Prelude: What are tensors, forms, and the exterior derivative?

If we just jumped into the next few sections, many students (me included!) might say "why in the world are we doing this?" only to discover many chapters (or years!) later that this stuff is actually useful. So, to save you possible mental strain during the next few years, we shall give a leisurely review of elementary vector calculus and show you that you already know the main ideas of tensors, forms, and the exterior derivative. So get a drink, put your feet up, and enjoy.

- **Tensors.** A **tensor** is just a fancy name for a multi-linear real or complex-valued function of vectors,[10] a concept you've already seen throughout elementary calculus and in your physics classes. In the examples below we work with vectors at the tangent space of $\mathbb{R}^3$ at some fixed point that we omit for simplicity. We remark that we could consider tensors in higher dimensions than three, but we don't want to think too hard for now; we'll have our full load of hard work the next section!

Example 2.25. Without a doubt, the easiest examples of tensors are the linear functions of vectors, which is to say, elements of the dual space of vectors; we'll talk about multi-linear functions in the next two examples. We already talked about the dual space in Section 2.3. For example, recall that dx, dy, dz are the dual vectors to $\partial_x, \partial_y, \partial_z$ (at any point in $\mathbb{R}^3$) and we identify

$$\vec{i} \longleftrightarrow \frac{\partial}{\partial x} \quad , \quad \vec{j} \longleftrightarrow \frac{\partial}{\partial y} \quad , \quad \vec{k} \longleftrightarrow \frac{\partial}{\partial z};$$

see our discussion in Section 2.3.1 where we did this for $\mathbb{R}^2$ (for $\mathbb{R}^3$ we just add the vector $\vec{k}$ and another partial ∂_z). Under this identification, dx, dy, dz is the dual basis to $\vec{i}, \vec{j}, \vec{k}$. Hence, for $v = v_1\,\vec{i} + v_2\,\vec{j} + v_3\,\vec{k}$, we have

$$dx(v) = v_1 \quad , \quad dy(v) = v_2 \quad , \quad dz(v) = v_3.$$

In words, $dx(v)$ is the $\vec{i}$ component of v, $dy(v)$ is the $\vec{j}$ component of v, and $dz(v)$ is the $\vec{k}$ component of v. The cotangent vectors dx, dy, dz are examples of **one-tensors** because they eat one vector and return a number. They are also called **one-forms**.

Here is another familiar example.

Example 2.26. Recall that the **cross product** of vectors v and w is defined by

$$v \times w := (v_2\,w_3 - v_3\,w_2)\vec{i} + (v_3\,w_1 - v_1\,w_3)\vec{j} + (v_1\,w_2 - v_2\,w_1)\vec{k}.$$

for $v = v_1\,\vec{i} + v_2\,\vec{j} + v_3\,\vec{k}$ and $w = w_1\,\vec{i} + w_2\,\vec{j} + w_3\,\vec{k}$. The components of the cross-product are tensors; for example, consider the first component

$$\alpha(v, w) := v_2\,w_3 - v_3\,w_2.$$

We can also write α as a determinant:

$$\alpha(v, w) = \det \begin{bmatrix} v_2 & v_3 \\ w_2 & w_3 \end{bmatrix}.$$

This function is **bilinear** because for any vectors u, v, w and $a \in \mathbb{R}$

$$\alpha(au + v, w) = a\,\alpha(u, w) + \alpha(v, w) \ , \ \ \alpha(u, av + w) = a\,\alpha(u, v) + \alpha(u, w).$$

---

[10] Actually, a tensor includes a multi-linear function with codomain a vector space (not just $\mathbb{R}$ or $\mathbb{C}$), but we focus on real or complex values only for simplicity.

Thus, α is a tensor and specifically α is an example of a **two-tensor** because α operates on two vectors to give a number. Notice that α switches signs when its arguments are switched:

$$\alpha(v, w) = -\alpha(w, v);$$

this is because the determinant changes sign when the columns (or rows) are switched. Because α alternates signs, the tensor α is called an **alternating** or **anti-symmetric** tensor. We also call α a **two-form**. Writing

$$v \times w = \alpha(v, w)\,\vec{i} + \beta(v, w)\,\vec{j} + \gamma(v, w)\,\vec{k},$$

we see (by a similar analysis with β and γ) that all the component functions of the cross product are two-forms.

Here is another example you've already heard of.

Example 2.27. Recall that the **scalar triple product** of vectors u, v, w is defined by

$$f(u, v, w) = u \cdot (v \times w),$$

where $v \times w$ is the cross product of v and w and $u \cdot (v \times w)$ is the dot product of u with $v \times w$. The scalar triple product represents the signed volume of the parallel piped formed by u, v, and w; the sign is positive if u, v, and w are "right-handed" otherwise the sign is negative. Written out in gory details, we have

$$f(u, v, w) = u_1(v_2\,w_3 - v_3\,w_2) + u_2(v_3\,w_1 - v_1\,w_3) + u_3(v_1\,w_2 - v_2\,w_1).$$

We can also write the scalar triple product as a determinant:

$$f(u, v, w) = \det \begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix}.$$

The scalar triple product function is **tri-linear** in the sense that for all vectors u, v, w, z and $a \in \mathbb{R}$, we have

$$\begin{aligned} f(au + v, w, z) &= a\,f(u, w, z) + f(v, w, z), \\ f(u, av + w, z) &= a\,f(u, v, z) + f(u, w, z), \\ f(u, v, aw + z) &= a\,f(u, v, w) + f(u, v, z). \end{aligned}$$

Thus, f is an example of a **three-tensor**. Notice that f switches signs whenever any two of its arguments are switched:

$$f(u, v, w) = -f(v, u, w)\ ,\ f(u, v, w) = -f(u, w, v)\ ,\ f(u, v, w) = -f(w, v, u);$$

this is because the determinant changes sign whenever two of its rows (or columns) are switched. Because f alternates signs, f is either called an alternating or anti-symmetric three-tensor, or simply a **three-form**.

Another famous example of a tensor is the **dot product** of vectors; we'll let you see why the dot product of two vectors is a 2-tensor and specifically why it's called a **symmetric** 2-tensor.

A **k-tensor** is simply a function f of k vectors such that f is linear in each argument. A **k-form** is simply a k-tensor that changes sign whenever two of its arguments are switched.

> *Summary: You already know tensors and forms, you just probably never called them that!*

Just from this survey of $\mathbb{R}^3$ you can image that if tensors and forms are so prevalent in Euclidean space how much more they should be important for manifolds, which are generalizations of Euclidean space.

• **The wedge product.** Note that although the triple scalar product has a cross product in its definition, it also has a dot product in it as well, which is of a very different nature than the cross product, and hence the cross product and triple scalar product seem "different." One of the biggest goals in math and physics to find a framework that gives a unified treatment of seemingly "different" objects. For example, the Atiyah-Singer index theorem puts many of the topological-geometric theorems into a single framework and this won Atiyah and Singer the Abel prize in 2004. Physicists nowadays are trying unify the strong, weak, electromagnetic, and gravitational forces into one unified force called the "Grand Unified Theory"; whoever accomplishes such a thing is guaranteed a nobel prize. We shall put the cross product and triple scalar product into one framework and even though this won't give us any prize, it'll make us happy, a prize in itself!

This trick here, which took many years to discover, is to change the cross product just very slightly ... then something magical happens. Recall that the cross product has the following properties:

$$\vec{i}\times\vec{i}=0 \quad , \quad \vec{j}\times\vec{j}=0 \quad , \quad \vec{k}\times\vec{k}=0\ , \tag{2.29}$$

and the cross product is anti-commutative,

$$\vec{i}\times\vec{j}=-\vec{j}\times\vec{i} \quad , \quad \vec{j}\times\vec{k}=-\vec{k}\times\vec{i} \quad , \quad \vec{i}\times\vec{k}=-\vec{k}\times\vec{i}\ . \tag{2.30}$$

Of course, we know from vector calculus that $\vec{i}\times\vec{j}=\vec{k}$, $\vec{j}\times\vec{k}=\vec{i}$, and $\vec{k}\times\vec{i}=\vec{j}$, so the identities in (2.30) revert to the trivial relations $\vec{k}=\vec{k}$, $\vec{i}=\vec{i}$, and $\vec{j}=\vec{j}$.

Now, the aforementioned trick is the following: Let us define another "multiplication" on vectors, which is called the **wedge product** and is denoted by ∧ to distinguish it from the cross product ×, satisfying the *same* properties as (2.29) and (2.30), but without the condition that the product of two vectors be a vector; the product will be "itself," a new entity. Thus, let us formally[11] define

$$\vec{i}\wedge\vec{i}=0 \quad , \quad \vec{j}\wedge\vec{j}=0 \quad , \quad \vec{k}\wedge\vec{k}=0\ , \tag{2.31}$$

and require the wedge product to be anti-commutative,

$$\vec{i}\wedge\vec{j}=-\vec{j}\wedge\vec{i} \quad , \quad \vec{j}\wedge\vec{k}=-\vec{k}\wedge\vec{i} \quad , \quad \vec{i}\wedge\vec{k}=-\vec{k}\wedge\vec{i}\ . \tag{2.32}$$

This multiplication is said to be **anti-commutative** because commuting produces a minus sign. Note that $\vec{i}\wedge\vec{j}$ does *not* equal $\vec{k}$, the entity $\vec{i}\wedge\vec{j}$ is just itself, a "symbolic" vector denoting the wedge product of $\vec{i}$ and $\vec{j}$. We assume that the multiplication given by the wedge product satisfies all the rules of arithmetic that we can think of except of course commutativity: it is associative, distributive over multiplication by real numbers and addition of vectors, and so on. We can also wedge more than three vectors, for example, using (2.32), we have

$$\vec{k}\wedge\vec{i}\wedge\vec{j}=-\vec{i}\wedge\vec{k}\wedge\vec{j}=\vec{i}\wedge\vec{j}\wedge\vec{k}=-\vec{j}\wedge\vec{i}\wedge\vec{k}=\vec{j}\wedge\vec{k}\wedge\vec{i}=-\vec{k}\wedge\vec{j}\wedge\vec{i}. \tag{2.33}$$

---

[11] A "formal computation" in mathematics usually means something like "a symbolic manipulation of an expression without paying attention to rigor nor to meaning". This is very different to the use of "formal" in everyday life, one meaning of which is something like "to perform an act in a proper or correct manner"!

Here, $\vec{k}\wedge\vec{\jmath}\wedge\vec{\imath}$ is to be thought of as a new symbolic vector, not equal to $\vec{\imath}$, $\vec{\imath}\wedge\vec{\jmath}$, etc, but of course equal to all the other vectors in the list (2.33). A nice exercise for you to think about is that the wedge of four or move vectors in $\mathbb{R}^3$ is zero; thus, the wedge product is only interesting if we wedge at most three vectors.

The wedge product admittedly looks "artificial" and non-rigorous, but it will be shown to have some amazing properties in this section especially when we talk about grad, curl, and div later, so in some sense the results justifies its artificial nature. If you're uncomfortable with this, just accept it for now and in fact, later we shall make sense of what exactly the wedge product is (it turns out that, for example, $\vec{\imath}\wedge\vec{\jmath}$ is just a 2-form but this isn't obvious now).

Back to our goal: To give a unified treatment of the cross product and the scalar triple product. First the cross product. Let $v = v_1\,\vec{\imath} + v_2\,\vec{\jmath} + v_3\,\vec{k}$ and $w = w_1\,\vec{\imath} + w_2\,\vec{\jmath} + w_3\,\vec{k}$. Then wedging v and w, we get

$$\begin{aligned} v\wedge w &= (v_1\,\vec{\imath} + v_2\,\vec{\jmath} + v_3\,\vec{k})\wedge(w_1\,\vec{\imath} + w_2\,\vec{\jmath} + w_3\,\vec{k}) \\ &= v_1\,\vec{\imath}\wedge(w_1\,\vec{\imath} + w_2\,\vec{\jmath} + w_3\,\vec{k}) + v_2\,\vec{\jmath}\wedge(w_1\,\vec{\imath} + w_2\,\vec{\jmath} + w_3\,\vec{k}) \\ &\qquad + v_3\,\vec{k}\wedge(w_1\,\vec{\imath} + w_2\,\vec{\jmath} + w_3\,\vec{k}). \end{aligned}$$

To simplify this, we first use that $\vec{\imath}\wedge\vec{\imath} = 0$, $\vec{\jmath}\wedge\vec{\jmath} = 0$, and $\vec{k}\wedge\vec{k} = 0$ to obtain

$$\begin{aligned} v\wedge w = (v_1\,w_2\,\vec{\imath}\wedge\vec{\jmath} + v_1\,w_3\,\vec{\imath}\wedge\vec{k}) + (v_2\,w_1\,\vec{\jmath}\wedge\vec{\imath} + v_2\,w_3\,\vec{\jmath}\wedge\vec{k}) \\ + (v_3\,w_1\,\vec{k}\wedge\vec{\imath} + v_3\,w_2\,\vec{k}\wedge\vec{\jmath}). \end{aligned}$$

Second, using that $\vec{\imath}\wedge\vec{\jmath} = -\vec{\jmath}\wedge\vec{\imath}$, $\vec{\jmath}\wedge\vec{k} = -\vec{k}\wedge\vec{\imath}$, and $\vec{\imath}\wedge\vec{k} = -\vec{k}\wedge\vec{\imath}$, we can write this as

$$\begin{aligned} v\wedge w = v_1\,w_2\,\vec{\imath}\wedge\vec{\jmath} - v_1\,w_3\,\vec{k}\wedge\vec{\imath} - v_2\,w_1\,\vec{\imath}\wedge\vec{\jmath} + v_2\,w_3\,\vec{\jmath}\wedge\vec{k} \\ + v_3\,w_1\,\vec{k}\wedge\vec{\imath} - v_3\,w_2\,\vec{\jmath}\wedge\vec{k}. \end{aligned}$$

Finally, combining like terms, we obtain

$$v\wedge w = (v_2\,w_3 - v_3\,w_2)\,\vec{\jmath}\wedge\vec{k} + (v_3\,w_1 - v_1\,w_3)\,\vec{k}\wedge\vec{\imath} + (v_1\,w_2 - v_2\,w_1)\,\vec{\imath}\wedge\vec{\jmath}. \tag{2.34}$$

By convention we keep the order of $\vec{\imath}$, $\vec{\jmath}$, $\vec{k}$ "circular" as shown in the following diagram:

$$\begin{tikzcd} \vec{\imath} \arrow[rr] & & \vec{\jmath} \arrow[dl] \\ & \vec{k} \arrow[ul] & \end{tikzcd} \tag{2.35}$$

Now the cross product of v and w is

$$v\times w = (v_2\,w_3 - v_3\,w_2)\vec{\imath} + (v_3\,w_1 - v_1\,w_3)\vec{\jmath} + (v_1\,w_2 - v_2\,w_1)\vec{k}.$$

Thus, if in (2.34) we replace $\vec{\jmath}\wedge\vec{k}$ with $\vec{\imath}$, $\vec{k}\wedge\vec{\imath}$ with $\vec{\jmath}$, and $\vec{\imath}\wedge\vec{\jmath}$ with $\vec{k}$ (that is, the "next" vector in the diagram (2.35)) then (2.34) is exactly the cross product of v and w! This isn't really that surprising because, after all, the cross product also satisfies similar anti-commutativity properties as the wedge product, so we can do the same computation as we did above substituting × for ∧ everywhere and come out with the same answer.

The real difference comes with the scalar triple product. Let $u = u_1\,\vec{i} + u_2\,\vec{j} + u_3\,\vec{k}$ be another vector. It turns out that the scalar triple product of u, v, w is just $u \wedge v \wedge w$. To see this, we analyze

$$\begin{aligned} u \wedge v \wedge w = {} & (u_1\,\vec{i} + u_2\,\vec{j} + u_3\,\vec{k}) \wedge \\ & \Big((v_2\,w_3 - v_3\,w_2)\,\vec{j} \wedge \vec{k} + (v_3\,w_1 - v_1\,w_3)\,\vec{k} \wedge \vec{i} + (v_1\,w_2 - v_2\,w_1)\,\vec{i} \wedge \vec{j} \Big). \end{aligned}$$

When we wedge $u = u_1\,\vec{i} + u_2\,\vec{j} + u_3\,\vec{k}$ with the first term $(v_2\,w_3 - v_3\,w_2)\,\vec{j} \wedge \vec{k}$ of $v \wedge w$, only the $\vec{i} \wedge \vec{j} \wedge \vec{k}$ term remains because

$$\vec{j} \wedge \vec{j} \wedge \vec{k} = 0 \quad \text{and} \quad \vec{k} \wedge \vec{j} \wedge \vec{k} = -\vec{j} \wedge \vec{k} \wedge \vec{k} = 0.$$

There are similar statements when we wedge $u = u_1\,\vec{i} + u_2\,\vec{j} + u_3\,\vec{k}$ with the second and third terms of $v \wedge w$. Thus, only keeping the nonzero terms, we get

$$\begin{aligned} u \wedge v \wedge w = u_1(v_2\,w_3 - v_3\,w_2)\,\vec{i} \wedge \vec{j} \wedge \vec{k} + u_2(v_3\,w_1 - v_1\,w_3)\,\vec{j} \wedge \vec{k} \wedge \vec{i} \\ + u_3(v_1\,w_2 - v_2\,w_1)\,\vec{k} \wedge \vec{i} \wedge \vec{j}. \end{aligned}$$

Using anti-commutativity, we see that

$$\vec{j} \wedge \vec{k} \wedge \vec{i} = -\vec{j} \wedge \vec{i} \wedge \vec{k} = \vec{i} \wedge \vec{j} \wedge \vec{k} \quad , \quad \vec{k} \wedge \vec{i} \wedge \vec{j} = -\vec{i} \wedge \vec{k} \wedge \vec{j} = \vec{i} \wedge \vec{j} \wedge \vec{k}.$$

Thus,

$$\begin{aligned} & u \wedge\ v \wedge w \\ & = \Big(u_1(v_2\,w_3 - v_3\,w_2) + u_2(v_3\,w_1 - v_1\,w_3) + u_3(v_1\,w_2 - v_2\,w_1) \Big) \vec{i} \wedge \vec{j} \wedge \vec{k} \qquad (2.36) \\ & = \left(\det \begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix} \right) \vec{i} \wedge \vec{j} \wedge \vec{k}. \end{aligned}$$

The expression in the parentheses is exactly the scalar triple product of u, v, w! Thus, up to the symbolic factor $\vec{i} \wedge \vec{j} \wedge \vec{k}$, the wedge product of u, v, w is exactly the scalar triple product of u, v, w!

Summarizing what we've found so far:

$$\text{the cross product of } v, w \quad \text{“} = \text{”} \quad v \wedge w,$$

$$\text{the scalar triple product of } u, v, w \quad \text{“} = \text{”} \quad u \wedge v \wedge w.$$

From this point of view, we have found a “unified” framework that treats both the cross product and the scalar triple product: the cross product is just wedging two vectors while the scalar triple product is just wedging three vectors! They only differ by the number of wedges one does! Although innocent it may seem, make no mistake this “unification” is really a very deep result, especially as we consider ...

- **Grad, curl, and div.** We now explain the origin of the exterior derivative that we'll introduce for manifolds later. Recall that the **gradient** of a function f on $\mathbb{R}^3$ is defined by

$$\nabla f := \partial_x f\,\vec{i} + \partial_y f\,\vec{j} + \partial_z f\,\vec{k}.$$

The **curl** of a vector field $F = P\,\vec{i} + Q\,\vec{j} + R\,\vec{k}$, where P, Q, R are functions on $\mathbb{R}^3$, is defined by

$$\operatorname{curl} F := (\partial_y R - \partial_z Q)\vec{i} + (\partial_z P - \partial_x R)\vec{j} + (\partial_x Q - \partial_y P)\vec{k}. \qquad (2.37)$$

Finally, the **divergence** of F is

$$\operatorname{div} F := \partial_x P + \partial_y Q + \partial_z R. \tag{2.38}$$

The gradient, curl, and divergence look, at first sight, very different. However, many years ago someone (I don't know who) made a deep observation: Grad, curl, and div are actually one and the same! As we already mentioned, it's a big thing in math and physics to unify seemingly different ideas so we can imagine the elation people had when the main operators of vector calculus were unified. It turns out that they are actually the same differential operator

$$\nabla = \partial_x \vec{i} + \partial_y \vec{j} + \partial_z \vec{k}$$

as we now shall see. Of course, we can see that the gradient is just ∇ applied to a function:

$$\nabla(f) = (\partial_x \vec{i} + \partial_y \vec{j} + \partial_z \vec{k})(f) = \partial_x f \vec{i} + \partial_y f \vec{j} + \partial_z f \vec{k}.$$

Now what about the curl? Formally applying ∇ to $F = P\vec{i} + Q\vec{j} + R\vec{k}$ we get

$$\begin{aligned} \nabla(F) &= \nabla\Big(P\vec{i} + Q\vec{j} + R\vec{k}\Big) \\ &= (\nabla P)\vec{i} + (\nabla Q)\vec{j} + (\nabla R)\vec{k}. \end{aligned} \tag{2.39}$$

Now, ∇P is a vector, namely the vector (field)

$$\nabla P = \partial_x P \vec{i} + \partial_y P \vec{j} + \partial_z P \vec{k}$$

and so is $\vec{i}$ (similar for ∇Q and $\vec{j}$, and also ∇R and $\vec{k}$), so how do we interpret multiplication of vectors in the expression (2.39)? Well, we had success before with the wedge product, we interpret multiplication in (2.39) as wedge product:

$$\nabla(F) = (\nabla P) \wedge \vec{i} + (\nabla Q) \wedge \vec{j} + (\nabla R) \wedge \vec{k}.$$

Thus, multiplying (really wedging) ∇P and $\vec{i}$, ∇Q and $\vec{j}$, and ∇R and $\vec{k}$, and remembering that $\vec{i} \wedge \vec{i} = 0$, $\vec{j} \wedge \vec{j} = 0$, and $\vec{k} \wedge \vec{k} = 0$, and that $\vec{i} \wedge \vec{j} = -\vec{j} \wedge \vec{i}$, $\vec{j} \wedge \vec{k} = -\vec{k} \wedge \vec{i}$, and $\vec{i} \wedge \vec{k} = -\vec{k} \wedge \vec{i}$, we see that

$$\begin{aligned} \nabla(F) &= (\partial_y P \vec{j} \wedge \vec{i} \;+\; \partial_z P \vec{k} \wedge \vec{i}) + (\partial_x Q \vec{i} \wedge \vec{j} + \partial_z Q \vec{k} \wedge \vec{j}) \\ &\qquad\qquad + (\partial_x R \vec{i} \wedge \vec{k} + \partial_y R \vec{j} \wedge \vec{k}) \\ &= -\partial_y P \vec{i} \wedge \vec{j} \;+\; \partial_z P \vec{k} \wedge \vec{i} + \partial_x Q \vec{i} \wedge \vec{j} - \partial_z Q \vec{j} \wedge \vec{k} \\ &\qquad\qquad - \partial_x R \vec{k} \wedge \vec{i} + \partial_y R \vec{j} \wedge \vec{k}. \end{aligned}$$

By convention we keep the order of $\vec{i}$, $\vec{j}$, $\vec{k}$ "circular":

$$\begin{tikzcd} \vec{i} \arrow[rr] & & \vec{j} \arrow[dl] \\ & \vec{k} \arrow[ul] & \end{tikzcd}$$

Finally, combining like terms we get

$$\nabla(F) = (\partial_y R - \partial_z Q)\vec{j} \wedge \vec{k} + (\partial_z P - \partial_x R)\vec{k} \wedge \vec{i} + (\partial_x Q - \partial_y P)\vec{i} \wedge \vec{j}. \tag{2.40}$$

Now this is in exact agreement with our definition (2.37) of curl by replacing $\vec{j} \wedge \vec{k}$ with $\vec{i}$, $\vec{k} \wedge \vec{i}$ with $\vec{j}$, and $\vec{i} \wedge \vec{j}$ with $\vec{k}$!

Lastly, we come to the divergence. The trick here is to write a vector field $F = P\vec{\imath} + Q\vec{\jmath} + R\vec{k}$ in the way considered in (2.40) above by replacing $\vec{\imath}$, $\vec{\jmath}$, and $\vec{k}$ with $\vec{\jmath}\wedge\vec{k}$, $\vec{k}\wedge\vec{\imath}$, and $\vec{\imath}\wedge\vec{\jmath}$, respectively:

$$F = P\,\vec{\jmath}\wedge\vec{k} + Q\,\vec{k}\wedge\vec{\imath} + R\,\vec{\imath}\wedge\vec{\jmath}.$$

Then formally applying ∇ to F term by term and using $\wedge$ are our interpretation of multiplication, we get

$$\begin{aligned}\nabla(F) &= \nabla(P\,\vec{\jmath}\wedge\vec{k} + Q\,\vec{k}\wedge\vec{\imath} + R\,\vec{\imath}\wedge\vec{\jmath})\\ &= (\nabla P)\wedge\vec{\jmath}\wedge\vec{k} + (\nabla Q)\wedge\vec{k}\wedge\vec{\imath} + (\nabla R)\wedge\vec{\imath}\wedge\vec{\jmath}\\ &= \partial_x P\,\vec{\imath}\wedge\vec{\jmath}\wedge\vec{k} + \partial_y Q\,\vec{\jmath}\wedge\vec{k}\wedge\vec{\imath} + \partial_z R\,\vec{k}\wedge\vec{\imath}\wedge\vec{\jmath},\end{aligned}$$

where we used that all the terms with two repeated vectors zero out. Finally, using the relations

$$\vec{\imath}\wedge\vec{\jmath} = -\vec{\jmath}\wedge\vec{\imath} \quad , \quad \vec{\jmath}\wedge\vec{k} = -\vec{k}\wedge\vec{\imath} \quad , \quad \vec{\imath}\wedge\vec{k} = -\vec{k}\wedge\vec{\imath}\ ,$$

we have

$$\vec{\jmath}\wedge\vec{k}\wedge\vec{\imath} = -\vec{\jmath}\wedge\vec{\imath}\wedge\vec{k} = \vec{\imath}\wedge\vec{\jmath}\wedge\vec{k},$$

and

$$\vec{k}\wedge\vec{\imath}\wedge\vec{\jmath} = -\vec{\imath}\wedge\vec{k}\wedge\vec{\jmath} = \vec{\imath}\wedge\vec{\jmath}\wedge\vec{k},$$

and hence,

$$\begin{aligned}\nabla(F) &= \partial_x P\,\vec{\imath}\wedge\vec{\jmath}\wedge\vec{k} + \partial_y Q\,\vec{\imath}\wedge\vec{\jmath}\wedge\vec{k} + \partial_z R\,\vec{\imath}\wedge\vec{\jmath}\wedge\vec{k}\\ &= (\partial_x P + \partial_y Q + \partial_z R)\,\vec{\imath}\wedge\vec{\jmath}\wedge\vec{k}.\end{aligned}$$

Without the term $\vec{\imath}\wedge\vec{\jmath}\wedge\vec{k}$ and comparing to (2.38), we can see that $\nabla(F)$ is exactly the divergence of F!

• **The unification of grad, curl, div: A summary.** OK, let's summarize the main ideas of our unification process. Let $\vec{\imath}$, $\vec{\jmath}$, and $\vec{k}$ have a multiplication "$\wedge$" governed by the rules

$$\vec{\imath}\wedge\vec{\imath} = 0 \quad , \quad \vec{\jmath}\wedge\vec{\jmath} = 0 \quad , \quad \vec{k}\wedge\vec{k} = 0\ ,$$

and

$$\vec{\imath}\wedge\vec{\jmath} = -\vec{\jmath}\wedge\vec{\imath} \quad , \quad \vec{\jmath}\wedge\vec{k} = -\vec{k}\wedge\vec{\imath} \quad , \quad \vec{\imath}\wedge\vec{k} = -\vec{k}\wedge\vec{\imath}\ .$$

Define

$$\nabla = \partial_x\,\vec{\imath} + \partial_y\,\vec{\jmath} + \partial_z\,\vec{k}.$$

Then ∇ applied to a function is

$$\nabla(f) = \partial_x f\,\vec{\imath} + \partial_y f\,\vec{\jmath} + \partial_z f\,\vec{k},$$

the gradient, which is not so surprising. Also, we saw that ∇ applied to a vector of the form

$$F = P\,\vec{\imath} + Q\,\vec{\jmath} + R\,\vec{k}$$

is

$$\nabla(F) = (\partial_y R - \partial_z Q)\,\vec{\jmath}\wedge\vec{k} + (\partial_z P - \partial_x R)\,\vec{k}\wedge\vec{\imath} + (\partial_x Q - \partial_y P)\,\vec{\imath}\wedge\vec{\jmath},$$

which is basically the curl of F. Finally, ∇ applied to a vector of the form

$$F = P\,\vec{\jmath}\wedge\vec{k} + Q\,\vec{k}\wedge\vec{\imath} + R\,\vec{\imath}\wedge\vec{\jmath}. \tag{2.41}$$

is

$$\nabla(F) = (\partial_x P + \partial_y Q + \partial_z R)\,\vec{\imath} \wedge \vec{\jmath} \wedge \vec{k}, \tag{2.42}$$

which is basically the divergence of F. We now generalize this to manifolds.

• **Relations to manifolds.** At the end of Section 2.3 in our discussions concerning the cotangent space, we saw that on any arbitrary abstract manifold there is a God-given operator that emulates the gradient of a function in Euclidean space — the differential of a function. For a quick review, recall that the differential of a smooth function f is a completely natural (coordinate independent) object defined in (2.22), which in local coordinates $(\mathcal{U}, x)$ on an n-dimensional manifold M happens to look like an "n-dimensional gradient":

$$df = \partial_{x_1} f\, dx_1 + \partial_{x_2} f\, dx_2 + \cdots + \partial_{x_n} f\, dx_n.$$

Therefore, in local coordinates this "gradient" operator is just

$$d = \partial_{x_1} dx_1 + \partial_{x_2} dx_2 + \cdots + \partial_{x_n} dx_n. \tag{2.43}$$

For example, in $\mathbb{R}^3$, this is just

$$d = \partial_x\, dx + \partial_y\, dy + \partial_z\, dz,$$

which is exactly the gradient except we have to replace dx with $\vec{\imath}$, dy with $\vec{\jmath}$, and dz with $\vec{k}$. Because the differential is intrinsically defined on any manifold, we should really study the differential rather than the gradient $\nabla = \partial_x\,\vec{\imath} + \partial_y\,\vec{\jmath} + \partial_z\,\vec{k}$, the object studied in elementary vector calculus. (Actually, many geometers consider it to be a historical mistake that vector calculus is taught using tangent spaces instead of cotangent spaces!)

Thus, as we did with $\vec{\imath}, \vec{\jmath}, \vec{k}$, we should define a multiplication "$\wedge$" on dx, dy, dz governed by the rules

$$dx \wedge dx = 0 \quad , \quad dy \wedge dy = 0 \quad , \quad dz \wedge dz = 0 \ ,$$

and

$$dx \wedge dy = -dy \wedge dx \quad , \quad dy \wedge dz = -dz \wedge dx \quad , \quad dx \wedge dz = -dz \wedge dx.$$

Let α be a section of the cotangent space of $\mathbb{R}^3$ (a one-form):

$$\alpha = P\, dx + Q\, dy + R\, dz,$$

where P, Q, R are smooth functions on $\mathbb{R}^3$. Since dx, dy, dz and $\vec{\imath}, \vec{\jmath}, \vec{k}$ have identical multiplication properties, following word-for-word the computation used to find (2.41), if we apply d to α we obtain

$$\begin{aligned} d\alpha &= d\big(P\, dx + Q\, dy + R\, dz\big) \\ &= (dP) \wedge dx + (dQ) \wedge dy + (dR) \wedge dz \\ &= (\partial_y R - \partial_z Q)\, dy \wedge dz + (\partial_z P - \partial_x R)\, dz \wedge dx + (\partial_x Q - \partial_y P)\, dx \wedge dy, \end{aligned}$$

which is the "differential" version of the curl of a vector field. Here, $d\alpha$ is called a "(smooth differential) two-form." Finally, d applied to a two-form

$$\alpha = P\, dy \wedge dz + Q\, dz \wedge dx + R\, dx \wedge dy,$$

following word-for-word the computation used to find (2.42), we obtain

$$\begin{aligned} d\alpha &= d\big(P\,dy\wedge dz + Q\,dz\wedge dx + R\,dx\wedge dy\big)\\ &= (dP)\wedge\, dy\wedge dz + (dQ)\wedge dz\wedge dx + (dR)\wedge dx\wedge dy)\\ &= (\partial_x P + \partial_y Q + \partial_z R)\,dx\wedge dy\wedge dz, \end{aligned}$$

which is the "differential" version of the divergence of a vector field. Here, $d\alpha$ is called a "(smooth differential) three-form." When the differential d in (2.43) operates on smooth k-forms we call d the **exterior derivative**. In particular, for $\mathbb{R}^3$, the exterior derivative in a sense contains the gradient, curl, and divergence! In Section 2.7 we shall study the exterior derivative in great detail for general manifolds.

• **What exactly is the wedge product?** Now comes the obvious question: What in the world is $dx\wedge dy$, $dx\wedge dy\wedge dz$, and the other wedge products? When I was an undergrad, I took a class in differential forms and we were never told! So, I will reveal the mystery!

Let us start with something we already know: dx, dy, dz. Focusing on dx, we know that at any fixed point $p\in\mathbb{R}^3$ (see Example 2.25)

$$dx : T_p\mathbb{R}^3 \to \mathbb{R}$$

is the map

$$dx(v) = v_1 \quad\text{for}\quad v = v_1\,\vec{\imath} + v_2\,\vec{\jmath} + v_3\,\vec{k};$$

that is,

$$dx(v) \text{ is the } \vec{\imath} \text{ component of } v.$$

Similarly, $dy(v) = v_2$ and $dz(v) = v_3$. This is old news. Now what about $dx\wedge dy$? Well, since dx and dy are defined on $T_p\mathbb{R}^3$, one "obvious" candidate is to have $dx\wedge dy$ be defined on $T_p\mathbb{R}^3\times T_p\mathbb{R}^3$; so let us consider

$$dx\wedge dy : T_p\mathbb{R}^3\times T_p\mathbb{R}^3 \to \mathbb{R}.$$

Now the question is: What is

$$(dx\wedge dy)(v,w) = \ ? \quad\text{for vectors } v, w.$$

To derive what this might possibly be, recall that for a vector v, $dx(v)$ is the $\vec{\imath}$ component of v, $dy(v)$ is the $\vec{\jmath}$ component of v, and $dz(v)$ is the $\vec{k}$ component of v — the point is to associate dx with $\vec{\imath}$, dy with $\vec{\jmath}$, and dz with $\vec{k}$. Therefore, in analogy, one "obvious" candidate for $(dx\wedge dy)(v,w)$ is that

$$(dx\wedge dy)(v,w) \text{ is the } \vec{\imath}\wedge\vec{\jmath} \text{ component of } v\wedge w.$$

Recall from (2.34) that

$$v\wedge w = (v_2\,w_3 - v_3\,w_2)\,\vec{\jmath}\wedge\vec{k} + (v_3\,w_1 - v_1\,w_3)\,\vec{k}\wedge\vec{\imath} + (v_1\,w_2 - v_2\,w_1)\,\vec{\imath}\wedge\vec{\jmath}.$$

Therefore, it seems natural to *define* $(dx\wedge dy)(v,w)$ as

$$(dx\wedge dy)(v,w) := v_1\,w_2 - v_2\,w_1,$$

for all $v = v_1\,\vec{\imath} + v_2\,\vec{\jmath} + v_3\,\vec{k}$ and $w = w_1\,\vec{\imath} + w_2\,\vec{\jmath} + w_3\,\vec{k}$. By Example 2.26, with this definition, we know that $dx\wedge dy$ is a two-form! Analogously,

$$(dy\wedge dz)(v,w) \text{ is the } \vec{\jmath}\wedge\vec{k} \text{ component of } v\wedge w$$

and

$$(dz \wedge dx)(v, w) \text{ is the } \vec{k} \wedge \vec{i} \text{ component of } v \wedge w.$$

Therefore, in view of the formula above for $v \wedge w$, it makes sense to *define*

$$\begin{aligned} (dy \wedge dz)(v, w) &:= v_2\, w_3 - v_3\, w_2 \\ (dz \wedge dx)(v, w) &:= v_3\, w_1 - v_1\, w_3. \end{aligned} \tag{2.44}$$

Actually, we can do better than this: we can define the wedge product of any two one-forms! As motivation consider, for example, the formula

$$(dx \wedge dy)(v, w) = v_1\, w_2 - v_2\, w_1.$$

Since $v_1 = dx(v)$, $w_2 = dy(w)$, $w_1 = dx(w)$, and $v_2 = dy(v)$, we can write

$$(dx \wedge dy)(v, w) = dx(v)\, dy(w) - dy(v)\, dx(w).$$

This motivates the following definition: Given any one-forms $\alpha : T_p\mathbb{R}^3 \to \mathbb{R}$ and $\beta : T_p\mathbb{R}^3 \to \mathbb{R}$, we define

$$\alpha \wedge \beta : T_p\mathbb{R}^3 \times T_p\mathbb{R}^3 \to \mathbb{R}$$

as the map

$$(\alpha \wedge \beta)(v, w) := \alpha(v)\, \beta(w) - \beta(v)\, \alpha(w) = \det \begin{bmatrix} \alpha(v) & \alpha(w) \\ \beta(v) & \beta(w) \end{bmatrix}. \tag{2.45}$$

Since α and β are linear, $\alpha \wedge \beta$ is certainly bilinear so is a two-tensor, and since

$$(\alpha \wedge \beta)(w, v) = \det \begin{bmatrix} \alpha(w) & \alpha(v) \\ \beta(w) & \beta(v) \end{bmatrix} = -\det \begin{bmatrix} \alpha(v) & \alpha(w) \\ \beta(v) & \beta(w) \end{bmatrix} = -(\alpha \wedge \beta)(v, w),$$

$\alpha \wedge \beta$ is a two-form. Moreover, the definition shows that $\alpha \wedge \alpha = 0$ and

$$(\alpha \wedge \beta)(v, w) := \det \begin{bmatrix} \alpha(v) & \alpha(w) \\ \beta(v) & \beta(w) \end{bmatrix} = -\det \begin{bmatrix} \beta(v) & \beta(w) \\ \alpha(v) & \alpha(w) \end{bmatrix} = -(\beta \wedge \alpha)(v, w),$$

therefore $\alpha \wedge \beta = -\beta \wedge \alpha$. Notice that the definition (2.45) agrees with the definition of $dy \wedge dz$ and $dz \wedge dx$ in (2.44).

Now how do we define $dx \wedge dy \wedge dz$? Well, proceeding by analogy, we should consider $dx \wedge dy \wedge dz$ as a map

$$dx \wedge dy \wedge dz : T_p\mathbb{R}^3 \times T_p\mathbb{R}^3 \times T_p\mathbb{R}^3 \to \mathbb{R}$$

and more specifically, this map has to be (based on our motivation above)

$$(dx \wedge dy \wedge dz)(u, v, w) \text{ is the } \vec{i} \wedge \vec{j} \wedge \vec{k} \text{ component of } u \wedge v \wedge w.$$

By formula (2.36), we conclude that we should *define*

$$(dx \wedge dy \wedge dz)(u, v, w) := \det \begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix}.$$

Since we can write this as

$$(dx \wedge dy \wedge dz)(u, v, w) := \det \begin{bmatrix} dx(u) & dx(v) & dx(w) \\ dz(u) & dy(v) & dy(w) \\ dz(u) & dz(v) & dz(w) \end{bmatrix},$$

this motivates the following definition: Given any one-forms $\alpha : T_p\mathbb{R}^3 \to \mathbb{R}$, $\beta : T_p\mathbb{R}^3 \to \mathbb{R}$, and $\gamma : T_p\mathbb{R}^3 \to \mathbb{R}$, we define

$$\alpha \wedge \beta \wedge \gamma : T_p\mathbb{R}^3 \times T_p\mathbb{R}^3 \times T_p\mathbb{R}^3 \to \mathbb{R}$$

as the map

$$(\alpha \wedge \beta \wedge \gamma)(u, v, w) := \det \begin{bmatrix} \alpha(u) & \alpha(v) & \alpha(w) \\ \beta(u) & \beta(v) & \beta(w) \\ \gamma(u) & \gamma(v) & \gamma(w) \end{bmatrix}. \tag{2.46}$$

Using this definition and the fact that the determinant changes whenever any two columns or any two rows are switched, one can check that $\alpha \wedge \beta \wedge \gamma$ is a three-form and the following identities hold:

$$\alpha \wedge \beta \wedge \gamma = -\beta \wedge \alpha \wedge \gamma = \beta \wedge \gamma \wedge \alpha = \cdots,$$

a minus sign is introduced whenever to adjacent terms are switched.

Summarizing: The definitions (2.45) and (2.46) give completely rigorous definitions of the wedge product of one-forms! So, we have now made the computations in the previous section "**Relations to manifolds**" completely rigorous!

Now what about $\vec{i} \wedge \vec{j}$ and the other wedge products of vector fields that we did before? We can define their wedge products by duality. For example, let $v, w \in T_p\mathbb{R}^3$ and let us consider $v \wedge w$. We define $v \wedge w$ as a map

$$v \wedge w : T_p^*\mathbb{R}^3 \times T_p^*\mathbb{R}^3 \to \mathbb{R}$$

defined by

$$(v \wedge w)(\alpha, \beta) := (\alpha \wedge \beta)(v, w) \quad \text{for all} \quad (\alpha, \beta) \in T_p^*\mathbb{R}^3 \times T_p^*\mathbb{R}^3.$$

Of course, we already know what $(\alpha \wedge \beta)(v, w)$ means so $(v \wedge w)(\alpha, \beta)$ is well-defined. One can check that $v \wedge w$ defined in this way is a two-form on the vector space $T_p^*\mathbb{R}^3$ and that $w \wedge v = -v \wedge w$. Similarly, we can define the wedge product of three vectors. The end result is: If we define the wedge product in this rigorous way, then all the computations we did in this section actually work!

Enough for motivation, let's get down to real work.

2.5. The tensor algebra

In this section we study the tensor algebra needed in future sections.

SECTION OBJECTIVES: THE STUDENT WILL BE ABLE TO ...

- explain what tensors are.
- describe a basis for tensors.
- know algebraic properties of direct sums and tensor products.
- state the universal property of the tensor product.

2.5.1. Tensors and double duals. To define tensors, there are two roads to follow: The high road for sophisticated folks (involving free vector spaces and a little abstract nonsense)[12] and the low road for the common "pedestrian" folks like me (involving familiar notions such as linear maps). We shall follow the low road in the main text and leave the high road for the exercises; see Problems 2 and 3.

Let $V_1, \ldots, V_k$ be finite-dimensional $\mathbb{K}$ ($= \mathbb{R}$ or $\mathbb{C}$) vector spaces. A map

$$f : V_1 \times V_2 \times \cdots \times V_k \to \mathbb{K}$$

---

[12] A mathematical argument might be called "abstract nonsense" if it involves long-winded theoretical steps that hold because "they do" either because they are built in to a definition, they involve some universal property, a diagram forces the statement to hold, etc.

is called a **tensor**, or in more familiar language **multi-linear**, if it is linear in each factor, that is, for any $j = 1, \dots, k$, we have

$$\begin{aligned} f(v_1, \dots, v_{j-1}, a\, v_j + w_j, v_{j+1}, \dots, v_k) = a\, f(v_1, \dots, v_{j-1}, v_j, v_{j+1}, \dots, v_k) \\ + f(v_1, \dots, v_{j-1}, w_j, v_{j+1}, \dots, v_k). \end{aligned}$$

Thus, as remarked in the previous section, "tensor" is just a fancy word for multi-linear, something you are already familiar with. Here are some examples mentioned in the previous section.

Example 2.28. If $k = 1$, a tensor $f : V \to \mathbb{K}$ is simply a linear map, which means that $f \in V^*$, the dual space. Another name for an element of V^* is a **one-form**.

Example 2.29. If $k = 2$, a tensor is a map $f : V \times W \to \mathbb{K}$ that satisfies

$$f(a\, v_1 + v_2, w) = a\, f(v_1, w) + f(v_2, w)$$

and

$$f(v, a\, w_1 + w_2) = a\, f(v, w_1) + f(v, w_2);$$

note that such a map is usually called a **bilinear** map.

Example 2.30. If V is a real vector space, recall that a map $g : V \times V \to \mathbb{R}$ is an **inner product** means that g is a bilinear map, and moreover, it's symmetric and positive definite:

a) $g(v, w) = g(w, v)$ for all $v, w \in V$;
b) $g(v, v) \geq 0$ for all $v \in V$ and $g(v, v) = 0$ iff $v = 0$.

Being bilinear, g is a tensor and since g is also symmetric it's called a symmetric tensor. We'll study inner products in Section 2.8.

In the usual way we can add tensors and multiply them by scalars and still remain in the class of tensors; e.g. if f and g are tensors, $a \in \mathbb{K}$, and $v \in V_1 \times \cdots \times V_k$, then

$$(f + g)(v) := f(v) + g(v)\ , \quad (af)(v) := a\, f(v).$$

We denote by

$$V_1^* \otimes V_2^* \otimes \cdots \otimes V_k^*$$

the vector space of all such multi-linear maps and it is called the **tensor product** of $V_1^*, \dots, V_k^*$. Note that if $k = 1$ so we only have a single vector space V, then the tensor product is denoted by V^*, the same notation for dual space, which of course is consistent with the fact that tensors in this case are exactly elements of the dual space of V as we saw in the first example above.

Elements $\alpha_1 \in V_1^*, \dots, \alpha_k \in V_k^*$ define an element of $V_1^* \otimes \cdots \otimes V_k^*$ as follows: We define

$$\alpha_1 \otimes \cdots \otimes \alpha_k : V_1 \times \cdots \times V_k \to \mathbb{K}$$

by

$$\boxed{(\alpha_1 \otimes \cdots \otimes \alpha_k)(v_1, \dots, v_k) := \alpha_1(v_1) \cdot \alpha_2(v_2) \cdots \alpha_k(v_k)} \tag{2.47}$$

for all $(v_1, \dots, v_k) \in V_1 \times \cdots \times V_k$. Since $\alpha_1, \dots, \alpha_k$ are linear, one can check that $\alpha_1 \otimes \cdots \otimes \alpha_k$ is multi-linear and hence gives an element of $V_1^* \otimes \cdots \otimes V_k^*$, which we call the **tensor product** of $\alpha_1, \ldots, \alpha_k$. It's easy to check that the tensor product (2.47) is multi-linear; that is, linear in each factor: For any $1 \leq j \leq k$,

$$\begin{aligned}\alpha_1 \otimes \cdots \otimes \alpha_{j-1} \otimes (a\alpha_j + \beta_j) \otimes \alpha_{j+1} \otimes \cdots \otimes \alpha_k = \\ a\,(\alpha_1 \otimes \cdots \otimes \alpha_j \otimes \cdots \otimes \alpha_k) \;+\; \alpha_1 \otimes \cdots \otimes \beta_j \otimes \cdots \otimes \alpha_k.\end{aligned} \tag{2.48}$$

Example 2.31. Recall from (2.45) of the previous section that given any one-forms $\alpha : T_p\mathbb{R}^3 \to \mathbb{R}$ and $\beta : T_p\mathbb{R}^3 \to \mathbb{R}$, we define

$$\alpha \wedge \beta : T_p\mathbb{R}^3 \times T_p\mathbb{R}^3 \to \mathbb{R}$$

as the map

$$(\alpha \wedge \beta)(v, w) := \alpha(v)\,\beta(w) - \beta(v)\,\alpha(w).$$

In tensor notation, this simply means

$$\alpha \wedge \beta := \alpha \otimes \beta - \beta \otimes \alpha.$$

Summarizing what we've done so far:

$$\boxed{\begin{gathered} V_1^* \otimes V_2^* \otimes \cdots \otimes V_k^* = \big\{\, \mathbb{K}\text{-valued multi-linear maps on } \; V_1 \times \cdots \times V_k \,\big\}, \\ \alpha_i \in V_i^* \quad \Longrightarrow \quad \alpha_1 \otimes \alpha_2 \otimes \cdots \otimes \alpha_k \in V_1^* \otimes V_2^* \otimes \cdots \otimes V_k^*. \end{gathered}}$$

In words,

multi-linear maps on the product of vector spaces define
elements of the tensor product of the *dual* vector spaces.

A couple remarks: (1) This (going between vector spaces and duals) may seem confusing at first, however our definition is one of the more elementary standard definitions and moreover it's extremely user-friendly. (2) It seems like we have only defined the tensor product of dual vector spaces! In particular, what does

$$V_1 \otimes \cdots \otimes V_k$$

mean? According to our definition, we need $V_1 = W_1^*, \ldots, V_k = W_k^*$, then

$$V_1 \otimes V_2 \otimes \cdots \otimes V_k = \{\mathbb{K}\text{-valued multi-linear maps on } \; W_1 \times \cdots \times W_k\}.$$

It turns out that $W_1 = V_1^*, \ldots, W_k = V_k^*$, that is, $V_1 = (V_1^*)^*, \ldots, V_k = (V_k^*)^*$. In the following proposition we prove that for any finite-dimensional vector space V, we have $V = (V^*)^*$ or more succinctly, $V = V^{**}$. (Here, "=" means "canonically isomorphic".) Thus, our summary above becomes

$$\boxed{\begin{gathered} V_1 \otimes V_2 \otimes \cdots \otimes V_k = \big\{\, \mathbb{K}\text{-valued multi-linear maps on } \; V_1^* \times \cdots \times V_k^* \,\big\}, \\ v_i \in V_i \quad \Longrightarrow \quad v_1 \otimes v_2 \otimes \cdots \otimes v_k \in V_1 \otimes V_2 \otimes \cdots \otimes V_k. \end{gathered}}$$

PROPOSITION 2.18. *A finite-dimensional vector space is naturally*[13] *isomorphic to its double dual. That is, $V \cong V^{**}$ for any finite-dimensional vector space V.*

---

[13] The words "natural" and "canonical" used in mathematical statements (when not directly referring to category theory) convey a sense of uniqueness in that the statements are independent of any special choices or "coordinates". For example, consider the statement that a given n-dimensional real vector space is isomorphic to $\mathbb{R}^n$. The isomorphism here is in general *not* natural because it usually depends on the choice of basis for the vector space; if it didn't depend on any choices, then it would be natural. Naturally isomorphic vector spaces are regarded "equal" and we usually don't distinguish the two; for example we do not distinguish between $\mathbb{R}^3$ and $(\mathbb{R} \times \mathbb{R}) \times \mathbb{R}$.

Proof. For a vector space V, consider the map

$$V \ni v \mapsto \tilde{v} \in V^{**}, \tag{2.49}$$

where

$$\tilde{v} : V^* \to \mathbb{K} \quad \text{is the map} \quad \tilde{v}(\alpha) := \alpha(v) \quad \text{for all } \alpha \in V^*. \tag{2.50}$$

Note that $\tilde{v} : V^* \to \mathbb{K}$ is linear, so $\tilde{v} \in V^{**}$. Also note that F is linear because if $a \in \mathbb{K}$ and $v, w \in V$, then for any $\alpha \in V^*$,

$$\widetilde{(a\,v + w)}(\alpha) := \alpha(a\,v + w) = a\,\alpha(v) + \alpha(w) = a\,\tilde{v}(\alpha) + \tilde{w}(\alpha),$$

which shows that $\widetilde{(a\,v + w)} = a\,\tilde{v} + \tilde{w}$. We shall prove that (2.49) is an isomorphism.

First of all, by Proposition 2.14 we know that the dimension of the dual space of any finite-dimensional vector space is the same as the original vector space; hence,

$$\dim V^{**} = \dim V^* = \dim V.$$

Thus, we just have to prove injectivity of (2.49), then we get surjectivity for free (by the dimension or rank theorem from linear algebra). To prove injectivity, assume that $\tilde{v} = 0$, which means

$$\tilde{v}(\alpha) := \alpha(v) = 0 \quad \text{for all } \alpha \in V^*;$$

we shall prove that $v = 0$ too. The easiest (and perhaps only) way to prove $v = 0$ is through some type of basis argument. Thus, let $\{v_j\}_{j=1}^n$ be a basis of V. Let $\{v_i^*\}$ denote the dual basis defined by $v_i^*(v_j) = \delta_{ij}$. Now write v in the basis $\{v_j\}$,

$$v = a_1\,v_1 + a_2\,v_2 + \cdots + a_n\,v_n.$$

Take any $j = 1, \ldots, n$ and observe that since $\tilde{v} = 0$,

$$0 = \tilde{v}(v_j^*) = v_j^*(v) = a_j.$$

Since j was arbitrary, $v = 0$ and our proof is complete. ☐

We shall not distinguish between V and V^{**} so given $v \in V$ we shall denote the element $\tilde{v} \in V^{**}$ simply as v. In other words, we identify a vector $v \in V$ with the map

$$\boxed{\text{“}v : V^* \to \mathbb{K}\text{” defined by } \; v(\alpha) := \alpha(v) \text{ for all } \alpha \in V^*.} \tag{2.51}$$

More generally, if two vector spaces V and W are canonically isomorphic, we shall use the same notation for an element $v \in V$ with its corresponding element in W. This convention rarely causes confusion and it simplifies life considerably.

2.5.2. Dimension and the universal property. We now compute the dimension of the tensor product of vector spaces.

Theorem 2.19. *Let $V_1, V_2, \ldots, V_k$ be finite-dimensional vector spaces. Then the tensor product $V_1 \otimes V_2 \otimes \cdots \otimes V_k$ is finite-dimensional with*

$$\dim(V_1 \otimes V_2 \otimes \cdots \otimes V_k) = \dim V_1 \cdot \dim V_2 \cdots \dim V_k.$$

In fact, taking any bases for $V_1, V_2, \ldots, V_k$, let us say $\{u_i\}_{i=1}^{n_1}$ is a basis for V_1, $\{v_j\}_{j=1}^{n_2}$ is a basis for V_2, … etc. …, $\{w_\ell\}_{\ell=1}^{n_k}$ is a basis for V_k, then the vectors

$$\{\, u_i \otimes v_j \otimes \cdots \otimes w_\ell \mid i = 1, \ldots, n_1, \; j = 1, \ldots, n_2, \; \ldots, \; \ell = 1, \ldots, n_k\}$$

form a basis for $V_1 \otimes V_2 \otimes \cdots \otimes V_k$.

Proof. In other words, a basis for $V_1 \otimes \cdots \otimes V_k$ is obtained by taking all possible tensor products of the basis vectors for $V_1, \ldots, V_k$. Note that there are exactly $n_1 \cdot n_2 \cdots n_k = \dim V_1 \cdot \dim V_2 \cdots \dim V_k$ elements in the set $\{ u_i \otimes v_j \otimes \cdots \otimes w_\ell \}$, therefore our theorem is proved once we prove that this set is a basis.

To simplify notation we shall prove this theorem for $k = 2$. The proof for general k is not harder only notationally ugly so we focus on the $k = 2$ case for aesthetic reasons. In this case, let $\{v_i\}$ and $\{w_j\}$ be bases for finite-dimensional vector spaces V and W. We shall prove that $\{v_i \otimes w_j\}$ is a basis for $V \otimes W$; this implies in particular that $\dim(V \otimes W) = \dim V \cdot \dim W$.

To prove that $\{v_i \otimes w_j\}$ is linearly independent, assume that

$$\sum a_{ij}\, v_i \otimes w_j = 0 \quad \text{in} \quad V \otimes W. \tag{2.52}$$

This means, by definition of $V \otimes W$, that

$$\sum a_{ij}\, v_i \otimes w_j : V^* \times W^* \to \mathbb{K}$$

is the zero map. Let $\{v_i^*\}$ and $\{w_j^*\}$ be the dual bases of $\{v_i\}$ and $\{w_j\}$, respectively. Then fixing k, ℓ and applying both sides of the equality (2.52) to $(v_k^*, w_\ell^*) \in V^* \times W^*$, we obtain

$$0 = \sum a_{ij}\, (v_i \otimes w_j)(v_k^*, w_\ell^*) = \sum a_{ij}\, v_i(v_k^*)\, w_j(w_\ell^*) = \sum a_{ij}\, \delta_{ik}\delta_{j\ell} = a_{k\ell}.$$

Thus, all the a_{ij}'s in (2.52) are zero and hence $\{v_i \otimes w_j\}$ is linearly independent.

To prove that $\{v_i \otimes w_j\}$ spans $V \otimes W$, let $f \in V \otimes W$, which means that $f : V^* \times W^* \to \mathbb{K}$ is bilinear. We claim that

$$f = \sum_{i,j} f(v_i^*, w_j^*)\, v_i \otimes w_j.$$

To see this, define $g = \sum_{i,j} f(v_i^*, w_j^*)\, v_i \otimes w_j$ and observe that

$$g(v_k^*, w_\ell^*) = \sum_{i,j} f(v_i^*, w_j^*)\, v_i(v_k^*)\, w_j(w_\ell^*) = \sum_{i,j} f(v_i^*, w_j^*)\, \delta_{ik}\, \delta_{j\ell} = f(v_k^*, w_\ell^*).$$

Therefore, f and g have the same values on the pairs of basis vectors $\{(v_i^*, w_j^*)\}$. By (multi-)linearity, f and g must have the same values on all vectors. Thus, $f = g$ and our proof is complete. □

We ended our proof using a fact akin to (2.21):

$$\boxed{\text{A multi-linear map is completely determined by its values on a basis.}} \tag{2.53}$$

The tensor product has a useful "universal property" that will be used several times in the sequel; see Problem 3 for more on this property.

Theorem 2.20 (The universal property). *Let $V_1, \ldots, V_k$ be finite-dimensional $\mathbb{K}$-vector spaces and let W be a (not necessarily finite-dimensional) $\mathbb{K}$-vector space. Then given an arbitrary multi-linear map*

$$f : V_1 \times \cdots \times V_k \to W,$$

there exists a unique linear map

$$\tilde{f} : V_1 \otimes \cdots \otimes V_k \to W$$

such that $\tilde{f}(v_1 \otimes \cdots \otimes v_k) = f(v_1, \ldots, v_k)$ for all $(v_1, \ldots, v_k) \in V_1 \times \cdots \times V_k$.

PROOF. To simplify notation we prove this theorem for $k = 2$. The proof for general k is only notationwise harder but not conceptually. Thus, let

$$f : U \times V \to W$$

be a bilinear map. We need to show there exists a unique linear map

$$\tilde{f} : U \otimes V \to W$$

such that $\tilde{f}(u \otimes v) = f(u, v)$ for all $(u, v) \in U \times V$. The easiest way to define $\tilde{f}$ is through bases, although by the uniqueness part of the theorem, the function $\tilde{f}$ is actually completely basis independent. So, let $\{u_i\}$ and $\{v_j\}$ be bases for U and V, respectively. Then we *define*

$$\tilde{f}(u_i \otimes v_j) := f(u_i, v_j)$$

on the basis $\{u_i \otimes v_j\}$ of $U \otimes V$ and extend $\tilde{f}$ linearly to all of $U \otimes V$; here we use the basic fact that a linear map is determined by its values on a basis. If you'd rather get your hands dirty and want an explicit formula for $\tilde{f}$, here it is: Any element $\xi \in U \otimes V$ can be written uniquely as

$$\xi = \sum a_{ij}\, u_i \otimes v_j \in U \otimes V$$

for some constants $a_{ij} \in \mathbb{K}$. We *define*

$$\tilde{f}(\xi) := \sum a_{ij}\, f(u_i, v_j) \in W. \tag{2.54}$$

Let $(u, v) \in U \times V$; we shall prove that $\tilde{f}(u \otimes v) = f(u, v)$. To this end, write $u = \sum a_i u_i$ and $v = \sum b_j v_j$ and use the multi-linearity of the tensor product (see (2.48)) to get

$$u \otimes v = \sum a_i\, b_j\, u_i \otimes v_j.$$

Hence,

$$\tilde{f}(u \otimes v) := \sum a_i\, b_j\, f(u_i, v_j).$$

On the other hand, by multi-linearity of f, we also have

$$f(u, v) = f\left(\sum a_i u_i, \sum b_j v_j\right) = \sum a_i\, b_j\, f(u_i, v_j).$$

Thus, $\tilde{f}(u \otimes v) = f(u, v)$. Finally, for uniqueness, suppose that there were a linear map $g : U \otimes V \to W$ such that $g(u \otimes v) = f(u, v)$ for all $(u, v) \in U \times V$. Then for $\xi = \sum a_{ij}\, u_i \otimes v_j$, we have

$$g(\xi) = \sum a_{ij}\, g(u_i \otimes v_j) = \sum a_{ij}\, f(u_i, v_j),$$

which of course is just $\tilde{f}(\xi)$. ☐

2.5.3. Properties of the tensor product. If we think of the tensor product as "multiplication" of vector spaces, here are some algebraic properties for the "multiplication" of vector spaces.

THEOREM 2.21. *For any finite-dimensional vector spaces U, V, W, we have the following natural isomorphisms:*

(i) $V \otimes W \cong W \otimes V$.
(ii) $(U \otimes V) \otimes W \cong U \otimes (V \otimes W) \cong U \otimes V \otimes W$.
(iii) $(V \otimes W)^* \cong V^* \otimes W^*$.

In other words, tensoring is commutative, associative, and commutes with taking duals. Similar statements hold for any finite number of vector spaces.

Proof. We'll leave *(i)* to you (it's the easiest of the three) and prove *(ii)* and *(iii)*. To prove *(iii)*, we shall define an isomorphism from $V^* \otimes W^* \to (V \otimes W)^*$. Let $f \in V^* \otimes W^*$, which means that

$$f : V \times W \to \mathbb{K} \quad \text{is multi-linear.}$$

By the universal property of the tensor product, there exists a unique linear map

$$\tilde{f} : V \otimes W \to \mathbb{K} \quad \text{such that } \tilde{f}(v \otimes w) = f(v, w) \text{ for all } v \in V,\ w \in W.$$

In particular, since $\tilde{f} : V \otimes W \to \mathbb{K}$ is linear, we know that $\tilde{f} \in (V \otimes W)^*$. Hence, we get a map

$$V^* \otimes W^* \ni f \mapsto \tilde{f} \in (V \otimes W)^*. \tag{2.55}$$

It can be checked that this map is linear; we shall prove that this map is an isomorphism. Since

$$\begin{aligned}\dim(V \otimes W)^* = \dim V \otimes W &= \dim V \cdot \dim W \\ &= \dim V^* \otimes \dim W^* = \dim(V^* \otimes W^*),\end{aligned}$$

we just have to show that the map (2.55) is injective (or surjective). To prove injectivity, assume that $\tilde{f} = 0$. This implies, in particular, that

$$\tilde{f}(v \otimes w) = f(v, w) = 0 \quad \text{for all } v \in V,\ w \in W.$$

Of course, this implies that $f = 0$, so the map (2.55) is an isomorphism. We can also prove that the map (2.55) is an isomorphism by proving that it takes basis vectors to basis vectors; we shall use this method of proof next.

To prove *(ii)*, we'll just prove that $(U \otimes V) \otimes W \cong U \otimes V \otimes W$. Consider the map

$$f : U \times V \times W \to (U \otimes V) \otimes W \quad \text{defined by} \quad f(u, v, w) := (u \otimes v) \otimes w$$

for all $u \in U$, $v \in V$, and $w \in W$. This map is certainly multi-linear, so by the universal property there exists a unique linear map

$$\tilde{f} : U \otimes V \otimes W \to (U \otimes V) \otimes W \quad \text{such that} \quad \tilde{f}(u \otimes v \otimes w) = (u \otimes v) \otimes w.$$

Note that if $\{u_i\}$, $\{v_j\}$, and $\{w_k\}$ are bases for U, V, W, respectively, then by Theorem 2.19 we know that $\{u_i \otimes v_j \otimes w_k\}$ and $\{(u_i \otimes v_j) \otimes w_j\}$ are bases for $U \otimes V \otimes W$ and $(U \otimes V) \otimes W$, respectively. Clearly $\tilde{f}$ takes basis vectors to basis vectors and hence is an isomorphism. □

We now generalize the tensor product of one-forms in (2.47) to arbitrary tensors. Let $f \in V_1 \otimes \cdots \otimes V_k$ and $g \in W_1 \otimes \cdots \otimes W_\ell$. Then we define the **tensor product** of f and g as the element $f \otimes g \in V_1 \otimes \cdots \otimes V_k \otimes W_1 \otimes \cdots \otimes W_\ell$ defined by

$$(f \otimes g)(\alpha, \beta) := f(\alpha) \cdot g(\beta) \quad \text{for all} \ \ \alpha \in V_1^* \times \cdots \times V_k^*\ ,\ \beta \in W_1^* \times \cdots \times W_k^*.$$

This definition is a generalization of the definition (2.47). Notice that since f and g are multi-linear, $f \otimes g$ is also. Thus,

$$\begin{aligned} &f \in V_1 \otimes \cdots \otimes V_k\ ,\ g \in W_1 \otimes \cdots \otimes W_\ell \quad \Longrightarrow \\ &\qquad\qquad f \otimes g \in V_1 \otimes \cdots \otimes V_k \otimes W_1 \otimes \cdots \otimes W_\ell. \end{aligned} \tag{2.56}$$

We can, in the obvious way, generalize this process to define the tensor product of any number of tensors not just two.

If $V = V_1 = V_2 = \cdots = V_k$, then we write

$$V^{\otimes k} := \underbrace{V \otimes V \otimes \cdots \otimes V}_{k\ V\text{'s}}.$$

An element of $V^{\otimes k}$ is called a **k-tensor** and k is the **degree** of the tensor. Then the map (2.56) gives rise to a product map

$$\otimes : V^{\otimes k} \times V^{\otimes \ell} \to V^{\otimes (k+\ell)}.$$

Tensoring tensors has the following properties.

PROPOSITION 2.22. *For tensors* $f \in U_1 \otimes \cdots \otimes U_j$, $g \in V_1 \otimes \cdots \otimes V_k$, $h \in W_1 \otimes \cdots \otimes W_\ell$, *we have*

(i) $(f \otimes g) \otimes h = f \otimes (g \otimes h) = f \otimes g \otimes h$.
(ii) $(f + g) \otimes h = f \otimes h + g \otimes h$ *(we assume that* $j = k$ *and each* $U_i = V_i$*) .*
(iii) It is NOT true in general that $f \otimes g = g \otimes f$.

Thus, tensoring tensors is associative and distributive but not commutative. Similar statements hold for the tensor product of any finite number of tensors.

PROOF. We'll leave *(i)* and *(ii)* to you (just use the definition (2.56) and for *(i)* you need the analogous definition for the tensor product of three tensors and you'll also need Property *(ii)* of Theorem 2.21). It's easy to find simple examples for *(iii)*. For instance, let V be a finite-dimensional vector space with $\dim V > 1$ and let $\{v_i\}$ be a basis for V. Then as elements of $V \otimes V$, we claim that

$$v_1 \otimes v_2 \neq v_2 \otimes v_1.$$

Indeed, if $\{v_i^*\}$ denotes the dual basis, then, for example,

$$(v_1 \otimes v_2)(v_1^*, v_2^*) = v_1(v_1^*) \cdot v_2(v_2^*) = 1 \cdot 1 = 1,$$

while

$$(v_2 \otimes v_1)(v_1^*, v_2^*) = v_2(v_1^*) \cdot v_1(v_2^*) = 0 \cdot 0 = 0.$$

□

2.5.4. Direct sum of vector spaces. So far we have defined the tensor product of vector spaces, looked at bases for this space, studied the universal property, and defined the tensor product of tensors and studied some its algebraic properties. We now study the direct sum of vector spaces.

Let V and W be vector spaces. In order to do "algebra" on vector spaces we need a "multiplication" and an "addition". The "multiplication" shall be tensoring: $V \otimes W$, some properties of which are in Theorem 2.21. For "addition" we define the direct sum of V and W. The **direct sum** of V and W, denoted by $V \oplus W$, is the Cartesian product $V \times W$ with addition and scalar multiplication defined as follows: If $(v, w) \in V \times W$, $(v', w') \in V \times W$, and $a \in \mathbb{K}$, then

$$(v, w) + (v', w') := (v + v', w + w') \ , \quad a(v, w) := (av, aw).$$

The zero vector in $V \oplus W$ is $(0, 0)$. Notice that we can consider V as a subspace of $V \oplus W$, namely as the set of vectors of the form $(v, 0)$ where $v \in V$. Similarly, we can consider W as the subspace of $V \oplus W$ consisting of vectors of the form $(0, w)$ where $w \in W$. Unless it might cause confusion, we usually write "v" for $(v, 0)$ where $v \in V$ and "w" for $(0, w)$ where $w \in W$.

We remark that one can also define the direct sum of any finite number of vector spaces by taking the Cartesian product of the vector spaces and defining addition and scalar multiplication in a similar manner as we did above.

Proposition 2.23. *For finite-dimensional vector spaces V, W, the direct sum $V \oplus W$ is finite dimensional with*

$$\dim(V \oplus W) = \dim V + \dim W.$$

In fact, if $\{v_i\}$ and $\{w_j\}$ are bases for V and W, respectively, then $\{v_i, w_j\}$ is a basis for $V \oplus W$, where v_i is really $(v_i, 0)$ and w_j is really $(0, w_j)$. A similar statement holds for the direct sum of any finite number of vector spaces.

Proof. To prove independence, assume that

$$\sum a_i(v_i, 0) + \sum b_j(0, w_j) = 0 \;\;\big(\text{where } 0 = (0,0)\big).$$

Using the definition of addition and scalar multiplication on $V \oplus W$, we can rewrite this as

$$\Big(\sum a_i v_i \;,\; \sum b_j w_j\Big) = (0,0).$$

Hence, $\sum a_i v_i = 0$ and $\sum b_j w_j = 0$, and therefore, since $\{v_i\}$ and $\{w_j\}$ are bases for V and W, we must have $a_i = 0$ and $b_j = 0$ for all i, j. To prove that $\{v_i, w_j\}$ span $V \oplus W$, let $(v, w) \in V \oplus W$. Then $v \in V$ and $w \in W$ so we can write $v = \sum a_i v_i$ and $w = \sum b_j w_j$ for some $a_i, b_j \in \mathbb{K}$. Using the definition of addition and scalar multiplication on $V \oplus W$, it follows that

$$(v, w) = \sum a_i(v_i, 0) + \sum b_j(0, w_j),$$

and our proof is complete. ☐

If we think of direct sum as "addition" of vector spaces, then the following theorem describes some "algebraic" properties of the direct sum.

Theorem 2.24. *For finite-dimensional vector spaces U, V, W, we have the following natural isomorphisms:*

(i) $V \oplus W \cong W \oplus V$.
(ii) $(U \oplus V) \oplus W \cong U \oplus (V \oplus W) \cong U \oplus V \oplus W$.
(iii) $(U \oplus V) \otimes W \cong (U \otimes W) \oplus (V \otimes W)$.
(iv) $(V \oplus W)^* \cong V^* \oplus W^*$.

In other words, direct product is commutative, associative, tensoring is distributive over direct sum, and direct sum commutes with taking duals. Similar statements hold for any finite number of vector spaces.

Proof. It is instructive to prove many of these properties yourself so we'll only prove the last two.

To prove *(iii)*, consider the map

$$f : (U \oplus V) \times W \to (U \otimes W) \oplus (V \otimes W)$$

defined by

$$f((u, v), w) := \big((u \otimes w)\,,\,(v \otimes w)\big) \quad \text{for all } u \in U,\ v \in V, \text{ and } w \in W.$$

This map is easily checked to be multi-linear, so by the universal property there exists a unique linear map

$$\tilde{f} : (U \oplus V) \otimes W \to (U \otimes W) \oplus (V \otimes W)$$

such that

$$\tilde{f}((u,v)\otimes w) = \big((u\otimes w)\,,\,(v\otimes w)\big) \quad \text{for all } u\in U,\ v\in V, \text{ and } w\in W.$$

If $\{u_i\}$, $\{v_j\}$, and $\{w_k\}$ are bases for U, V, and W, respectively, then by Theorem 2.19 and Proposition 2.23, $\{(u_i,0)\otimes w_k,(0,v_j)\otimes w_k\}$ is a basis for $(U\oplus V)\otimes W$ and $\{(u_i\otimes w_k,0),(0,v_j\otimes w_k)\}$ is a basis for $(U\otimes W)\oplus(V\otimes W)$. Notice that $\tilde{f}$ takes basis vectors to basis vectors and hence is an isomorphism.

Now let's prove *(iv)*. Define a map $V^*\oplus W^*\to(V\oplus W)^*$ as follows. Let $(\alpha,\beta)\in V^*\oplus W^*$ so that $\alpha:V\to\mathbb{K}$ and $\beta:W\to\mathbb{K}$ are linear. Then define (using the same notation) $(\alpha,\beta):V\oplus W\to\mathbb{K}$ by

$$(\alpha,\beta)(v,w) := \alpha(v)+\beta(w) \quad \text{for all } \ (v,w)\in V\oplus W.$$

Since α and β are linear, the map $(\alpha,\beta):V\oplus W\to\mathbb{K}$ is also linear, so $(\alpha,\beta)\in(V\oplus W)^*$. We shall prove that the map $V^*\oplus W^*\ni(\alpha,\beta)\mapsto(\alpha,\beta)\in(V\oplus W)^*$ is an isomorphism. One can prove that this map takes basis vectors to basis vectors, but we're tired of such proofs, so here's a direct proof of bijectivity.

To prove injectivity, assume that $(\alpha,\beta)=0$ in $(V\oplus W)^*$, which means that

$$(\alpha,\beta)(v,w) := \alpha(v)+\beta(w)=0 \quad \text{for all } \ (v,w)\in V\oplus W.$$

Taking $w=0$, we conclude that $\alpha(v)=0$ for all $v\in V$, which means that $\alpha=0$. Then, we must have $\beta(w)=0$ for all $w\in W$, which means that $\beta=0$. Hence, as an element of $V^*\oplus W^*$, we have $(\alpha,\beta)=(0,0)=0$.

To prove surjectivity, let $f\in(V\oplus W)^*$, that is, $f:V\oplus W\to\mathbb{K}$ is linear. Define $\alpha:V\to\mathbb{K}$ and $\beta:W\to\mathbb{K}$ by

$$\alpha(v):=f(v,0) \ \text{ for all } v\in V \quad , \quad \beta(w):=f(0,w) \ \text{ for all } w\in W.$$

Since f is linear it follows that α and β are both linear, so $(\alpha,\beta)\in V^*\oplus W^*$. Now we claim that as an element of $(V\oplus W)^*$, we have $(\alpha,\beta)=f$. To see this, let $(v,w)\in V\oplus W$ be arbitrary. Then,

$$(\alpha,\beta)(v,w) := \alpha(v)+\beta(w) = f(v,0)+f(0,w) = f((v,0)+(0,w)) = f(v,w).$$

Hence, $(\alpha,\beta)=f$ and we are done. ☐

2.5.5. Homomorphisms and tensor products. For finite-dimensional vector spaces V and W, we define

$$\hom(V,W) := \big\{ \text{linear maps } f:V\to W \big\},$$

the set of homomorphisms or linear maps from V to W. If $V=W$, we denote $\hom(V,V)$ by $\hom(V)$. You've studied $\hom(V,W)$ in linear algebra class so you already know various facts concerning this space. For example, if $n=\dim V$ and $m=\dim W$, then you know that $\hom(V,W)$ is isomorphic to $\mathbb{K}^{m\times n}$, the space of $m\times n$ matrices with entries in $\mathbb{K}$. Such an isomorphism is obtained as follows. Let $\{v_j\}_{j=1}^n$ and $\{w_i\}_{i=1}^m$ be bases for V and W, respectively. If $f\in\hom(V,W)$, then we can write

$$f(v_j)=\sum_{i=1}^m a_{ij}\,w_i\ , \quad j=1,\dots,n, \tag{2.57}$$

for unique constants a_{ij}. Then an isomorphism is $\hom(V,W)\ni f\mapsto[a_{ij}]\in\mathbb{K}^{m\times n}$. In particular,

$$\dim\big(\hom(V,W)\big) = n\cdot m = \dim V\cdot\dim W.$$

This isomorphism is not canonical because it depends on the choice of bases. However, it turns out that there is a very beautiful and useful canonical isomorphism between $\hom(V, W)$ and tensor produces.

THEOREM 2.25. *For any finite-dimensional vector spaces V, W, there are natural isomorphisms*

$$\boxed{\hom(V, W) \cong W \otimes V^* \cong V^* \otimes W.}$$

PROOF. Since $W \otimes V^* \cong V^* \otimes W$ we just have to prove that $\hom(V, W) \cong V^* \otimes W$. Define a map $T : V^* \times W \to \hom(V, W)$ as follows: If $(\alpha, w) \in V^* \times W$, then $T(\alpha, w) : V \to W$ is the map

$$T(\alpha, w)(v) := \alpha(v) \cdot w \quad \text{for all} \quad v \in V.$$

The map T is easily checked to be bilinear, therefore by the universal property of the tensor product, there exists a unique linear map $\tilde{T} : V^* \otimes W \to \hom(V, W)$ such that for $\alpha \in V^*$ and $w \in W$,

$$\tilde{T}(\alpha \otimes w)(v) = \alpha(v)\, w \text{ for all } v \in V. \tag{2.58}$$

We shall prove that $\tilde{T}$ is an isomorphism. Since $\dim \hom(V, W) = \dim V \cdot \dim W = \dim(V^* \otimes W)$, we just have to prove that $\tilde{T}$ is injective or surjective. Let's do surjectivity, so let $f \in \hom(V, W)$. With respect to bases $\{v_j\}$ and $\{w_i\}$ of V and W, respectively, let $[a_{ij}]$ be the matrix of f defined by (2.57). Then define

$$\xi := \sum a_{ij}\, v_j^* \otimes w_i \in V^* \otimes W,$$

where $\{v_j^*\}$ is the dual basis to $\{v_j\}$. We claim that $\tilde{T}(\xi) = f$. To see this, observe that by (2.58), for any k,

$$\tilde{T}(\xi) v_k = \sum a_{ij}\, v_j^*(v_k)\, w_i = \sum a_{ij}\, \delta_{jk}\, w_i = \sum_{i=1}^{m} a_{ij}\, w_i.$$

In view of (2.57), we have $\tilde{T}(\xi) = f$. ☐

As a consequence of the proof of this theorem (see (2.58)) we know how the isomorphism works: For $\alpha \otimes w \in V^* \otimes W$, omitting henceforth the " $\tilde{T}$ " we consider $\alpha \otimes w$ as the element of $\hom(V, W)$ defined by

$$(\alpha \otimes w)(v) = \alpha(v)\, w \quad \text{for all } v \in V.$$

Similarly, the isomorphism between $W \otimes V^*$ works as follows: For $w \otimes \alpha \in W \otimes V^*$ we consider $w \otimes \alpha \in \hom(V, W)$ as the map

$$(w \otimes \alpha)(v) = \alpha(v)\, w \quad \text{for all } v \in V.$$

EXERCISES 2.5.

1. Here are various (unrelated) exercises related to tensor products.
    - (i) Prove some of the statements we didn't prove in Theorems 2.21 and 2.24.
    - (ii) We usually consider "1" as a canonical basis of $\mathbb{K}$ ($= \mathbb{R}$ or $\mathbb{C}$). Prove that if V is a finite-dimensional $\mathbb{K}$-vector space, then $\mathbb{K} \otimes V$ and V are canonically isomorphic. Thus, $\mathbb{K}$ acts like an "multiplicative identity".
    - (iii) Let $f \in \hom(V, W)$, which we identify with $V^* \otimes W$, and let $\{v_j\}$ and $\{w_i\}$ be bases for V and W, respectively. Prove that $[a_{ij}]$ is the matrix of f with respect to these bases as defined in (2.57) if and only if

$$f = \sum a_{ij}\, v_j^* \otimes w_i \ , \quad \text{where} \ \ a_{ij} = w_i^*(f(v_j)).$$

(iv) Prove that if V is one-dimensional, then every linear map $f : V \to V$ is given by multiplication by an element of $\mathbb{K}$; that is, $f \in \hom(V)$ if and only if

$$f(v) = a\,v$$

for some $a \in \mathbb{K}$. In particular, $\hom(V)$ is canonically isomorphic to $\mathbb{K}$.

2. In this problem we look at a fancy and common definition of $V_1 \otimes \cdots \otimes V_k$. We first recall how to make any set into a vector space. Let X be a set and $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Then the **free vector space** over $\mathbb{K}$ is the set defined as

$$\mathrm{Fr}(X) := \{\, f : X \to \mathbb{K} \,|\, f(x) \neq 0 \;\text{ for only finitely many } x \in X \,\}.$$

$\mathrm{Fr}(X)$ is a vector space over $\mathbb{K}$ with the operations defined as usual: If $f, g \in \mathrm{Fr}(X)$ and $a \in \mathbb{K}$, then $f + g$ and af are the maps $x \mapsto f(x) + g(x)$ and $x \mapsto a\,f(x)$, both of which are easily seen to be in $\mathrm{Fr}(X)$.

(i) For each $\alpha \in X$ define (using the same notation) the function $\alpha : X \to \mathbb{K}$ by $\alpha(x) = 1$ if $x = \alpha$ otherwise $\alpha(x) = 0$. Prove that $f \in \mathrm{Fr}(X)$ if and only if we can write

$$f = a_1\,\alpha_1 + \cdots + a_n\,\alpha_n, \quad \text{for unique elements } \alpha_i \in X\,,\; a_i \in \mathbb{K}. \tag{2.59}$$

Thus, $\mathrm{Fr}(X)$ can be thought of (and is often defined) as the set of all "formal sums of elements of X."

(ii) Let $V_1, \ldots, V_k$ be finite-dimensional $\mathbb{K}$-vector spaces and consider the free vector space $\mathrm{Fr}(V_1 \times \cdots \times V_k)$ where we consider $V_1 \times \cdots \times V_k$ as just a set with no vector space structure. Let I be the subspace spanned by all elements of the form

$$(v_1, \ldots, a\,v_j + b\,v_j', \ldots, v_k) - a(v_1, \ldots, v_j, \ldots, v_k) - b(v_1, \ldots, v_j', \ldots, v_k),$$

where we are using the "formal sum notation" (2.59). Only the j-th slots are different in the three terms. Define

$$F : \mathrm{Fr}(V_1 \times \cdots \times V_k) \to V_1 \otimes \cdots \otimes V_k$$

as follows: If $u = (v_1, \ldots, v_k) \in V_1 \times \cdots \times V_k$, we define

$$F(u) := v_1 \otimes \cdots \otimes v_k,$$

and if $u = a_1\,u_1 + \cdots + a_n\,u_n \in \mathrm{Fr}(V_1 \times \cdots \times V_k)$ is written (uniquely) as in (2.59) where each $u_i \in V_1 \times \cdots \times V_k$ and $a_i \in \mathbb{K}$, then we define

$$F(u) := a_1\,F(u_1) + \cdots + a_n\,F(u_n),$$

where each $F(u_n)$ has already been defined. Prove that $F(u) = 0$ for all $u \in I$. In particular, F induces a map (using the same notation F) on the quotient:

$$F : \mathrm{Fr}(V_1 \times \cdots \times V_k)/I \to V_1 \otimes \cdots \otimes V_k.$$

(iii) Prove that F defines an isomorphism:

$$\mathrm{Fr}(V_1 \times \cdots \times V_k)/I \cong V_1 \otimes \cdots \otimes V_k$$

such that $\mathrm{Fr}(V_1 \times \cdots \times V_k)/I \ni [(v_1, \ldots, v_k)] \mapsto v_1 \otimes \cdots \otimes v_k \in V_1 \otimes \cdots \otimes V_k$.

Note 1: $V_1 \otimes \cdots \otimes V_k$ is usually *defined* as $\mathrm{Fr}(V_1 \times \cdots \times V_k)/I$ and $v_1 \otimes \cdots \otimes v_k$ is usually *defined* as the equivalence class $[(v_1, \ldots, v_k)] \in \mathrm{Fr}(V_1 \times \cdots \times V_k)/I$.

Note 2: The space $\mathrm{Fr}(V_1 \times \cdots \times V_k)/I$ is well-defined even when $V_1, \ldots, V_k$ are infinite-dimensional. This gives a way to define $V_1 \otimes \cdots \otimes V_k$ for vector spaces that are not necessarily finite-dimensional.

3. (**Universal mapping property**) Let $V_1, \ldots, V_k$ be vector spaces, not necessarily finite-dimensional. Let V be a vector space and let $\varphi : V_1 \times \cdots \times V_k \to V$ be a multilinear map. Suppose that (V, φ) has the following property: If $f : V_1 \times \cdots \times V_k \to W$

is a multi-linear map into some vector space W, then there exists a unique *linear* map $\tilde{f} : V \to W$ such that the following diagram commutes:

$$\begin{tikzcd} V_1 \times \cdots \times V_k \arrow[r, "\varphi"] \arrow[dr, "f"'] & V \arrow[d, "\tilde{f}"] \\ & W \end{tikzcd} , \qquad \text{that is,} \quad f = \tilde{f} \circ \varphi.$$

The function f is said to **factor through** φ. Thus, we can see that (V, φ) produces a linear map $\tilde{f}$ given a multi-linear map f. The pair (V, φ) is said to have the **universal mapping property** or is **universal** for multi-linear maps on $V_1 \times \cdots \times V_k$; “universal” because it linearizes arbitrary multi-linear maps on $V_1 \times \cdots \times V_k$.

(i) Prove that if (V, φ) and (V', φ') both have the universal mapping property for multi-linear maps on $V_1 \times \cdots \times V_k$, then there is an isomorphism $\psi : V \to V'$ such that $\varphi' = \psi \circ \varphi$. The argument that you use is a prime example of “generalized abstract nonsense”.

(ii) Let $\varphi : V_1 \times \cdots \times V_k \to \mathrm{Fr}(V_1 \times \cdots \times V_k)/I$ be the quotient map and define $V_1 \otimes \cdots \otimes V_k := \mathrm{Fr}(V_1 \times \cdots \times V_k)/I$. Prove that $(V_1 \otimes \cdots \otimes V_k, \varphi)$ has the universal mapping property for multi-linear maps on $V_1 \times \cdots \times V_k$.

In this sense, the tensor product $V_1 \otimes \cdots \otimes V_k$ “solves” the **universal mapping problem** for multi-linear maps on $V_1 \times \cdots \times V_k$.