow \Big(\bigoplus\nolimits_{i\in I} M_i\Big) \otimes_R N\\ (x, n) &\longmapsto (\delta_{ij} x)_i \otimes n\end{aligned}$$

are all bilinear and give rise to an inverse. □

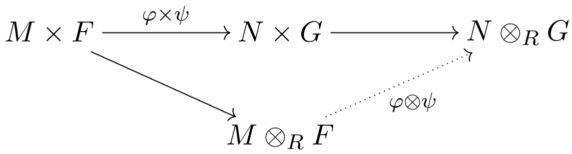
1.7 Corollary. If $F = \bigoplus_{i\in I} \langle f_i \rangle$ and $G = \bigoplus_{j\in J} \langle g_j \rangle$ are free R-modules, then

$$G \otimes_R F = \bigoplus_{(i,j)\in I\times J} \langle f_i \otimes g_j \rangle$$

is free. In particular, $\mathrm{rank}_R(F \otimes_R G) = \mathrm{rank}_R(F) \cdot \mathrm{rank}_R(G)$ which is symbolically true if either module has infinite rank.

Proof. This follows directly from the above. □

1.8 Remark. If $\varphi : M \to N$ and $\psi : F \to G$ are homomorphisms of R-modules, we obtain



which is given by

$$\begin{aligned}\varphi \otimes \psi : M \otimes_R F &\longrightarrow N \otimes_R G\\ m \otimes f &\longmapsto \varphi(m) \otimes \psi(f)\end{aligned}$$

Clearly, $\varphi \otimes \psi$ is surjective if both φ and ψ are.

1.9 Proposition. If F is a free R-module and $\varphi : M \hookrightarrow N$ is an injective homomorphism of R-modules, then $\varphi \otimes \mathrm{id}_F$ is injective as well.

Proof. If $F \cong R^{\oplus n}$, say, then

$$\begin{array}{ccc} M \otimes_R F & \xrightarrow{\varphi\otimes\mathrm{id}_F} & N \otimes_R F \\ \| & & \| \\ M^{\oplus n} & \xhookrightarrow{\varphi^{\oplus n}} & N^{\oplus n} \end{array}$$

commutes, which already proves the claim. □

1.10 Corollary. Let S be an R-algebra and M some R-module. Then, $S \otimes_R M$ has the structure of an S-module via the scalar multiplication $a \cdot (b \otimes m) := ab \otimes m$.

If M is free and $\iota : R \hookrightarrow S$, the map

$$\begin{array}{ccc} M & \longrightarrow & S \otimes_R M \\ m & \longmapsto & 1 \otimes m \end{array}$$

is an injective homomorphism of R-modules.

Proof. The first statement is obvious and to see the second statement, we simply note that the above map is the composition

$$M \xrightarrow{\sim} R \otimes_R M \xrightarrow{\iota\otimes\mathrm{id}_M} S \otimes_R M$$ □

1.1 Example. Assume that V is some $\mathbb{R}$-vectorspace. Then, $\mathbb{C} \otimes_{\mathbb{R}} V$ is a complex vectorspace of the same (complex) dimension: Pick some basis $b_1, \dots, b_n$ of V and consider the vectors $1 \otimes b_i \in \mathbb{C} \otimes_{\mathbb{R}} V$. Certainly any $\alpha \otimes v$ can be represented as a $\mathbb{C}$-linear combination of them and

$$0 = \sum\nolimits_{j=1}^{n} (\alpha_j + \beta_j i) \cdot (1 \otimes b_j) = \sum\nolimits_{j=1}^{n} \alpha_j \cdot (1 \otimes b_j) + \sum\nolimits_{j=1}^{n} \beta_j \cdot (i \otimes b_j)$$

implies that $\alpha_j = \beta_j = 0$ for all j by 1.7.

2 Differential Geometry - Revenge of the Sith

(!Todo Basis -> Basis für Boxtensor = Otensor)

2.1 A quick and easy Path

2.1 Global Assumption. For the entire section let K be a field, $V, V_1, \dots, V_k, W$ be finite dimensional vector spaces over K of dimensions

$$n := \dim V, \qquad \forall 1 \leq i \leq k : n_i := \dim V_i, \qquad m := \dim W.$$

We assume that $B := (b_1, \ldots, b_n)$ is a basis of V, $B_{(i)} := (b_1^{(i)}, \ldots, b_{n_i}^{(i)})$ are bases of V_i, $1 \leq i \leq k$, $C := (c_1, \ldots, c_m)$ is a basis of W.
We denote by $\operatorname{Hom}(V, W)$ the space of linear maps $V \to W$, by $V' := \operatorname{Hom}(V, K)$ the dual space of V, by $B' := (b^1, \ldots, b^n)$ the basis dual to B and by $B'_{(i)} = (b^1_{(i)}, \ldots, b^k_{(i)})$ the dual bases of V_i.

2.2 Definition (multlinear)**.** A map

$$F : V_1 \times \ldots \times V_k \to W$$

is *k-fold multilinear*, if it is linear in all its arguments. The space of all those maps is denoted by

$$\operatorname{Mult}(V_1, \ldots, V_k; W).$$

Clearly $\operatorname{Mult}(V; W) = \operatorname{Hom}(V, W)$. We use the notation $\operatorname{Bil}(V, W) := \operatorname{Mult}(V, W; K)$.

2.3 Definition (Tensor)**.** A map

$$F \in \operatorname{Mult}(V_1, \ldots, V_k; K) =: \operatorname{Mult}(V_1, \ldots, V_k)$$

is a *tensor (on $V_1, \ldots, V_k$)*.

2.4 Definition (Tensor product)**.** Let $(W_1, \ldots, W_l)$ be another tuple of vector spaces. For any $F \in \operatorname{Mult}(V_1, \ldots, V_k)$, $G \in \operatorname{Mult}(W_1, \ldots, W_l)$, the map

$$\begin{aligned} F \boxtimes G : & V_1 \times \ldots \times V_k \times W_1 \times \ldots \times W_l \to K, \\ & (x_1, \ldots, x_k, y_1, \ldots, y_l) \mapsto F(x_1, \ldots, x_k) G(y_1, \ldots, y_l), \end{aligned}$$

is the *tensor product between F and G*. This defines a map

$$\boxtimes : \operatorname{Mult}(V_1, \ldots, V_k) \times \operatorname{Mult}(W_1, \ldots, W_l) \to \operatorname{Mult}(V_1, \ldots, V_k, W_1, \ldots, W_l).$$

This is a very general definition. Usually people do not consider tensors on an arbitrary tuple of vector spaces $(V_1, \ldots, V_k)$, but restrict their attention to the following special type.

2.5 Definition (Tensor of type (k, l))**.** Let $k, l \in \mathbb{N}_0$ be arbitrary. A map

$$F \in \operatorname{Mult}(\underbrace{V^*, \ldots, V^*}_{l \text{ copies}}, \underbrace{V, \ldots, V}_{k \text{ copies}})$$

is a *tensor of type (k, l) on V*. We say F is *k-fold covariant and l-fold contravariant.* We define

$$T_l^k(V) := \{F \mid F \text{ is a tensor of type } (k, l) \text{ on } V\}$$

and use the notation

$$\begin{aligned} T^k(V) &:= T_0^k(V), \\ T_l(V) &:= T_l^0(V). \end{aligned}$$

The direct sum

$$T(V) := \bigoplus_{k,l \in \mathbb{N}_0} T^k_l(V)$$

is the *tensor algebra over* V.

2.6 Convention. When working with tensors, it is sometimes useful to group the arguments by type to avoid getting lost in them. If we denote

$$\Omega := (\omega^1, \dots, \omega^l) \in (V^*)^l, \qquad X := (x_1, \dots, x_k) \in V^k,$$

we may express the action of a tensor $F \in T^k_l(V)$ by

$$F(\Omega, X) := F(\omega^1, \dots, \omega^l, x_1, \dots, x_k).$$

In case we need two tensors or two such groups of vectors and covectors, we will usually call the second ones $H := (\eta^1, \dots, \eta^{l'}) \in (V^*)^{l'}$, $Y := (y_1, \dots, y_{k'}) \in V^{k'}$.
In case $k = l$, we will denote the action of a tuple of covectors to a tuple of vectors by

$$\Omega(X) := \omega^1(x_1) \dots \omega^k(x_k) = (\omega^1 \boxtimes \dots \boxtimes \omega^k)(x_1, \dots, x_k).$$

2.7 Remark. One could specialize 2.4 to tensors of type (k, l) and call this a tensor product. The only inconvenience here is the order of the arguments: Let $k, k', l, l' \in \mathbb{N}$, $F \in T^k_l(V)$, $G \in T^{k'}_{l'}(V)$, $\Omega \in (V^*)^l$, $H \in (V^*)^{l'}$, $X \in V^k$, $Y \in V^{k'}$. Then the tensor product from 2.4 would be defined by

$$F \boxtimes G(\Omega, X, H, Y) = F(\Omega, X)G(H, Y).$$

According to our definition, $F \boxtimes G \notin T^{k+k'}_{l+l'}(V)$ and the only reason for this to fail is, that $F \boxtimes G$ has to act on (Ω, H, X, Y) instead of (Ω, X, H, Y). Since there no agreement on the order of arguments of tensors in the literature anyway one may decide to just live with that and accept this slight discrepancy. Alternatively one could complicate the definition of a tensor of type (k, l) even further by requireing it to act on any permutation of $(V^*)^l \times V^k$. We choose to adopt the convention from [1] and just change the definition of the tensor product that case.

2.8 Definition. With the notation from 2.7 above, the map defined by

$$(\Omega, H, X, Y) \mapsto F(\Omega, X)G(H, Y)$$

is called the (k, l)-*tensor product*. Usually we will just call it *tensor product* as well and we will also denoted by $\boxtimes$. This defines maps

$$\boxtimes \in \operatorname{Mult}(T^k_l(V), T^{k'}_{l'}(V), T^{k+k'}_{l+l'}(V)), \qquad \boxtimes \in \operatorname{Mult}(T(V), T(V); T(V)).$$

This should cause no confusion since it differs only slightly. In its finest beauty, this map looks like

$$\begin{aligned} &F \boxtimes G(\omega^1, \dots, \omega^l, \eta^1, \dots, \eta^{l'}, x_1, \dots, x_k, y_1, \dots, y_{k'}) \\ &\quad = F(\omega^1, \dots, \omega^l, x_1, \dots, x_k)G(\eta^1, \dots, \eta^{l'}, y_1, \dots, y_{k'}). \end{aligned}$$

2.9 Remark. Notice this tensor product is a product beween two elements of two spaces, not between two entire spaces. The notation $\boxtimes$ is very uncommon for the tensor product. Usually the map defined above is also denoted $\otimes$. We will investigate the relations between $\otimes$ and $\boxtimes$ later in 2.23.

2.2 Attack of the Indices

Tensors have been condemned by the Jedi, because working with tensors leads to indices, indices lead to index wars and index wars lead to the dark side of the force - to us! A true Sith Lord is not afraid of indices, embraces their dark powers and ridicules those who are to weak to deal with them.
Well... at least that is what he tells the others. But between you and me: When it comes to tensors one should really think about notation and don't make things even more complicated than necessary. A standard tool to simplify notation is the so called *Einstein summation convention.* Depending on what book you read, the only thing you might be told about this convention is that it is "in power". Since a convention is usually something more than one people agree on, we should try to state it here as clearly as possible. This is much more comprehensive than required for this article, but could be helpful for those who want to delve deeper in the literature.

2.10 Convention (Einstein Summation)**.** If in an equation an index occurs exactly once as a subscript and exactly once as a superscript, we sum over this index starting from 1 to the dimension of the space.

This convention has obviously been invented by a phisicist (guess who?). Of course not every object occuring in some formula is something you can just "sum". In a far less general, but much more mathematically precise way, one may state this convention as follows.

2.11 Convention (Einstein Summation, Take 2)**.** Let $x_1, \ldots, x_k \in V$ and $\lambda^1, \ldots, \lambda^k \in K$. We define

$$\lambda^i x_i := \sum_{i=1}^{n} \lambda^i x_i.$$

In almost all cases the $x_1, \ldots, x_k$ are a basis $(b_1, \ldots, b_n)$. So the primary application is to simplify expressions that occur when expanding vectors with respect to some basis, i.e.

$$x = x^i b_i = \sum_{i=1}^{n} x^i b_i,$$

where the $x^i \in K$ are the components of x.

Here you already see the next problem with this convention: Mathematical objects may be godgiven, but their indices are manmade. Theoretically you are perfectly free to index objects with subscripts or with superscripts. In order to utilize this convention in a senseful manner, there are usually hidden conventions in power that ensure this.

2.12 Convention (Index Positioning)**.** An index is to be put as a subscript or a superscript according to table 1. Thus, the positioning of the index indicates whether or

	vector space V	dual space V^*
subscript	vectors $x_1, \ldots, x_k$	co-components $\omega_1, \ldots, \omega_n$
superscript	components $y^1, \ldots, y^n$	co-vectors $\eta^1, \ldots, \eta^l$
expansion	$x = x^i b_i \in V$	$\omega = \omega_i b^i \in V^*$

Table 1: Index Positioning

not we work on the space or on the dual space (yes, a dual space is a space as well, but we ignore this here.) As you can see, there is another important convention concerning expansions: Components of a vector are written with the same letter, but they have an index.

2.13 Convention (Matrices and Homomorphisms)**.** Let $f \in \mathrm{Hom}(V, W)$. Then the coordinate matrix $A \in K^{m \times n}$ usually has two indices as a subscript namely $A = A_{ij} = A_{\text{column,row}}$. It is defined by the relation

$$f(b_j) = \sum_{i=1}^{m} A_{ij} c_i.$$

As you might guess, this is awful when one wants to utilize Einstein summation convention. Therefore we change this convention and write

$$A = A^i_j = A^{\text{row}}_{\text{column}}.$$

Then Einstein summation works and we may write things like

$$f(b_j) = A^i_j c_i, \qquad f(x) = f(x^j b_j) = x^j f(b_j) = x_j A^i_j c_i.$$

2.14 Convention (Inverse Matrices)**.** To make things worse, sometimes we do stick to the old convention, even if we want to use Einstein Summation. This is often done when there is an invertible matrix $A \in K^{n \times n}$. Then $A = A_{\text{column,row}}$ is the matrix itself and $A^{-1} =: A^{\text{column,row}}$ is the inverse. Theoretically this could cause conflicts, but it is usually clear from the context, which convention is used. If this disgusts you, notice that you can now do cool things like

$$(AB)_{ik} = A^i_k B^k_j, \qquad A_{ik} A^{kj} = \delta^j_i.$$

2.15 Remark (Typical failures)**.** Despite all this, there are several situations, where Einstein Summation usually failes.

- Subspaces: If $U \subset V$ is a subspace, the expression $x = x^i b_i$ makes no sense, because it is not clear what "the dimension of the space" shall be. If it is your very point that $x \in U$, i.e. that the last k components of x for instance are zero, one should not use Einstein Summation.

- Scalar products: If V has an inner product $\langle \_, \_ \rangle$ and $b_1, \dots, b_n$ is an orthonormal basis, Einstein Summation always fails at the last step of

$$\langle x, y \rangle = \langle x^i b_i, y^j b_j \rangle = x^i y^j \langle b_i, b_j \rangle = x^i y^j \delta_{ij} = \sum_{i=1}^{n} x^i y^i,$$

thus we have to put a sum there.

- Components: If you really want to consider the number $x_i b_i$ for one fixed i, then one should absolutely avoid writing this as $x^i b_i$ in that context.

Because of the possible confusion you are perfectly free to condemn Einstein summation. But you are not free to be able to understand it or not, since otherwise you will not be able to read a big amount of literature concerning tensors. We will use it in the following.

2.3 Feel the Force

Now we will systematically study tensors. Our ultimate goal is to understand the relations bewteen the (k, l)-tensor produkt $\boxtimes$ and the tensor product $\otimes$ of the Jedi.

Remember the following crucial fact from linear algebra.

2.16 Lemma (Dual Expansion)**.** Any vector $x' \in V'$ has a unique representation

$$x' = x'_i b^i = x'(b_i) b^i.$$

Proof. It is clear that any vector has a unique expansion $x' = \lambda_i b^i$, so we only have to determine the λ_i. By definition of a dual basis

$$x'(b_j) = \lambda_i b^i(b_j) = \lambda_i \delta^i_j = \lambda_j.$$

□

2.17 Lemma (Canonical identification of the bidual)**.** The map $\iota : V \to V''$

$$x \mapsto \left(x' \mapsto x'(x) \right),$$

is an isomorphism. Its inverse is given by

$$\forall x'' \in V'' : \iota^{-1}(x'') = x''(b^i) b_i.$$

Proof. Notice that $\iota(x) : V' \to K$. Since V is finite dimensional,

$$\dim V'' = \dim V' = \dim V$$

is finite as well. Therefore the dual basis B' is in fact a basis (which can be wrong if V would be inifite dimensional). We calculate one the one hand

$$\forall x \in V : (\iota^{-1} \circ \iota)(x) = \iota(x)(b^i) b_i = b^i(x) b_i = x^i b_i = x,$$

and also

$$\forall x'' \in V'' : \forall x' \in V' : (\iota \circ \iota^{-1})(x'')(x') = x'(\iota^{-1}(x'')) = x'(x''(b^i)b_i) = x''(b^i)x'(b_i)$$
$$= x''(b^i)x'_i = x''(x'_i b^i) = x''(x').$$

This proves the statement. □

2.18 Convention. Without further reference we will identify V'' with V using 2.17 above. In particular we will "apply" vectors $x \in V$ to covectors $x' \in V'$ by $x(x') := x'(x)$. Furthermoore we regard any basis of V as a basis of V''. With this in mind 2.16 can be formulated for V as: Any vector $x \in V$ has a unique expansion

$$x = x(b^i)b_i.$$

2.19 Convention. For any multi-index $\nu = (\nu_1, \ldots, \nu_k)$, we will denote

$$B^{\otimes\nu} := b^{\nu_1}_{(1)} \otimes \ldots \otimes b^{\nu_k}_{(k)}, \qquad B_{\otimes\nu} := b^{(1)}_{\nu_1} \otimes \ldots \otimes b^{(k)}_{\nu_k},$$
$$B^{\nu} := (b^{\nu_1}_{(1)}, \ldots, b^{\nu_k}_{(k)}), \qquad B_{\nu} := (b^{(1)}_{\nu_1}, \ldots, b^{(k)}_{\nu_k}),$$

and the same for $\boxtimes$. We will employ the convention that the Einstein summation convention holds for multi-indices as well. If $\mu = (\mu_1, \ldots, \mu_k)$ is another multi-index, we will use the *Kronecker-Delta for multi-indicex*

$$\delta^{\nu\mu} := \delta^{\nu}_{\mu} := \delta_{\nu\mu} := \delta^{\nu_1}_{\mu_1} \ldots \delta^{\nu_k}_{\mu_k}.$$

2.20 Lemma (Multi Expansion)**.** For any ν, μ

(i). $B^{\boxtimes\nu}(B_\mu) = \delta^{\nu}_{\mu}$.

(ii). $B_{\boxtimes\nu}(B^\mu) = \delta^{\mu}_{\nu}$.

Proof. For part (i), we simply calculate

$$B^{\boxtimes\nu}(B_\mu) = b^{\nu_1}_{(1)} \boxtimes \ldots \boxtimes b^{\nu_k}_{(k)}(b^{(1)}_{\mu_1}, \ldots, b^{(k)}_{\mu_k}) = b^{\nu_1}_{(1)}(b^{(1)}_{\mu_1}) \ldots b^{\nu_k}_{(k)}(b^{(k)}_{\mu_k}) = \delta^{\nu_1}_{\mu_1} \ldots \delta^{\nu_k}_{\mu_k} = \delta^{\nu}_{\mu}$$

and part (ii) follows equivalently. □

2.21 Theorem (Tensor Bases)**.** The set

$$[\{B^{\boxtimes\nu} \in \operatorname{Mult}(V_1, \ldots, V_k) \mid \nu \in \mathbb{N}_0^k\} \tag{2.1}$$

is a basis of $\operatorname{Mult}(V_1, \ldots, V_k)$. Any tensor F has a unique expansion $F = F(B_\nu)B^{\boxtimes\nu}$.

Proof. To prove linear independence, assume there are scalars $\lambda_\nu \in K$, such that $0 = \lambda_\nu B^{\boxtimes\nu}$. For any μ, this implies

$$0 = \lambda_\nu B^{\boxtimes\nu}(B_\mu) \overset{2.20}{=} \lambda_\nu \delta^{\nu}_{\mu} = \lambda_\mu.$$

Thus all $\lambda_\mu = 0$.

We now calculate for any $F \in \mathrm{Mult}(V_1, \dots, V_k)$ and any $X = (x_1, \dots, x_k) \in V_1 \times \dots V_k$ that

$$\begin{aligned} F(X) = F(x_1, \dots, x_k) &= F(x_1^{\nu_1} b_{\nu_1}^{(1)}, \dots, x_k^{\nu_k} b_{\nu_k}^{(k)}) \\ &= x_1^{\nu_1} \dots x_k^{\nu_k} F(b_{\nu_1}^{(1)}, \dots, b_{\nu_k}^{(k)}) \overset{2.16}{=} F(B_\nu) B^{\boxtimes \nu}(X). \end{aligned}$$

Thus, (2.1) is a system of generators, proving our claim. □

2.22 Definition (Notational Trick)**.** Now we define $\boxtimes$ for dual spaces: Notice that for any two spaces V and W, we may write the dual spaces by $V' = \mathrm{Mult}(V)$, $W' = \mathrm{Mult}(W)$. Therefore the tensor product is defined and determines a map

$$\boxtimes : \mathrm{Mult}(V) \times \mathrm{Mult}(W) \to \mathrm{Mult}(V, W).$$

Since $\boxtimes$ is bilinear, its image $\boxtimes(V', W') \subset \mathrm{Mult}(V, W)$ need not be a vector space. But

$$V' \boxtimes W' := \mathrm{Lin}(\boxtimes(V', W')) \overset{2.21}{=} \mathrm{Mult}(V, W)$$

is a vector space.

2.23 Theorem (Tensors on dual spaces)**.** The map

$$\begin{aligned} V_1' \times \dots \times V_k' \times W &\longrightarrow \mathrm{Mult}(V_1, \dots, V_k, W) \\ (H, w) &\longmapsto \Big(X \mapsto H(X) w \Big) \end{aligned}$$

is itself k-fold multlinear. By the universal property 1.1 of the tensor product it descends to a map

$$\varphi : V_1' \otimes \dots \otimes V_k' \otimes W \to \mathrm{Mult}(V_1, \dots, V_k; W).$$

This map satisfies

$$\varphi(\eta^1 \otimes \dots \otimes \eta^k \otimes w)(x_1, \dots, x_k) = (\eta^1 \boxtimes \dots \boxtimes \eta^k)(x_1, \dots, x_k) w \tag{2.2}$$

and is a canonical isomorphism[2].

In particular

(i). $V_1' \otimes \dots \otimes V_k' \cong V_1' \boxtimes \dots \boxtimes V_k' = \mathrm{Mult}(V_1, \dots, V_k)$, in wich case

$$\varphi(\eta^1 \otimes \dots \otimes \eta^k) = \eta^1 \boxtimes \dots \boxtimes \eta^k.$$

(ii). $V' \otimes W' \cong V' \boxtimes W' = \mathrm{Bil}(V, W)$,

[2]The word "canonical" here means that it may be written down without an explicit reference to bases, although we will require bases to prove that it is an isomorphism. By the way, this is the reason why we have to work over finite-dimensional vector spaces instead of R-modules.

(iii). $V' \otimes W \cong \mathrm{Mult}(V, W) = \mathrm{Hom}(V, W)$,

(iv). $V' \otimes V \cong \mathrm{End}(V)$.

By choosing the basis from 2.1 and employing convention 2.19, the inverse of φ may be explicitely written as

$$\forall F \in \mathrm{Mult}(V_1, \ldots, V_k, W) : \varphi^{-1}(F) = B^{\otimes \rho} \otimes F(B_\rho).$$

Proof. Equation (2.2) just unwinds the convention (2.6) and applies the universal property of the tensor product 1.1. It is obviously written down without a reference to any basis.
We have to check that φ^{-1} is in fact an inverse for φ. Clearly both maps are linear, so it suffices to check this on a basis. So take arbitrary

$$\begin{aligned} v = B^{\otimes \nu} \otimes c_\mu \in V_1^* \otimes \ldots \otimes V_k^* \otimes W, \\ F \in \mathrm{Mult}(V_1, \ldots, V_k, W), \\ (X, w) = (x_1, \ldots, x_k, w) \in V_1 \times \ldots \times V_k \times W, \end{aligned}$$

and calculate

$$\begin{aligned} (\varphi^{-1} \circ \varphi)(v) &= B^{\otimes \rho} \otimes \big(\varphi(B^{\otimes \nu} \otimes c_\mu)(B_\rho)\big) \\ &= B^{\otimes \rho} \otimes \left(B^{\boxtimes \nu}(B_\rho) c_\mu\right) \overset{2.20}{=} B^{\otimes \rho} \otimes \left(\delta^\nu_\rho c_\mu\right) = v, \\ (\varphi \circ \varphi^{-1})(F)(X) &= \varphi\left(B^{\otimes \rho} \otimes F(B_\rho)\right)(X) \\ &= B^{\boxtimes \rho}(X) F(B_\rho) = F(B^{\boxtimes \rho}(X) B_\rho) \overset{2.21}{=} F(X). \end{aligned}$$

Thus, we are now done. □

This theorem is often stated in a much less general form.

2.24 Corollary.

(i). The map

$$\begin{array}{ccc} V' \times W' & \longrightarrow & V' \boxtimes W' \\ (v', w') & \longmapsto & \big((v, w) \mapsto v'(v) w'(w)\big) \end{array}$$

descends to an isomorphism $V' \otimes W' \to V' \boxtimes W'$.

(ii). The map

$$\begin{array}{ccc} V' \times W & \longrightarrow & \mathrm{Hom}(V, W) \\ (v', w) & \longmapsto & \big((v, w) \mapsto v'(v) w\big) \end{array}$$

descends to an isomorphism $V' \otimes W \to \mathrm{Hom}(V, W)$.

Proof. These are special cases of the map constructed in 2.23. □

We are not yet able to identify tensor products of spaces themselves, but only on their duals. Lemma 2.17 guides us the way to solve that.

2.25 Definition. For any two vector spaces V, W, we define

$$V \boxtimes W := V'' \boxtimes W'' = \mathrm{Mult}(V', W').$$

Notice that the underlying spaces to define $\boxtimes$ are V' and W' here instead of V and W. Therefore the notation

$$\forall (v, w) \in V \otimes W : v \boxtimes w := \iota_V(v) \boxtimes \iota_W(w)$$

makes perfect sense.

2.26 Corollary. We obtain

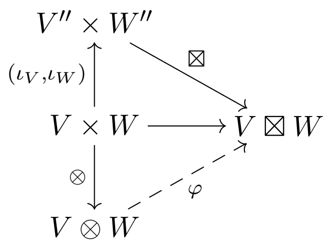
$$V_1 \otimes \ldots \otimes V_k \otimes W \cong V_1'' \otimes \ldots \otimes V_k'' \otimes W \cong \mathrm{Mult}(V_1', \ldots, V_1'; W).$$

Proof. This is just a combination of 2.23 and 2.17. □

Let's exhibit this situation for two spaces in more detail.

2.27 Corollary. The map $V \otimes W \to V \boxtimes W$, $v \otimes w \mapsto v \boxtimes w$, is an isormorphism.

Proof. This follows from 2.23 and 2.25.
But because this is so cool, lets recapitulate the construction of this map and unwind the definitions in a bit more detail: By 2.17 there are isomorphisms $\iota_V : V \to V''$, $\iota_W : W \to W''$. Therefore $(\iota_V, \iota_W) : V \times W \to V'' \times W''$ is an isomorphism as well. We consider $V'' = (V')' = \mathrm{Mult}(V')$ and consequently $\boxtimes : V'' \times W'' \to V \boxtimes W = \mathrm{Mult}(V', W')$ is bilinear. The universal property of the tensor product yields a commutative diagram



By definition the map φ satisfies for any $(v, w) \in V \times W$, $(v', w') \in V' \times W'$

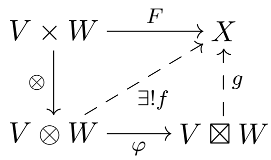
$$\varphi(v \otimes w)(v', w') = (\boxtimes \circ (\iota_V, \iota_W))(v, w)(v', w') = \iota_V(v)(v')\iota_W(w)(w') = v(v')w(w') = v'(v)w'(w),$$

which reveals φ to coincide with 2.23 in that particular situation. □

To make this perfectly clear:

2.28 Corollary. The tuple $(V \boxtimes W, \boxtimes)$ is a tensor product of V and W in the sense of 1.1

Proof. Let $\varphi : V \otimes W \to V \boxtimes W$ be the isomorphism from 2.27. Assume $F : V \times W \to X$ is bilinear. By the universal property of the tensor product



there exists a unique f, such that $\otimes \circ f = F$. Now define $g := f \circ \varphi^{-1}$. □

References

[1] Lee, John M.: "Riemannian manifolds. Introduction to Curvature", Graduate Texts in Mathematics, 176. Springer-Verlag, New York, 1997